This is to certify that the
dissertation entitled

MAKING SENSE OF THE LINKS: FROM GOVERNMENT
POLICY TO STUDENT ACHIEVEMENT

presented by

MARJORIE CAROL RHODES WALLACE

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Curriculum, Teaching, and Education Policy

Major Professor's Signature

11/4/05

Date

*MSU is an Affirmative Action/Equal Opportunity Institution*

# MAKING SENSE OF THE LINKS:
# FROM GOVERNMENT POLICY TO STUDENT ACHIEVEMENT

By

Marjorie Carol Rhodes Wallace

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Teacher Education

2005

**ABSTRACT**

MAKING SENSE OF THE LINKS:
FROM GOVERNMENT POLICY TO STUDENT ACHIEVEMENT

By

Marjorie Carol Rhodes Wallace

Using the Beginning Teacher Preparation Study (BTPS) in Connecticut and Tennessee and the National Assessment of Educational Progress in Mathematics and Reading 1999-2000 (NAEP), this study provides evidence on the strength of relationships and relative contributions of four important elements in the education system to student achievement in mathematics and reading: teacher preparation, teacher characteristics, professional development, and teacher practice. Results from structural equation modeling (SEM) indicate that teacher practices, teacher characteristics, and professional development across teachers are generally linked to teacher practice and student achievement in consistent ways. They have moderate to large direct effects on teacher practice and small but significant indirect effects on student achievement in mathematics and reading ($\delta$ = .05). Findings vary based on curriculum and policy context.

Hierarchical linear modeling (HLM) results indicate that teacher characteristics, professional development, and teacher practices can differentially affect student achievement by ethnicity and gender. Differential effects on gain scores (estimated at the state level) portray effects on learning while effects on achievement status confound teacher effects with prior differences in achievement at the national level. Black mean achievement varies across teachers in mathematics (Connecticut, NAEP 1996) and

reading (Connecticut, NAEP 1998). The mean for Black (Connecticut reading) varies across teachers but teacher variables do not account for variation. Female mean mathematics achievement varies across teachers in two data sets. Hispanic mean mathematics (NAEP 2000) and Black female mean mathematics achievement (NAEP 1996) vary across teachers in one data set each. Teacher variables exhibiting differential effects (increasing achievement for one group while simultaneously decreasing achievement for another) are teacher drill and lecture activities, sensemaking activities, knowledge, preparation, reform orientation, and professional development for mathematics. For reading, the list includes traditional activities, computer activities, reading preparation, assessment preparation, and professional development. The differential effects sometimes split by ethnicity and sometimes by gender.

In the three-level HLM for Connecticut, effects of teacher preparation programs do not vary significantly across teachers when student gains are the outcome. Using SEM, teacher preparation programs as a whole have effects on student achievement but those effects cannot be attributed to individual teacher preparation programs using HLM.

To my family and friends, especially my husband, Charles E. Wallace, and my mother, Marian J. Rhodes (1922-1995). Who knew?

# ACKNOWLEDGEMENTS

This research was supported by a grant from the American Educational Research Association which receives funds for its “AERA Grants Program” from the National Center for Education Statistics and the Office of Education Sciences (US Department of Education), and the National Science Foundation under NSF Grant #REC-9980573. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

A great many individuals deserve thanks for their input and support of this work. The first acknowledgment goes to my family and friends who continued to ask, “How is the dissertation coming?” even when they knew they would not want to hear the answer. Over the years, my children patiently waited when I was late to pick them up, late for school events, and late to say good night. Now that Casey and Charlie are in graduate school and Lesley working toward her bachelors, they understand the level of commitment involved as well as the rewards. I couldn't be more proud and I thank them for their love, support, and tolerance. Chuck, my husband, kept me grounded and wouldn't let me give up golf even when the dissertation preoccupied nearly every waking and sleeping minute. Without him, I would have imploded long ago.

I had the best committee ever assembled: Robert Floden, Mark Reckase, Kenneth Frank, Mary Kennedy, and Kimberly Maier. There is no way to really thank them for their vast knowledge, encouragement, and support. Special kudos to Bob and Kim who modeled very important lessons in mentoring and patience. Additional thanks go to Gail Nutter and Lynne Frechen who made it possible for me to do this.

Back in the past, thanks to Dr. Stephen Raudenbush, Dr. Alex Sedlacek, and my colleagues from LAMMP who surprised me with a new way of seeing the world, related and multilevel.

And last but far from least, without dissertation support from American Educational Research Association, I would not have had time to start or finish the dissertation. It was a surprise and a great pleasure to get to know Dr. Jeremy Kilpatrick, my AERA board mentor. Jeanie Murdock (I have to wonder how she manages all the fledgling researchers year after year) took some of the sting out of having to admit that my revision aspirations outpaced the time available as the grant period stretched and stretched again. Now that I am finished, I hope this and my future work lives up to the honor you all have bestowed on me.

## TABLE OF CONTENTS


LIST OF TABLES ........ xi

LIST OF FIGURES ........ xvii

INTRODUCTION ........ 1

1. PREVIOUS RESEARCH ........ 7
    - Accountability and Assessment. ........ 7
    - Teacher Preparation ........ 8
    - Teacher Characteristics ........ 9
    - Teacher Professional Development. ........ 10
    - Teacher Practice ........ 10
    - Ongoing BTPS studies ........ 11

2. EDUCATION POLICY ENVIRONMENTS ........ 13
    - State Level ........ 13
        - Connecticut ........ 15
        - Tennessee ........ 17
    - National Level ........ 18

3. CONCEPTUAL MODEL ........ 21

4. DATA SETS ........ 24
    - State Data ........ 24
        - Connecticut Teacher and Student Samples ........ 25
        - Tennessee Teacher and Student Samples ........ 26
    - National Data ........ 27
        - NAEP Teacher and Student samples ........ 27
        - Scale comparisons for BTPS and NAEP ........ 28

5. LATENT VARIABLES AND THEIR INDICATORS ........ 33
    - Teacher Preparation ........ 33
    - Teacher Characteristics ........ 33
    - Professional Development ........ 37
    - Teacher Practices ........ 38
    - Student Achievement ........ 39

6. METHODS & PROCEDURES ........ 41

7. STRUCTURAL EQUATION MODELING RESULTS ........ 47
    - Connecticut Base Model for Mathematics ........ 48

Connecticut Base Model for Reading . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
Tennessee Mathematics and Reading Results . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
NAEP Mathematics and Reading Models . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 58
Comparisons Across All Data Sets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 68

8. RESULTS FROM THE HLM ANALYSIS OF CONNECTICUT $4^{th}$-$8^{th}$ GRADE MATHEMATICS GAINS . . . . . . . . . . . . . . 74
I. Setting the Stage: Variation in Mathematics Achievement Gains by Student Demographic Groups and Grade Level . . . . . . . . . . . . . . . . . . . . . . . . . . 76
II. Effects of Teacher Practices, Teacher Characteristics, and Professional Development on Student Achievement Gains . . . . . . . . . . . . . . . . . . . . . 88
III. IHE Teacher Preparation, Teachers, and Student Achievement . . . . . . . . . 115
Results Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 127

9. RESULTS FROM THE HLM ANALYSIS OF CONNECTICUT $4^{th}$-$8^{th}$ READING GAINS . . . . . . . . . . . . . . . . . . . . . . . 133
I. Setting the Stage: Variation in Reading Achievement Gains by Student Demographic Groups . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 135
II. Differential Effects of Teacher Practices, Teacher Characteristics, and Professional Development on Student Reading Achievement Gains . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 145
III. IHE Teacher Preparation, Teachers, and Student Achievement . . . . . . . . . 152
Results Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 158

10. RESULTS OF THE HLM ANALYSIS FOR 1996 NAEP $4^{TH}$ GRADE MATHEMATICS . . . . . . . . . . . . . . . . . . . . . . 161
I. Setting the Stage: Variation in Mathematics Achievement Gains by Student Demographic Groups and SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 163
II. Differential Effects of Teacher Characteristics and Teacher Practices on Student Achievement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 174

11. RESULTS OF THE HLM ANALYSIS FOR 2000 NAEP $4^{TH}$ GRADE MATHEMATICS . . . . . . . . . . . . . . . . . . . . . . 196
I. Setting the Stage: Variation in Mathematics Achievement by Student Demographic Groups and SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 198
II. Differential Effects of Teacher Characteristics and Teacher Practices on Student Achievement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 209

12. RESULTS OF THE HLM ANALYSIS FOR 1998 NAEP $4^{TH}$ GRADE READING . . . . . . . . . . . . . . . . . . . . . . . . . . 218
I. Setting the Stage: Variation in Reading Achievement Gains by Student Demographic Groups and SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 220
II. Differential Effects of Teacher Characteristics and Teacher Practices on Student Achievement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 231

13. RESULTS OF THE HLM ANALYSIS
FOR 2000 NAEP 4$^{TH}$ GRADE READING . . . . . 244
I. Setting the Stage: Variation in NAEP 2000 Reading Achievement Status by Student Demographic Groups and SES . . . . . 246
II. Differential Effects of Teacher Characteristics and Teacher Practices on Student Achievement . . . . . 257

14. COMPARISON OF HLM RESULTS
FOR MATHEMATICS AND READING ACROSS DATA SETS . . . . . 264

15. DISCUSSION AND IMPLICATIONS . . . . . 277
The Conceptual Model . . . . . 277
SEM Analysis . . . . . 279
HLM Analysis . . . . . 280
Gains Versus Achievement Status . . . . . 281
State Versus National Data Sets . . . . . 281
Relationships Across Models . . . . . 282
Contexts . . . . . 283
Differential Effects . . . . . 285
Implications for the Present and Future . . . . . 289

APPENDICES . . . . . 293
APPENDIX A: Correlation Tables . . . . . 294
APPENDIX B: Miscellaneous Tables . . . . . 305

REFERENCES . . . . . 307

## LIST OF TABLES

4. DATA SETS . . . . . 24
Table 4.1 Sample Comparisons by Data Set and Analysis . . . . . 24
Table 4.2 BTPS Scales and Properties . . . . . 29
Table 4.3 NAEP 1996 Mathematics Scales . . . . . 30
Table 4.4 NAEP 2000 Mathematics Scales . . . . . 30
Table 4.5 1998 NAEP 4th Grade Reading Scales . . . . . 31
Table 4.7 2000 NAEP 4th Grade Reading Scales . . . . . 31

6. METHODS & PROCEDURES . . . . . 41
Table 6.1 Analysis Matrix . . . . . 43
Table 6.2 Variable Availability Across Data Sets . . . . . 44

7. STRUCTURAL EQUATION MODELING RESULTS . . . . . 47
Table 7.1 Connecticut Standardized Total Effects for Mathematics . . . . . 50
Table 7.2. Connecticut Standardized Total Effects for Reading . . . . . 55
Table 7.3 Comparison of $r^2$ for Mathematics and Reading Constructs Across Data Sets . . . . . 63
Table 7.4. Standardized Total Effects for Mathematics Across Data Sets . . . . . 64
Table 7.5. Standardized Total Effects for Reading Across Data Sets . . . . . 66

8. RESULTS FROM THE HLM ANALYSIS OF CONNECTICUT $4^{th}$-$8^{th}$ GRADE MATHEMATICS GAINS . . . . . 74
Table 8.1 Description of Variables Used in the Connecticut Mathematics Analyses . . . . . 75
Table 8.2 Connecticut's Average Mathematics Gain 1999-2000 . . . . . 77
Table 8.3 Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Ethnicity . . . . . 79
Table 8.4 Predicted Mathematics Gains by Student Demographic Group . . . . . 80
Table 8.5 Regression Coefficients for Connecticut $4^{th}$ Grade Mathematics Gain 1999-2000 by Gender . . . . . 81
Table 8.6 Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Gender and Ethnicity . . . . . 82
Table 8.7 Predicted Mathematics Gains by Student Demographic Group with Differences by Gender . . . . . 84
Table 8.8 Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Gender, Ethnicity, Grade . . . . . 86
Table 8.9 Predicted Mathematics Gains by Student Demographic Group and Grade . . . . . 87
Table 8.10 Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Gender . . . . . 91
Table 8.11 Regression Coefficients for Average Gains by Gender Predicted by Average District SES . . . . . 93

Table 8.12 Variance Decomposition for Gender by District SES . . . . . . . . . . . . 94
Table 8.13 Variance Decomposition for Gender Controlling SES - Teacher Practices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
Table 8.14 Variance Decomposition by Gender for Teacher Characteristics Controlling for SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 98
Table 8.15 Variance Decomposition for Gains by Gender Controlling for SES - Teacher Professional Development . . . . . . . . . . . . . . . . . . . . . . . . . . . . 100
Table 8.16 Regression Coefficients for Gains by Ethnicity 1999-2000 . . . . . . 104
Table 8.17 Regression Coefficients for Black/White Gains by Average District SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 105
Table 8.18 Variance Decomposition for Black/White Gains by SES . . . . . . . . 106
Table 8.19 Variance Decomposition for White/Black Controlling SES - Teacher Practices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 109
Table 8.20 Variance Decomposition for Black/White Controlling for SES - Teacher Characteristics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 110
Table 8.21 Variance Decomposition for Black/White Controlling for SES - Teacher Professional Development . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
Table 8.22 Description of Teacher Preparation Variables . . . . . . . . . . . . . . . . 117
Table 8.23 Connecticut's Average Mathematics Knowledge by Teacher Preparation Institutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 118
Table 8.24 Connecticut's Average Mentoring Frequency by Teacher Preparation Institutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 119
Table 8.25 Average Mathematics Knowledge Predicted by Program Characteristics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 120
Table 8.26 Average Mentoring Frequency Predicted by Program Characteristics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 120
Table 8.27 Average Mathematics Knowledge Distributed Across Teacher Preparation Institutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 122
Table 8.28 Distribution of Predicted Mathematics Knowledge across Teacher Preparation Programs (low to high) . . . . . . . . . . . . . . . . . . . . . . . . . . 123
Table 8.29 Average Mentoring Frequency Distributed Across Teacher Preparation Institutions (low to high) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 124
Table 8.30 Distribution of Mentoring Frequency Across Teacher Preparation Programs (low to high) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 125
Table 8.31 Teacher Preparation Programs Ranked on Teacher Graduates' Mathematics Knowledge and Mentoring Frequency . . . . . . . . . . . . . . . 131

9. RESULTS FROM THE HLM ANALYSIS OF CONNECTICUT $4^{th}$-$8^{th}$ READING GAINS . . . . . . . . . . . . . . . . . . . . . . . 133
Table 9.1 Description of Variables Used in the Connecticut Reading Analyses . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 134
Table 9.2 Connecticut's Average Reading Gain 1999-2000 . . . . . . . . . . . . . . 136
Table 9.3 Regression Coefficients for Connecticut Reading Gain 1999-2000 by Ethnicity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 138
Table 9.4 Predicted Reading Gains by Student Demographic Group . . . . . . . . 139

Table 9.5 Regression Coefficients for Reading Gains by Gender (CT) ...... 140
Table 9.6 Regression coefficients for Connecticut Reading Gain 1999-2000 by Gender and Ethnicity ........................................ 141
Table 9.7 Predicted Reading Gain by Student Demographic Groups with Differences by Gender ........................................ 142
Table 9.8 Regression Coefficients for Connecticut Reading Gain 1999-2000 by Gender, Ethnicity, and Grade ................................ 144
Table 9.9 Predicted Mathematics Gains by Student Demographic Group and Grade ........................................................ 144
Table 9.10 Regression Coefficients for Average Connecticut Reading Gain 1999-2000 Across Teachers ................................ 148
Table 9.11 Regression Coefficients for Average Reading Gain by Average District SES ................................................ 149
Table 9.12 Variance in Average Reading Gains Explained by SES ......... 150
Table 9.13 Variance Decomposition Table for Average Reading Gains Across Teachers in Connecticut .................................... 151
Table 9.14 Unconditional Models for Levels of Teacher Variables Across Teacher Preparation Institutions ............................ 154
Table 9.15 Connecticut's Average English Preparation by Teacher Preparation Institutions ................................................ 155
Table 9.16 Regression Coefficients for Teacher English Preparation by Teacher Preparation Institutions ................................ 156
Table 9.17 Teacher English Preparation by 13 Connecticut Teacher Preparation Programs (low to high) ................................ 156

10. RESULTS OF THE HLM ANALYSIS FOR 1996 NAEP $4^{TH}$ GRADE MATHEMATICS ..................... 161
Table 10.1 Description of Variables Used in the NAEP 1996 $4^{th}$ Grade Mathematics Analyses ............................................ 162
Table 10.2 NAEP 1996 Average Mathematics Achievement .............. 164
Table 10.3 Regression Coefficients for NAEP $4^{th}$ Grade Mathematics 1996 by Ethnicity ................................................ 166
Table 10.4 Predicted Mathematics Achievement Student Demographic Group 167
Table 10.5 Regression coefficients for NAEP Mathematics Achievement 1996 by gender ................................................... 168
Table 10.6 Regression Coefficients for NAEP $4^{th}$ Grade Mathematics Achievement 1996 by Gender and Ethnicity .................... 169
Table 10.7 Predicted Mathematics Achievement by Student Demographic Group with Differences by Gender .......................... 170
Table 10.8 Regression Coefficients for NAEP 4th Mathematics Achievement 1996 by Gender, Ethnicity, and SES .......................... 172
Table 10.9 Predicted Mathematics Achievement by Student Demographic Group and SES ................................................ 173
Table 10.10 Regression Coefficients for NAEP Mathematics Achievement 1996 by Gender ............................................... 177

Table 10.11 Regression Coefficients for NAEP Mathematics Achievement 1996 by Gender and Average Classroom SES . . . . . . . . . . . . . . . . . . . . . . . 179
Table 10.12 Variance Decomposition for Male/Female Controlling SES - Teacher Practices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 181
Table 10.13 Variance Decomposition for Male/Female Controlling for SES - Teacher Characteristics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 182
Table 10.14 Regression Coefficients for 1996 NAEP 4th Mathematics Achievement by Ethnicity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 184
Table 10.15 Regression Coefficients for Black/White Achievement Predicted by Classroom SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 186
Table 10.16 Variance Decomposition for Black/White achievement by SES . 187
Table 10.17 Teacher Practices - Variance Decomposition for Average White/Black Achievement Across Teachers Controlling SES . . . . . . . 189
Table 10.18 Variance Decomposition for Black/White controlling for SES - Teacher Characteristics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 191
Table 10.19 Variance Decomposition for Black Female Controlling for SES - Teacher Practice, Characteristics, and Professional Development . . . . 193

11. RESULTS OF THE HLM ANALYSIS FOR 2000 NAEP 4TH GRADE MATHEMATICS . . . . . . . . . . . . . . . . . . . . . . 196
Table 11.1 Description of Variables Used in the 2000 NAEP 4th Grade Mathematics Analyses . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 197
Table 11.2 2000 NAEP 4th Grade Average Mathematics Achievement . . . . . . 199
Table 11.3 Regression Coefficients for 2000 NAEP 4th Grade Mathematics by Ethnicity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 201
Table 11.4 Predicted Mathematics Achievement Student Demographic Group 202
Table 11.5 Regression Coefficients for 2000 NAEP Mathematics Achievement by Gender . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 203
Table 11.6 Regression Coefficients for 2000 NAEP 4th Grade Mathematics Achievement by Gender and Ethnicity . . . . . . . . . . . . . . . . . . . . . . . . 204
Table 11.6 Predicted Mathematics Achievement by Student Demographic Group with Differences by Gender . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 205
Table 11.8 Regression Coefficients for 2000 NAEP 4th Mathematics Achievement by Gender, Ethnicity, and SES . . . . . . . . . . . . . . . . . . . . 207
Table 11.9 Predicted Mathematics Achievement by Student Demographic Group and SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 208
Table 11.10 Regression Coefficients for 2000 NAEP 4th Mathematics Achievement by Ethnicity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 212
Table 11.11 Regression Coefficients for Black/White Achievement Predicted by Average Classroom SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 214
Table 11.12 Variance Decomposition for Hispanic/White Achievement by SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 215
Table 11.13 Variance Decomposition for Hispanic/White controlling SES - Teacher Practices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 216

12. RESULTS OF THE HLM ANALYSIS
FOR 1998 NAEP 4$^{TH}$ GRADE READING ........................... 218
Table 12.1 Description of Variables Used in the 1998 NAEP 4$^{th}$ Grade Reading Analyses ........................................... 219
Table 12.2 1998 NAEP 4$^{th}$ Grade Average Reading Achievement .......... 221
Table 12.3 Regression Coefficients for 1998 NAEP 4$^{th}$ Grade Reading by Ethnicity ........................................... 223
Table 12.4 Predicted Reading Achievement Student Demographic Group ... 224
Table 12.5 Regression Coefficients for 1998 NAEP Reading Achievement by Gender ........................................... 225
Table 12.6 Regression Coefficients for 1998 NAEP 4$^{th}$ Grade Reading Achievement by Gender and Ethnicity ........................ 226
Table 12.7 Predicted Reading Achievement by Student Demographic Group with Differences by Gender ..................................... 227
Table 12.8 Regression Coefficients for 1998 NAEP 4th Reading Achievement by Gender, Ethnicity, and SES ............................... 229
Table 12.9 Predicted Reading Achievement by Student Demographic Group and SES ........................................... 230
Table 12.10 Regression Coefficients for 1998 NAEP 4$^{th}$ Reading Achievement by Ethnicity ........................................... 234
Table 12.11 Regression Coefficients for Black/White Achievement Predicted by Average District SES ..................................... 236
Table 12.12 Variance Decomposition for Black/White Achievement by SES . 237
Table 12.13 Variance Decomposition for White/Black controlling SES - Teacher Practices ..................................... 239
Table 12.14 Variance Decomposition for Black/White Controlling for SES - Teacher Characteristics ..................................... 240
Table 12.15 Variance Decomposition for Black/White Controlling for SES - Teacher Professional Development ............................ 242

13. RESULTS OF THE HLM ANALYSIS
FOR 2000 NAEP 4$^{TH}$ GRADE READING ........................... 244
Table 13.1 Description of Variables Used in the 2000 NAEP 4$^{th}$ Grade Reading Analyses ........................................... 244
Table 13.2 2000 NAEP 4$^{th}$ Grade Average Reading Achievement .......... 247
Table 13.3 Regression Coefficients for 2000 NAEP 4$^{th}$ Grade Reading by Ethnicity ........................................... 249
Table 13.4 Predicted Reading Achievement Student Demographic Group .... 250
Table 13.5 Regression Coefficients for 2000 NAEP Reading Achievement by Gender ........................................... 251
Table 13.6 Regression Coefficients for 2000 NAEP 4$^{th}$ Grade Reading Achievement by Gender and Ethnicity ........................ 252
Table 13.7 Predicted Reading Achievement by Student Demographic Group with Differences by Gender ..................................... 253
Table 13.8 Regression Coefficients for 2000 NAEP 4th Grade Reading

Achievement by Gender, Ethnicity, and SES . . . . . . . . . . . . . . . . . . . . 255
Table 13.9 Predicted Reading Achievement by Student Demographic Group and SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 256
Table 13.10 Regression Coefficients for 2000 NAEP 4th Grade Reading Achievement by Ethnicity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 259
Table 13.11 Regression Coefficients Mean Achievement Predicted by Average District SES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 260
Table 13.12 Variance Decomposition for Mean Achievement by SES . . . . . . . 261
Table 13.13. Teacher Effects on Mean Achievement Controlling SES . . . . . . 263

14. COMPARISON OF HLM RESULTS FOR MATHEMATICS AND READING ACROSS DATA SETS . . . . . . . . . 264
Table 14.1 Comparison of Variation Across Data Sets . . . . . . . . . . . . . . . . . . . 264
Table 14.2 Group Means Which Varied Across Teachers (x) for All Data Sets 266
Table 14.3 Variance in Mean Achievement by Ethnicity Explained by Teacher Variables Across All Data Sets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 267
Table 14.4 Variance in Mean Achievement by Gender Explained by Teacher Variables . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 273

APPENDIX A: Correlation Tables . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 294
Table A.1 Connecticut 2000 4th-8th Mathematics (n=168) . . . . . . . . . . . . . . . 295
Table A.2 Connecticut 2000 4th-8th Reading (n=168) . . . . . . . . . . . . . . . . . . . 297
Table A.3 Tennessee 2000 4th-8th Mathematics and Reading Correlations (n=442) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 299
Table A.4 1996 4th Grade NAEP Mathematics Correlations (n=584 teachers) 301
Table A.5 NAEP 2000 4th Mathematics Correlations (n=1029) . . . . . . . . . . . . 302
Table A.6 NAEP 1998 Reading (n=757) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 303
Table A.7 NAEP Reading 2000 (n=867) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 304

APPENDIX B: Miscellaneous Tables . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 305
Table B.1 Student Demographics - Dummy Coding . . . . . . . . . . . . . . . . . . . . 306

## LIST OF FIGURES

3. CONCEPTUAL MODEL .............................................. 21
Figure 3.1 Conceptual Model .......................................... 21

7. STRUCTURAL EQUATION MODELING RESULTS ....................... 47
Figure 7.1 Structural Equation Model for Mathematics in Connecticut and Tennessee ............................................. 49
Figure 7.2 Structural Equation Model for Reading in Connecticut and Tennessee ............................................. 54
Figure 7.3 1996 NAEP $4^{th}$ Grade Mathematics Structural Equation Model .... 59
Figure 7.4 2000 NAEP $4^{th}$ Grade Mathematics Structural Equation Model .... 60
Figure 7.5 1998 NAEP $4^{th}$ Grade Reading Structural Equation Model ........ 61
Figure 7.6 2000 NAEP $4^{th}$ Grade Reading Structural Equation Model ........ 62

14. COMPARISON OF HLM RESULTS
FOR MATHEMATICS AND READING ACROSS DATA SETS ......... 264
Figure 14.1 Conceptual Model (repeated) ............................ 278
Figure 14.2 Three Level HLM model ................................ 280

# INTRODUCTION

Ideas about the importance of assessment, accountability, standards, and teacher quality to student achievement in education are changing. Although education reform has often foundered due to a lack of system-wide cohesion, the focus of current federal policy is on system-wide accountability tied to student scores. The underlying premise is that the system is or can be made cohesive, that pushing button A will trigger button B to produce the desired result, but is it really so simple? Can accountability based on students' adequate yearly progress (AYP) really force the system to improve? Is there evidence that teacher preparation, teacher characteristics, professional development, and/or teacher practice are sufficiently connected to K-12 student outcomes so that change in one is related to change in the other?

The lack of research which ties the elements of the education system to student achievement in a single model has fueled arguments on both sides of the education controversy. Those who believe that the current system is "broken," that high barriers and low standards prohibit the best applicants from entering teaching focus on an expensive system that doesn't produce increased student achievement (Angus, 2001; Paige, 2002; Ravitch, 2002). Those who believe that the lack of highly qualified teachers is tied to the weakening of certification standards, an underfunded public education system, and lack of teacher professionalization, also point to low student achievement as the end result (Darling-Hammond & Youngs, 2002). In either case, empirical evidence that teacher preparation, teacher characteristics, professional development, and teacher practices are or are not related to each other and to student achievement is lacking.

No one denies that education in the United States is complicated by a world of competing expectations, demands, and separately administered but intertwined levels of responsibility. A long list of research, both qualitative and quantitative, attests to a resistant and fragmented education system fraught with uncertainty and conflicting goals, where students, teachers, administrators, and districts work toward different ends (Cusick, 1992; Floden & Buchmann, 1993; Labaree, 1997; Lortie, 1975; Muncey & McQuillan, 1996). At the local level, principals and teachers may or may not understand and follow state standards and practice guidelines (Spillane, 2005; Youngs & King, 2002). Teachers, especially beginning teachers, often default to the way they were taught when they were students {Lortie, 1975 #57}. When teachers do seek to improve their practice based on research results, evidence is often lacking on what constitutes high-quality research and what practices actually improve student scores {What Works Clearinghouse, 2002 #488}. Conflicting goals, resistence, and lack of proven teaching practices surely place time on the side of education tradition (Tyack & Cuban, 1995).

Teacher shortages and underfunded state and district budgets tangle the quest for teacher quality with the need to fill classrooms. In some cases at the school district level, this means hiring teachers who lack appropriate credentials or who come from alternative certification programs that are sometimes less rigorous routes into teaching {Darling-Hammond, 1999 #498}. In other cases, teaching positions are left unfilled while class size climbs. Additionally, rising costs curtail academic support programs, giving teachers and administrators more to do with less.

Politicians are under great pressure to prove they are spending the taxpayers money wisely in times of shrinking budgets and term limits. They cannot afford to wait for positive news about failing schools in their districts so many are embracing the push for accountability tied to standards, assessment, and teacher quality at all levels. Under President Clinton, the Title II State Report Card (Reauthorization of HEA 1998)(US Department of Education, 1998a) was the first federal legislation aimed at improving teacher quality and teacher preparation by requiring all states to report teacher certification pass rates by teacher preparation institution to the Department of Education for public distribution. It strongly encouraged communication between teacher preparation programs and state government, as well as alignment among teacher preparation program content, teacher certification examinations, and K-12 student standards and assessments.

Even newer is the federal law that requires all states to develop their own accountability systems for public education. "No Child Left Behind" (NCLB) Act of 2001, formerly known as ESEA, mandated that all states test annually in mathematics and reading for all students in grades 3-8. NCLB, the flagship of President Bush's education policy, requires that states align student tests and teacher professional development with state standards. By the end of the 2006 school year, all teachers must be "highly qualified." {, 2002 #257}

While the federal government has gradually increased its presence in public education since the Elementary and Secondary School Act (ESEA) of 1988 (US Department of Education, 1998b), the most recent NCLB assessment and accountability provisions have proven far-reaching and expensive. States must balance increased student testing costs against programs and services as federal funding dwindles.

The idea that large-scale assessments and accountability will improve the quality of instruction and increase student learning for *all* students enjoys widespread support (Amrein & Berliner, 2002; Robert L Linn, Baker, & Betebenner, 2002; Traub, 2002). The pressure and stakes are high for all participants in the education system but it is unclear how strongly the teacher education, teacher quality, teacher professional development, teacher practices are related to the fifth component, K-12 student scores, or what policy instruments help or hinder student gains.

Qualitative and quantitative research connects teacher preparation, teacher characteristics, and professional development to teacher practice, but consistent findings are rare. There is little empirical research that specifically traces the effects from teacher preparation, teacher characteristics, or professional development through teacher practice to student achievement. The most convincing evidence supporting the idea that alignment among standards, professional development, and teacher practice can improve student achievement is the Cohen and Hill (D. K. Cohen & Hill, 2001) study of mathematics reform in California. They found that teachers who had the opportunity to study the reform and received professional development in the supported mathematics framework improved student scores on the state test.

The history of education suggests a system resistant to change and lacking consistent connections, yet accountability, which relies on a connected system, is already in place demanding swift, positive results for all children. Can student improvement be attributed to changes in teacher practice? Can changes in teacher practice be traced to changes in teacher preparation and professional development? Can changes in student scores be attributed to teacher personal and professional characteristics? Do all children benefit from new teaching practices associated with the standards movement or does this reform, like so many other efforts, unwittingly benefit the advantaged groups in our society?

To shed light on these concerns, this research examines the effects of teacher preparation, teacher characteristics, professional development, and teaching practice on student mathematics and reading achievement using data from the "Beginning Teacher Preparation Survey (BTPS)"(Valli, 2000) and consecutive administrations of the National Assessment of Educational Progress (NAEP). Employing a combination of qualitative and quantitative methods, the project asks the following series of questions at both state and national levels:

1) What characteristics of teacher preparation programs are strongly related to:
    - Teacher practice?
    - Student math and literacy achievement?
    - Differences in achievement across student demographic groups?
2) What characteristics of teacher professional development are strongly related to:
    - Teacher practice?
    - Student math and literacy achievement?
    - Differences in achievement across student demographic groups?
3) What teacher characteristics are strongly related to:
    - Teacher practice?
    - Student math and literacy achievement?
    - Differences in achievement across student demographic groups?

4) When all three are included in the model, what are their relative contributions to:

- Teacher practice?
- Student math and literacy achievement?
- Differences in achievement across student demographic groups?

Sections 1-2 review research on the constructs of interest, their relationships to each other and student achievement, as well as related education policy at the state and national levels. The third section develops the conceptual model that includes the four constructs and their hypothesized connections to student achievement in mathematics and reading, This is followed by a description of data sets, scales, and variables. Section 6 presents the methods and procedures used in the analyses followed by the section on structural equation modeling (SEM) which determines the relative importance of teacher preparation, teacher characteristics and professional development on teacher practice and on subsequent student achievement. Sections 8-14 present the results for hierarchical linear modeling (HLM) analyses which specifically looks at differential outcomes of teacher variables on average achievement by gender and ethnicity across data sets. Finally, Section 15 discusses overarching research results and implications for future research.

# 1. PREVIOUS RESEARCH

In the emerging atmosphere of accountability linked to student assessments, the research focused on five high profile targets for education policy: teacher preparation at institutions of higher education (hereafter referred to as teacher preparation or teacher preparation program), teacher quality, teacher professional development, teacher practices, and their relationships to student gains/achievement.

**Accountability and Assessment**. Large scale accountability systems which hold students, teachers, schools, and districts accountable for student scores on state assessments are relatively new to American education. Accountability systems must answer three basic questions: who is to be held accountable, for what will they be held accountable, and what will be the consequences (Goertz, Duffy, & with Le Floch, 2001). NCLB (2001) holds students, schools, districts and states accountable for the pass rates of all students on the required state student assessment; HEA (1998) holds states and teacher preparation programs accountable for the pass rates of their students state teacher certification examinations. States face the threat of bad publicity on "report cards" and possible withdrawal of federal funding if they, or those under them, do not comply with the law. In addition to negative publicity and loss of funds, failing K-12 schools face a series of increasing consequences beginning with "being identified" for improvement and culminating in takeover or reconstitution. Schools of education face adverse publicity on the State Report Card, the loss of accreditation, and funding.

The idea that large-scale assessments and accountability will improve the quality of instruction and increase student learning for *all* students enjoys widespread support

(Amrein & Berliner, 2002; Robert L. Linn, 2001; "No Child Left Behind Act of 2001," 2002; Paige, 2002, 2003, 2004; Traub, 2002). Unfortunately, it is far easier to agree with policy than to implement it. It is well documented that policy is modified by local interpretations and contexts during the implementation process (David K. Cohen, 1990; D. K. Cohen & Spillane, 1993; Kingdon, 1995; Tyack & Cuban, 1995); modifications are more likely to occur the further one gets from the source of authority and closer to the point of delivery (Lipsky, 1980).

The pressure and stakes are high for all participants in the education system but it is unclear how strongly the four components (teacher education, teacher quality, teacher professional development, teacher practices) are related to the fifth component, K-12 student scores, or what policy instruments help or hinder student gains. In part, the lack of large scale databases which span all five measures tied to student gains/achievement has limited the research community's ability to tackle such questions. This is not to say that the relationships among teacher preparation, teacher characteristics, professional development, and teacher practice have not been studied, but rather that only a small proportion of studies have extended those relationships to include student achievement or gains.

**Teacher Preparation**. Policy makers have focused on teacher education as a pressure point to improve K-12 student learning although they lack strong evidence about what teacher education practices should be supported (M. M. Kennedy, 1999; S. Wilson, Darling-Hammond, & Berry, 2000; S. M. Wilson, Floden, & Ferrini-Mundy, 2001). Nonetheless, it is important to discover which elements or combinations of elements actually make a difference in teaching practice and student learning. If government wishes to push the educational system by requiring higher student scores, then it is imperative to know which factors contribute to the desired student achievement outcomes for all students.

Wilson, Floden and Ferrini-Mundy (S. M. Wilson et al., 2001) examined fifty-seven rigorous, peer-reviewed studies that addressed the key issues in initial teacher preparation. Little research directly assessed prospective teacher subject matter or pedagogical knowledge and then evaluated the relationships between teacher preparation and student learning. There were inconsistent among those that did. The review called for research which was more specific about the features of program content and quality.

**Teacher Characteristics**. Teacher characteristics include both professional and personal attributes. High quality teachers make a large difference in student learning (Darling-Hammond, 2000; Ferguson, 1991; Sanders & Rivers, 1996; Spellings, 2005). Indicators of teacher quality typically include scores on tests, degrees, course work, and certification status. Wayne and Youngs (Wayne & Youngs, 2003) reviewed twenty-one studies that connected teacher characteristics to student achievement. The authors found that positive relationships existed between college ratings of teachers' undergraduate institutions and K-12 student achievement. The review could not determine the importance of degrees and course work in general, although mathematics achievement in high school was consistently positive. One cited study (Eberts & Stone, 1984) found no relationship between number of college level mathematics courses in the last three years and 4$^{th}$ grade math achievement. Certification studies also indicated that subject specific qualifications matter but there were differences in interpretation across studies.

Other studies have looked at the impact of teacher personal characteristics on student achievement. Barton & Oja (Barton & Oja, 1999), found that efficacy, teachers' beliefs that they can affect student achievement, is small but positively related to student performance although others have reported mixed results (Brian Rowan, Chiang, & Miller, 1997; Valli, Reckase, & Raths, 2003).

**Teacher Professional Development**. Somewhat more research was done to connect in-service teacher preparation to student achievement. Kennedy (M.M. Kennedy, 1998) and Cohen & Hill (D. K. Cohen & Hill, 2001) focused on the relationship between teacher professional development/in-service teacher education and student achievement. Kennedy found twelve studies that used student achievement as the outcome measure for the effectiveness of professional development. She noted the lack of research on teacher professional development, intermediate outcomes such as teacher practices, and their ultimate effect on student outcomes.

Cohen & Hill's study (2001) centered on the implementation of mathematics reform in California. Their findings indicated that although learning opportunities varied among teachers, certain types of learning opportunities influenced practice, the amount of professional development influenced practice, using the student test as a learning opportunity for teachers influenced practice, and certain types of teacher learning opportunities affected student achievement. Student achievement was the dependent measure; teachers' practice was both an intermediate dependent measure of policy enactment and a direct influence on students' performance; teachers were the key connection between policy and practice.

**Teacher Practice**. While there is evidence that teacher practice is related to student achievement, there is little evidence that teacher practice actually increases student achievement (Supovitz, 2001). In his review of the literature, Supovitz found strong evidence, replicated across contexts, that connected teacher professional development to teacher practice. There were a handful of studies that connected teacher practice to student achievement. The results from those studies showed that, although positive, the impact of teacher practice on student achievement was hard to detect.

As in mathematics, experts disagree about what teacher reading practices are most desirable. The effects of teacher reading practices may be more difficult to interpret since they are confounded by issues of culture and power (Delpit, 1988; Foster & Peele, 1999; Heath, 1983; Hirsch, 1996). Although most educators agree that a combination of traditional and reform instructional strategies produce the best overall results for most children (Camilli, Vargas, & Yurecko, 2003; Pearson, 1996), it is also possible that specific instructional practices benefit student demographic groups differentially. The National Reading Panel (NRP)(National Reading Panel, 2000) found that direct instruction in phonics was especially useful for children who had not developed these skills during the initial phases of reading acquisition. This lends support to Delpit's idea that when instruction relies heavily on implicit components, it may be difficult or impossible for children from different cultures and ethnic backgrounds to succeed unless these hidden components are addressed directly. While more children show achievement gains when a blend of instructional practices is used those gains may not be significant for specific demographic subgroups of children.

**Ongoing BTPS studies**. Important work on the connections among teacher preparation, professional development, and student achievement in mathematics has been done by the developers of BTPS. Valli, Reckase, and Raths (Valli et al., 2003) focused on teacher education's relationship to student gains in Tennessee. They found modest but significant relationships among teacher preparation program variables, program outcomes, teaching practice and student mathematics achievement. They also noted a negative effect of preparation to teach mathematics on student achievement, possibly due to the study's grade level span or the inexperience of beginning teachers to assess how well prepared they, in fact, were. In a later article (Reckase, 2002), university faculty were found to be critical components of teacher preparation for mathematics instruction in Tennessee when student math achievement was the outcome.

The literature points to a need for more studies that include student scores on state assessments as the indicator of the effectiveness of teacher preparation, teacher quality, professional development, and teacher practice. Now that policy makers, politicians, educators, and parents believe that these connections can realistically be made and measured, they will continue to insist that states, teacher preparation institutions, researchers, districts, schools, teachers, and students use K-12 student scores as the litmus test of system effectiveness. In the hopes of providing baseline evidence, it is this study's challenge to make sense of the links between government policy and student achievement in a complex, pre-NCLB world.

## 2. EDUCATION POLICY ENVIRONMENTS

Government policy affects education at all levels. Three state level policy choices believed to influence education systems within states are the decisions to institute low stakes versus high stakes testing, centralized versus decentralized responsibility for education, and level of per pupil spending (Fuhrman, 2001; Greenwald, Hedges, & Laine, 1996; Sanders & Rivers, 1996). Connecticut and Tennessee approached improvement in K-12 student achievement from decidedly different angles while existing in the same national environment.

In 1999-2000, the national environment was rapidly changing from one that traditionally left matters of education to the states to one that exerted strong pressure on states to improve student achievement in prescribed ways. It is important to note that many of the national changes had been introduced and were being discussed but had not yet taken effect. This section describes the policy environments at the state and national levels during the 1990s and early 2000s.

**State Level:**

Connecticut and Tennessee made major changes to education during the 1990s. The most striking difference in their policy environments was system focus. Connecticut concentrated on inputs such as raising teacher quality with low stakes K-12 assessment and Tennessee focused on outputs by instituting a high-stakes accountability system based on K-12 student gains. State focus affected how pre-service teachers were prepared, how in-service teachers taught and were professionally supported, as well as affecting the content that students were expected to know and demonstrate.

There were serious disparities between Connecticut and Tennessee that played a part in teacher quality and student achievement. Connecticut outspent Tennessee and Connecticut students outperformed their Tennessee counterparts on NAEP. In many ways, Connecticut had the financial luxury to focus on inputs. An Education Week analysis of 1999-2000 SASS data (Education Counts Research Center, 2002) revealed that 99.9% of Connecticut students attended schools in districts that have per pupil expenditures at or above the national median ($5,385) compared to 25.3% students in Tennessee. Willis Hawley (1991), then director of the Center for Education Policy at Vanderbilt, commented that the Tennessee tax structure didn't yield enough revenue to make massive changes to education. TVAAS (Sanders & Horn, 1998) offered a reasonable and cost-effective alternative to Connecticut's centralized, multifaceted reforms.

During this period, the courts mandated that both states address educational equity based on district funding differences. In Sheff v. O'Neill ("Sheff v O'Neill," 1996), the Connecticut state supreme court ruled that poor, urban schools were receiving an inferior and unequal education. In Tennessee ("Tennessee Small School Systems v. McWherter," 1993), the court ruled that wealthier urban districts were advantaged in comparison to poor rural districts.

To fill in background for the state policy environments, additional information from the Education Week analysis of 1999-2000 SASS data indicated that Connecticut and Tennessee stacked up in the following ways: Both states were reported to have similar teacher/staff ratios and about the same percentage of total education dollars spent on teachers. Tennessee was more likely to have teachers without a major or minor in the subject and/or without certification. While both states offered educational assistance to teachers in the form of loans, scholarships and the waiving of license fees, Connecticut additionally offered housing assistance (loans, low interest mortgages, tax credits) targeted toward subject area shortages and high need schools.

In terms of achievement, 32% of Connecticut 4$^{th}$ grade students scored at proficient or higher on NAEP mathematics compared to 18% in Tennessee. 6% of black 4$^{th}$ graders were proficient or higher in Connecticut compared to 4% in Tennessee and 9% of hispanic 4$^{th}$ grades were proficient or higher in both state. A discussion of state specific education policy follows.

**Connecticut:**

The BTPS survey took place during a time of comprehensive, standards-based education reform in Connecticut. The ambitious goals, outlined in Connecticut's Comprehensive Plan for Education 1996-2000, aimed "to set and meet high standards for the performance of teachers and administrators leading to and evidenced by improved student learning" (Sergi, Commissioner of Eduction,1998). Expectations and supports for teachers undergirded expectations for student learning. There was a concerted effort to align all levels of the system with K-12 standards.

The emphasis on teacher quality included higher standards for entry into teaching. Along with the increased expectations for teachers came increased teacher salaries (S. Wilson et al., 2000). All prospective teachers had to pass both the Connecticut Competency Examination for Prospective Teachers (CONNCEPT) which tested basic knowledge and skills in reading, writing and mathematics plus either the National Teacher Examination (NTE) or the Connecticut Elementary Certification Test

(CONNECT) demonstrating subject matter proficiency. To further ensure teacher quality, all prospective teachers had to pass PRAXIS II content area examinations before graduation.

Teacher preparation programs were not only required to obtain National Council for the Accreditation of Teacher Education (NCATE) approval but also incorporate instruction on the aligned policy instruments into their programs. These instruments included the Common Core of Teaching (CCT) and the Code of Responsibility for Teachers aimed at teachers and the Common Core of Learning (CCL), the Connecticut Mastery Test (CMT), and the Connecticut Academic Performance Test (CAPT) which impact K-12 students.

Professional development in Connecticut came under the heading BEST, Beginning Educator Support and Training. During their first year, all beginning teachers were assessed six times using the Connecticut Competency Instrument (CCI), an observation instrument designed to measure effective teaching. District-based mentoring and state supported professional development were available to support new teachers undergoing observations. Teachers unable to meet CCI standards were refused provisional certification and prohibited from teaching.

The CMT, a criterion-referenced test aligned with K-12 state standards, measured student achievement against standards-based performance levels. As a low stakes assessment, it was meant to inform student progress and teacher practice.

NAEP results from 1998 showed that Connecticut students were among the best in the nation on reading. There were, however, documented achievement gaps between white and minority students. To correct the problem, a 1996 Supreme Court decision

(Sheff v. O'Neill) determined that a 1909 law created a system of economic and racial segregation. Consequently, the state began a number of new initiatives to address inequity as it overhauled education. The state increased school funding by fifty million dollars (1996-97). Districts were mandated to develop minority recruitment plans and create magnate schools to increase movement across district lines.

Neither education nor policy is static so it was not surprising that policy environment changed during the study year. Just as the final pieces of Connecticut's Comprehensive Plan for Education 1996-2000 were implemented, Greater Expectations: Connecticut's Comprehensive Plan for Education 2001-2005 emerged.

**Tennessee:**

At the core of Tennessee education reform was the adoption of Tennessee Value-Added Assessment System (TVAAS) in 1992. System focus was on the product of the educational experience rather than the process by which it was to be achieved (Sanders & Horn, 1998, p 339). TVAAS defined academic progress in terms of yearly student gains on the Tennessee Competency Assessment Program (TCAP). Teachers and schools were rated on whether their students make more or less progress than a typical student is expected to make in a given subject and grade, after adjusting for the prior achievement of each child (Archer, 1999). Rewards and sanctions were in place for those teachers who either exceeded or failed to meet the minimum progress.

The State Board of Education adopted new professional licensure standards for all teacher candidates in 1998. The standards were aligned with the Interstate New Teacher Assessment and Support Consortium (INTASC).

School-based decision making was encouraged and was tied to student academic progress. Teacher professional development activities and teacher practices were also expected to be results-driven. TVAAS allowed teachers and schools to use whatever methods resulted in achieving student academic progress (Sanders & Horn, 1998). The state role was to enhance the capacity of each school and district to take responsibility for the professional development of it's own personnel. By focusing on test results, other elements in the system were expected to fall into line.

**National Level:**

NAEP has been the Nation's Report Card since 1969. It's purpose from inception was to reflect current educational and assessment practices and measure change reliably over time. Mandated by Congress, the voluntary assessment was informational with no stakes attached. Rather than representing one state's content standards, it included nationally representative content identified by the Council of Chief State School Officers (CCSSO). The National Assessment Governing Board (NAGB) hired ACT, Inc., to identify appropriate achievement standards for each subject and grade tested.

While arguments and plans for the improvement of K-12 education raged at all levels of government, the authority to act in these matters rested with state and local government. In 1998, public and special interest group pressure, such as the National Commission for Teaching and America's Future (NCTAF) which focused in issues of teacher quality (National Center for Research on Teaching and Learning, 1993), combined to create a new expectation that the federal government could improve the quality of K-12 education by improving the quality of teacher preparation. This pressure opened a window that eased the way for Title II and NCLB reporting requirements to come into existence.

Title II of the 1998 Amendments to Higher Education Act (HEA) of 1965 was designed to recruit new teachers to areas with shortages of qualified personnel; to prepare them in both subject matter and pedagogical knowledge; and to gather information on state and university policies and procedures related to teacher quality. Section 207 of this amendment mandated that all states and institutions of higher education (IHEs) receiving any funds directly or indirectly through the HEA must provide the Secretary of Education with data on teacher standards and preparation. In addition to the reporting of pass rates on state teacher examinations by subject and by population subgroups, other information in the state report was to be determined by the Secretary of Education in conjunction with the National Center for Educational Statistics (NCES).

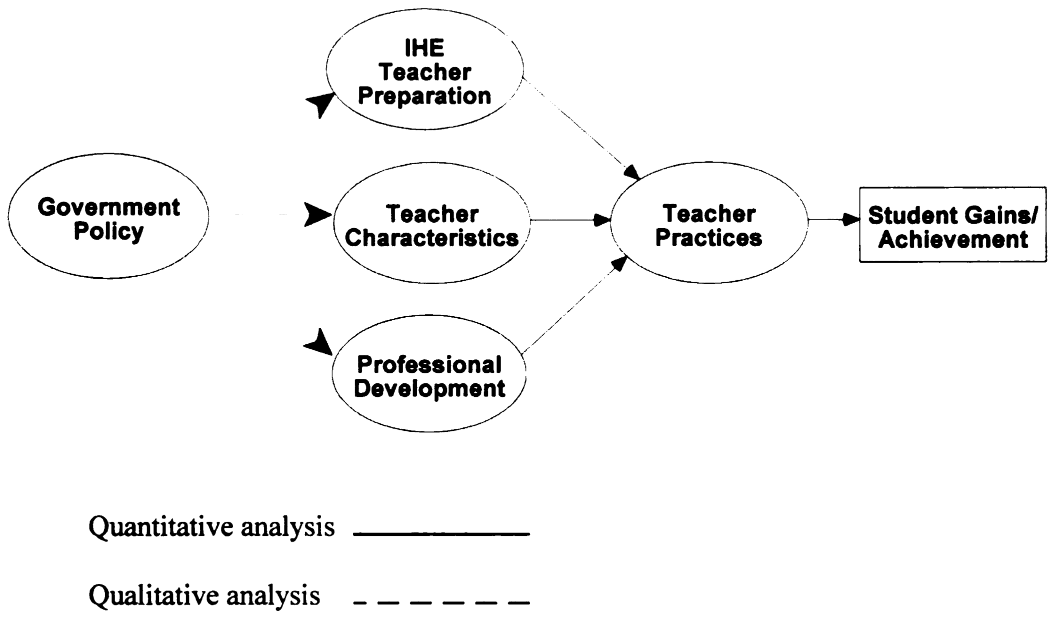
Virtually all states, territories, and the District of Columbia received HEA funds directly or indirectly and therefore had to provide this information on their teacher preparation programs. With the possible exception of a few, small, private IHEs, most colleges and universities enrolled students who received federal financial assistance authorized under this Act. Consequently, information on nearly all US teacher preparation programs was gathered and reported to the Secretary of Education in 2001. The Initial Report of the Secretary on the Quality of Teacher Preparation (otherwise known as the State Report card) was voluntarily collected in the Spring of 1999 to provide a baseline measure of teacher quality. While reporting was not mandatory during the BTPS survey year, the initial report card heralded impending changes to relationship between the federal government and states.

Measures of state and federal policy were not available in the data sets used for this study so it was not possible to link most policies to the model statistically. The one exception was that teacher pass rates from the Initial State Report Card could be appended to the Connecticut data and tied to IHE teacher preparation program name.

# 3. CONCEPTUAL MODEL

In order to answer the questions posed in this study, it was necessary to develop a conceptual model that depicted the hypothesized relationships among government policy, teacher preparation, teacher characteristics, professional development, teaching practice, and student achievement. It was posited, as did BTPS, that there were direct linear relationships between teacher preparation, teacher characteristics, professional development and teacher practice, as well as a direct linear relationship between teacher practice and student achievement. Given the standards-based climates of Connecticut and Tennessee, this linear relationship was extended from government policy through teacher preparation, teacher characteristics, and professional development to teacher practice and from teacher practice through to student achievement as the following model suggests.

**Figure 3.1** Conceptual Model



**Elements of the model:**

**Government policy** refers to state policies for teacher qualifications and professional development, standards alignment, equity, and state assessments. The state policy environment is set in a federal policy environment which, in 1999-2000, influenced state education policies primarily using inducements in the form of grants (McDonald & Elmore, 1987).[1] Dashed lines represent the suspected effects of policies that can be reviewed but not tested in this study.

**Teacher preparation** encompasses the quality of teachers' formal pre-service teacher education experiences. **Teacher characteristics** cover individual qualities, both personal and professional, that teachers bring to their teaching positions. Examples include subject matter credentials, efficacy (belief that they can make a difference for students), and specialized training in areas like assessment and diversity. **Professional development** includes time spent in professional development as well as its perceived impact. **Teacher practices** for both reading and math include traditional and reform activities. An example of a traditional activity would be mathematics drill and lecture activities while group and project activities would represent a reform activity. Depending on the data set used, the outcome variable for **student achievement** is either student gains over the course of one year (1999-2000) or the plausible value scores from the 4th grade NAEP reading (1998, 2000) or 4th grade mathematics (1996, 2000). Note that only teacher practice is hypothesized to be directly related to gains, that the variables in the model must go through student classroom experiences to affect student achievement.

---

[1] In 2005, the model would be the same but federal No Child Left Behind policies are now believed to affect state education policy strongly and in direct ways.

Solid lines represent the hypothesized paths tested in this study.

Full descriptions of these variables follow in Section 5, Latent Variables and Their Indicators. First, however, come descriptions of the BTPS and NAEP data sets.

# 4. DATA SETS

The study used six existing databases, two from BTPS and four from NAEP. All offered similar variables spanning the policy to student achievement continuum. Samples varied by the analysis performed and are presented in Table 2.

**Table 4.1** Sample Comparisons by Data Set and Analysis

| | Math | | | | Reading | | | |
|---|---|---|---|---|---|---|---|---|
| | CT | TN | NAEP 4th 2000 | NAEP 4th 1996 | CT | TN | NAEP 4th 2000 | NAEP 4th 1998 |
| SEM Teacher n | 168 | 442 | 1029 | 584 | 168 | 442 | 867 | 757 |
| HLM Student n | 1550 | X | 6205 | 4409 | 90 | X | 4548 | 4303 |
| HLM Teacher n | 81 | X | 1245 | 584 | 1776 | X | 918 | 828 |

The Connecticut and Tennessee SEM correlation matrices included all teachers and average student gains for both mathematics and reading. Due to the presence of individual student data nested withing teachers in Connecticut, it was possible to split the mathematics and reading samples for separate HLM analysis.

AM Software, necessary for creating weighted NAEP correlation matrices, used listwise deletion for missing values. HLM, which also needs complete data, used a refined set of variables based on findings from the SEM analyses. This allowed some cases to be recovered. Consequently, NAEP *n* for HLM is larger than that for SEM. More information on methods and procedures are included in Chapter 6.

***State data.*** NPEAT researchers conducted BTPS during the 1999-2000 school year in Connecticut and Tennessee. There were three related goals (Valli, Reckase, Raths, & Rennert-Ariev, 2001):

1. To develop a survey that would provide information from beginning teachers about their preservice and induction experiences; their teaching

beliefs, knowledge and practices; and their teaching context

2. To develop theoretically and empirically based constructs of teacher preparation, induction, teacher knowledge and beliefs, and teaching practices

3. To determine the relationships among three sets of variables: a) teacher preparation, b) teaching knowledge, practice, and beliefs, and c) student learning.

The survey was designed to study the linear relationship of teacher preparation and professional development to teacher knowledge, practice, and beliefs to student achievement. Valli and colleagues targeted 3$^{rd}$ through 8$^{th}$ grade math and reading/language arts teachers in their first three years of practice in public schools.

BTPS selected the states of Connecticut and Tennessee based on their national reputations for innovative K-12 educational practices. The focus on student outcomes required pre- and post-test student scores which could be directly tied to specific teachers. BTPS took advantage of an existing sample (n=168) selected for a prior study of beginning teachers by the Connecticut State Department of Education (CSDE). 100% of the CSDE sample responded to the BTPS questionnaire. The Tennessee sample consisted of 442 beginning math and reading teachers from 3rd-8th grade. Since all teachers in Tennessee have student gain scores attached to their performance, researchers compiled mailing lists to select their sample.

**Connecticut Teacher and Student Samples.** Teacher samples were compared to a national profile of beginning teachers, the 1998 Teacher Survey on Professional Development and Training conducted by NCES. While the study does not plan to generalize from the Connecticut and Tennessee samples to the rest of the United States, it is interesting to note that Connecticut and Tennessee teachers were demographically much like US teachers overall with a few exceptions: they were more likely to be white, female (no high school teachers were included in the survey), and possess masters degrees. There were two major differences between Connecticut and Tennessee: 22% of the Connecticut sample was male in comparison to in 7.6% in Tennessee. The Tennessee sample was similar to the national profile on diversity (21% minority) but the Connecticut sample was predominantly white (94.5%) (Valli, Reckase et al., 2001). The outcome variable, student achievement, was defined as student gain scores from the beginning to the end of the school year based on two administrations of the state assessment. In Connecticut, demographic information and scores exist for approximately 2500 students nested within the 168 teachers. The primary sampling unit for the survey was teachers.

**Tennessee Teacher and Student Samples**. Due to confidentiality concerns, Tennessee did not include individual student scores tied to teachers as part of the BTPS data. After aggregating students by teacher using the value-added framework, Dr. William Sanders, director of the TVAAS program, provided student outcome correlations to BTPS researchers who used them to create a full correlation matrix for entry into SEM. Specific demographic information at the student level was absent.

The BTPS data included variables that did not exist in NAEP but that were nevertheless important for studying the effects of teacher preparation on teacher practice and student achievement. Not only were factual measures available for teacher preparation, knowledge, and practice, BTPS also included attitudinal measures. For example, BTPS researchers found that a teacher's sense of self-efficacy was an important predictor of student gains in their preliminary multivariate analyses.

***National Data.*** NAEP regularly assesses a nationally representative sample of students in grades 4, 8, and 12 in most subjects. While students are the primary sampling unit, teachers fill out questionnaires regarding their education, certification, and teaching practices for each class that includes assessed students. At the national level, gain scores do not exist for the same samples of children and classrooms; however, one purpose of NAEP and its complex design is to serve as an effective barometer of change over time. Two consecutive administrations of $4^{th}$ grade math (Math:1996 & 2000) and two consecutive administrations of $4^{th}$ grade reading assessments (Reading: 1998 & 2000) provide a measure of change in student achievement and teaching practice over time.

**NAEP Teacher and Student samples.** For all four NAEP data sets, the public school sample with no accommodations (psamp) was selected. The national weight (origwt) was applied to produce national samples of students. The teacher sample exists only as part of the student record so the data for teachers who have multiple students assessed appeared multiple times at the student level. In order to create a teacher level sample, teacher variables were aggregated by teacher booklet. Consequently, the teacher sample is not random. Analyses using these data must be interpreted as random samples of students who have teachers with certain characteristics or who employ certain practices.

The single most important characteristic of both BTPS and NAEP data is the existence of a measure of student achievement that is nested within specific teachers. Teachers, then, provide the all-important link between the four latent variables and student achievement. These six data sets provide state and national assessment results as a measure upon which to explore the effects of teacher characteristics, professional development, teaching practices, and, in the case of Connecticut, teacher preparation.

**Scale comparisons for BTPS and NAEP**. No comparison of multiple data sets can be completed without a comparison of scales that are used. BTPS data included many scales important to this study. It was necessary to create similar scales for NAEP data. Table 4.2 presents the scales created for BTPS. Tables 4.3 through 4.7 contain information on similar scales created for NAEP mathematics and reading.

**Table 4.2** BTPS Scales and Properties

| Beginning Teacher Preparation Survey Scales | # of Items | Range scores | Mean | Standard Deviation | Internal consistency |
|---|---|---|---|---|---|
| **IHE Teacher Preparation** | | | | | |
| Faculty Characteristics | 6 | 9-30 | 15.20 | 4.9 | 0.77 |
| Program Quality | 4 | 5-25 | 17.70 | 4.10 | 0.73 |
| Field Experience | 4 | 4-20 | 15.20 | 3.70 | 0.77 |
| **Teacher characteristics** | | | | | |
| ***MATH*** | | | | | |
| Traditional Orientation | 4 | 4-20 | 13.60 | 3.00 | 0.64 |
| Reform Orientation | 4 | 9-20 | 15.30 | 2.20 | 0.45 |
| Math Preparation | 6 | 0-12 | 3.50 | 1.80 | 0.70 |
| Math Instruction Prep | 13 | 14-65 | 45.50 | 11.20 | 0.96 |
| Math knowledge | 6 | 9-30 | 24.30 | 17.80 | 0.90 |
| ***BOTH*** | | | | | |
| Efficacy | 7 | 9-35 | 28.37 | 29.00 | 0.66 |
| Diversity Prep | 7 | 7-35 | 19.70 | 6.20 | 0.92 |
| Assessment Preparation/Variety | 8 | 8-40 | 29.00 | 6.80 | 0.95 |
| ***READING*** | | | | | |
| Balanced Literacy Orientation | 8 | 10-40 | 34.60 | 3.30 | 0.65 |
| Reading/Writing Preparation* | 1 | 0-4 | | | 0.80 |
| Reading/Writing Instruction Preparation | 12 | 15-60 | 41.80 | 10.00 | 0.94 |
| Literacy knowledge | 10 | 3-50 | 37.10 | 7.10 | 0.90 |
| **Professional Development** | | | | | |
| Mentoring worth | 11 | 11-33 | 23.70 | 7.80 | 0.96 |
| Professional development frequency | 10 | 10-50 | 21.60 | 8.30 | 0.90 |
| Professional development impact | 12 | 12-72 | 46.20 | 11.90 | 0.90 |
| **Teacher Practice** | | | | | |
| Math Drill and Lecture activities | 6 | 11-30 | 26.00 | 2.80 | 0.64 |
| Math group and Project activities | 7 | 8-35 | 20.00 | 5.10 | 0.79 |
| Math sensemaking activities | 7 | 18-35 | 29.30 | 3.60 | 0.75 |
| ***READING*** | | | | | |
| Reform activities | 8 | 15-40 | 30.60 | 5.00 | 0.82 |
| Traditional Activities | 7 | 13-35 | 31.70 | 3.20 | 0.70 |
| Literacy materials | 4 | 4-20 | 10.10 | 4.00 | 0.68 |

* only one variable available

**Table 4.3** NAEP 1996 Mathematics Scales

| | # of Items | Range scores | Mean | Standard Deviation | Alpha |
|---|---|---|---|---|---|
| **Teacher characteristics** | | | | | |
| Reform Orientation | 4 | | 11.53 | 2.17 | 0.67 |
| Math Preparation | 2 | | | | |
| Math Instruction Prep | 2 | | 7.49 | 0.9 | 0.89 |
| Math knowledge | 4 | | 3.25 | 1.14 | 0.74 |
| Diversity Prep | 3 | | 1.58 | 1.15 | 0.68 |
| **Professional Development** | | | | | |
| Professional development frequency* | 1 | | | | |
| **Teacher Practice** | | | | | |
| Math group and Project activities | 4 | 4-16 | 11.97 | 2.69 | 0.74 |
| Math sensemaking activities | 5 | 5-20 | 10.35 | 2.74 | 0.73 |
| Emphasize Group & Project Assessment | | 3-12 | | | 0.69 |

* only one variable available

**Table 4.4** NAEP 2000 Mathematics Scales

| | # of Items | Range scores | Mean | Standard Deviation | Alpha |
|---|---|---|---|---|---|
| **Teacher characteristics** | | | | | |
| Math Preparation | 6 | | 38.67 | 6.66 | 0.61 |
| Math Instruction Prep | 8 | | 27.88 | 3.38 | 0.85 |
| NCTM standards knowledge* | 1 | | | | |
| State standards knowledge* | 1 | | | | |
| **Professional Development** | | | | | |
| Professional development frequency* | 1 | | | | |
| **Teacher Practice** | | | | | |
| Math group and Project activities | 3 | 4-16 | 9.45 | 2.21 | 0.74 |
| Math sensemaking activities | 6 | 5-20 | 13.53 | 3.07 | 0.70 |
| Math Drill & Lecture Activities | 3 | 4-9 | 8.63 | 0.78 | 0.67 |

* only one variable available

**Table 4.5** 1998 NAEP 4th Grade Reading Scales

| | # of Items | Range scores | Mean | Standard Deviation | Alpha |
|---|---|---|---|---|---|
| **Teacher characteristics** | | | | | |
| Diversity Prep | 2 | | 3.97 | 1.24 | 0.69 |
| Assessment Preparation/Variety | 11 | | 29.01 | 5.29 | 0.71 |
| Reading/Writing Preparation | 4 | | 16.55 | 8.27 | 0.92 |
| Reading/Writing Instruction Prep | 9 | | 23.54 | 3.26 | 0.85 |
| Technology preparation | 3 | | 5.68 | 1.56 | 0.71 |
| **Professional Development** | | | | | |
| Professional development frequency* | 1 | | | | |
| **Teacher Practice** | | | | | |
| Reform activities | 14 | | 38.16 | 6.53 | 0.84 |
| Traditional Activities | 12 | | 37.57 | 3.79 | 0.64 |
| Computer Activities | | | | | |

* only one variable available

**Table 4.7** 2000 NAEP 4th Grade Reading Scales

| | # of Items | Range scores | Mean | Standard Deviation | Alpha |
|---|---|---|---|---|---|
| **Teacher characteristics** | | | | | |
| Diversity Prep | 3 | | 6.19 | 1.68 | 0.72 |
| Assessment Preparation/Variety | 6 | | 16.35 | 3.19 | 0.63 |
| Reading/Writing Preparation | 4 | | 17.3 | 8.61 | 0.92 |
| Reading/Writing Instruction Prep | 8 | | 21.04 | 3.05 | 0.86 |
| Technology preparation | 4 | | 7.72 | 2.12 | 0.8 |
| **Professional Development** | | | | | |
| Professional development frequency* | 1 | | | | |
| **Teacher Practice** | | | | | |
| Reform activities | 5 | | 14.83 | 2.81 | 0.74 |
| Traditional Activities | 3 | | 9.44 | 1.84 | 0.75 |
| Computer Activities | 3 | | 7.17 | 2.96 | 0.83 |

* only one variable available

Internal consistency using Cronbach's Alpha was computed for NAEP scales in SPSS. In most cases, the number of items and therefore the range of the scales is considerably smaller for NAEP but sample sizes are much larger than BTPS. Additionally, results are intended to reflect national samples of students and their teachers rather than smaller state samples.

NAEP alphas ranged from .61 to .92, with eight in the .60s, ten in the .70s, seven in the .80s, and two in the .90s. BTPS alphas ranged from .45 to .96, with five in the 60s, seven in the 70s, two in the 80s, and nine in the nineties. It is reasonable that an instrument such as BTPS designed to capture these particular scales would have better reliabilities overall than NAEP which created for other purposes. In spite of these differences in intent, the NAEP scales performed reasonably well and in similar fashion to the BTPS scales across analyses.

## 5. LATENT VARIABLES AND THEIR INDICATORS

Three key latent variables (teacher preparation, professional development, and teacher characteristics) were hypothesized to be related to a fourth latent variable, teacher practices. Each latent variable had three or more indicator variables, a requirement for analysis stability in SEM. A discussion of the latent variables and their indicators follows.

**Teacher Preparation**. As of the dates of this study, both Connecticut and Tennessee required graduation from an accredited IHE teacher preparation program to be licensed as a teacher. While the importance of IHE teacher preparation as the only route to teaching is under debate, the qualities of IHE teacher preparation programs which are related to K-12 teacher practice and subsequent student gains has not been studied. In order to look at these connections, BTPS asked teachers to rate their preparation programs on a number of factors. Three scales are included as indicators of teacher preparation: faculty characteristics, field experiences, and program quality. These constructs are unique to BTPS and one of the characteristics that makes the data so attractive.

The *faculty characteristics* scale was designed to measure whether or not teacher preparation faculty practices are consistent with the ideals they espouse in the classroom. *Field experience* is an attendee rating of the quality of the internship experience. The *program quality* scale includes items related to the foundation and strength of disciplinary preparation as well as a rating by graduates of their teacher program quality.

**Teacher Characteristics.** Indicators of teacher characteristics include personal and professional qualities of individual teachers: subject matter, pedagogical, and diversity preparation; teacher efficacy and orientation; and assessment preparation/variety. Teachers with greater subject matter knowledge tend to stress conceptual understanding and ask higher order questions (Ball, 1997; Grossman, 1989) These teacher understandings are particularly relevant when the goal is student understanding (Borko & Putnam, 1996). Pedagogical preparation is also related to student understanding (Shulman, 1987). It allows teachers to maximize student learning by applying subject specific strategies for teaching the subject. Scales for subject matter and pedagogical preparation are included for both mathematics and reading below.

Scales used for both mathematics and reading - Teachers will always have prior knowledge and beliefs that filter their learning and teaching (Borko & Putnam, 1996; Lortie, 1975). *Teacher efficacy* is one such belief. The Valli et al. (Valli, Raths, Reckase, & Rennert-Ariev, 2001) work using the BTPS in Tennessee found that teacher efficacy, a scale which assessed the belief that a teacher can make a difference for all students, was related to increased student achievement. It has also been related to improvement in teaching practice and student motivation and performance (Barton & Oja, 1999).

The *diversity preparation* scale includes measures on that how well teachers believe they are prepared to teach students from other cultures and/or with different languages. When teachers choose instructional strategies, they must also be able to tailor that knowledge to the diverse learners represented in their classroom (Borko & Putnam, 1996; Shulman, 1987).

The National Research Council (1999) argued the need for matching assessments with student learning goals. They claimed that effective teachers learn about students' learn about students' understanding and assess students often. *Assessment preparation* covers teacher readiness to evaluate student learning, to align assessments and expectations, and to use assessments to guide decisions about teaching.

Mathematics only - Teacher characteristics applicable only to mathematics preparation includes traditional orientation, reform orientation, mathematics preparation, mathematics knowledge, and knowledge of national and state mathematics standards.

The data includes two orientation scales, traditional and reform. Mathematics *reform orientation* is measured by asking whether the respondents agreed with items such as "a teacher's primary goal is to help students achieve a deeper conceptual understanding of mathematics." Mathematics *traditional orientation* reflects agreement with items such as "a teacher's primary goal is to help students learn mathematical terms and master computational skills." The scales are not mutually exclusive and both reflect elements of the NCTM standards (National Council of Teachers of Mathematics, 2000) which advocate a mixture of conceptual and computational instruction for students.

The *mathematics knowledge* scale measures teachers' self-reported understanding in five areas of mathematics (general knowledge; fractions and decimals; estimation and measurement; patterns, functions, and algebraic thinking; statistics and probability; and geometric concepts) with ratings from "poor" to "excellent." Teacher self-reports of mathematics knowledge correlate with student achievement (Robert L Linn & Baker, 1998)

The *preparation for teaching mathematics* scale combines responses ("not at all" to "very well prepared") on a series of questions regarding a teacher's pedagogical training in computations, problem solving, reasoning, assessment, etc. A separate scale combines scores for mathematics major, passed mathematics course work, and possession of mathematics license. These are indicators of a teacher's pre-service *mathematics preparation.*

In NAEP 4$^{th}$ grade mathematics analyses, *knowledge of NCTM standards* and *knowledge of state standards knowledge* were included in the analyses. Knowledge of NCTM standards is believed to result in teaching with a combination of conceptual and computational methods which will improve student achievement. According to Cohen and Hill, knowledge of state standards helps teachers align their instruction to the goals of their state which result in better scores on assessment (D. K. Cohen & Hill, 2001). Both are single variables based on a 4 point Likert scale from "no knowledge" to "very knowledgeable."

Reading Only- There is some overlap between reading, English, and language arts in these scales. Teacher characteristics included in the reading analyses are balanced literacy instruction, reading preparation, reading instruction preparation, literacy knowledge, and technology preparation related to reading/writing instruction. According the National Reading Panel, it is critically important to know what teacher characteristics influence successful reading instruction (National Reading Panel, 2000). The Panel targeted a combination of reading comprehension techniques.

*Balanced literacy orientation* reflects a moderate orientation toward reading instruction which includes emphasis on the mechanics of reading and writing as well as on comprehension. *Reading preparation* combines scores for reading/English major, passed courses in subject matter course work, and holding a license in reading/English/language arts.

*Reading/writing instruction preparation* measures how well teachers are prepared to teach vocabulary, oral and silent reading, comprehension activities, etc. Knowledge of language, literacy development, and writing processes as well as knowledge of literature is included in *Literacy knowledge*. *Technology preparation*, a scale encompassing teachers' preparation to use computers and software for reading/writing instruction, is included in the NAEP analyses.

**Professional Development.** Professional development and continuing education allow the teacher to hone current skills and knowledge, as well as keep abreast of new knowledge, theories, and methods (Borko & Putnam, 1996). Professional development has the ongoing potential to support and improve the existing teaching force. The National Reading Panel (National Reading Panel, 2000) found that in-service professional development produced significantly higher student achievement. For many in-service teachers, professional development was the first and only interface with computer technology in the classroom, one of the most important instructional innovations of the last decade (Prawat & Worthington, 1998; Rich, Sun, Velu, & Wallace, 1998).

Professional development is defined as a combination of mentoring experiences that teachers received and professional development they attended since completing formal teacher preparation. Scales for professional development impact, professional development frequency, and mentoring worth were developed as part of the original BTPS study. Of the three, only professional development frequency was available in NAEP Math and Reading across study years.

Mentoring scales from BTPS are highly correlated with each other $r^2$ =.936,

$p<.001$), a strong indication that the scales are measuring the same construct. Only *mentoring frequency* was chosen for the analyses since it had the fewest missing cases. *Professional development impact* includes ratings of how helpful teachers found various activities commonly covered in professional development, such as using a wider range of strategies to help students learn or managing the classroom more effectively. How much time teachers spent in professional development activities is measured by the *professional development frequency* scale. This scale is available across all data sets.

**Teacher Practices**. Teacher practices in mathematics and reading are assumed to be the most important influence on student learning. It is in the classroom where pedagogical content knowledge, the combination of subject matter knowledge and pedagogical skills, come into play. According to Shulman (1987), pedagogical content knowledge is more elusive than subject matter knowledge but equally important. It is the knowledge base, that allows a teacher to synthesize disciplinary knowledge and pedagogical knowledge into a cohesive lesson that students can understand. Pedagogical content knowledge also enables the teacher to elicit and/or circumvent common misconceptions surrounding the subject matter (Borko & Putnam, 1996).

<u>Mathematics Only.</u> As mentioned earlier, the 2000 NCTM Standards recommend a combination of strategies (both conceptual and computational) as important to student mathematics achievement. BTPS scales for teacher mathematics practices covered three important types of activities which include both conceptual and computational elements: mathematics group and project activities, mathematics drill and lecture activities, and mathematics sense-making activities. Frequencies for items measuring each activity were added together to form the three teacher practice scales.

*Mathematics group and project activities* combine items such as working in small groups to solve math problems and working on mathematical investigations that span several days. Items like lecturing or presenting information and drilling students on computational skills are included in *mathematics drill and lecture activities*. The *mathematics sense making activities* scale reflects the frequency of conceptual activities that teachers employ in the classroom to help students make sense of mathematics (i.e., having students make conjectures, leading discussions on ways students solve particular problems, applying math to real world problems, modeling different learning strategies, etc.)

The scale for drill and lecture activities is not available in NAEP Mathematics 1996. Since SEM requires three indicator variables, *emphasis on group and project assessment* was chosen since the other possible variables measuring teacher practices had already been incorporated into scales.

<u>Reading Only</u>. Reading practices measure the frequency that teachers use reform activities, traditional activities, computer activities, and literacy materials in there classrooms. *Reform activities* include comprehension activities, metacognitive routines for independent reading, oral and written responses to literature, etc. *Traditional activities* encompass activities such as oral and silent reading and reading vocabulary. Most teachers use both types of activities in the classroom. The scale for *literacy materials* is only available in BTPS and shows the range of textbook use from "none" to "a great amount." *Computer activities* in the NAEP analyses provides a connection from teacher technology preparation to student achievement.

**Student Achievement.** Data sets in this study have two different outcome measures for student mathematics and reading, gain scores in BTPS and multiple plausible values in NAEP. While student *gain scores* have been considered somewhat suspect due to the many influences besides teaching that can affect student scores over the course of a single year, gains over three years have been successfully employed as a barometer of growth in Tennessee (Sanders & Horn, 1998). Each student essentially serves as his own control group eliminating many family and societal factors external to the classroom. The Connecticut State Department of Education (CSDE) supplied two scores on the Connecticut Mastery Test per student, one from the beginning and one from the end of the 1999-2000 school year for all teachers involved in the BTPS. Gains scores were then computed subtracting pre-test results from post-test results.

The outcome variable in NAEP is a measure of achievement status, a one time score. NAEP is unusual in that no student completes the entire assessment. Using a complex system of weighting and matrix sampling, five *plausible values* representing a range of possible scores are created for each student based on like student patterns found in the data. For the SEM analyses, the five plausible values (MMPCM1-5 for mathematics or RRPCM1-5 for reading) are included as indicator variables for the outcome variable, student achievement. HLM handles the five plausible values simultaneously by running five separate analyses and then averaging the results.

# 6. METHODS & PROCEDURES

***Overview.*** All analyses were based on the conceptual model (Figure 1) and repeated separately for reading and mathematics across six data sets. Two primary types of data analysis software were used to confirm constructs, measure relative relationships to student achievement, determine differential effects by student demographics groups, and compare results for mathematics and reading across policy contexts: structural equation modeling (SEM) using AMOS 4.01 (Arbuckle, 1994-1999) and hierarchical linear modeling (HLM) using HLM 6.01h (Raudenbush, Bryk, & Congdon, 2005).

SEM was especially useful for analyses conducted at the teacher level with aggregated mean student gains/achievement standardized by grade as the outcome. It allowed estimation of the relationships among the latent variables and their indicators in order to validate constructs. SEM also provided standardized direct effects for teacher preparation, teacher characteristics, and professional development on teacher practices and indirect effects on student gains/achievement simultaneously. This answered questions regarding the relative strength of the relationships among the latent variables. The results from the SEM analyses were then compared across data sets.

Education data often has two- or sometimes three-level structures. In this study, students were nested within teachers who were then nested within teacher preparation programs. HLM was particularly appropriate because it solved the unit of analysis problem: each level in the structure was formally represented by its own submodel (Raudenbush & Bryk, 2002). HLM did not require that student data be aggregated to the teacher level, as in SEM, or teacher data to be replicated for each student at level one, as found in NAEP data. HLM also allowed the selection of the level of generalization for the analyses. This was particularly important for NAEP since the students were randomly sampled at level one, not the teachers at level two. Results can then be interpreted for a random sample of students who have teachers with certain characteristics.

Given the complexities of the educational system and the many intervening circumstances that impact teachers and students inside and outside the classroom, effect sizes are expected to be small. The primary criterion for judging effect size in education research, however, is its potential value for informing or benefitting education practice (Camilli et al., 2003; McCartney & Dearing, 2002; McCartney & Rosenthal, 2000). BTPS researchers anticipated effects sizes of no more than .05 over the course of one year and this is a potentially significant finding (Reckase, 2002; Valli et al., 2003).

Due to limitations in variable availability, it was not possible to run every model with every data set. Great care was taken to make scales across data sets as similar as possible. Table 4 shows the analyses that were performed and Table 5 indicates the variables available in each data set. Correlation matrices for all SEM models may be found in Appendix A.

**Table 6.1** Analysis Matrix

| | | CT | | TN | | NAEP 1996 | NAEP 1998 | NAEP 2000 | NAEP 2000 |
|---|---|---|---|---|---|---|---|---|---|
| Analysis | Unit of analysis | Math | Read | Math | Read | Math | Read | Math | Read |
| **SEM** | Teacher | x | x | x | x | x | x | x | x |
| **HLM** | Lvl 1 Student | x | x | NA | NA | x | x | x | x |
| | Lvl 2 Teacher | x | x | NA | NA | x | x | x | x |
| | Lvl 3 Teacher Preparatio n Program | x | x | NA | NA | NA | NA | NA | NA |

Correlation matrices served as input for all SEM analyses since the student achievement data from Tennessee was only available in this form. HLM analyses were not conducted with Tennessee data since there were no individual student scores that could be nested within teachers. Only the data from Connecticut was suitable for the three level analysis.

**Table 6.2** Variable Availability Across Data Sets

| Scales/Variables Available by Dataset | Math | | | | Read | | | |
|---|---|---|---|---|---|---|---|---|
| | CT | TN | NAEP4th 2000 | NAEP4th 1996 | CT | TN | NAEP4th 2000 | NAEP4th 1998 |
| **IHE Teacher Preparation** | | | | | | | | |
| Faculty Characteristics | x | x | NA | NA | x | x | NA | NA |
| Program Quality | x | x | NA | NA | x | x | NA | NA |
| Field Experience | x | x | NA | NA | x | x | NA | NA |
| **Teacher characteristics** | | | | | | | | |
| ***MATHEMATICS*** | | | | | | | | |
| Traditional Orientation | x | x | N/A | N/A | | | | |
| Reform Orientation | x | x | N/A | x | | | | |
| Math Preparation | x | x | x | x | | | | |
| Math Instruction Prep | x | x | x | x | | | | |
| Math knowledge | x | x | N/A | x | | | | |
| NCTM standards knowledge | * | * | x | * | | | | |
| State standards knowledge | * | * | x | * | | | | |
| ***BOTH*** | | | | | | | | |
| Efficacy | x | x | N/A | N/A | x | x | NA | NA |
| Diversity Prep | x | x | N/A | x | x | x | x | x |
| Assessment Preparation/Variety | x | x | N/A | N/A | x | x | x | x |
| ***READING*** | | | | | | | | |
| Balanced Literacy Orientation | | | | | x | x | NA | x |
| Reading/Writing Preparation | | | | | x | NA | x | x |
| Reading/Writing Instruction Prep | | | | | x | x | x | x |
| Literacy knowledge | | | | | x | x | NA | NA |
| Technology preparation | | | | | * | * | x | x |
| **Professional Development** | | | | | | | | |
| Mentoring worth | x | x | NA | NA | x | x | NA | NA |
| Professional development freq | x | x | x | x | x | x | x | x |
| Professional development impact | x | x | NA | NA | x | x | NA | NA |
| **Teacher Practice** | | | | | | | | |
| ***MATHEMATICS*** | | | | | | | | |
| Math Drill and Lecture activities | x | x | x | N/A | | | | |
| Math group and Project activities | x | x | x | x | | | | |
| Math sensemaking activities | x | x | x | x | | | | |
| Emphasize Group & Project Assessment | N/A | N/A | N/A | x | | | | |
| ***READING*** | | | | | | | | |
| Reform activities | | | | | x | x | x | x |
| Traditional Activities | | | | | x | x | x | x |
| Computer Activities | | | | | * | * | x | x |
| Literacy materials | | | | | x | x | x | x |
| **Gains/Achievement** | | | | | | | | |
| Gain on State Test | x | x | NA | NA | x | x | NA | NA |
| Plausible Values 1-5 | NA | NA | x | x | NA | NA | x | x |

* Computer activities were available but were not included for CT and TN

**Methodological issues unique to NAEP data analyses.** Methodological issues that need to be addressed when dealing with NAEP data are the appropriate uses of design weights and multiple plausible values. NAEP oversamples students from various categories to ensure sufficient sample size for within group analyses. Without proper weighting, these groups will be over-represented. Unfortunately, AMOS does not adjust for weighting. Using *origwt*[2] as the design weight, weighted correlation matrices were generated in AM Software (J. Cohen, 2002), a statistical package created specifically for use with NCES data sets.[3] These weighted correlation matrices were then suitable for input into SEM as excel spreadsheets.

In NAEP, students do not take all portions of the test but their probable total scores are imputed and captured by five multiple plausible values. Each of the five plausible values became an indicator variable for the latent variable for student achievement in the AMOS analyses. HLM 6 includes special estimation settings for design weights and multiple plausible values.[4] The program runs the same equation five times using each multiple plausible value as an outcome and then averages the results.

**Methodological issues unique to the Connecticut analyses.** Variation between schools in many areas (socioeconomic status, funding, urbanicity, teacher quality, etc.) is known to account for large portions of variation in student scores (D. K. Cohen & Hill,

---

[2]This is the weight for the national sample using *psamp,* the public school sample, no testing accommodations.

[3]The software and documentation for this program may be downloaded at http://am.air.org/.

[4]For descriptions of these processes, please see "Three additional features of HLM" at http://www.ssicentral.com/hlm/example6-2.html.

2001; Raudenbush, Kasim, Eamsukkawat, & Miyazaki, 1996). BTPS collected information on student race, grade, and gender but no measure of student SES. Consequently, district percent free and reduced lunch from the 1999-2000 school year was added to the teacher data as a measure of district socioeconomic status. The information came from the Connecticut Department of Education website.[5]

[5]Data originally comes from the Common Core of Data, NCES.

## 7. STRUCTURAL EQUATION MODELING RESULTS

The SEM analysis is not only intended to address the strengths and weaknesses of the conceptual model across data sets, but also the research questions about relationships among teacher preparation, teacher characteristics, and professional development and their relative contributions to teacher practice and subsequent student achievement in mathematics and reading.

Teachers are the unit of analysis for SEM analyses. Student gains for Connecticut and Tennessee have been aggregated to the teacher level and standardized by grade to control for grade level effects. The five plausible values in NAEP have been aggregated by teacher for these analyses. No standardization by grade was necessary since all students are in the $4^{th}$ grade. The five aggregated plausible values serve as indicator variables for a latent variable labeled student achievement.

Using a recursive hybrid structural equation model, it is possible to judge the adequacy of the conceptual model as a base for SEM analysis, determine the size of the relationships among the latent and indicator variables, and assess their indirect and direct effects on student achievement/gains across two curriculum contexts, mathematics and reading. The hybrid model includes measurement models for each of the latent variables and their indicators as well as the overall structural model based on the conceptual model (Figure 1). To assess the adequacy of the measurement models, variance explained ($r^2$) indicates the extent to which each indicator variable measures its underlying construct (Byrne, 1998). Generally, $r^2 > .35$ is considered large, $r^2 > .15$ is moderate, and $r^2 > .02$ is small in structural equation modeling (J. Cohen, 1988). Overall model fit for the structural model is determined using cmin ($\chi^2$)/df <3, GFI (Joreskog)>.90, Pratio (parsimony index>.60), and RMSEA<.05.

Results are reported as standardized effects to make it easier to compare results across models. AMOS uses Monte Carlo estimation when data is input as a correlation matrix. The procedure also calculates bias corrected significance levels and confidence intervals. While the measurement and structural models are expected to retain high goodness-of-fit properties across data sets, it is also expected that the models will show sensitivity to policy and curriculum contexts at the state and national levels. Alternative models were analyzed in Connecticut, but only the best-fitting model is presented below. The identical model was used for the Tennessee analyses.

**Connecticut Base Model for Mathematics**

There are four latent exogenous variables indicated by ovals, each with at least three observed endogenous variables (indicated by rectangles). Disturbance terms (ex., e GPA for group and project activities) are included for all endogenous variables. One coefficient for each latent variable is constrained to 1 as required by SEM. In the case of reform and traditional orientation, both were set to 1 since they are believed to exert equal influence on teacher characteristics. Teacher preparation, teacher characteristics, and professional development are correlated with each other as indicated by the curved arrows.

Figure 7.1 Structural Equation Model for Mathematics in Connecticut and Tennessee

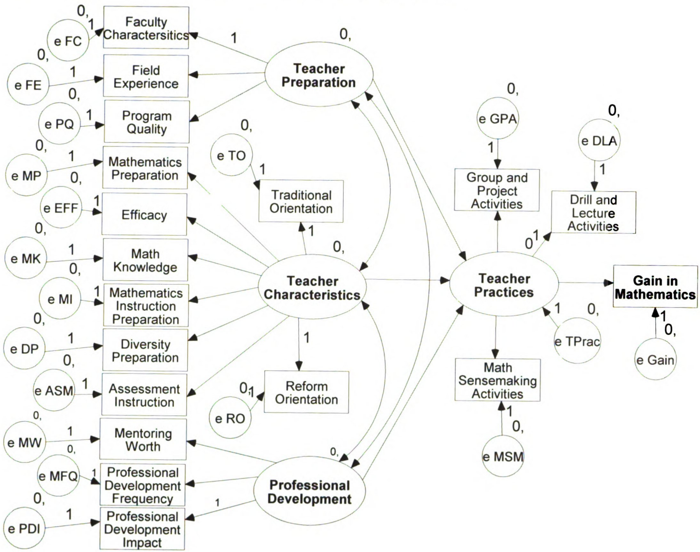


Table 7.1 presents the standardized total effects for the Connecticut mathematics model. Variance explained is indicated by $r^2$. Tennessee results follow in Table 7.3.

**Table 7.1** Connecticut Standardized Total Effects for Mathematics

| | $r^2$ | Teacher Preparation | Teacher Characteristics | Professional Development | Teacher Practices |
|---|---|---|---|---|---|
| Teacher Preparation | | | | | |
| Faculty Characteristics | 0.314 | 0.560 * | | | |
| Program Quality | **0.592** | 0.769 * | | | |
| Field Experience | 0.200 | 0.447 ** | | | |
| Teacher Characteristics | | | | | |
| Traditional Orientation | 0.004 | | 0.063 * | | |
| Reform Orientation | 0.009 | | 0.095 * | | |
| Math Preparation | 0.013 | | 0.113 | | |
| Math Instruction Preparation | **0.630** | | 0.794 * | | |
| Diversity Preparation | 0.313 | | 0.559 ** | | |
| Assessment Preparation | **0.578** | | 0.760 ** | | |
| Math Knowledge | 0.155 | | 0.393 * | | |
| Efficacy | 0.115 | | 0.339 * | | |
| Professional Development | | | | | |
| Mentoring worth | **0.372** | | | 0.610 * | |
| Professional Development Frequency | 0.303 | | | 0.551 * | |
| Professional Development Impact | **0.776** | | | 0.881 ** | |
| Teacher Practices | 0.326 | -0.660 * | 1.017 ** | 0.166 | |
| Math Drill & Lecture Activities | 0.114 | -0.223 * | 0.344 ** | 0.056 | 0.338 ** |
| Math Group & Project Activities | **0.691** | -0.549 * | 0.845 ** | 0.138 | 0.831 * |
| Mathsensemaking | **0.371** | -0.402 * | 0.619 ** | 0.101 | 0.609 ** |
| Standardized Math Gains | 0.002 | -0.028 | 0.043 | 0.007 | 0.042 |
| Fit Indices | | cmin/df | NFI | RMSEA | PRATIO |
| | | 2.908*** | 0.984 | 0.107 | 0.760 |

p<.01*, p<.001**, r2>.35 is bold

**Discussion of Connecticut mathematics results**:

***Measurement model***. Given the complexity of the education system, the theory and research undergirding the conceptual model and constructs, and the rather small sample for Connecticut, the four measurement models were not respecified based on $r^2$ values since overall model fit was high. The construct for teacher preparation accounts for 59.2% of the variance in program quality, 31.4% of the variance in faculty characteristics, and 20% of the variation in field experiences. Of the eight teacher characteristics indicators, the largest $r^2$ are reported for mathematics instruction preparation (pedagogy) which explains 63% of the variation and preparation for assessment which explains 57.8% of the variation in teacher characteristics in Connecticut.

Math preparation, the indicator of subject matter preparation, does not explain any of the variance in Connecticut teacher characteristics ($r^2$ =.013)· Similarly low $r^2$ using Tennessee data was noted in Valli, Reckase, & Raths (Valli et al., 2003). Reform orientation and traditional orientation are candidates for deletion due to low $r^2$. These three variables, however, are too conceptually important to discard and they will also be included in the later HLM analyses. Consequently, they are kept for comparison purposes.

The construct for professional development explains 77% of the variation in professional development impact, 37.2% of mentoring worth, and 30% of the variation in professional development frequency. The construct for teacher practice explains 69.1% of the variation in a teachers use of group and project activities and 37.1% of the variance in math sensemaking activities ($r^2$=.371). The percent of variance explained for drill and lecture activities is small (11.4%).

On the whole, the latent variables account for moderate to large amounts of variance in this model and the literature supports the connections.

***Structural model.*** The model for Connecticut could be improved. For the purposes of this research, however, it was deemed more important to rely on theory to develop a model that would work well across multiple data sets, rather than tailor it too specifically for a single state. The overall fit of the Connecticut mathematics model is good (cmin($\chi^2$)/df <2.908, p<.000), GFI>.984, Pratio>.760). RMSEA is a little high (.107). Additional relationships among the indicator variables for professional development and the indicator variables for the other latent variables are the most likely source of misfit. The correlation between teacher preparation and teacher characteristics is high (r=.87) suggesting that these may not be separate constructs for this data.

. ***Effects.*** Generally effect sizes above .10 are considered small, above .30 are considered moderate, and above .50 are considered large (Kline, p. 98). This is a relative rather than absolute rule of thumb. As mentioned earlier, an effect of 5% is deemed potentially significant over the course of one academic year (Sanders & Rivers, 1996; Valli, Raths, & Rennert-Ariev, 2001; Valli et al., 2003).

Teacher preparation and teacher characteristics have large effects on teacher practice. The link between teacher preparation and classroom practice, however, was negative ($\delta$=-.66), corroborating the findings in (Valli et al., 2003). They suggest that elementary teachers, the largest proportion of teachers in the BTPS sample, are not as likely to have majored in mathematics at the undergraduate or graduate level since it was not a requirement for teaching elementary mathematics in 1999. The effect of teacher characteristics on teacher practices is $\delta$=1.017. Professional development produces a small, positive effect on teacher practice ($\delta$=.166). Teacher practice has a small, positive direct mathematics effect on gains ($\delta$=.042). The negative direct effects of teacher preparation on teacher practice translated into a small, negative indirect effect on student achievement ($\delta$=-.028).

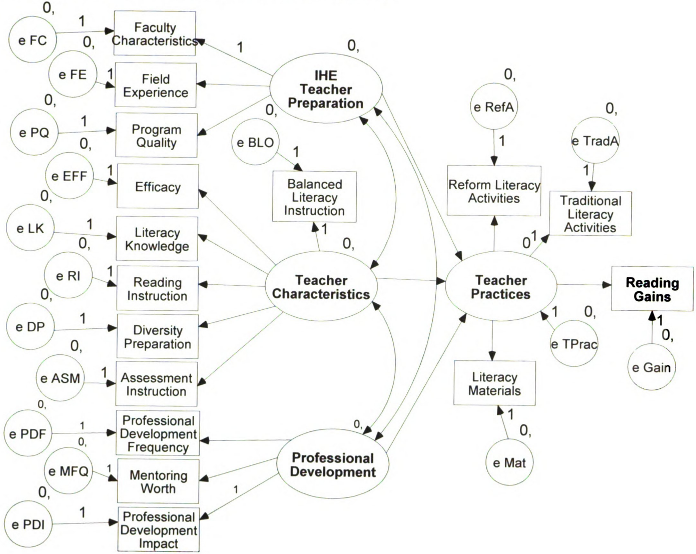
The hybrid structural equation model has good measurement and structural fit.

Effect sizes are reasonable given theory and prior research. The model was adapted to analyze reading in Connecticut.

**Connecticut base model for reading**

Once it was clear that the structural equation model (Figure 2) adequately fit the mathematics data, the same structural model was fit to Connecticut reading data. Elements of the education system, represented by the latent variables, are presumed to be related in similar ways across subject matters in a strongly centralized state such as Connecticut. The measurement model is slightly different since it was necessary to substitute indicator variables that were pertinent to reading.

**Figure 7.2** Structural Equation Model for Reading in Connecticut and Tennessee



Results of the SEM analysis for reading in Connecticut are presented in Table 7.2. Results for Tennessee follow in Table 7.3.

**Table 7.2.** Connecticut Standardized Total Effects for Reading

| | $r^2$ | Teacher Preparation | Teacher Characteristics | Professional Development | Teacher Practices |
|---|---|---|---|---|---|
| Teacher Preparation | | | | | |
| Faculty Characteristics | **0.368** | 0.607 * | | | |
| Program Quality | **0.461** | 0.679 * | | | |
| Field Experience | 0.231 | 0.481 * | | | |
| Teacher Characteristics | | | | | |
| Balanced Literacy Orientation | 0.025 | | 0.157 * | | |
| Reading Instruction Preparation | **0.832** | | 0.912 ** | | |
| Diversity Preparation | **0.350** | | 0.591 ** | | |
| Assessment Preparation | **0.697** | | 0.835 * | | |
| Literacy Knowledge | 0.324 | | 0.569 * | | |
| Efficacy | 0.067 | | 0.259 * | | |
| Professional Development | | | | | |
| Mentoring Worth | 0.301 | | | 0.549 | |
| Professional Development Frequency | 0.094 | | | 0.307 * | |
| Professional Development Impact | **0.795** | | | 0.891 ** | |
| Teacher Practices | 0.144 | 0.274 | 0.010 | 0.163 | |
| Reform Literacy Activities | **0.438** | 0.181 | 0.006 | 0.108 | 0.662 * |
| Traditional Literacy Activities | **0.359** | 0.164 | 0.006 | 0.098 | 0.599 * |
| Literacy Materials | 0.006 | 0.021 | 0.001 | 0.012 | 0.075 ** |
| Reading Gains | 0.081 | 0.078 | 0.003 | 0.047 | 0.285 * |

| Fit Indices | cmin/df | NFI | RMSEA | PRATIO |
|---|---|---|---|---|
| | 3.264*** | 0.985 | 0.116 | 0.721 |

p<.01*, p<.001**, r2>.35 is bold

**Connecticut Reading Results**

***Measurement Model.*** Teacher preparation explains less of the variance in program quality (46.1%) but more of the variance in faculty characteristics (36.8%) and field experience (23.1%) than in the CT mathematics analysis. Teacher characteristics explain 83.2% of the variance in reading instruction preparation and 69.7% of the variance in assessment preparation, both larger than the $r^2$ for corresponding mathematics variables. Professional development impact accounts for the largest amount of variance

(79.5%) in CT reading. Teacher practice explains 43.8% of the variance in reform reading activities and 35.9% of the variance in traditional reading activities. No variance is explained for literacy materials.

***Structural model***. The overall fit of the CT reading model is good (cmin($\chi^2$)/df <3.264, p<.000), GFI>.985, Pratio>.721) although the RMSEA is high (.116). A review of the residual covariance matrix shows additional unanalyzed relationships among indicators for each of the latent variables. The correlation between teacher preparation and teacher charactersitics is high (r=.83) indicating these may not be separate constructs for this data. In spite of this, the Connecticut models for reading and mathematics have many good properties and appear to respond similarly in the analyses.

***Effects.*** When looking at the pattern of effect sizes in the structural model for reading, teacher preparation, teacher characteristics, and professional development have only moderate effects on teacher practice. The indirect effects of teacher practice on student reading gains are small but closer to the predicted level of $\delta$=.05 than for mathematics. One notable difference by subject matter is the direct effect of teacher practice on reading gains ($\delta$=.285) compared to mathematics ($\delta$=.042).

The hybrid structural equation model fits both mathematics and reading data well indicating that the latent variables are reasonably well-measured given theory and the structural model is performing as anticipated across the Connecticut education system. Effects appear sensitive to subject matter context. The next step is to confirm the model in the remaining data sets. Results for Tennessee mathematics and reading come first followed by the results for NAEP.

**Tennessee Mathematics and Reading Results:**

Tennessee models for mathematics and reading are identical to the Connecticut models presented in Figures 7.1 and 7.2. Table 7.3 includes $r^2$ and standardized effects for mathematics and reading for Tennessee and NAEP. Tables 7.4 and 7.5 which present results for all data sets, include the estimates of direct and indirect effects for mathematics and reading in Tennessee.

***Measurement Model.*** In general, there is more variation explained across constructs in Tennessee than in Connecticut. Teacher preparation program explains more variance in faculty characteristics (mathematics=75%, reading=76.3%), program quality (mathematics=49.5%, reading=48.8%), and field experience (mathematics=27.8%, reading=27.3%) in mathematics and reading than in CT. Teacher characteristics explain the large amounts of variation in the variables for preparation to teach a particular subject (mathematics=66.7%, reading=71%), diversity preparation (mathematics=45%, reading=42.4%), and assessment preparation (mathematics=63.7%, reading=68.8%), all components of teachers' educational background. Professional development explains the large amounts of variation in professional development impact (mathematics=86.2%, reading=73.9%), followed by mentoring worth (mathematics=26.8%, reading=29.3%), and professional development frequency (mathematics=24%, reading=28.6%)

***Structural Model.*** The overall fit of the TN reading model was good (cmin($\chi^2$)/df <4.962, p<.000), GFI>.974, Pratio>.76) and the RMSEA is lower (.095) than CT although still above the desired level of .05. As in CT, a review of the residual covariance matrix shows additional unanalyzed relationships among indicators for each of the latent variables. The correlation between teacher preparation and teacher characteristics is again high (mathematics r = .87; reading r = .85) but similar to the results for Connecticut. The Tennessee models for reading and mathematics have many good properties and responded similarly to the Connecticut analyses.

***Effects.*** The patterns of coefficient sign and size in the structural model for Tennessee begin to show differences from the Connecticut mathematics and reading models. For example, the effect of teacher mathematics preparation on teacher practice is considerably smaller although still negative (CT $\delta$=-.660, $p<.05$; TN $\delta$=-.111). A more complete discussion of these differences will follow in the data analysis results across data sets.

The same hybrid structural equation model fit mathematics and reading data well in both Connecticut and Tennessee indicating that the latent variables were reasonably well-measured given theory and the structural model was performing as anticipated across the states. Effects appeared sensitive to subject matter context. The next step was to run similar models using NAEP data.

**NAEP Mathematics and Reading Models.**

Figures 5 to 8 present the models for NAEP 4$^{th}$ grade mathematics and reading. NAEP questionnaires changed between administrations and subject matter from 1996 to 2000. Consequently it was not possible to use the same variables across all NAEP models. Reasonable substitutions were made when possible but sometimes different variables all together were incorporated to fulfill specifications necessary to run SEM.[6]

Comparisons of NAEP results with those of Connecticut and Tennessee follow in Tables 7.3-7.6. Results are discussed as part of the cross data analysis.

---

[6]Each latent variable needs to have at least three indicator variables.

**Figure 7.3** 1996 NAEP 4th Grade Mathematics Structural Equation Model

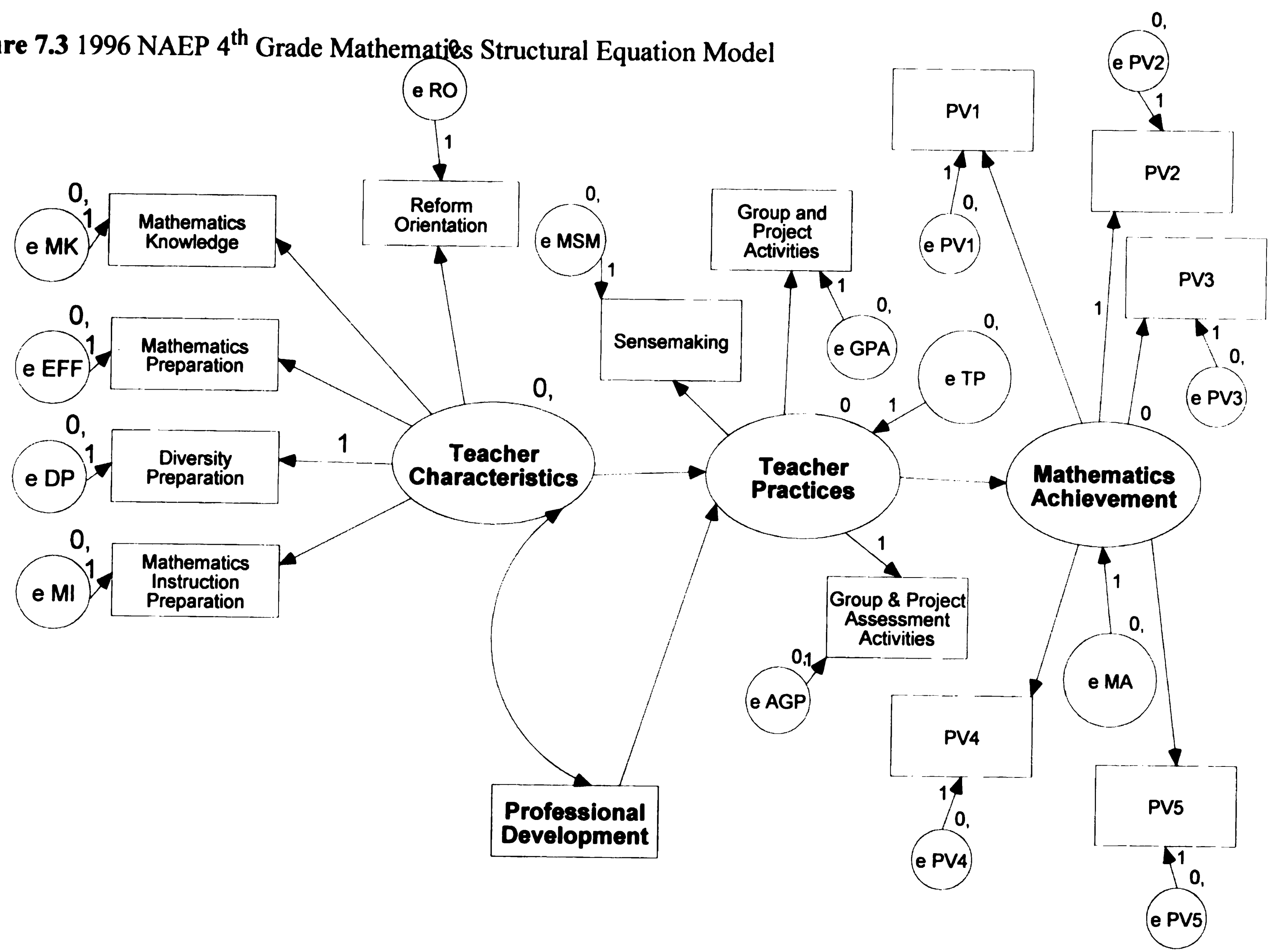


**Figure 7.4** 2000 NAEP 4th Grade Mathematics Structural Equation Model

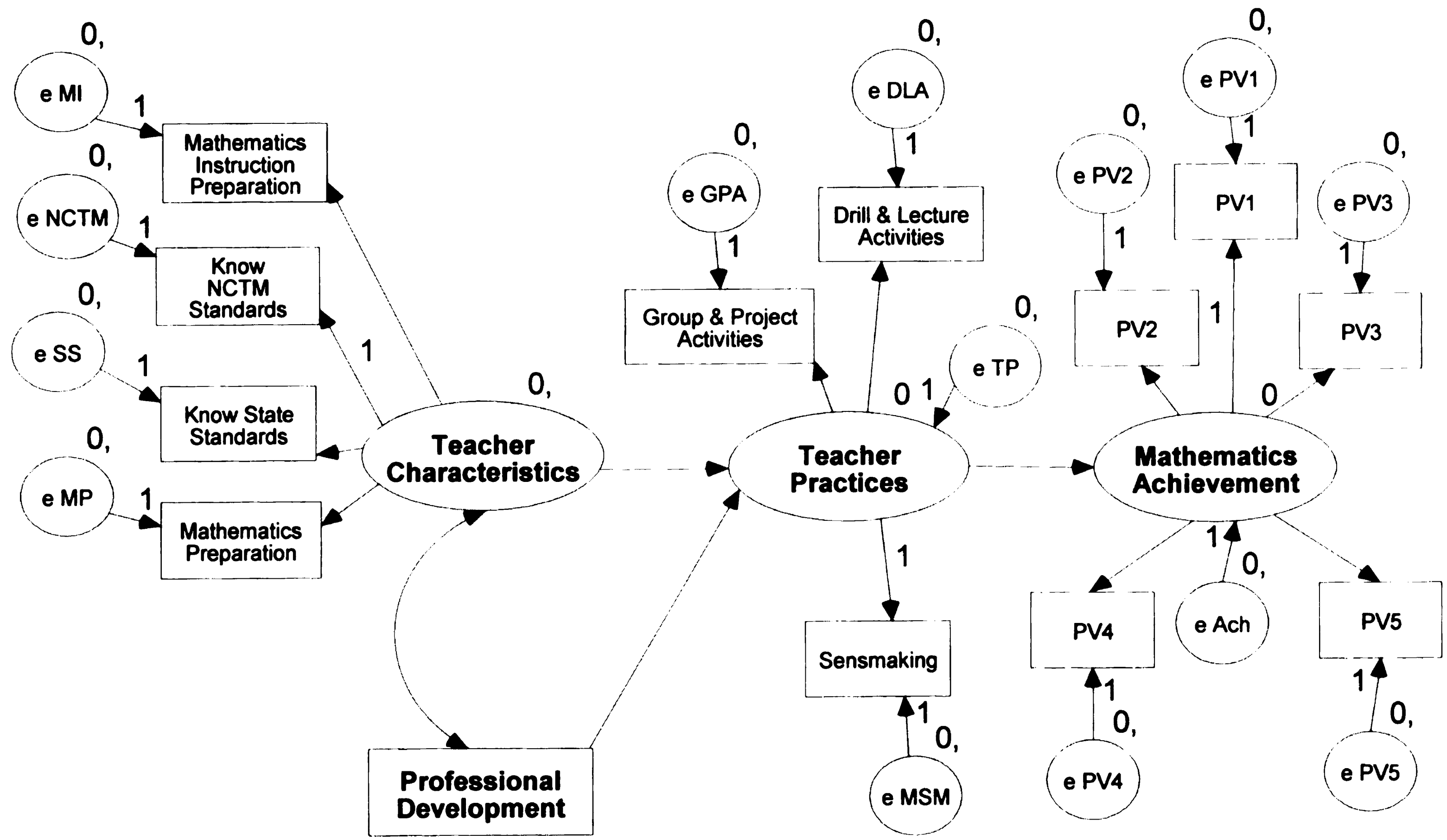


**Figure 7.5** 1998 NAEP 4th Grade Reading Structural Equation Model

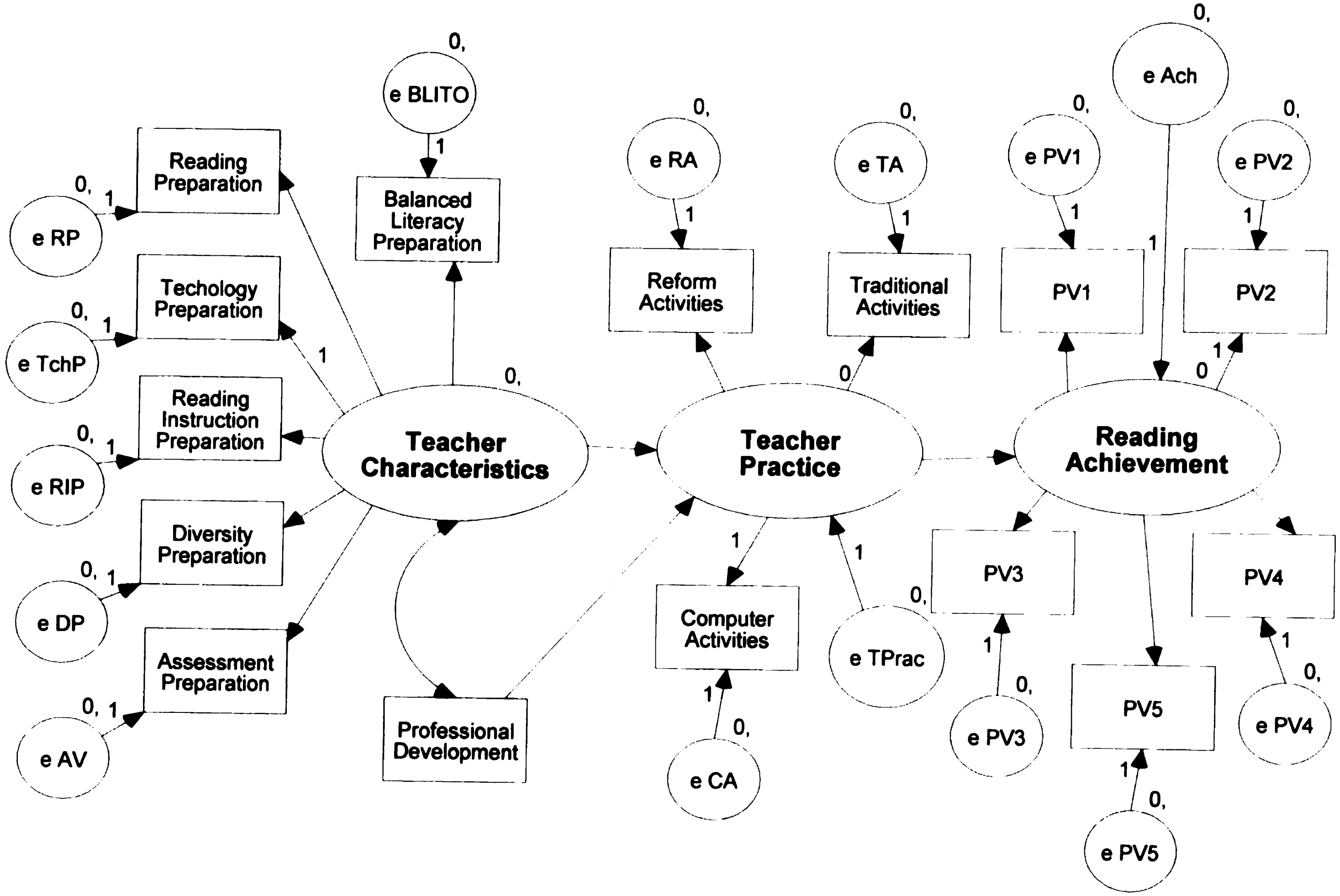


**Figure 7.6.** 2000 NAEP 4th Grade Reading Structural Equation Model

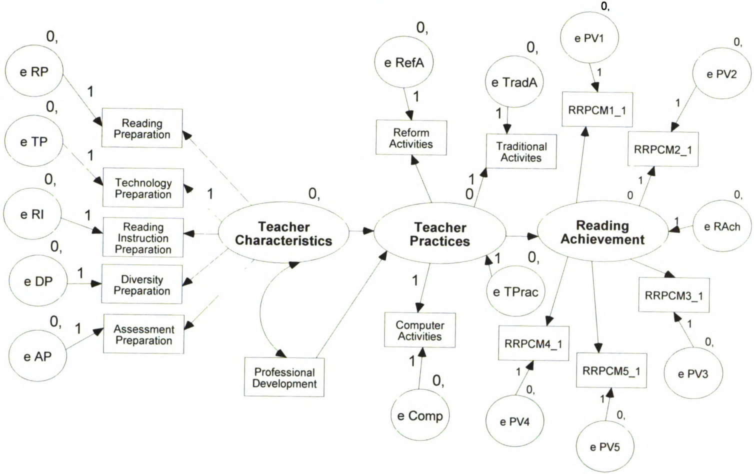


**Table 7.3** Comparison of $r^2$ for Mathematics and Reading Constructs Across Data Sets

| Teacher n<br>Variables<br>**MATHEMATICS - $r^2$** | 168<br>CT | 442<br>TN | 1029<br>NAEP<br>2000 | 584<br>NAEP<br>1996 | Teacher n<br>Variables<br>**READING - $r^2$** | 168<br>CT | 442<br>TN | 867<br>NAEP<br>2000 | 757<br>NAEP<br>1998 |
|---|---|---|---|---|---|---|---|---|---|
| **IHE Teacher Preparation** | | | | | **IHE Teacher Preparation** | | | | |
| Faculty Characteristics | 0.314 | 0.495 | | | Faculty Characteristics | 0.368 | 0.488 | | |
| Program Quality | **0.592** | **0.750** | | | Program Quality | **0.461** | **0.763** | | |
| Field Experience | 0.200 | 0.278 | | | Field Experience | 0.231 | 0.273 | | |
| **Teacher characteristics** | | | | | **Teacher characteristics** | | | | |
| Traditional Orientation | 0.004 | 0.020 | | | Balanced Literacy Orientation | 0.025 | 0.044 | | 0.176 |
| Reform Orientation | 0.009 | 0.037 | | 0.164 | Reading/Writing Preparation | | | 0.019 | 0.017 |
| Math Preparation | 0.013 | 0.000 | 0.027 | 0.088 | Technology Prep | | | 0.215 | 0.171 |
| Math Instruction Prep | **0.630** | **0.667** | **0.392** | **0.392** | Reading/Writing Inst Prep | **0.832** | **0.710** | **0.364** | **0.231** |
| Diversity Prep | **0.313** | **0.450** | | **0.540** | Diversity Prep | **0.350** | **0.424** | **0.339** | **0.187** |
| Assessment Prep | **0.578** | **0.637** | | | Assessment Prep | **0.697** | **0.688** | **0.325** | **0.361** |
| Math Knowledge | 0.155 | 0.075 | | 0.076 | Literacy Knowledge | **0.324** | **0.304** | | |
| Efficacy | 0.115 | 0.055 | | | Efficacy | 0.067 | 0.050 | | |
| NCTM standards knowledge | | | 0.306 | | | | | | |
| State standards knowledge | | | 0.314 | | | | | | |
| **Professional Development** | | | | | **Professional Development** | | | | |
| Mentoring worth | 0.372 | 0.268 | | | Mentoring worth | 0.301 | 0.293 | | |
| Professional development freq | 0.303 | 0.240 | | | Professional development freq | 0.094 | 0.286 | | |
| Professional development impact | **0.776** | **0.862** | | | Professional development impact | 0.795 | 0.739 | | |
| **Teacher Practice** | **0.326** | **0.237** | **0.296** | **0.397** | **Teacher Practice** | | **0.244** | **0.542** | **0.937** |
| Math Drill and Lecture activities | 0.114 | 0.080 | 0.006 | | Reform Activities | **0.438** | **0.534** | **0.333** | **0.704** |
| Math group and Project activities | 0.691 | 0.270 | 0.360 | 0.404 | Traditional Activities | **0.359** | **0.350** | **0.704** | **0.544** |
| Math sensemaking activities | 0.371 | 0.944 | 0.655 | 0.638 | Computer Activities | | | 0.168 | 0.115 |
| Emphasize GPA assessment | | | | 0.606 | Literacy materials | 0.006 | 0.061 | | |
| **Math Gains/Achievement** | **0.002** | **0.029** | **0.012** | **0.003** | **Reading Achievement/Gains** | **0.081** | **0.000** | **0.000** | **0.016** |
| PV1 | | | 0.945 | 0.966 | PV1 | | | 0.900 | 0.913 |
| PV2 | | | 0.944 | 0.961 | PV2 | | | 0.925 | 0.904 |
| PV3 | | | 0.951 | 0.953 | PV3 | | | 0.914 | 0.910 |
| PV4 | | | 0.945 | 0.954 | PV4 | | | 0.914 | 0.906 |
| PV5 | | | 0.953 | 0.966 | PV5 | | | 0.924 | 0.905 |

**Table 7.4 Standardized Total Effects for Mathematics Across Data Sets**

| | IHE Teacher Preparation | | | | Teacher characteristics | | | | Professional development | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Teacher n | 168 | 442 | 1029 | 584 | 168 | 442 | 1029 | 584 | 168 | 442 | 1029 | 584 |
| | | | NAEP 4th | NAEP 4th | | | NAEP 4th | NAEP 4th | | | NAEP 4th | NAEP 4th |
| Variables | CT | TN | 2000 | 1996 | CT | TN | 2000 | 1996 | CT | TN | 2000 | 1996 |
| **IHE Teacher Preparation** | | | | | | | | | | | | |
| Faculty Characteristics | 0.560* | 0.704* | x | x | | | | | | | | |
| Program Quality | 0.769* | 0.866* | x | x | | | | | | | | |
| Field Experience | 0.447** | 0.527* | x | x | | | | | | | | |
| **Teacher characteristics** | | | | | | | | | | | | |
| Traditional Orientation | | | | | 0.063* | 0.140** | | | | | | |
| Reform Orientation | | | | | 0.095* | 0.191** | | 0.405* | | | | |
| Math Preparation | | | | | 0.113 | 0.010** | 0.166 | 0.297* | | | | |
| Math Instruction Prep | | | | | 0.794* | 0.817 | 0.621* | 0.275* | | | | |
| Diversity Prep | | | | | 0.559** | 0.671** | | 0.735** | | | | |
| Assessment Prep | | | | | 0.760** | 0.798** | | | | | | |
| Math Knowledge | | | | | 0.393* | 0.275** | | 0.626* | | | | |
| Efficacy | | | | | 0.339* | 0.233** | | | | | | |
| NCTM standards knowledge | | | | | | | 0.554* | | | | | |
| State standards knowledge | | | | | | | 0.560* | | | | | |
| **Professional Development** | | | | | | | | | | | | |
| Mentoring worth | | | | | | | | | 0.610* | 0.518** | | |
| Professional development frequency | | | | | | | | | 0.551* | 0.490** | | |
| Professional development impact | | | | | | | | | 0.881** | 0.928* | | |
| **Teacher Practice** | **-0.660*** | **-0.111** | **x** | **x** | **1.017**** | **0.514*** | **0.421**** | **0.442*** | **0.166** | **0.118*** | **0.238**** | **0.339**** |
| Math Drill and Lecture activities | -0.223 | -0.031 | x | x | 0.344** | 0.145** | -0.032* | | 0.056 | 0.033* | -0.018* | |
| Math group and Project activities | -0.549* | -0.057 | x | x | 0.845** | 0.267** | 0.252** | 0.281* | 0.139 | 0.061* | 0.143** | 0.215* |
| Math sensemaking activites | -0.402* | -0.107 | x | x | 0.619** | 0.499** | 0.341** | 0.353* | 0.101 | 0.115~ | 0.193* | 0.271** |
| Emphasize GPA assessment | | | x | x | | | | 0.344* | | | | 0.264* |
| **Math Gains/Achievement** | **-0.028** | **-0.019** | **x** | **x** | **0.043** | **0.087** | **0.046**** | **0.022**** | **0.007** | **0.020*** | **0.026**** | **0.017**** |
| PV1 | x | x | x | x | x | x | 0.044** | 0.022** | x | x | 0.025** | 0.017 |
| PV2 | x | x | x | x | x | x | 0.044** | 0.022** | x | x | 0.025** | 0.017 |
| PV3 | x | x | x | x | x | x | 0.044** | 0.022** | x | x | 0.025** | 0.017 |
| PV4 | x | x | x | x | x | x | 0.044** | 0.022** | x | x | 0.025** | 0.017 |
| PV5 | x | x | x | x | x | x | 0.044** | 0.022** | x | x | 0.025** | 0.017 |

**Table 7.4. (cont'd).**

| | Teacher Practice | | | |
|---|---|---|---|---|
| Teacher n | 168 | 442 | 1029 | 584 |
| | | | NAEP 4th | NAEP 4th |
| Variables | CT | TN | 2000 | 1996 |
| **IHE Teacher Preparation** | | | | |
| Faculty Characteristics | | | | |
| Program Quality | | | | |
| Field Experience | | | | |
| **Teacher characteristics** | | | | |
| Traditional Orientation | | | | |
| Reform Orientation | | | | |
| Math Preparation | | | | |
| Math Instruction Prep | | | | |
| Diversity Prep | | | | |
| Assessment Prep | | | | |
| Math Knowledge | | | | |
| Efficacy | | | | |
| NCTM standards knowledge | | | | |
| State standards knowledge | | | | |
| **Professional Development** | | | | |
| Mentoring worth | | | | |
| Professional development frequency | | | | |
| Professional development impact | | | | |
| **Teacher Practice** | | | | |
| Math Drill and Lecture activities | 0.338** | 0.282** | -0.075* | |
| Math group and Project activities | 0.831* | 0.520* | 0.600** | 0.636* |
| Math sensemaking activites | 0.609** | 0.972** | 0.809* | 0.798* |
| Emphasize GPA assessment | | | | 0.778* |
| **Math Gains/Achievement** | **0.042** | **0.170*** | **0.108**** | **0.051** |
| PV1 | x | x | 0.105** | 0.050** |
| PV2 | x | x | 0.105** | 0.050** |
| PV3 | x | x | 0.106** | 0.049** |
| PV4 | x | x | 0.105** | 0.050** |
| PV5 | x | x | 0.106** | 0.049** |

**Significance**
*p<.05
**p<.01
***p<.001
~<.07

| **Fit Indices** | **CT** | **TN** | **1996** | **2000** |
|---|---|---|---|---|
| cmin/df | 2.908*** | 4.962*** | 3.391*** | 1.463** |
| NFI | 0.984 | 0.974 | 0.996 | 0.997 |
| Pratio | 0.760 | 0.760 | 0.681 | 0.705 |
| RMSEA | 0.107 | 0.095 | 0.048 | 0.028 |

**Table 7.5 Standardized Total Effects for Reading Across Data sets**

| | IHE Teacher Preparation | | | | Teacher characteristics | | | | Professional development | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Teacher n | 168 | 442 | 867 | 757 | 168 | 442 | 867 | 757 | 168 | 442 | 867 | 757 |
| | CT | TN | NAEP 4th | NAEP 4th | CT | TN | NAEP 4th | NAEP 4th | CT | TN | NAEP 4th | NAEP 4th |
| Variables | | | 2000 | 1998 | | | 2000 | 1998 | | | 2000 | 1998 |
| **IHE Teacher Preparation** | | | | | | | | | | | | |
| Faculty Characteristics | 0.607* | 0.698* | x | x | | | | | | | | |
| Program Quality | 0.679* | 0.873** | x | x | | | | | | | | |
| Field Experience | 0.481* | 0.522** | x | x | | | | | | | | |
| **Teacher characteristics** | | | | | | | | | | | | |
| Balance Literacy Orientation | | | | | 0.157* | 0.211* | | 0.419** | | | | |
| Reading/Writing Preparation (mjrs) | | | | | 0.113 | | 0.138* | 0.129* | | | | |
| Technolgy Prep | | | | | | | 0.463* | 0.413* | | | | |
| Reading/Writing Instruction Prep | | | | | 0.912** | 0.843** | 0.603* | 0.481** | | | | |
| Diversity Prep | | | | | 0.591** | 0.651** | 0.582* | 0.432* | | | | |
| Assessment Prep | | | | | 0.835** | 0.830** | | | | | | |
| Assessment Variety | | | | | | | 0.570** | 0.601** | | | | |
| Literacy Knowledge | | | | | 0.569 | 0.551* | | | | | | |
| Efficacy | | | | | 0.259* | 0.225* | | | | | | |
| **Professional Development** | | | | | | | | | | | | |
| Mentoring worth | | | | | | | | | 0.549~ | 0.541 | | |
| Professional development frequency | | | | | | | | | 0.307* | 0.535* | | |
| Professional development impact | | | | | | | | | 0.891** | 0.86* | | |
| **Teacher Practice** | **0.274** | **0.367** | **x** | **x** | **0.010** | **-0.127** | **0.727**** | **0.991*** | **0.163** | **0.118*** | **0.026** | **-0.071~** |
| Reform Activities | 0.181 | 0.268 | x | x | 0.006 | -0.093 | 0.61** | 0.831** | 0.108 | 0.237** | 0.022** | -0.059~ |
| Traditional Activites | 0.164 | 0.219 | x | x | 0.006 | -0.075 | 0.420** | 0.731* | 0.098 | 0.192* | 0.015** | -0.052~ |
| Computer Activities | | | x | x | | | 0.298 | 0.336* | | | 0.011** | -0.024* |
| Literacy materieals | 0.021 | 0.090 | x | x | 0.001 | -0.031 | | | 0.012 | 0.080** | | |
| **Reading Achievement/Gains** | **0.078** | **-0.006** | **x** | **x** | **0.003** | **0.002** | **-0.085** | **-0.123**** | **0.047** | **-0.005** | **0.000** | **0.009*** |
| PV1 | x | x | x | x | x | x | -0.008** | -0.118** | x | x | 0.000 | 0.008* |
| PV2 | x | x | x | x | x | x | -0.008** | -0.117** | x | x | 0.000 | 0.008* |
| PV3 | x | x | x | x | x | x | -0.008** | -0.118** | x | x | 0.000 | 0.008* |
| PV4 | x | x | x | x | x | x | -0.008** | -0.118** | x | x | 0.000 | 0.008* |
| PV5 | x | x | x | x | x | x | -0.008** | -0.117** | x | x | 0.000 | 0.008* |

**Table 7.5 (cont'd).**

| | Teacher Practice | | | |
|---|---|---|---|---|
| Teacher n | 168 | 442 | 867 | 757 |
| | CT | TN | NAEP 4th | NAEP 4th |
| Variables | | | 2000 | 1998 |
| **IHE Teacher Preparation** | | | | |
| Faculty Characteristics | | | | |
| Program Quality | | | | |
| Field Experience | | | | |
| **Teacher characteristics** | | | | |
| Balance Literacy Orientation | | | | |
| Reading/Writing Preparation (mjrs) | | | | |
| Technolgy Prep | | | | |
| Reading/Writing Instruction Prep | | | | |
| Diversity Prep | | | | |
| Assessment Prep | | | | |
| Assessment Variety | | | | |
| Literacy Knowledge | | | | |
| Efficacy | | | | |
| **Professional Development** | | | | |
| Mentoring worth | | | | |
| Professional development frequency | | | | |
| Professional development impact | | | | |
| **Teacher Practice** | | | | |
| Reform Activities | 0.662* | 0.731* | 0.839** | 0.839** |
| Traditional Activites | 0.599** | 0.591* | 0.577** | 0.737** |
| Computer Activities | | | 0.410** | 0.339** |
| Literacy materieals | 0.075 | 0.246* | | |
| **Reading Achievement/Gains** | **0.285*** | **-0.016** | **0.012** | **-0.125**** |
| PV1 | x | x | 0.011 | -0.119** |
| PV2 | x | x | 0.011 | -0.118* |
| PV3 | x | x | 0.011 | -0.119* |
| PV4 | x | x | 0.011 | -0.119* |
| PV5 | x | x | 0.011 | -0.119* |

**Significance**
*p<.05
**p<.01
***p<.001
~<.07

| **Fit Indices** | **CT** | **TN** | **1996** | **2000** |
|---|---|---|---|---|
| cmin/df | 3.264*** | 3.678*** | 4.057*** | 3.538*** |
| NFI | 0.985 | 0.984 | 0.994 | 0.993 |
| Pratio | 0.721 | 0.721 | 0.714 | 0.725 |
| RMSEA | 0.116 | 0.078 | 0.059 | 0.059 |

**Comparisons Across All Data Sets**:

Connecticut and Tennessee results are based on correlation matrices developed from identical variables and scales from the BTPS Questionnaire. Both use identical models presented for mathematics (Figure 2) and reading (Figure 3). NAEP models, also based on correlation matrices, vary from the Connecticut and Tennessee models as well as each other and are shown in Figures 4-8.

Tables 7.3 through 7.5 compare $r^2$ and direct and indirect effects from all SEM analyses. Effects are standardized for comparison purposes although it is important to note that variables/scales and their standard errors are not always the same across data sets (see Tables 4.2-4.7). A discussion of the model confirmation, goodness of fit, and standardized direct and indirect effects across all six data sets is included in this section.

**Model Confirmation across data sets**

***Measurement model.*** As anticipated, model-fit improves with increased sample size. The fit indices indicate that all models have good fit across data sets. In general, the measurement models show that latent variables explain large amounts of variance ($r^2$) in indicator variables in Connecticut and Tennessee compared to moderate variance in NAEP. The latent variable for teacher practice in mathematics explains comparable variance across data sets ($r^2$=.237 to .397). For reading, the results for teacher practice are more varied ($r^2$ =.144 to .937).

***Structural model.*** The four latent variables performed reasonably well in the structural portion of the model for in Connecticut and Tennessee and have already been discussed in the preceding sections. The overall fit for the mathematics analyses based on the conceptual model is good across data sets (cmin($\chi^2$)/df <1.463 to 4.962, p<.00;

GFI>.974 to .997; Pratio>.681 to .760). While RMSEA is slightly high for Connecticut and Tennessee mathematics (.095 to .107), it is well within acceptable range for NAEP mathematics (.028 to .048). It is possible that possible the latent variables for teacher preparation and teacher characteristics are not measuring distinctly different constructs in Connecticut and Tennessee. Teacher preparation programs are not measured in NAEP and professional development is measured by a single exogenous variable rather than a latent variable with three indicators.

The reading analyses had similar goodness of fit statistics (cmin($\chi^2$)/df <3.264 to 4.057, p<.000; GFI>.984 to .994; Pratio>.714 to .725) None of the reading analyses had RMSEAs under .05 although NAEP data sets are close (.059). It is again possible that the latent variables for teacher preparation and teacher characteristics for reading in Connecticut and Tennessee do not measure separate constructs. Teacher preparation programs are not measured in NAEP reading data and professional development is measured by a single exogenous variable as in NAEP mathematics.

***Direct Effects.*** As discussed earlier, the direct effects of teacher preparation on teacher mathematics practice is very large and negative in Connecticut ($\delta$=-.660). In Tennessee the effect is smaller but still negative ($\delta$=-.111). There is no variable for teacher preparation in NAEP. The direct effects of teacher characteristics on teacher mathematics practices are large and positive, exceeding $\delta$=.421 for all data sets. The direct effects of professional development on teacher mathematics practices are positive and small to moderate ($\delta$=.118 to .339) across data sets. NAEP 1996 is the only data set showing professional development effects on teacher mathematics practices in the moderate range ($\delta$=.339).

The direct effect of teacher preparation on teacher reading practice is small and positive in Connecticut (δ=.274). In Tennessee the direct effect is moderate and positive (δ=.367). The direct effects of teacher characteristics on teacher practice in reading is inconsistent across data sets. The effects are negligible to large as well as both negative and positive depending on the data set. In Connecticut, the direct effect is negligible (δ=.010); in Tennessee, the direct effect is small and negative (δ=-127); in NAEP 1998, the effect is nearly perfect and positive (δ=.991); and in NAEP 2000, the effect is positive and very large (δ=.727).

The direct effect of professional development on teacher reading practice is also mixed. At the state level, the effect is small and positive for both Connecticut (δ=.163) and Tennessee (δ=.118). At the national level using NAEP reading, the effects of professional development are less than δ=.05. The effect is negative in NAEP 1998 and positive in NAEP 2000.

The direct effects of teacher practice on student mathematics achievement are positive ranging from δ=.042 to .170. For reading, the direct effects of teacher practice on achievement are mixed and range from δ=-.125 to .285.

***Indirect effects***. The indirect effects of teacher preparation, teacher characteristics, and professional development on average student achievement in mathematics and reading across data sets are small (δ=-.123 to .087) but often significant at δ>|.05|.

**Summary of SEM Results**

In addition to the stability of the hybrid SEM model across policy environments and curriculum contexts, effect sizes are sufficiently varied to provide a venue for comparing results across data sets.

**1.** The conceptual model (Figure 1) served well as a base for the SEM analyses.

**2.** It is possible to trace the pathways from teacher preparation, teacher characteristics, and teacher professional development to teacher practice and subsequent student achievement in mathematics and reading across multiple data sets. Many of the direct and indirect effects on student achievement were significant at the predetermined threshold of $\delta>.05$.

**3.** It is possible to assess the relative contributions of teacher preparation, teacher characteristics, and professional development to teacher practice and subsequent student achievement. Of the variables measured in this data, teacher characteristics have the largest indirect effects on teacher practice followed by professional development. Teacher preparation played a larger role than professional development in Connecticut and Tennessee but it is not possible to extend this analysis to NAEP. The only variable modeled to have direct effects on student achievement was teacher practice.

**a.** *Teacher preparation* at the university level has large negative direct effects on teacher mathematics practice in Connecticut and small negative direct effects on teacher mathematics practice in Tennessee. The direct effects of teacher preparation on teacher reading practice are positive and small in Connecticut and positive and moderate in Tennessee.

The indirect effects of teacher preparation on student mathematics achievement are very small and negative for both Connecticut and Tennessee. The indirect effects of teacher preparation on average reading achievement in Connecticut are positive and significant at $\delta>.05$. In Tennessee the effect of teacher preparation on average reading achievement is negligible.

Connections among teacher preparation, teacher practice, and student achievement are similar for mathematics in both states. The negative relationships among the latent variables in Connecticut are larger than those in Tennessee but both ultimately have negative effects on student mathematics achievement. The effects of the latent variables for reading vary by state but the consistently positive connections in Connecticut are worth noting.

**b**. *Teacher characteristics* have mostly large, positive direct effects on teacher practices employed in the classroom for both mathematics and reading across data sets. The indirect effects of teacher characteristics on mathematics are small although positive across data sets. The indirect effects of teacher characteristics on student reading achievement are inconsistent and negligible ($\delta<.01$) in both Connecticut and Tennessee. For NAEP reading 1998 and 2000, the indirect effects of teacher characteristics on average achievement are negative and exceed $\delta>.05$. The links among teacher characteristics, teacher practices, and student achievement work similarly for mathematics across state and national contexts (the same relative size and direction) but vary across those same contexts for reading.

**c.** *Professional development* has small to moderate, positive direct effects on teacher mathematics practices across all data. Professional development also has small, positive direct effects on reading practices in Connecticut and Tennessee. These effects are smaller and, in one case, negative in NAEP. In general, professional development has positive but very little indirect effect on average student achievement. The strongest effects of professional development on teacher practices and student achievement occur in reading in Connecticut ($\delta$=.047).

**d.** *Teacher practice* has positive, direct effects on average student mathematics achievement across all data. The results for reading are mixed. Teacher reading practices in Connecticut have a positive, direct effect ($\delta$=.285) while teacher reading practices in NAEP 1998 have a negative direct effect ($\delta$=-.125) on average student reading achievement.

**4.** Curriculum context makes a difference. Mathematics, for the most part, has a connected structure where each element is consistently related throughout the system and across data sets. With the exception of teacher preparation, the effects are positive and related to increases in average student achievement. When it comes to reading, the structure does not work in the same way across data sets. Connecticut is the only data set with consistent positive effects throughout the system related to increases in average student reading gains.

**5.** Policy contexts are difficult to discuss in more than speculative fashion. Changes in policy may or may not be related to differences in results across policy environments. Variables to test such hypotheses do not exist in this data.

# 8. RESULTS FROM THE HLM ANALYSIS OF CONNECTICUT $4^{th}$-$8^{th}$ GRADE MATHEMATICS GAINS

The HLM analysis is intended to address the research questions about the relationships among student achievement and

- Teacher characteristics and practices
- Characteristics of professional development
- Characteristics of teacher preparation

The analysis also is intended to address the ways in which these relationships vary by student demographic group, including gender, ethnicity, and grade. Tennessee is not included in the HLM analyses because there is no individual measure of student achievement available in the BTPS data, only correlation coefficients based on aggregation of student scores within teacher.

To set the context for addressing these questions, results are presented first which describe the differences in student achievement (actually achievement gains in CT, achievement status in NAEP data[7]) across student demographic groups. The analyses yield results in terms of regression coefficients. To aid in understanding what the coefficients mean, predicted achievement (gain over one year) is computed for selected combinations of student characteristics.

The district percent free and reduced lunch[8] was assigned to each teacher as a

---

[7] Nye, Konstantopoulos, & Hedges {Nye, 2004 #402} use "gains" to indicate the difference between pre- and post-test scores and "achievement status" to indicate a one-time administration score such as NAEP (p. 17).

[8] 1999-2000 district data from Connecticut Department of Education website.

measure of classroom socioeconomic status since there was no individual student indicator available in the BTPS data.[9]

The full set of mathematics variables for CT is given in Table 8.1.[10] These are the same variables used in the earlier structural equation models.

**Table 8.1** Description of Variables Used in the Connecticut Mathematics Analyses.

| Type | Variable Abbreviation | Variable Description |
|---|---|---|
| | | **Level 1 - students** |
| Continuous (Outcome) | mdiff | Mathematics gain, difference between pre- and post-test scores (mean=4.80, sd=9.70) |
| Dichotomous | dblack | Black (Black=1, else=0; mean=.14, sd=.34) |
| | dhispanic | Hispanic (Hispanic=1, else=0; mean=.09, sd=.29) |
| | dallother | All other ethnicities (all other=1, else=0; mean=.09, sd=.29) |
| | dfem | Female (female=1, else=0; mean=.52, sd=.50) |
| | dgrade4 | Grade 4 (grade 4=1, else=0; mean=2.30, sd=1.98) |
| | dgrade6 | Grade 6 (grade 6=1, else=0; mean = 1.08, sd=2.74) |
| Effects coded | egenxbl | Effect of Black female; interaction of eBlack x egender where Black =1, White=-1, otherwise=0 multiplied by female=1, male=-1 (mean=.01, sd=.90) |
| | egenxhisp | Effect of Hispanic female; interaction of eHispanic x egender where Hispanic=1, White=-1, otherwise=0 multiplied by female=1, male=-1 (mean=.01, sd=.88) |
| | egenxoth | Effect of All other female; interaction of eAllother x egender where All other =1, White=-1, otherwise=0 multiplied by female=1, male=-1 (mean=.01, sd=.88) |

[9] This causes some disaggregation bias since the schools in which teachers teach are likely to vary within districts on this variable and students are likely to vary on SES within teachers. Bias is less problematic, however, than leaving out an indicator of SES for these analyses.

[10] All variables used in the SEM analysis are listed here and all were tested. Results for teacher level variables that did not explain variance in average student gains across teachers are not reported in all sections.

**Table 8.1** (cont'd).

| | | |
|---|---|---|
| | | **Level 2 - teachers** |
| Continuous | Z-scores | Sum of interview responses rescaled to have mean=0, sd=1 |
| Teacher Practices | zmthgpa | Frequency of group and project activities |
| | zmthdla | Frequency of drill and lecture activities |
| | zmthsmkg | Frequency of sensemaking activities |
| Teacher Characteristics | zmoref | Teacher's reform orientation |
| | zmthknow | Self report of mathematics knowledge |
| | ztrad | Teacher's traditional orientation |
| | zmathprep | Major/coursework/license in mathematics, mathematics education |
| | zmathinst | Preparation for mathematics instruction, pedagogical coursework |
| | zdivprep | Preparation for diversity |
| | zefficacy | Teacher's sense of efficacy, that a teacher can make a difference in students' learning |
| | zassessprep | Preparation for assessment |
| Professional Development | zpdi | professional development impact (self report) |
| | zpdf | professional development frequency |
| | zmentfrq | mentoring frequency |
| Continuous | pctfrlu | Percent district free and reduced lunch (mean=.21,sd=.16) |

**I. Setting the Stage: Variation in Mathematics Achievement Gains by Student Demographic Groups and Grade Level** (Results at the Student Level)

In this section, sets of background variables are added to look at changes in mathematics gains in Connecticut. Model 1 predicts mean student gains across the sample and partitions the variance between students and teachers where students are nested within teachers. Models 2-5 predict differences in mean gains across gender, ethnicity, and grade.

Mean Gains. To set the context for answering the questions about connections between student achievement and teacher variables, results are presented first for differences in student achievement gains across student demographic groups. The level one and level two equations for the model are given followed by their estimated regression coefficients in Tables 8.2.

***Model 1.*** *Mean Gains*

Level 1 - Students

$$Y_{ij} = \beta_{0j} + e_{ij} \qquad (8.1)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

where $Y_{ij}$ is the mathematics gain for student *i* in teacher *j*'s classroom, $\beta_{0j}$ is the intercept or mean students' mathematics gain within teacher j's classroom, $e_{ij}$ is the level 1 error, $\gamma_{00}$ is the mean value of classroom gains for all teachers, and $r_{0j}$ is the level 2 error.

**Table 8.2** Connecticut's Average Mathematics Gain 1999-2000

| | Model 1 |
|---|---|
| Intercept ($\gamma_{00}$) | 4.85(.40)*** |
| LVL 1 Variance ($\sigma^2$) | 85.88 |
| LVL 2 Variance ($\tau_{00}$) | 8.42*** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

The predicted average gain score for 4th-8th grade mathematics students for this sample of students in Connecticut is 4.85 points ($\gamma_{00}$=4.85, sd=.40) when no other variables are included in the model.

$\sigma^2$ and $\tau_{00}$ are used to compare variation in the errors at the teacher level with variation at the student level using the formulas:

$$\frac{\sigma^2}{\sigma^2 + \tau_{00}} = \textit{level one variance} \qquad \frac{\tau^2{}_{00}}{\sigma^2 + \tau_{00}} = \textit{level two variance} \qquad (8.2)$$

91% of the variance in gains for Model 1 is within teachers (between students) leaving 9% of the variance in gains between teachers. That is to say, there is ten time more variation within teachers than between teachers.

Mean Gains by Demographic Group. Next, demographic variables are added to form Models 2, 3 and 4. Model 5 introduces the control variables for grade level. The equations for Model 2, ethnic group, are presented first (Eq.8.3) To interpret the coefficients, the estimated gain for ethnicity can be computed by using the appropriate set of 1s and 0s for the dummy variables. As an example, Equations 8.4 and 8.5 demonstrate the computation for obtaining Black student gains. Estimated gains by ethnicity are presented in Table 8.3.[11]

***Model 2.*** *Ethnic Groups*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij}\ e_{ij};$$

Level 2 - Teachers (8.3)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

---

[11] Note that the category for All other includes Asian and those students who did not report ethnicity. Often these smaller groups are folded into the intercept since they are not significantly different from the intercept. In this case, the coefficient for all other was nearly significant at p<.053. Additionally, it was important to maintain the meaning of the intercept as a specific group, in this case white students, before analyzing the differential effects of teacher variables.

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

The regression coefficients for predicting student achievement gains from ethnic background are found in Table 8.3.

**Table 8.3** Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Ethnicity

| | Model 2 |
|---|---|
| Intercept ($\gamma_{00}$) | 4.05(.42)*** |
| Black ($\gamma_{01}$) | 2.64(1.06)* |
| Hispanic ($\gamma_{02}$) | 2.89(1.04)** |
| All other ($\gamma_{03}$) | 1.53(.80)~ |
| LVL 1 Variance ($\sigma^2$) | 85.32 |
| LVL 2 Variance ($\tau_{00}$) | 7.64*** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

To make it easier to interpret the results from the gains by ethnicity group analysis, these coefficients can be used to compute the gain for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[12] For example, to determine the predicted gains for a black student who would have ones for the intercept and black the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} \qquad (8.4)$$

---

[12] A table showing ones and zeros for all dummy variables is provided in the Appendix, p. 295.

Level 2 - Teachers

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 4.05$$

$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = 2.63$$

When the betas from Equation 8.4, are substituted in the level one equation (Eq. 8.4), the result for black gain is produced (Eq. 8.5).

$$\hat{Y}_{ij} = (4.05) + (2.63)*1 \quad (8.5)$$

$$= 6.68$$

Predicted mathematics gains for all demographic groups are presented in Table 8.4. This table shows that the model predicts that all non-white student groups gain more in mathematics than white students. The coefficients for White, Black, and Hispanic students are significantly different from zero. The coefficient for all other is nearly significant at $p<.053$.

**Table 8.4** Predicted Mathematics Gains by Student Demographic Group

| Ethnic Group | Gains |
|---|---|
| White | 4.05 |
| Black | 6.68 |
| Hispanic | 6.94 |
| All other | 5.58 |

In the next model, gender predicts average mathematics gains.

***Model 3.*** *Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{4j}\, dfem_{ij} + e_{ij};$$

Level 2 - Teachers (8.6)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{4j} = \gamma_{10}$$

The regression coefficients for predicting student achievement gains from gender are found in Table 8.5.

**Table 8.5** Regression Coefficients for Connecticut 4th Grade Mathematics Gain 1999-2000 by Gender

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | 4.84(.51)*** |
| Female ($\gamma_{04}$) | .03(.52) |
| LVL 1 Variance ($\sigma^2$) | 85.93 |
| LVL 2 Variance ($\tau_{00}$) | 8.42*** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

These coefficients indicate that male mean gains in mathematics are predicted to be 4.84 points compared to 4.87 for females. The coefficient for female is not significantly different from zero.

***Model 4**. Combined Gender and Ethnicity*

Model 4 tests whether or not there is an interaction between gender and ethnicity. Equation 8.7 computes coefficients for combinations of gender and ethnicity in CT. Effects coding for gender x ethnicity produces the same fixed effects as dummy coding.[13]

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij} + e_{ij}; \quad (8.7)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

[13] When interactions are effects-coded, it facilitates statistical testing of the differences by gender by ethnicity variables found in Section 2. (Anderson, p.22, web notes)

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

The regression coefficients for predicting student achievement gains from ethnic background and gender are found in Table 8.6.

**Table 8.6** Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Gender and Ethnicity

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | 3.72(.65)*** |
| Black ($\gamma_{01}$) | 2.63(1.08)* |
| Hispanic ($\gamma_{02}$) | 2.76(1.09)* |
| All other ($\gamma_{03}$) | 1.49(.84)~ |
| Female ($\gamma_{04}$) | .66(.89) |
| Effect Black x female ($\gamma_{05}$) | -.35(.63) |
| Effect Hispanic x female ($\gamma_{06}$) | .93(.71) |
| Effect All other x female ($\gamma_{07}$) | -.02(.70) |
| | |
| LVL 1 Variance ($\sigma^2$) | 85.35 |
| LVL 2 Variance ($\tau_{00}$) | 7.70*** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

To make it easier to interpret the results from the gains by demographic group analysis, these coefficients can be used to compute the gain for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[14] For example, to determine the predicted gains for a black female student who would have ones for the intercept, black, female and black female, the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} eblackxefem_{ij}$$

Level 2 - Teachers (8.8)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 3.72$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = 2.63$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = .66$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = -.35$$

When the level two betas from Equation 8.8, are substituted in the level one equation (Eq. 8.8), the result for black female gain is produced (Eq. 8.9).

$$\hat{Y}_{ij} = (3.72) + (2.63)*1 + (0.66)*1 + (-0.35)*1 \quad (8.9)$$
$$= 6.66$$

Predicted mathematics gains for all subsets of students by demographic groups are presented in Table 8.6. The table shows that the model predicts all female students are predicted to gain more than their male counterparts.

---

[14] A table showing ones and zeros for all dummy variables is provided in the Appendix.

**Table 8.7** Predicted Mathematics Gains by Student Demographic Group with Differences by Gender

| | Males | Females | Difference in Gains by Gender |
|---|---|---|---|
| White | 3.72 | 4.38 | + .66 |
| Black | 6.35 | 6.66 | + .31 |
| Hispanic | 6.48 | 8.07 | +2.52 |
| All other | 5.21 | 5.85 | + .62 |

Score differences for males and females are predicted to be less than 1 point for White, Black, and All other students. Hispanic females are estimated to gain 2.52 points more than Hispanic males. Coefficients for gender x ethnicity effects are not significantly different from zero over the course of one year.

***Gender, Ethnicity, and Grade Level***

Another known source of variation in student achievement is grade level. In the earlier SEM analyses, the outcome variable for gains was standardized by grade level. In the upcoming NAEP analyses, only one grade is represented in each analyses. Consequently, dummy variables for grade level are introduced at level one in the Connecticut analyses to control for grade level. The reference category for Model 5 is therefore white male students in $8^{th}$ grade.

***Model 5.*** **Gender, Ethnicity, and Grade Level**

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + \beta_{4j}\,dfem_{ij} \quad (8.10)$$

$$+ \beta_{5j}egenxbl_{ij} + \beta_{6j}egenxhisp_{ij} + \beta_{7j}egenxoth_{ij}$$

$$+ \beta_{8j}dgrade4_{ij} + \beta_{9j}dgrade6_{ij} + e_{ij};$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

The regression coefficients for predicting student achievement gains from ethnic background, gender, and grade level are found in Table 8.8.

**Table 8.8** Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Gender, Ethnicity, Grade

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | .06(.98) |
| Black ($\gamma_{01}$) | 2.36(1.09)* |
| Hispanic ($\gamma_{02}$) | 2.29(1.07)* |
| All other ($\gamma_{03}$) | 1.44(.80)~ |
| Female ($\gamma_{04}$) | .64(.88) |
| Effect Black x female ($\gamma_{05}$) | -.33(.62) |
| Effect Hispanic x female ($\gamma_{06}$) | 1.00(.70) |
| Effect All other x female ($\gamma_{07}$) | -.10(.69) |
| 4th grade ($\gamma_{08}$) | .92(.26)** |
| 6th grade ($\gamma_{09}$) | .94(.19)*** |
| LVL 1 Variance ($\sigma^2$) | 85.28 |
| LVL 2 Variance ($\tau_{00}$) | 5.11*** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

As demonstrated in prior sections, the computational example to determine the predicted gains for a black female student in the 6th grade who would have ones for the intercept, black, female, black female, and grade 6, would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} egenxbl_{ij} + \hat{B}_{9j} dgrade6$$

Level 2 - Teachers (8.11)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = .06$$

$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = 2.36$$

$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = .64$$

$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = -.33$$

$$\hat{\beta}_{9j} = \hat{\gamma}_{90} = .94$$

When the betas from Level 2, Equation 8.11, are substituted in the Level 1 equation, the result for black female gain is produced (Eq. 8.12).

$$\hat{Y}_{ij} = (.06) + (2.36)*1 + (0.64)*1 + (-0.33)*1 + (.94*1) \quad (8.12)$$
$$= 3.67$$

Predicted mathematics gains for all subsets of students by demographic group and grade are presented in Table 8.9. The estimated gains associated with the $4^{th}$ and $6^{th}$ grade years average just under one point each and are the largest, significant predictor of gains for all students. Note that estimated gains during the eighth grade year are considerably smaller than gains made at Grades 4 and 6 for all ethnicities.

**Table 8.9** Predicted Mathematics Gains by Student Demographic Group and Grade

| | | Grade 4 (+.92) | Grade 6 (+.94) | Grade 8 |
|---|---|---|---|---|
| Males | White | .98 | 1.00 | .06 |
| | Black | 3.34 | 3.36 | 2.42 |
| | Hispanic | 3.27 | 3.29 | 2.35 |
| | All other | 2.42 | 2.44 | 1.50 |
| Females | White | 1.62 | 1.06 | .70 |
| | Black | 3.65 | 3.67 | 2.73 |
| | Hispanic | 4.91 | 4.92 | 3.99 |
| | All other | 2.96 | 2.98 | 2.04 |

The table also indicates that the model predicts that all non-white student groups gain more in mathematics than white students once ethnicity, gender, and grade level are taken into consideration. Hispanic students show the largest differences by gender with girls predicted to gain 1.64 points more than boys; however, coefficients for dfemale and the interaction terms (gender x ethnicity) are non-significant.

*Summary of Section 1.* Developing the level one equation for students. As the level one base equation for mathematics in Connecticut developed, average yearly gains were first predicted with a simple unconditional model. This model indicated that gain variance was ten times larger within teachers (between students) than across teachers at level 2. Variables known to affect mean gains (ethnicity, gender, and grade) were added to the level one equation revealing that while differences in gains exist within teachers, only the coefficients for ethnicity and grade level are significant predictors of yearly mathematics gains when gender is controlled. For the Connecticut mathematics data, all groups in all grades are expected to have higher average gains than White students.

**II. Effects of Teacher Practices, Teacher Characteristics, and Professional Development on Student Achievement Gains** (Results at the teacher level)

This section addresses the questions:

1. Does mathematics achievement gains for demographic groups vary across teachers?
2. How much variation in student mathematics gains do variables about teacher practices, teacher characteristics, and professional development explain?
3. Is the amount of variance explained by teacher characteristics and practices different for different student demographic groups?

The analyses in this section first test if the achievement gains for the various demographic groups vary across teachers as the unit of analysis shifts from students to teachers. There is a brief discussion of those demographic group gains which vary across teachers followed by changes to the results when district SES is controlled. This establishes the level two context for a series of analyses which look at the effects of teacher practice, teacher characteristics, and professional development on the variance of achievement gains by demographic groups.

**Variance in Connecticut mathematics achievement across teachers**

In order to determine if there is significant variance in mathematics achievement by demographic groups across teachers, it is important to understand how significant variance is indicated and its impact on the level two equation. When main effects variables are dummy coded, not only does $\beta_0$ represent the mean for white male achievement, but the variance of $\beta_0$, known as $\tau_{00}$, is the variance of white male achievement (see Raudenbush, p.34). Nonsignificant error terms for variance ($r_{00}$, $r_{10}$, ...$r_{x0}$) as determined by $\chi^2$ indicate that group mean achievement does not vary across teachers so the error terms for those demographic groups may be dropped from the model. Although the nonsignificant error terms are dropped, the fixed level two equations for those demographic groups remain in the analysis.

Analyses indicated that average gains for female and Black varied significantly across teachers ($H_a$: $\tau_{..0} \neq 0$). Black females, Hispanic males and females, as well as All other males and females, did not vary significantly from the intercept (H: $\tau_{..0}=0$) so their error terms have been eliminated. Their variance is then captured as part of the error term for the intercept ($r_{00}$). In other words, the variance is no different from that of White.

As teacher variables are added at Level 2, different proportions of variance are explained depending on the student demographic group whose mean achievement is being predicted. Each group with significant variance (Female, Black) in its gains across teachers is addressed below in a series of separate analyses which include teacher practices, teacher characteristics, and teacher professional development. The results from the analysis of student gains by gender across teachers are followed by the results from the analysis of gains by White and Black students across teachers.

**Effects of teacher level variables on mean gains by gender**

While the fixed effects model for gender, Model 3, predicted average gain scores, a random slopes and intercepts model, Model 6, is necessary to test whether or not the effects of male and female vary significantly across teachers. If they vary, do teacher practices, characteristics, or professional development have differential effects on the average mathematics gains for one gender as opposed to the other? If variance is significant across teachers, it is then possible to determine how much of the variation can be explained by a specific level two variables.

***Model 6.*** *Unconditional Model for Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} \quad (8.13)$$
$$+ \beta_{8j} dgrade4_{ij} + \beta_{9j} dgrade6_{ij} + e_{ij};$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \quad (8.14)$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$
$$\beta_{4j} = \gamma_{40} + r_{4j}$$
$$\beta_{8j} = \gamma_{70}$$
$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average student gains across teachers

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, average Black student gains across teachers

$\gamma_{20j}$ = intercept, average Hispanic student gains across teachers

$\gamma_{30j}$ = intercept, average All other student gains across teachers

$\gamma_{40j}$ = intercept, average female student gains across teachers

$r_{4j}$ = the error term or unique teacher effects

$\gamma_{80j}$ = intercept, average grade 4 gains across teachers

$\gamma_{90j}$ = intercept, average grade 6 gains across teachers

The estimated level two regression coefficients for Model 6 are given in Table 8.10.

**Table 8.10** Regression Coefficients for Connecticut Mathematics Gain 1999-2000 by Gender

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | .50(1.04) |
| Black ($\gamma_{10}$) | 2.46(1.06)* |
| Hispanic ($\gamma_{20}$) | 2.63(.99)** |
| All other ($\gamma_{30}$) | 1.50(.76)* |
| Female ($\gamma_{40}$) | -.08(.54) |
| Grade 4 ($\gamma_{80}$) | .88(.29)** |
| Grade 6 ($\gamma_{90}$) | .92(.20)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 84.04 |
| LVL 2 Intercept ($\tau_{00}$) | 9.33*** |
| Female ($\tau_{40}$) | 5.09* |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

The effect of female varies significantly across teachers ($\tau_{40}$=5.09, p<.05) as does the effect of male ($\tau_{00}$=9.33, p<.001). Coefficients from this analysis serve as the base for the variance decompositions for gender found in Tables 8.13-8.16. Variance decompositions systematically present the amount of variance explained by each additional variable included at level two.

***Model 7.*** *SES and Gender*

Level 1- Students - Same as Equation 8.13

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, pctfrlu + r_{0j} \tag{8.16}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + \gamma_{41}\, pctfrlu + r_{4j}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average student gains across teachers

$\gamma_{01j}$ = the change in average gains associated with a one unit increase in district free and reduced lunch

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{41j}$ = the change in average female student gains associated with one unit increase in average district free and reduced lunch

$r_{4j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for Model 6 are given in Table 8.11.

**Table 8.11** Regression Coefficients for Average Gains by Gender Predicted by Average District SES

| | Gender |
|---|---|
| Intercept ($\gamma_{00}$) | .33(1.06) |
| Free and reduced lunch ($\gamma_{01}$) | -6.44(3.94) |
| Black ($\gamma_{10}$) | 2.59(1.07)* |
| Hispanic ($\gamma_{20}$) | 2.77(1.06)* |
| All other ($\gamma_{30}$) | 1.46(.77)~ |
| Female ($\gamma_{40}$) | -.03(.54) |
| Free and reduced lunch ($\gamma_{41}$) | 7.90(4.25)~ |
| Grade 4 ($\gamma_{80}$) | .90(.29)** |
| Grade 6 ($\gamma_{90}$) | .94(.20)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 84.02 |
| LVL 2 Variance ($\tau_{00}$) | 8.76*** |
| ($\tau_{40}$) | 3.91* |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Table 8.11 shows that being in a district with average SES (21% free and reduced lunch) has a negative effect on boys' average mathematics gains across teachers ($\gamma_{02}$ = -6.44, sd=3.94) and a positive effect on girls' average gains across teachers ($\gamma_{12}$ = 7.90, sd=4.25,p<.066). There are two districts in this data that have over 50% (+2 standard deviations) free and reduced lunch. Boys would be expected to lose -12.5 points ($Y_{ij}$=.33 +2 x (-6.44)) while girls would be expected to gain 3.6 points ($Y_{ij}$=-12.5 +.33 + (-.03) + 2*(7.9) = 16.1) in those same districts over the course of the year.

How much do the average gains for males and females vary across average SES districts? To answer this question, the following equation calculates the amount of variance explained by average district SES across teachers.[15] For example, to find the

[15] A three-level analysis, students nested within teachers nested within districts, was conducted for the ten Connecticut districts that were represented by 5 or more

variance in female gains explained by district SES, substitute the estimates for $\tau_{40}$ from Table 8.12 into Equation 8.18.

$$\frac{\tau_{40}(\text{model 6}) - \tau_{40}(\text{model 7})}{\tau_{40}(\text{model 6})} = \frac{5.09 - 3.91}{5.09} = .232 = 23.2\% \qquad (8.17)$$

This means that 23.2% of the variance in the average female gains is explained by district SES across teachers. Using the same process for $\tau_{00}$, the results are presented at the bottom of Tables 8.12.-8.15.

**Table 8.12** Variance Decomposition for Gender by District SES

| | | SES |
|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lunch |
| Intercept ($\gamma_{00}$) | .50(1.04) | 0.33(1.06) |
| SES($\gamma_{02}$) | | -6.44(3.94) |
| Female ($\gamma_{10}$) | -.08(.54) | -.03(.54) |
| SES($\gamma_{12}$) | | 7.90(4.25)~ |
| Intercept Variance ($\tau_{00}$) | 9.33*** | 8.76*** |
| Dfem Variance ($\tau_{40}$) | 5.09* | 3.91* |
| ($\sigma^2$) | 84.04 | 84.02 |
| % variance reduction male | | 6.1% |
| % variance reduction female | | 23.2% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

One standard deviation increase in district free and reduced lunch accounts for a -6.44 point drop in mathematics gains for boys. In the same district, a one standard deviation increase in district free and reduced lunch signals an increase of +7.90 points for girls. The effect of being in an average SES district is the opposite for boys and girls across teachers. SES, therefore, has a differential effect on boys and girls. Although

teachers. There was no significant variance across districts for this sample so it was determined that the two-level model controlling for SES was the most appropriate.

there is more variance in average achievement gains for boys across teachers, SES accounts for only a small percentage of the variance in boys gains (6.1%) compared to girls average gains (23.2%) across teachers.

***Gender and Teacher Effects***

Now that the context is set at level two, it is time to determine the effects of teacher variables on the variance of average gains for boys and girls across teachers. The next analyses track variance explained by teacher level variables controlling for SES and also determine whether or not particular teacher variables explain different proportions of variance in the average achievement gains for males and females. One teacher variable at a time is added to the level two equation, EQ 8.16, forming an equation template (Eq. 8.18). The term "*variable x*" is exchanged for a single teacher variable for Mathematics Practice (math sensemaking, drill and lecture, or group and project activities), Teacher Characteristic (mathematics preparation, preparation for mathematics instruction, mathematics knowledge, reform and traditional orientation, assessment preparation, or efficacy), or Professional Development (professional development impact, mentoring frequency, or professional development frequency).

***Model 8***. Template for Level 2 equations

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, pctfrlu + \gamma_{02..}\, variable\ x + r_{0j} \qquad (8.18)$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + \gamma_{41}\, pctfrlu + \gamma_{42..}\, variable\ x + r_{4j}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{02..j}$ = change in average in male student gains associated with a one unit change in "*variable x*" holding SES constant at the district average

and

$\gamma_{42..j}$ = change in average female student gains associated with one unit change in "*variable x*" holding SES constant at the district average

**Gender and Teacher Practices.** The variance decomposition for teacher practices is located in Table 8.14. Teacher practices explain more of the variation across teachers in average gains for girls than boys in average SES districts.

**Table 8.13** Variance Decomposition for Gender Controlling SES - Teacher Practices

| | | SES | Practice | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lunch | Sensemaking | Drill&Lecture | Group&Project |
| Intercept ($\gamma_{00}$) | .50(1.04) | 0.33(1.06) | 0.40(1.07) | 0.33(1.07) | 0.34(1.06) |
| ($\gamma_{01}$) | | | 0.70(.45) | -0.30(.48) | 0.52(.49) |
| SES($\gamma_{02}$) | | -6.44(3.94) | -6.53(3.94) | -6.57(3.96) | -6.18(3.98) |
| Dfem ($\gamma_{10}$) | -.08(.54) | -.03(.54) | 0.00(.53) | -0.04(.53) | 0.01(.53) |
| ($\gamma_{11}$) | | | -0.98(.48)* | 0.60(.53) | -1.06(.59)~ |
| SES($\gamma_{12}$) | | 7.90(4.25)~ | 8.04(4.26)~ | 8.16(4.26)~ | 7.58(4.39)~ |
| Intercept Variance ($\tau_{00}$) | 9.33*** | 8.76*** | 8.54*** | 8.80*** | 8.53** |
| Dfem Variance ($\tau_{40}$) | 5.09* | 3.91* | 3.60* | 3.37* | 3.18~ |
| ($\sigma^2$) | 84.04 | 84.02 | 83.93 | 84.15 | 84.10 |
| % variance reduction male | | 6.1% | 8.5% | 5.7% | 8.6% |
| % variance reduction female | | 23.2% | 29.3% | 33.8% | 37.5% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Variance in boys' gains explained by math practice variables across teachers controlling for average SES ranges from 5.7% to 8.6%. Although variation in teacher practices across teachers explains up to 8.6% of the variance in gains for boys, it explains 23%-38% of the variance in girls' gains controlling for SES. Variance explained by mathematics sensemaking activities across teachers adds a small amount to the variance already explained by SES in the average achievement gains for boys (a little over 2% more) while it explains an additional 6% of the variation in girls' gains across teachers. The change attributable to one standard deviation increase in frequency of a particular mathematics practice is the opposite for boys and girls in this data. Math sensemaking and group and project activities decrease female gains by approximately one point ($\gamma_{11}$=-.98 and $\gamma_{11}$= -1.06, respectively) while male gains increase by about one half of a point ($\gamma_{01}$=.70 and $\gamma_{01}$=.53, respectively). Drill and lecture activities increase female gains ($\gamma_{11}$=.60) while decreasing average gains for males ($\gamma_{01}$= -.30). Teacher mathematics practices have differential effects on gains by gender.

**Gender and Teacher Characteristics.** The variance decomposition for teacher characteristics (Table 8.14) answers questions regarding how much of the variance in average gains for boys and girls can be explained by specific teacher characteristics controlling for SES across teachers. While there are small percentages of variance explained by teacher characteristics across teachers for boys, up to 8 times as much variance is explained for girls by the same teacher characteristics. With the exceptions of math preparation, which is negatively related to average gains, and mathematics knowledge, which is near zero for both genders, the effect of teacher characteristics on the average gains of boys and girls across teachers is differential (in the opposite direction) once SES is controlled.

**Table 8.14** Variance Decomposition by Gender for Teacher Characteristics Controlling for SES

| | | SES | Teacher Characteristics | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | MthPrep | MthInst | MthKnow |
| Intercept ($\gamma_{00}$) | .50(1.04) | 0.33(1.06) | 1.51(1.10) | .30(1.08) | 0.05(1.11) |
| ($\gamma_{01}$) | | | -0.86(.46)~ | .43(.45) | -0.01(.47) |
| SES($\gamma_{02}$) | | -6.44(3.94) | -7.12(3.86)~ | -5.98(3.99) | -6.36(4.02) |
| Dfem ($\gamma_{10}$) | -.08(.54) | -.03(.54) | 0.03(.54) | -.02(.54) | -0.03(.53) |
| ($\gamma_{11}$) | | | -0.18(.45) | -.43(.71) | 0.89(.44)* |
| SES($\gamma_{12}$) | | 7.90(4.25)~ | 7.60(4.21)~ | 7.49(4.45)~ | 8.36(4.22)* |
| Intercept Var ($\tau_{00}$) | 9.33*** | 8.76*** | 8.83*** | 8.84*** | 9.07*** |
| Dfem Variance ($\tau_{40}$) | 5.09* | 3.91* | 4.40* | 4.32* | 3.97* |
| ($\sigma^2$) | 84.04 | 84.02 | 84.00 | 83.96 | 83.88 |
| % variance reduction male | | 6.1% | 5.4% | 5.3% | 2.8% |
| % variance reduction female | | 23.2% | 13.6% | 15.1% | 22.00% |

**Table 8.14** (cont'd).

| | | SES | Teacher Characteristics | | | |
|---|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | Reform Or | Trad Or | Assess | Efficacy |
| Intercept ($\gamma_{00}$) | .50(1.04) | 0.33(1.06) | 0.33(1.07) | .32(1.07) | 0.30(1.06) | 0.38(1.07) |
| ($\gamma_{01}$) | | | 0.58(.44) | .25(.48) | 0.29(.38) | 0.56(.43) |
| SES($\gamma_{02}$) | | -6.44(3.94) | -6.17(3.83) | -6.70(3.92)~ | -6.22(3.92) | -5.85(3.91) |
| Dfem ($\gamma_{10}$) | -.08(.54) | -.03(.54) | -0.02(.53) | -.02(.54) | -.01(.53) | -0.02(.52) |
| ($\gamma_{11}$) | | | -0.70(.47) | -.29(.59) | -0.73(.51) | -1.11(.54)* |
| SES($\gamma_{12}$) | | 7.90(4.25)~ | 7.66(4.17)~ | 8.19(4.13)* | 7.27(4.32)~ | 6.78(4.22) |
| Intercept Var ($\tau_{00}$) | 9.33*** | 8.76*** | 8.73** | 8.97*** | 8.71*** | 8.59** |
| Dfem Variance ($\tau_{40}$) | 5.09* | 3.91* | 3.89* | 4.47* | 3.56* | 2.97~ |
| ($\sigma^2$) | 84.04 | 84.02 | 83.96 | 83.95 | 84.12 | 84.15 |
| % variance reduction male | | 6.1% | 6.4% | 4.0% | 6.6% | 7.9% |
| % variance reduction female | | 23.2% | 23.6% | 12.1% | 30.1% | 41.7% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

For every one standard deviation increase in mathematics preparation, both boys (-.86) and girls (-.18) lose a partial point .The negative relationship between teacher preparation and student achievement is puzzling but inline with results from the earlier SEM analyses in this dissertation and with Valli and Reckase (Valli et al., 2003), and Rowan, Correnti, and Miller (B. Rowan, Correnti, & Miller, 2002). For teachers math knowledge, girls average gains increase by +.89 points for every standard deviation increase while boys gains stay nearly level. One standard deviation of teacher efficacy, however, decreases average gains for girls by -1.11 points.

Orientation, mathematics preparation, math instruction preparation, and mathematics knowledge variables controlling for SES do not explain more variance in the average gains for boys and girls across teachers than does SES alone. Teacher efficacy and preparation for assessment are the only two teacher characteristics that explain variation across teachers for boys (6.6-7.9%). These teacher variables explain 30.1%-

41.7% of the variation in gains for girls across teachers controlling for SES.

**Gender and Teacher Professional Development.** This analysis focuses on the differential effects of teacher professional development as measured by professional development impact, mentoring frequency, and professional development frequency. Professional development impact (self report) explains very little variance in average male or female gains across teachers. It predicts less well than SES alone. Mentoring frequency explains no more variation in boys gains across teachers than SES alone but it does explain 27.8% of the variance in girls gains. Professional development frequency explains 15-17% of the variance for both boys and girls gains across teachers controlling for SES. The variance decomposition for professional development follows in Table 8.15.

**Table 8.15** Variance Decomposition for Gains by Gender Controlling for SES - Teacher Professional Development

| | | SES | Professional Dev | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | PDImpact | MentFrq | PDFreq |
| Intercept ($\gamma_{00}$) | .50(1.04) | 0.33(1.06) | -0.03(1.04) | 0.33(1.03) | .12(.99) |
| ($\gamma_{01}$) | | | -0.34(.46) | -0.09(.50) | -.12(.05)* |
| SES($\gamma_{02}$) | | -6.44(3.94) | -6.79(3.94)~ | -6.53(3.99) | -6.13(3.69)~ |
| Dfem ($\gamma_{10}$) | -.08(.54) | -.03(.54) | -.02(.55) | 0.06(.52) | -.03(.53) |
| ($\gamma_{11}$) | | | -0.48(.51) | -1.41(.60)* | .02(.07) |
| SES($\gamma_{12}$) | | 7.90(4.25)~ | 7.35(4.22)~ | 6.35(3.98) | 7.79(4.22)~ |
| Intercept Var ($\tau_{00}$) | 9.33*** | 8.76*** | 8.86*** | 9.16*** | 7.68*** |
| Dfem Variance ($\tau_{40}$) | 5.09* | 3.91* | 4.40* | 3.53~ | 4.53* |
| ($\sigma^2$) | 84.04 | 84.02 | 83.92 | 83.63 | 83.94 |
| % variance reduction male | | 6.1% | 5.0% | 1.8% | 17.6% |
| % variance reduction female | | 23.2% | 13.6% | 27.8% | 15.5% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Mentoring frequency is negatively related to average achievement for girls across teachers. For every one standard deviation increase in mentoring frequency across teachers, girls lose -1.41 points compared to boys who remain nearly level. Mentoring frequency explains no more additional variation in average achievement for boys across teachers but it explains 27.8% of the variance in average gains for girls across teachers controlling for SES.

Professional development frequency has a negative relationship to boys mathematics gains across teachers controlling for SES. For every standard deviation increase in professional development frequency, boys lose -.12 points. Girls gains remain nearly level. Professional development frequency explains 11.5% of the variance in mathematics gains for boys across teachers while it adds nothing to the variance explained for average gains by girls across teachers.

**Summary of differential gender effects** - The effects of teachers on the average gains for male and females controlling for SES varies for certain teacher practices, charactersitics, and professional development. Teacher practice, characteristics, and professional development variables generally explain more of the variation that exists in average mathematics gains for girls across teachers in Connecticut once SES is controlled. SES explains 23.2% of the variation in female gains and 6.1% of the variation in male gains. The loss in gains for boys attributable to a one standard deviation increase in district free and reduced lunch (proxy for lower SES) is over 6 points while girls gain +7.9 points for a one standard deviation increase in free and reduced lunch.

The models for teacher practice explain 24-30% of the variation in average female gains across teachers when SES is controlled; the models for teacher characteristics explain 15-37% of the variation in female gains across teachers; and mentoring frequency explains 28% of the variation in female gains across teachers. The same models explain less than 7% of the variation in average achievement for males across teachers. One exception is the variable for professional development frequency which explains 15-17% of the variation in gains across both genders and is consistently, negatively related to gains controlling for SES.

Teacher variables that have differential effects on gains by gender (negative for girls and positive for boys) include two practices, math sensemaking activities and group and project activities, and one teacher characteristic, teacher efficacy. The teacher characteristic, math knowledge, is positive for girls and level for boys. In this data, the effects of all professional development variables are negative for both genders.

Three variables, mathematics sensemaking, group and project activities, and teacher efficacy signaled predicted losses of -.95 to -1.07 points for girls across teachers while predicting gains of +.52 to +.70 for boys. Teachers' professional development frequency signaled a decrease in gains for boys at -.12 points across teachers with girls remaining level. The only variable exhibiting a positive effect for girls was math knowledge which predicted an increase in girls' average gain of +.89 points for every standard deviation increase across teachers. Boys' average gain remains nearly level.

**Effects of Teacher Level Variables on Mean Gains by Black and White**

While the fixed effects model for ethnicity, Model 2, predicts that average gains vary significantly by ethnicity, a random slopes and intercepts model, Model 9, tests whether or not the average achievement gains for white and black students vary significantly across teachers. Other demographic groups did not vary significantly from Whites. In other words, do teacher practices, characteristics, or professional development variables have differential effects on the average mathematics gains of Black and/or White students across teachers? If variance is significant across teachers, it is then possible to determine how much of the variation in each groups' achievement can be explained by single level two teacher variable.

***Model 9.*** *Unconditional model for Black/White*

Level 1- Students - Same as Equation 8.13

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \tag{8.19}$$

$$\beta_{1j} = \gamma_{10} + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student gains/achievement in teacher j's classroom holding teacher practice constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, black student gains/achievement in teacher j's classroom all other predictors held constant

$r_{4j}$ = the error term or unique teacher effects

The estimated level two regression coefficients for Model 9 are given in Table 8.16.

**Table 8.16** Regression Coefficients for Gains by Ethnicity 1999-2000

| | Model 9 |
|---|---|
| Intercept ($\gamma_{00}$) | .43(.94) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ |
| Hispanic ($\gamma_{20}$) | 2.29(1.02)* |
| All other ($\gamma_{30}$) | 1.38(.76)~ |
| Grade 4 ($\gamma_{80}$) | .88(.25)** |
| Grade 6 ($\gamma_{90}$) | .98(.19)*** |
| | |
| LVL 1 Variance ($\sigma 2$) | 83.95 |
| LVL 2 Intercept ($\tau_{00}$) | 3.13** |
| Black ($\tau_{10}$) | 18.16** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

The effect of Black varies significantly across teachers ($\tau_{10}$=18.30, p<.01) as does the effect of White ($\tau_{00}$=3.13, p<.01). Selected coefficients from this analysis serve as the base for the variance decompositions found in Tables 8.18-8.21. Next district SES is added to the model to determine its effect on the average achievement gains for blacks and whites across teachers.

***Model 10.*** *Black, White achievement gains and SES*

Level 1- Students - Same as Equation 8.13

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\,pctfrlu + r_{0j} \qquad (8.21)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\,pctfrlu + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student gains/achievement in teacher j's classroom holding teacher practice constant

$\gamma_{01j}$ = the change in average gains/achievement associated with mean district free and reduced lunch, all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, black student gains/achievement in teacher j's classroom all other predictors held constant

$\gamma_{11j}$ = the change in average black student gains/achievement associated with one unit change in average district free and reduced lunch

$r_{1j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for Model 6 are given in Table 8.17.

**Table 8.17** Regression Coefficients for Black/White Gains by Average District SES

| | Model 10 |
|---|---|
| Intercept ($\gamma_{00}$) | .50(.85) |
| Free and reduced lunch ($\gamma_{01}$) | 1.39(2.26) |
| Black ($\gamma_{10}$) | 3.86(1.17)** |
| Free and reduced lunch ($\gamma_{11}$) | -19.14(6.49)** |
| Hispanic ($\gamma_{20}$) | 1.92(.95)~ |
| All other ($\gamma_{30}$) | 1.32(.83)~ |
| Grade 4 ($\gamma_{80}$) | .88(.24)** |
| Grade 6 ($\gamma_{90}$) | .98(.17)*** |
| | |
| LVL 1 Variance ($\sigma 2$) | 84.13 |
| LVL 2 Variance ($\tau_{00}$) | 3.05** |
| ($\tau_{10}$) | 9.70* |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

How much do teachers vary around the average mathematics gains for White and Black students in average SES districts? To answer this question the following equation computes the variance in black gains explained by average SES across teachers.[16] The estimated $\tau_{10}$ values from Tables 8.16 and 8.17 would be substituted in the following equation:

$$\frac{\tau_{10}(\text{model 6}) - \tau_{10}(\text{model 7})}{\tau_{10}(\text{model 6})} = \frac{18.16 - 9.7}{18.16} = .465 = 46.5\% \qquad (8.22)$$

This means that 46.5% of the variance in the predicted average achievement gains for Black students is explained by variance in district SES across teachers. Using the same process for the remaining taus, the results are presented at the bottom of Tables 8.18.-8.21.

**Table 8.18** Variance Decomposition for Black/White Gains by SES

| | | SES |
|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lunch |
| Intercept ($\gamma_{00}$) | .43(.94) | 0.50(.85) |
| SES($\gamma_{02}$) | | 1.39(2.26) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ | 3.86(1.17)** |
| SES($\gamma_{12}$) | | -19.14(6.49)** |
| Intercept Variance ($\tau_{00}$) | 3.13** | 3.05** |
| Black Variance ($\tau_{10}$) | 18.16** | 9.70* |
| Level 1 variance ($\sigma^2$) | 83.95 | 84.13 |
| % variance reduction White | | 2.2% |
| % variance reduction Black | | 45.9% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

[16] A three-level analysis, students nested within teachers nested within districts, was conducted for the ten Connecticut districts that were represented by 5 or more teachers. There was no significant variance across districts for this sample so it was determined that the two-level model controlling for SES was the most appropriate.

The table indicates that being in a district with average SES (21% free and reduced lunch) has a large, negative effect on Black mathematics gains ($\gamma_{12}$ = -19.14, sd=6.49). One standard deviation increase in district percentage of free and reduced lunch accounts for a -19.14 point drop in average mathematics gains for black students across teachers and explains a small increase of +1.39 points in White students average gain. SES, therefore, has a substantial differential effect on Black and White student gains. There are two districts in the Connecticut data that have over 50% free and reduced lunch (over +2 sd). This analysis predicts that Blacks would be expected to lose -28.28 points over the course of one year while White students are predicted to gain +2.78 points for a two standard deviation increase in free and reduced lunch.

SES explains almost 46% of the variance in average black gains across teachers while it accounts for only 2% of the variance in average white (non-black) gains across teachers.

**Ethnicity and Teacher Effects**

Now that the context is set for the analyses at level two, it is possible to track the variance explained by each teacher level variable controlling for SES and to determine its effect on average mathematics gains for Black and White across teachers. One variable at a time is added to the level two equation, EQ 8.23. The term "*variable x*" is exchanged for a single Mathematics Practice (math sensemaking, drill and lecture, or group and project activities), Teacher Characteristic (mathematics preparation, preparation for mathematics instruction, mathematics knowledge, reform and traditional orientation, assessment prep, or efficacy), or Professional Development (professional development impact, mentoring frequency, or professional development frequency).

***Model 11***. Template for Level 2 equations

Level 1 remains constant (Eq. 8.13)

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, pctfrlu + \gamma_{02..}\, variable\ x + r_{0j} \qquad (8.23)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, pctfrlu + \gamma_{12..}\, variable\ x + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{02..j}$ = change in average in average student gains associated with a one unit change in "*variable x*" holding SES constant

and

$\gamma_{12..j}$ = the change in average black student gains associated with one unit change in "*variable x*" holding SES constant

**Ethnicity and Teacher Practices**. A variance decomposition for teacher practices can be found in Table 8.19. The table indicates that very little additional variance in predicted gains for White or Black students is explained by teacher practices across teachers once SES is controlled.

**Table 8.19** Variance Decomposition for White/Black Controlling SES - Teacher Practices

| | | SES | Practice | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | Sensemaking | Drill&Lecture | Group&Project |
| Intercept ($\gamma_{00}$) | .43(.94) | 0.50(.85) | .62(.95) | .56(.89) | .49(.93) |
| ($\gamma_{01}$) | | | .18(.32) | -.27(.31) | .04(35) |
| SES($\gamma_{02}$) | | 1.39(2.26) | 1.34(2.11) | 1.45(2.11) | 1.41(2.08) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ | 3.86(1.17)** | 3.96(1.14)** | 3.77(1.15)** | 3.85(1.16)** |
| ($\gamma_{11}$) | | | 1.25(.95) | .74(.89) | .21(1.13) |
| SES($\gamma_{12}$) | | -19.14(6.49)** | -20.58(6.50)** | -18.63(6.91)** | -19.19(6.60)** |
| Intercept Var ($\tau_{00}$) | 3.13** | 3.05** | 3.12** | 3.02** | 3.17** |
| Black Var ($\tau_{40}$) | 18.16** | 9.70* | 9.54* | 10.54* | 11.21* |
| ($\sigma$2) | 83.95 | 84.13 | 84.13 | 84.15 | 84.09 |
| % variance reduction White | | 2.2% | .3% | 3.5% | 0% |
| % variance reduction Black | | 45.9% | 47.5% | 42.0% | 38.3% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Drill and lecture activities in average SES districts explains 3.5% of the variance in White achievement gains across teachers, slightly more than SES alone. It explains no additional variance for Black gains across teachers. One standard deviation increase in drill and lecture activities decreases white gains by -.27 points while simultaneously increasing black gains by +.74 points, an indicator of differential effects. For Black students, only mathsensmaking activities across teachers explain additional variance in average SES districts. The predicted change attributable to one standard deviation increase in frequency of sensemaking activities is +.16 point for Whites and +1.28 points for Blacks. Group and project activities do not explain variance in average gains across teachers for either group.

**Ethnicity and Teacher Characteristics.** The variance decomposition for teacher characteristics follows in Table 8.20. There are only two variables which explain more variance in the gains across teachers than SES alone - mathematics preparation and reform orientation. Mathematics preparation explains additional variation in gains for both Black and White Students across teachers while reform orientation only explains additional variance in gains for Black students.

**Table 8.20** Variance Decomposition for Black/White Controlling for SES - Teacher Characteristics

| | | SES | Teacher Characteristics | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | MthPrep | MthInst | MthKnow |
| Intercept ($\gamma_{00}$) | .43(.94) | 0.50(.85) | 1.62(.94) | .47(.96) | .28(.97)) |
| ($\gamma_{01}$) | | | -.92(.39)* | .26(.31) | .41(.33) |
| SES($\gamma_{02}$) | | 1.39(2.26) | .61(1.92) | 1.69(2.07) | 1.72(2.12) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ | 3.86(1.17)** | 3.15(1.21)* | 3.83(1.18)** | 3.97(1.17)** |
| ($\gamma_{11}$) | | | -1.77(1.09) | .01(.95) | .42(.68) |
| SES($\gamma_{12}$) | | -19.14(6.49)** | -19.20(6.34)** | -19.22(6.52)** | -19.46(6.53)** |
| Intercept Var ($\tau_{00}$) | 3.13** | 3.05** | 2.33** | 3.08** | 3.08** |
| Black Variance ($\tau_{40}$) | 18.16** | 9.70* | 7.64* | 10.80* | 11.21* |
| Level 1 Variance ($\sigma 2$) | 83.95 | 84.13 | 84.32 | 84.13 | 84.07 |
| % variance reduction White | | 2.2% | 25.6% | 1.6% | 1.3% |
| % variance reduction Black | | 45.9% | 57.9% | 40.5% | 37.5% |

| | | SES | Teacher Characteristics | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | Reform Or | Trad Or | Assess |
| Intercept ($\gamma_{00}$) | .43(.94) | 0.50(.85) | .50(.92) | .50(.93) | .50(.92) |
| ($\gamma_{01}$) | | | -.07(.27) | .07(.34) | -.15(.26) |
| SES($\gamma_{02}$) | | 1.39(2.26) | 1.49(2.12) | 1.32(2.21) | 1.23(2.08) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ | 3.86(1.17)** | 3.73(1.14)** | 3.84(1.24)** | 3.86(1.17)** |
| ($\gamma_{11}$) | | | 1.33(.86) | -.03(1.06) | .69(.80) |
| SES($\gamma_{12}$) | | -19.14(6.49)** | -17.34(6.02)** | -19.04(6.44)** | -18.73(6.82)** |
| Intercept Var($\tau_{00}$) | 3.13** | 3.05** | 3.15** | 3.17** | 3.19** |
| Black Var ($\tau_{40}$) | 18.16** | 9.70* | 8.09* | 10.81* | 9.97* |
| Level 1 Var ($\sigma 2$) | 83.95 | 84.13 | 84.16 | 84.11 | 84.11 |
| % variance reduction White | | 2.2% | 0% | 0% | 0% |
| % variance reduction Black | | 45.9% | 55.5% | 40.4% | 45.1% |

**Table 8.20** (cont'd).

| | Uncond | SES | Teacher Characteristics |
|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | | Free/Red Lu | Efficacy |
| Intercept ($\gamma_{00}$) | .43(.94) | 0.50(.85) | .50(.93) |
| ($\gamma_{01}$) | | | .00(.30) |
| SES($\gamma_{02}$) | | 1.39(2.26) | 1.39(2.04) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ | 3.86(1.17)** | 3.80(1.23)** |
| ($\gamma_{11}$) | | | .44(.98) |
| SES($\gamma_{12}$) | | -19.14(6.49)** | -18.86(6.36)** |
| Intercept Var ($\tau_{00}$) | 3.13** | 3.05** | 3.19** |
| Black Variance ($\tau_{40}$) | 18.16** | 9.70* | 10.79* |
| Level 1 Variance ($\sigma 2$) | 83.95 | 84.13 | 84.11 |
| % variance reduction White | | 2.2% | 0% |
| % variance reduction Black | | 45.9% | 40.6% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Note that for every standard deviation increase in mathematics preparation, both Black and White students are predicted to lose points and the loss is doubled for Black students (Black= -1.78, White =-.90). The negative relationship between teacher preparation and student achievement is puzzling but inline with results from the earlier SEM analyses in this dissertation and with Valli and Reckase (2000), and Rowan, Correnti, and Miller (2004).

Reform orientation controlling for SES across teachers explains 55% of the variation in Black gains over the course of one year while explaining no variance in the gains for White students. For every one standard deviation increase in teachers' reform orientation, Black students are expected to gain an additional +1.35 points over the course of one year. White students are expected to remain nearly level (-.08, sd=.27), another differential effect.

Although teachers' math knowledge explains significant variance in gains by gender, this is not true for Black and White gains across teachers in Connecticut. Both groups are predicted to increase gains by the same amount, about +.41 points, for every one standard deviation increase in mathematics knowledge across teachers.

**Ethnicity and Teacher Professional Development**. This analysis focuses on the differential effects of teacher professional development as measured by professional development impact, mentoring frequency, and professional development frequency. All professional development variables explain additional variance in mathematics achievement gains across teachers for White after SES is controlled. Only professional development frequency explains additional variance in Black gains across teachers once SES is controlled. The variance decomposition for professional development follows in Table 8.21.

**Table 8.21** Variance Decomposition for Black/White controlling for SES - Teacher Professional Development

| | | SES | Professional Dev | | |
|---|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | PDImpact | MentFrq | PD Frq |
| Intercept ($\gamma_{00}$) | .43(.94) | 0.50(.85) | -.19(.92) | .48(.86) | .24(.86) |
| ($\gamma_{01}$) | | | -.53(.27)~ | -.77(.32)* | -1.08(.24)*** |
| SES($\gamma_{02}$) | | 1.39(2.26) | .74(2.06) | .56(2.22) | 1.29(2.03) |
| Black ($\gamma_{10}$) | 1.92(1.04)~ | 3.86(1.17)** | 3.78(1.22)** | 3.62(1.18)** | 3.96(1.23)** |
| ($\gamma_{11}$) | | | -.03(.98) | -1.27(.94) | .90(.88) |
| SES($\gamma_{12}$) | | -19.14(6.49)** | -18.78(6.52)** | -19.82(6.59)** | -19.77(6.94)** |
| Intercept Variance($\tau_{00}$) | 3.13** | 3.05** | 2.83** | 2.58 | 1.74** |
| Black Variance ($\tau_{10}$) | 18.16** | 9.70* | 10.22* | 10.20 | 8.63* |
| ($\sigma$2) | 83.95 | 84.13 | 84.16 | 84.02 | 84.31 |
| % variance reduction White | | 2.2% | 9.9% | 17.5% | 44.4% |
| % variance reduction Black | | 45.9% | 43.0% | 43.7% | 52.5% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

For every standard deviation increase in professional development impact, Black students are predicted to lose nothing compared to White students who lose about half a point. Professional development impact in average SES districts explains 9.9% of the variation in White gains across teachers, an increase of 7.7% over SES alone.

For every one standard deviation increase in mentoring frequency in average SES districts, Black students lose -1.27 points. White students lose -.77 points for every standard deviation increase in mentoring frequency. Mentoring frequency explains 44.4% of the variation in average White gains across teachers controlling for SES and no additional variation over SES alone for Black gains across teachers.

Once SES is controlled, professional development frequency explains the most variance in the average gains for both Black and White students across teachers. Professional development frequency explains 44% of the variation in White gains across teachers and 53% of the variation in average Black gains. For every one standard deviation increase in professional development frequency, White gains are expected to drop by -1.08 points. Black gains are predicted to increase by +.90.

**Summary of differential effects by ethnicity** - In Connecticut, teachers vary significantly around the average mathematics gains for two groups, Black and White students. SES explains 45.9% of the variation in Black gains but only 2.2% of the variation in White gains. For every one standard deviation increase in free and reduced lunch, Black students lose -19.1 points.

Variables that explain additional variance in Black gains across teachers once SES is controlled are math sensemaking activities (an additional 1.5% of the variance explained), mathematics preparation (an additional 10.9%), reform orientation (an additional 9.1%), and professional development frequency (an additional 6.8%). Every one standard deviation increase in the teacher math sensmaking activities increases Black gains by +1.28 points over the course of one year; mathematics preparation decreases in Black gains by -1.78 points; reform orientation increases in Black gains of +1.35 points; and professional development frequency increases average Black gains by +.10 points.

Professional development variables explain the largest amounts of variance in White gains across teachers, an additional 7.7-42.7% once SES is controlled. Mathematics preparation explains an additional 23% of the variance in White gains across teachers after controlling for SES. The only teacher practice variable that adds slightly to variance explained for White student gains across teachers is drill and lecture activities (an additional 1.3%).

For every one standard deviation increase in the teacher variable for professional development impact, mentoring frequency, or professional development frequency, average White gains are expected to decrease from -.54 to -.75 points across teachers. For every standard deviation increase math drill and lecture activities, average White gains across teachers are predicted to decrease by -.29 points. For every standard deviation increase in mathematics preparation , White gains are also expected to decrease -.90 points.

There are indications that certain teacher variables, reform orientation and professional development frequency, have differential effects on average gains for Black and White students across teachers when SES is controlled. Generally, however, teacher variables predict losses for both groups or gains for both groups. Most of the variation explained in Black scores is explained by SES and much of the variation in White scores is explained by professional development variables.

## III. IHE Teacher Preparation, Teachers, and Student Achievement

This section addresses questions about university teacher preparation programs and their relationship to teacher practices, characteristics, professional development, and subsequent student achievement. That is to say, do Connecticut teacher preparation programs vary in the average levels of 4th-8th grade student gains, teacher practice, characteristics, or professional development they predict The analyses in this section use a smaller sample of teachers who indicated that they attended specific teacher preparation programs in Connecticut.

**Effects of teacher preparation programs on the mathematics gains of 4th-8th grade students.** The equation for the three level mathematics analysis with students nested within teachers nested within teacher preparation institutions is shown by Equation 8.23. There are 1098 students nested within 61 teachers nested within 13 identified teacher preparation institutions in Connecticut.[17].

$$\begin{aligned} Y_{ijk} &= \pi_{0jk} + e_{ijk}, \\ \pi_{0jk} &= \beta_{00k} + r_{0jk}, \\ \beta_{00k} &= \gamma_{000} + u_{00k} \end{aligned} \qquad (8.24)$$

Less than 1% of the variance in student mathematics gains in the unconditional three level model was located at level three, teacher preparation institution,. In other words, teacher preparation program effects are nearly identical when predicting 4th-8th grade student gains nested within their teacher graduates. Lack of variability at level three suggests that a simpler two level model is the better choice for further analysis

[17] Fixed effects are unbaised although standard errors are likely too small in small samples (Raudenbush, p. 281).

(Raudenbush et al., 2000, p237). No further investigation into variance explained in student gains across teachers by teacher preparation institution was possible for this data.

**Effects of teacher preparation on levels of teacher practice, characteristics, and professional development.** HLM analyses from Section 2 indicated that ten variables (math sensemaking, drill and lecture, group and project activities, mathematics preparation, assessment preparation, reform orientation, teacher efficacy, professional development impact and frequency, and mentoring frequency) controlling for SES explained variation in student gains by gender or ethnicity across teachers. A two level model with Connecticut teacher preparation institutions at level two predicting one teacher variable at a time at level one was constructed. The model included 61 teachers nested within 13 teacher preparation institutions.

Of the ten variables found to explain variance in average student gains by gender and ethnicity across teachers, only one, mentoring frequency, varies across teacher preparation programs. Although it did not explain variation in student gains across teachers, teachers' mathematics knowledge also varied across institutions so it is included below.

The set of new variables used in the Connecticut teacher preparation analysis follow in Table 8.22.

**Table 8.22** Description of Teacher Preparation Variables

| Type | Variable Abbreviation | Variable Description (Aggregated from teacher level) |
|---|---|---|
| Preparation Program Characteristics | facchar | Faculty characteristics (mean=24.37, sd=1.82) |
| | progqual | Program quality (mean=18.71, sd=1.93) |
| | fieldexp | field experience (mean=16.52, sd=1.45) |
| | rank | Program rank by pass rates on Praxis basic math (mean=6.28, sd=3.51) |

Effects of teacher preparation institutions on teachers' mathematics knowledge and mentoring frequency. In the next set of analyses, characteristics of teacher preparation programs predict levels of average teacher mathematics knowledge and frequency of mentoring.

Unconditional models predicting average mathematics knowledge and mentoring frequency are presented first followed by the results for each of four teacher preparation program variables, faculty characteristics, program quality, field experiences, and program rank. Program rank is based on teacher candidate pass rates on the ETS Praxis Basic Skills Exam, CBT: Mathematics, from 1998 as reported in the Initial Report of the Secretary on the Quality of Teacher Preparation (Riley, 2000), otherwise known as the State Report Card.

***Model 12.*** *Unconditional model for teacher variables across teacher preparation institutions*

Level 1 - Teacher

$$variable\ x_{ij} = B_0 + e_{ij} \quad (8.25)$$

Level 2 - IHE teacher preparation program

$$\beta_{0j} = \gamma_{00} + r_{0j} \tag{8.26}$$

The coefficients for mathematics knowledge are presented in Table 8.22 followed by the coefficients for mentoring frequency in Table 8.23. Mathematics knowledge

**Table 8.23** Connecticut's Average Mathematics Knowledge by Teacher Preparation Institutions

| | Model 12a |
|---|---|
| Intercept ($\gamma_{00}$) | 26.46(.66)*** |
| LVL 1 Variance ($\sigma^2$) | 11.89 |
| LVL 2 Variance ($\tau_{00}$) | 2.34*** |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Predicted average mathematics knowledge reported by teachers varies significantly across teacher preparation institutions in Connecticut. 83.6% of the variance in mathematics knowledge for Model 12 is within programs (between teachers) leaving 16.4% of the variance between programs. That is to say, there is five times more variation in average mathematics knowledge within programs than between programs. The average level of mathematics knowledge is 26.46 (sd=.63, p<.001) when no other variables are included in the model. This average scale score roughly corresponds to "good."

Results from the unconditional model for mentoring frequency are presented in Table 8.23. For this analysis, mentoring frequency is substituted for *variable x* in Eq. 8.24.

**Table 8.24** Connecticut's Average Mentoring Frequency by Teacher Preparation Institutions

| | Model 12b |
|---|---|
| Intercept ($\gamma_{00}$) | 33.38(2.31)*** |
| LVL 1 Variance ($\sigma 2$) | 143.16 |
| LVL 2 Variance ($\tau_{00}$) | 35.61* |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Predicted average mentoring frequency varies significantly across teacher preparation institutions in Connecticut. 80.1% of the variance in mentoring frequency for Model 12b is within programs (between teachers) leaving 19.9% of the variance between programs. That is to say, there is four times more variation within programs than between programs. The average level of mentoring frequency is 33.38 (sd=2.31, p<.001) when no other variables are included in the model. This roughly corresponds to a scale frequency of "a few times."

There are three rating scales for teacher preparation programs measured at the teacher level: faculty characteristics, program quality, and field experiences. These were aggregated from the teacher level to the program level to test if they explained variance in mathematics knowledge and mentoring frequency. The fourth variable, program rank, reflects teacher candidate pass rates on the Connecticut basic skill mathematics test from 1998.

***Model 13.*** *Base model to predict teacher variables by teacher preparation program characteristics*

Level 1 - Teacher

$$variable\ x_{ij} = B_0 + e_{ij} \tag{8.27}$$

Level 2 - IHE teacher preparation program

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, program\ characteristic\ y + r_{0j} \qquad (8.28)$$

where the *program characteristic y* is faculty characteristics, program quality, field experiences, or program rank.

The coefficients for mathematics knowledge are presented in Table 8.25 followed by the coefficients for mentoring frequency in Table 8.26.

**Table 8.25** Average Mathematics Knowledge Predicted by Program Characteristics

| | Unconditional | Faculty Characteristics | Program Quality | Field Experiences | Program Rank on CBT: Math |
|---|---|---|---|---|---|
| Intercept ($\gamma_{00}$) | 26.46(.66)*** | 26.55(.69)*** | 26.50(.70)*** | 26.49(.69)*** | 26.46(.70)*** |
| Prog Char ($\gamma_{01}$) | | .28(.48) | .29(.53) | .12(.57) | .04(.20) |
| LVL 1 Variance ($\sigma 2$) | 11.89 | 11.89 | 11.78 | 11.88 | 11.86 |
| LVL 2 Variance ($\tau_{00}$) | 2.34*** | 2.67* | 2.95* | 2.84* | 2.91* |
| Var exp by Prog Char at Lvl 2 | | 0% | 0% | 0% | 0% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

None of the teacher preparation program characteristics in this data explain variance in average mathematics knowledge reported by teachers across preparation programs.

**Table 8.26** Average Mentoring Frequency Predicted by Program Characteristics

| | Unconditional | Faculty Characteristics | Program Quality | Field Experiences | Rank on CBT: Math |
|---|---|---|---|---|---|
| Intercept ($\gamma_{00}$) | 33.28(2.42)*** | 32.99(2.14)*** | 33.45(2.30)*** | 33.55(2.32)*** | 33.55(2.48)*** |
| Prog Char ($\gamma_{01}$) | | -1.11(1.14) | 2.02(1.79) | 2.09(1.91) | -.53(.71) |
| LVL 1 Var ($\sigma 2$) | 143.16 | 143.47 | 145.72 | 145.57 | 143.44 |
| LVL 2 Var ($\tau_{00}$) | 35.61* | 38.79* | 28.73* | 29.41* | 37.90* |
| Var exp by Prog Char at Lvl 2 | | 0% | 19.3% | 17.4% | 0% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

Of the four measures for teacher preparation programs, two explain variance in average teacher mentoring frequency across teacher preparation institutions. Program quality explains 19.3% of the variance and field experiences explain 17.4% of the variation in teacher mentoring frequency across teacher preparation institutions. Both program quality and field experiences as rated by graduates of the program predict a 2 unit increase in mentoring frequency for every one standard deviation increase in program quality or field experiences. The mean of mentoring frequency roughly corresponds to "a few times."

**Ranking teacher preparation institutions on levels mathematics knowledge and mentoring frequency.** For the next analysis, 13 teacher preparation institutions are added to the model as dummy variables at level two. The intercept is allowed to vary. Model 13 answers the question, "Are different levels of specific teacher variables associated with particular teacher preparation institutions in Connecticut?" For these analyses, one of Connecticut's largest teacher preparation programs serves as the reference category and three small teacher preparation programs are combined into one dummy variable for "all other universities."

***Model 13.***

Level 1 - Teacher

$$variable\ x_{ij} = B_0 + e_{ij} \tag{8.29}$$

Level 2 - IHE teacher preparation program

$$\begin{aligned} B_{00} = \gamma_{00} + \gamma_{01}\ Univ1 + \gamma_{02}\ Univ2 + \gamma_{03}\ Univ3 + \gamma_{04} Univ4 \\ + \gamma_{05} Univ5 + \gamma_{06} Univ6 + \gamma_{07} Univ7 + \gamma_{08} Univ8 \\ + \gamma_{09} Univ9 + \gamma_{10} AllOthUniv \end{aligned} \tag{8.30}$$

where

$variable\ x_{ij}$ = one variable of two teacher variables (mathematics knowledge or mentoring frequency)

$r_i$ = the error term or unique teacher preparation program effects

Coefficients for mentoring frequency by teacher preparation institutions are presented in Tables 8.27.

**Table 8.27** Average Mathematics Knowledge Distributed Across Teacher Preparation Institutions

| | Unconditional | Mathematics Knowledge |
|---|---|---|
| Intercept ($\gamma_{00}$) | 26.46(.63)*** | 26.64(.70)*** |
| University 1 ($\gamma_{01}$) | | 6.80(2.87)~ |
| University 2 ($\gamma_{02}$) | | 5.25(2.98) |
| University 3 ($\gamma_{03}$) | | 4.67(2.80) |
| University 4 ($\gamma_{04}$) | | 3.00(2.63) |
| University 5 ($\gamma_{05}$) | | 6.33(2.80) |
| University 6 ($\gamma_{06}$) | | 4.00(3.14) |
| University 7 ($\gamma_{07}$) | | 1.68(2.87) |
| University 8 ($\gamma_{08}$) | | 5.55(2.72) |
| University 9 ($\gamma_{09}$) | | 8.00(3.44)~ |
| Three Other Universities ($\gamma_{30}$) | | 5.00(2.98) |
| LVL 1 Variance ($\sigma 2$) | 11.89 | 11.90 |
| LVL 2 Variance ($\tau_{00}$) | 2.34*** | 1.45 |
| Variance explained at Lvl 2 | | 38.0% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

The addition of dummy variables for 13 teacher preparation institutions explains 38% of the variance in average teachers' mathematics knowledge across programs. $\tau_{00}$ is no longer significantly different from zero. Fixed effects for University 1 may be computed using the following formula:

$$\hat{Y}_{mthknow} = \hat{\beta}_0 \tag{8.31}$$

$$\begin{aligned}\hat{\beta}_0 &= \hat{\gamma}_{00} + 1 * \hat{\gamma}_{01} \\ &= 26.64 + 1 * (6.80) \\ &= 33.44\end{aligned}$$

Table 8.28 presents the fixed effects for mathematics knowledge by teacher preparation program sorted from lowest to highest.

**Table 8.28** Distribution of Predicted Mathematics Knowledge across Teacher Preparation Programs (low to high)

| | Average Mathematics Knowledge | Approximate Scale Correspondence |
|---|---|---|
| Reference university | 26.63 | Good |
| University 7 ($\gamma_{07}$) | 28.31 | Good |
| University 4 ($\gamma_{04}$) | 29.63 | Good |
| University 6 ($\gamma_{06}$) | 30.63 | Excellent |
| University 3 ($\gamma_{03}$) | 31.30 | Excellent |
| University 2 ($\gamma_{02}$) | 31.58 | Excellent |
| Three Other Universities ($\gamma_{30}$) | 31.64 | Excellent |
| University 8 ($\gamma_{08}$) | 32.19 | Excellent |
| University 5 ($\gamma_{05}$) | 32.96 | Excellent |
| University 1 ($\gamma_{01}$) | 33.44 | Excellent |
| University 9 ($\gamma_{09}$) | 34.64 | Excellent |

There are small differences in expected teacher mathematics knowledge across teacher preparation programs but none of the coefficients are significantly different from zero. All programs are predicted to produce teachers, who on average, report good to excellent knowledge of mathematics.

Coefficients for mentoring frequency by teacher preparation institutions are presented in Tables 8.29.

**Table 8.29** Average Mentoring Frequency Distributed Across Teacher Preparation Institutions (low to high)

| | Unconditional | Mentoring Frequency |
|---|---|---|
| Intercept ($\gamma_{00}$) | 33.38(2.31)*** | 33.56(2.11)*** |
| University 1 ($\gamma_{01}$) | | -16.00(8.05) |
| University 2 ($\gamma_{02}$) | | -7.488(8.47) |
| University 3 ($\gamma_{03}$) | | -22.83(7.75) |
| University 4 ($\gamma_{04}$) | | -.33(6.94) |
| University 5 ($\gamma_{05}$) | | -9.99(7.75) |
| University 6 ($\gamma_{06}$) | | -13.33(9.14) |
| University 7 ($\gamma_{07}$) | | -3.40(8.05) |
| University 8 ($\gamma_{08}$) | | -7.49(7.36) |
| University 9 ($\gamma_{09}$) | | 10.00(10.35) |
| Three Other Universities ($\gamma_{30}$) | | -8.64(9.12) |
| | | |
| LVL 1 Variance ($\sigma 2$) | 143.16 | 141.45 |
| LVL 2 Variance ($\tau_{00}$) | 35.61* | .53 |
| Variance explained at Lvl 2 | | 98.5% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

The addition of dummy variables for 13 teacher preparation institutions explains nearly all of the variance in mentoring frequency across programs. $T_{00}$ is no longer significantly different from zero. Fixed effects for University 1 may be computed using the following formula:

$$
\begin{aligned}
\hat{Y}_{mentfrq} &= \hat{\beta}_0 \\
\hat{\beta}_0 &= \hat{\gamma}_{00} + 1 * \hat{\gamma}_{01} \\
&= 33.56 + 1 * (-16.00) \\
&= 17.56
\end{aligned}
\tag{8.32}
$$

Table 8.25 presents the fixed effects for mentoring frequency by teacher preparation program.

**Table 8.30** Distribution of Mentoring Frequency Across Teacher Preparation Programs (low to high)

| | Average Mentoring Frequency | Approximate Scale Correspondence |
|---|---|---|
| University 3 ($\gamma_{03}$) | 10.73 | Never |
| University 1 ($\gamma_{01}$) | 17.56 | Once |
| University 6 ($\gamma_{06}$) | 20.23 | Once |
| University 5 ($\gamma_{05}$) | 23.57 | More than once |
| Three Other Universities ($\gamma_{30}$) | 24.92 | More than once |
| University 8 ($\gamma_{08}$) | 26.07 | More than once |
| University 2 ($\gamma_{02}$) | 26.08 | More than once |
| University 7 ($\gamma_{07}$) | 30.16 | More than once |
| University 4 ($\gamma_{04}$) | 33.23 | A few times |
| Reference university | 33.56 | A few times |
| University 9 ($\gamma_{09}$) | 43.56 | Monthly |

There are differences in the expected frequency of teacher mentoring predicted by teacher preparation program but none of the coefficients are significantly different from zero.

***Summary of the HLM results for mathematics gains in Connecticut.***

Once the fixed effects for gains by ethnic and gender groups controlling for grade level were estimated at level one, it is possible to answer the main research questions regarding how those relationships changed when teacher practices, teacher characteristics, and professional development variables were added to the model. Fixed effects for students are presented first followed by the random effects of teachers on student gains. Last, the results for the analyses of student outcomes and teacher variables by teacher preparation institution in Connecticut are presented. The findings for mathematics gains in Connecticut can be summarized as follows:

## Results Summary

### Section I Results Summary. Contextual Effects of Gender, Ethnicity, and Diistrict SES on mathematics gains in Connecticut

**1.** Fixed effects at the student level show significant gains by gender, ethnic groups, and grade level. Girls gained more in mathematics than boys. Black, Hispanic, and All other students gained more in mathematics than White students over the course of one year. Coefficients for all groups are significantly different from zero with the exception of All other students which is nearly significant at $p<.053$. There are no gender by ethnicity interactions.

**2.** After grade level is introduced to the model, the patterns for gains by gender and ethnicity remain the same. Grade level, however, controlling for gender and ethnicity is a large predictor of mathematics gains and is highly significant at $p<.001$. $4^{th}$ grade and $6^{th}$ grade each account for nearly 1 additional point in mathematics gains over the course of the year. Curiously, being in Grade 8 controlling for gender and ethnicity does not increase gains over the year.

**3**. Most of the variation in Connecticut mathematics classroom is between students (91%). About 9% of the variation is located at the teacher level.

**4.** District SES varied significantly across teachers when predicting student mathematics gains by gender and ethnicity. Being in an average SES district (21% free and reduced lunch) was associated with a -6.44 loss in gains for boys when compared to girls and a -19.14 loss in gains for Black students when compared to White students. There are two Connecticut districts in this data that have more that 42% of it's students receiving free and reduced lunch. In these districts, the losses for boys and Black students are predicted to double.

Average district SES explains 23.2% of the variation in female gains and 6.1% of the variation in males gains across teachers. It affects male and female gains differentially – boys lose -6.44 points while girls gain +7.9 points for a one standard deviation increase in average district free and reduced lunch.

Average district SES explains 46.5% of the variation in Black gains across teachers but only 2.2% of the variation in White gains. Average district SES affects Black and White gains differentially – Black students are predicted to lose over -19 points in an average SES district while White students gain +1.4 points across teachers.

**Section II Results Summary. Teacher practice, characteristics, and professional development effects on student mathematics gains in Connecticut**

**4.** Average mathematics gains for boys and girls controlling for SES varies across teachers. There is twice as much variation in average male gains compared to female gains across teachers but teacher variables generally explain more of the variation that exists in average gains for girls across teachers once SES is controlled

. **a.** *Teacher practice* variables controlling for SES explain 5.7% to 8.6% of the variation in boys' average mathematics gains. The same practice variables explain 23% to 38% of the variance in girls' gains across teachers. Mathematics sensemaking, drill and lecture, and group and project activities have differential effects on average gains for boys and girls across teachers. Boys, on average, gain more than half a point for a one standard deviation increase in mathematics sensemaking and group and project activities while girls lose about one point over the course of the year. The opposite is true for mathematics drill and lecture activities. Girls gain about half a point and boys lose about one/third of a point for a one standard deviation increase in drill and lecture activities controlling for SES.

**b.** *Teacher characteristics* of reform orientation, preparation for assessment, and efficacy explain up to 41.7% of the variance in girls' gains across teachers controlling for SES, more variation than SES alone. The same variables only explain up to 7.9% of the variance in boys' gains across teachers. All three have differential effects on boys and girls across teachers controlling for SES. In general, girls lose up to 1 point over the course of the year while boys gain up to half a point for every standard deviation increase in these three teacher characteristics.

Mathematics preparation is negatively related to gains for both boys and girls. A one standard deviation increase in mathematics preparation results in a -.86 point loss for boys and a -.18 point loss for girls across teachers controlling for district SES. Similar results are reported by Valli and Reckase (2000) and Correnti, Rowan, and Miller (2002).

**c.** One teacher *professional development* variable, mentoring frequency, explains more variation in girls gains across teachers than SES alone. For a one standard deviation increase in mentoring frequency, boys remain about level (-.08) and girls lose -1.41 points. This explains about 28% of the variation in girls' average gains across teachers.

**5.** Average gains for Black and White students controlling for SES varied across teachers. There is 3 times as much variance in Black gains compared to White gains across teachers once SES is controlled. Average gains for Hispanic and all other students did not vary significantly from White students.

**a.** *Teacher practices* explain little about the variation in average mathematics achievement across teachers for Black and White students. The teacher practice variable for mathematics sensemaking controlling for SES explains 47.5% of the variation in Black average mathematics gains across teachers, approximately 2% more than SES alone. The same variable explains no additional variation in White gains. For every one standard deviation increase in sensemaking activities, Blacks gain +1.25 points and Whites gain +.18 points. Drill and lecture activities is the only variable which explains more variation in White gains than SES alone, about 2%. For every one standard deviation increase in drill and lecture activities, white students lose -.27 points and Blacks gain +.74 points, a differential effect.

**b.** *Teacher characteristics* for math preparation and reform orientation controlling for SES explain up to 57% of the variation in Black gains and up to 25.6% of White gains across teachers. Math preparation explains the largest amount of variation for both groups across teachers. It accounts for an additional 23% of the variation in average White gains and 12% of the additional variation in average Black gains over SES alone. For every one standard deviation increase in mathematics preparation, White students lose nearly a point and Black students lose nearly -2 points over the course of the year.

Reform orientation explains about 10% more variance in Black gains across teachers than SES alone. For every standard deviation increase in reform orientation, Black students gain +1.33 points across teachers. White students' average gains remain nearly level and it explains no additional variation.

**c.** Of the *professional development* variables, only professional development frequency explains additional variance in Black gains when SES is controlled, about 6% more. The effect of professional development frequency is differential. For every standard deviation increase in professional development frequency, Black students gain nearly a point while White students lose a point. Professional development impact and mentoring frequency explain additional variation in White gains across teachers, 7% to 40% more than SES alone. For every one standard deviation increase in professional development impact, White students lose half a point while Black student gains remain nearly level across teachers. For every standard deviation increase in mentoring frequency, White students lose -.77 points and Black students lose nearly -1.25 points.

**Section III Results Summary. University teacher preparation program effects on student mathematics gains and teacher variables.**

**6.** Average 4th-8th grade student *mathematics gains* produced by teachers who graduated from a specific program did not vary by teacher preparation programs.

**7.** Only one of the teacher variables that explained variance in average student scores across teachers also varied across teacher preparation programs, mentoring frequency. Both program quality (19.3%) and field experiences (17.4%) explained variance in average teacher mentoring frequency.

Although it did not explain variance in average student scores across teachers, mathematics knowledge varied across teacher preparation institutions. None of the measured university level teacher preparation variables (faculty characteristics, program quality, field experience, and program rank from the initial state report card) explained variance in average teacher knowledge across teacher preparation programs.

It is interesting to note that program rank based on pass rates reported in the Initial State Report Card did not explain variance in average mentoring frequency or mathematics knowledge.

**8.** It is possible to rank teacher preparation programs in Connecticut based on the fixed effects produced by a two level HLM model. Level two dummy variables for teacher preparation institutions predict levels of mathematics knowledge and mentoring frequency at level one. Table 8.31 presents program rankings on the average mathematics knowledge and mentoring frequency of its graduates.

**Table 8.31** Teacher Preparation Programs Ranked on Teacher Graduates' Mathematics Knowledge and Mentoring Frequency

| | Mathematics knowledge | Mentoring Frequency |
|---|---|---|
| University 0 | 1 | 10 |
| University 1 | 10 | 2 |
| University 2 | 6 | 6 |
| University 3 | 5 | 1 |
| University 4 | 3 | 9 |
| University 5 | 9 | 4 |
| University 6 | 4 | 3 |
| University 7 | 2 | 8 |
| University 8 | 8 | 7 |
| University 9 | 11 | 11 |
| Three Other Universities | 7 | 5 |

It is important to interpret this table with caution since institutions did not vary widely on mathematics knowledge which ranged from 26.63 (good) to 34.64 points (excellent). Mentoring frequency had a larger range from 10.73 (never) to 43.56 points (monthly).

## 9. RESULTS FROM THE HLM ANALYSIS OF CONNECTICUT $4^{th}$-$8^{th}$ READING GAINS

This section, like the previous section for mathematics, addresses the research questions about the relationships among student reading gains, and

- Teacher characteristics and practices
- Characteristics of professional development
- Characteristics of teacher preparation

The analysis also focuses on the ways in which these relationships vary by student demographic group, including gender, ethnicity, and grade. As before, Tennessee is not included in the HLM analysis because there is no individual measure of student achievement available in the BTPS data, only correlation coefficients based on the aggregation of student scores within teacher.

To set the context for addressing the questions, results are presented first which describe the differences in student reading gains in Connecticut across student demographic groups. The analyses yield results in terms of regression coefficients. To aid in understanding what the coefficients mean, predicted gains are computed for selected combinations of student characteristics. As was the case for mathematics, district percent free and reduced lunch is assigned to each teacher as a measure of classroom socioeconomic status since there was no individual student indicator available in the BTPS data.

The full set of reading variables for Connecticut used in these analyses can be found in Table 9.1.

**Table 9.1** Description of Variables Used in the Connecticut Reading Analyses.

| Type | Variable Abbreviation | Variable Description |
|---|---|---|
| | | **Level 1 - Students** |
| Continuous (Outcome) | readdiff | Reading gain, difference between pre- and post-test scores (mean = 3.40, sd = 7.71) |
| Dichotomous | dblack | Black (Black=1, else=0; mean=16.0, sd=.37) |
| | dhispanic | Hispanic (Hispanic=1, else=0; mean=.09, sd=.29) |
| | dallother | All other ethnicities (all other=1, else=0; mean=.08, sd=.27) |
| | dfem | Female (female=1, else=0; mean=.53, sd=.50) |
| | dgrade4 | Grade 4 (grade 4=1, else=0; mean=.43, sd=.50) |
| | dgrade6 | Grade 6 (grade 6=1, else=0; mean=.32, sd=.47) |
| Effects coded | egenxbl | Effect of Black female; interaction of eBlack x egender (Black =1, White=-1, otherwise=0 multiplied by female=1, male=-1; mean=.01, sd=.91) |
| | egenxhisp | Effect of Hispanic female; interaction of eHispanic x egender (Hispanic=1, White=-1, otherwise=0 multiplied by female=1, male=-1; mean=.01, sd=.87). |
| | egenxoth | Effect of All other female; interaction of eAllother x egender (All other =1, White=-1, otherwise=0 multiplied by female=1, male=-1; mean=01, sd=.87) |
| | | **Level 2 - Teachers** |
| Continuous | ztradlit | Traditional literacy activities (mean=0, sd=1) |
| Z-scores | zreform | Reform literacy activities (mean=0, sd=1) |
| | zlitmat | Literacy materials usage (mean=0, sd=1) |
| | zcomp | Computer activities (mean=0, sd=1) |
| | zdivprep | Preparation for diversity (mean=0, sd=1) |
| | zblito | Balanced literacy orientation (mean=0, sd=1) |
| | zEngprep | English preparation (mean=0, sd=1) |
| | zlitknow | Self report of literacy knowledge (mean=0, sd=1) |
| | zreadinst | Reading instruction, reading pedagogy (mean=0, sd=1) |
| | zassessprep | Preparation for assessment (mean=0, sd=1) |
| | zefficacy | Teacher sense of efficacy, that a teacher can make a difference in students' learning (mean=0, sd=1) |

**Table 9.1** (cont'd)

| | | |
|---|---|---|
| | zreadprep | Reading preparation (mean=0, sd=1) |
| | zpdimpact | Professional development impact, teacher's sense of the impact that professional development had on teaching (mean=0, sd=1) |
| | zpdfreq | Professional development frequency (mean=0, sd=1) |
| | zmentfrq | Mentoring frequency (mean=0, sd=1) |
| Continuous | pctfrlu | Percent district free and reduced lunch (mean=.22, sd=.17) |
| | | **Level 3 - Teacher preparation programs** |
| Continuous | facchar | Faculty characteristics (mean=23.99, sd=1.22) |
| | fieldexp | Field experiences (mean=16.32, sd=1.30) |
| | progqual | Program quality (mean=18.35, sd=1.50) |
| | rank | Rank on the State Report Card (2000), The Initial Report of the Secretary of Education on the Quality of Teacher Preparation based on the number of teachers passing the state certification examination (basic) (mean=6.67, sd=4.27) |

**I. Setting the Stage: Variation in Reading Achievement Gains by Student Demographic Groups** (Results at the Student Level)

In this section, sets of background variables are added to look at the changes in reading gains in Connecticut. Model 1 predicts mean student gains in reading across the sample and partitions the variance between students and teachers where students are nested within teachers. Models 2-5 predict differences in mean gains across gender and ethnicity.

Mean Gains. To set the context for answering the questions about the connections between student achievement and teacher variables, results are presented first for differences in student achievement gains across demographic groups. The level one and level two equations for the model are given followed by their estimated regression coefficients in Table 9.2.

***Model 1.*** *Mean Reading Gains*

Level 1 - Students

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (9.1)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

where $Y_{ij}$ is the reading gain for student i in teacher j's classroom, $\beta_{0j}$ is the intercept or a student's mean reading gain within teacher j's classroom, $e_{ij}$ is the level 1 error, $\gamma_{00}$ is the mean value of classroom gains for all teachers, and $r_{0j}$ is the level 2 error.

**Table 9.2** Connecticut's Average Reading Gain 1999-2000

| | Model 1 |
|---|---|
| Intercept ($\gamma_{00}$) | 3.44(.31)*** |
| LVL 1 Variance ($\sigma^2$) | 53.94 |
| LVL 2 Variance ($\tau_{00}$) | 5.76*** |

*** p<.001,** p<.01,* p<.05,~<.10

The predicted average gain for $4^{th}$-$8^{th}$ grade reading students for this sample of students in Connecticut is 3.44 points ($\gamma_{00}$=3.44, sd=.31) when no other variables are included in the model.

$\sigma^2$ and $\tau_{00}$ are used to compare variation in the errors at the teacher level with variation at the student level using the formulas:

$$\frac{\sigma^2_{ij}}{\sigma^2_{ij} + \tau_{00}} = \textit{level one variance} \qquad \frac{\tau_{00}}{\sigma^2_{ij} + \tau_{00}} = \textit{level two variance} \quad (9.2)$$

90.4% of the variance in gains for Model 1 is within teachers (between students) leaving 9.6% of the variance in gains between teachers. This means that there is nine times more variation within teachers than between teachers.

Mean Gains by Demographic Group. Student demographic variables are added to form Model 2, 3, and 4. Model 5 introduces the control variables for grade level. The equations for Model 2, ethnic group, are presented first (Eq. 9.3). To interpret the coefficients, the estimated gain for each combination of gender and ethnicity can be computed by using the appropriate set of 1s and 0s for the dummy variables. As an example, Equations 9.5 and 9.6 demonstrate the process for obtaining Black student gains. The estimated gains for ethnicity are presented in Table 9.3.

***Model 2.*** *Ethnic groups*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + e_{ij}; \quad (9.3)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \quad (9.4)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

The regression coefficients for predicting student achievement gains from ethnic background are found in Table 9.3.

**Table 9.3** Regression Coefficients for Connecticut Reading Gain 1999-2000 by Ethnicity

| | Model 2 |
|---|---|
| Intercept ($\gamma_{00}$) | 3.15(.32)*** |
| Black ($\gamma_{01}$) | -.06(.64) |
| Hispanic ($\gamma_{02}$) | 1.54(.66)* |
| All other ($\gamma_{03}$) | 1.67(.65)* |
| | |
| LVL 1 Variance ($\sigma^2$) | 53.62 |
| LVL 2 Variance ($\tau_{00}$) | 6.01*** |

*** p<.001,** p<.01,* p<.05,~<.10

To make it easier to interpret the results from the gains by demographic group analysis, the coefficients can be used to compute the gain for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[18] For example, to determine the predicted gains for a Black student who would have ones for the intercept and Black, the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} \tag{9.5}$$

Level 2 - Teachers

$$\begin{aligned} \hat{\beta}_{0j} &= \hat{\gamma}_{00} = 3.15 \\ \hat{\beta}_{1j} &= \hat{\gamma}_{10} = .06 \end{aligned} \tag{9.6}$$

When the betas from Equation 9.6 are substituted in the level one equation (Equation

[18] A table showing ones and zeros for all dummy variables is provided in the Appendix.

9.5), the result for Black gain is produced (Equation 9.7).

$$\hat{Y}_{ij} = (3.15) + (.06)*1 \qquad (9.7)$$

$$= 3.21$$

Predicted reading gains for all subsets of students by ethnicity are presented in Table 9.4. This table shows that the model predicts White students gain less in reading than the other groups. Coefficients for all groups but Black are significantly different from zero.

**Table 9.4** Predicted Reading Gains by Student Demographic Group

| Ethnic Group | |
|---|---|
| White | 3.15 |
| Black | 3.21 |
| Hispanic | 4.69 |
| All other | 4.82 |

In the next model, gender predicts reading gains.

***Model 3.*** *Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dfem + e_{ij}; \qquad (9.8)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \qquad (9.9)$$

$$\beta_{1j} = \gamma_{10}$$

The regression coefficients for predicting student achievement by gender are found in Table 9.5.

**Table 9.5** Regression Coefficients for Reading Gains by Gender (CT)

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | 3.18(.40)*** |
| Female ($\gamma_{01}$) | .42(.34) |
| LVL 1 Variance ($\sigma^2$) | 53.91 |
| LVL 2 Variance ($\tau_{00}$) | 5.77*** |

*** p<.001,** p<.01,* p<.05,~<.10

These coefficients indicate that males are predicted to gain 3.18 points in reading on average over the course of the year and females are predicted to gain 3.60 points. The coefficient for female is not significantly different from zero.

***Model 4.*** *Combined gender and ethnicity*

Model 4 tests whether or not there in an interaction between gender and ethnicity when predicting reading gains. Equation 9.10 computes coefficients for combinations of gender and ethnicity.

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij} + e_{ij}; \quad (9.10)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

The regression coefficients for predicting student achievement gains from ethnic background and gender are found in Table 9.6.

**Table 9.6** Regression coefficients for Connecticut Reading Gain 1999-2000 by Gender and Ethnicity

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | 2.77(.47)*** |
| Black ($\gamma_{01}$) | -.14(.66) |
| Hispanic ($\gamma_{02}$) | 1.36(.66)* |
| All other ($\gamma_{03}$) | 1.70(.61)** |
| Female ($\gamma_{04}$) | .81(.53) |
| Effect Black x female ($\gamma_{05}$) | .33(.44) |
| Effect Hispanic x female ($\gamma_{06}$) | .54(.48) |
| Effect All other x female ($\gamma_{07}$) | -.49(.47) |
| | |
| LVL 1 Variance ($\sigma^2$) | 53.58 |
| LVL 2 Variance ($\tau_{00}$) | 6.03*** |

*** p<.001,** p<.01,* p<.05,~<.10

To aid in interpreting these coefficients, an example using Black female students is presented next. The process is then repeated for all other groups using the coefficients from Table 9.6.

Level 1

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{\beta}_{1j} dblack_{ij} + \hat{\beta}_{4j}\, dfem_{ij} + \hat{\beta}_{5j} dblackfem_{ij} \quad (9.11)$$

Level 2

$$\begin{aligned} \hat{\beta}_{0j} &= 2.77 \\ \hat{\beta}_{1j} &= -.14 \\ \hat{\beta}_{4j} &= .81 \\ \hat{\beta}_{5j} &= .33 \end{aligned} \quad (9.12)$$

The estimated beta values are calculated first and then substituted into the level one equation (9.9). The sample computation for Black female is demonstrated in Equation 9.11.

$$\hat{Y}_{ij} = 3.15 + (-.86)*1 + (.04)*1 + (1.43)*1 \qquad (9.13)$$

$$= 3.76$$

Using this approach for each subgroup, expected reading gains by demographic groups over 1 year are presented in Table 9.7. The model predicts that girls gain more in reading than their male counterparts. Black males gain the least of all groups. The coefficient for Black is not significantly different from zero which implies that Black predicted gains are not statistically different from the intercept or White predicted gains for this sample.

**Table 9.7** Predicted Reading Gain by Student Demographic Groups with Differences by Gender

| | Males | Females | Difference in Gains by Gender |
|---|---|---|---|
| White Students | 2.77 | 3.65 | + .88 |
| Black Students | 2.63 | 3.84 | + 1.21 |
| Hispanic Students | 4.13 | 5.48 | + 1.35 |
| All other Students | 4.47 | 4.79 | + .32 |

Hispanic and All other students, groups which include the students most likely to have English as a second language, have higher average reading gains than their White and Black classmates by at least one point over the course of one year. Coefficients for the intercept, Hispanic, and All other are significantly different from zero.

***Gender, Ethnicity, and Grade Level***

Another known source of variation in student achievement is grade level. In the earlier SEM analyses, the outcome for gains was standardized by grade level. In the upcoming NAEP analyses, only one grade is represented in each analysis. Consequently, dummy variables for grade are introduced at level one in the Connecticut analyses to control for grade level. The reference category for Model 5 is therefore white, male students in the 8th grade.

***Model 5.*** *Gender, ethnicity, and grade level*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} \qquad (9.14)$$
$$+ \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij}$$
$$+ \beta_{8j} dgrade4_{ij} + \beta_{9j} dgrade6_{ij} + e_{ij};$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$
$$\beta_{4j} = \gamma_{40}$$
$$\beta_{5j} = \gamma_{50}$$
$$\beta_{6j} = \gamma_{60}$$
$$\beta_{7j} = \gamma_{70}$$
$$\beta_{8j} = \gamma_{70}$$
$$\beta_{9j} = \gamma_{70}$$

The regression coefficients for predicting student achievement gains from ethnic background, gender, and grade level are found in Table 9.8.

**Table 9.8** Regression Coefficients for Connecticut Reading Gain 1999-2000 by Gender, Ethnicity, and Grade

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | -.01(.65) |
| Black ($\gamma_{01}$) | -.35(.66) |
| Hispanic ($\gamma_{02}$) | 1.27(.64)* |
| All other ($\gamma_{03}$) | 1.67(.61)** |
| Female ($\gamma_{04}$) | .81(.53) |
| Effect Black x female ($\gamma_{05}$) | .39(.45) |
| Effect Hispanic x female ($\gamma_{06}$) | .50(.48) |
| Effect All other x female ($\gamma_{07}$) | -.51(.47) |
| 4th grade ($\gamma_{08}$) | 4.28(.72)*** |
| 6th grade ($\gamma_{09}$) | 2.91(.66)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 53.61 |
| LVL 2 Variance ($\tau_{00}$) | 3.25*** |

*** p<.001,** p<.01,* p<.05,~<.10

Predicted reading gains for all subsets of students by demographic group and grade are presented in Table 9.9.[19] There is a large, statistically significant difference in gains by grade with students in the 4th grade gaining the most, followed by those in the 6th grade, and with the smallest reading gain made by students during 8th grade.

**Table 9.9** Predicted Mathematics Gains by Student Demographic Group and Grade

| | | Grade 4 (+4.28) | Grade 6 (+2.91) | Grade 8 |
|---|---|---|---|---|
| Males | White | 4.27 | 2.90 | - .01 |
| | Black | 3.92 | 2.55 | - .36 |
| | Hispanic | 5.54 | 4.17 | 1.26 |
| | All other | 6.23 | 4.57 | 1.66 |
| Females | White | 5.08 | 3.71 | .80 |
| | Black | 5.12 | 3.75 | .84 |
| | Hispanic | 6.85 | 5.48 | 2.57 |
| | All other | 6.24 | 4.87 | 1.96 |

[19] To see a sample of the calculations performed to create Table 9.8, please refer to p. 35, Eq. 9.9-9.11.

Hispanic and All other students, those most likely to have English as a second language, are still expected to gain more in reading over the course of a year than White and Black students once grade is taken into account. Note that the coefficients for Hispanic and All other remain significantly different from zero for this analysis. Girls gain more than boys within each group.

*Summary of Section 1.* Developing the level one equation for students. As the level one one base model for reading in Connecticut developed, average yearly gains were first estimated with a simple unconditional model. The model revealed that average gain within teachers (between students) was nine times larger than across teachers at level two. Variables know to affect mean gains (ethnicity, gender and grade) were added to the level one equation indicating that while differences in gains exist, only the coefficients for ethnicity and grade level are significant predictors of reading gains when gender is controlled. By far, the most significant predictor of reading gains was grade level with scores declining as students progressed by grade. There are no significant interactions between gender and ethnicity. Additionally, the highest gains were predicted for students from ethnic categories that likely contained the most ESL students.

**II. Differential Effects of Teacher Practices, Teacher Characteristics, and Professional Development on Student Reading Achievement Gains** (Results at the teacher level)

This section addresses the following questions:

1. Do reading achievement gains for demographic groups vary across teachers?

2. How much variation in student reading gains do variables about teacher practices, teacher characteristics, and professional development explain?

3. Is the amount of variance explained by teacher practices, teacher characteristics, and professional development different for different student demographic groups?

Analyses in this section first test if the achievement gains for the various demographic groups vary across teachers. In other words, is the variance significantly different from zero.· There is a brief discussion of variance across teachers followed by changes to the results when district SES is controlled. This establishes the level two context for a series of analyses which look at the effects of teacher practice, teacher characteristics, and professional development on the variance of reading gains.

**Variance in Connecticut reading achievement across teachers**

Unlike mathematics, the $\chi^2$ test that HLM provides indicated that average reading gains for ethnicity and gender did not vary across teachers ($\tau_{00}$=0) but $\chi^2$ test for the effect of Black was based only 56 of 90 available teachers who had sufficient numbers of Black students in their classrooms. An alternate test to compute significance using change in deviance over degrees of freedom indicated p<.01.[20] In other words, average black gains do vary across teachers. A subsequent series of analyses for variance explained in Black average gains across teachers revealed that no variance in average reading gains for black students was explained by the selected teacher variables so the results are not included in this dissertation.

---

[20] Deviance: 12174-12164=|10|, degrees of freedom: 2-4=|-2|

Consequently, the next series of analyses focus on question two, how much variance is explained by teacher level variables on the average reading gains for all students in the sample.

**Effects of teacher level variables on mean reading gains in Connecticut**

The level one equation for ethnicity and grade is presented as Model 6. Note that the non-significant interaction and error terms for gender by ethnicity have been dropped. Only the intercept is allowed to vary. The dummy variable for female remains in the equation as a control.

***Model 6.*** *Unconditional model for average reading gains*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + \beta_{4j}\,dfem_{ij} + \beta_{8j}dgrade4_{ij} + \beta_{9j}dgrade6_{ij} + e_{ij}; \quad (9.15)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \quad (9.16)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average student gains for all teachers

$r_{0j}$ = the error term or unique teacher effects

The estimated level two regression coefficients for Model 6 are given in Table 9.10.

**Table 9.10** Regression Coefficients for Average Connecticut Reading Gain 1999-2000 Across Teachers

| | Model 6 |
|---|---|
| Intercept ($\gamma_{00}$) | .18(.65) |
| Black ($\gamma_{10}$) | -.27(.64) |
| Hispanic ($\gamma_{20}$) | 1.40(.63)* |
| All other ethnicities ($\gamma_{30}$) | 1.59(.64)* |
| Female ($\gamma_{40}$) | .43(.34) |
| $4^{th}$ grade ($\gamma_{80}$) | 4.26(.72)*** |
| $6^{th}$ grade ($\gamma_{90}$) | 2.89(.66)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 53.65 |
| LVL 2 Variance Intercept ($\tau_{00}$) | 3.25*** |

*** p<.001,** p<.01,* p<.05,~<.10

Selected coefficients from this analysis serve as the base for the variance decompositions found in Tables 9.12 and 9.13. Next, Model 7 expands Model 6 to include district SES (average free and reduced lunch).

***Model 7****. Average reading achievement gains and SES*

Level 1 - Students (same as Eq. 9.11)

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, pctfrlu + r_{0j} \qquad (9.17)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average student gains for all teachers

$\gamma_{01j}$ = change in average gains associated with mean district free and reduced lunch, all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for HLM Model 7 are given in Table 9.11.

**Table 9.11** Regression Coefficients for Average Reading Gain by Average District SES

| | Model 7 |
|---|---|
| Intercept ($\gamma_{00}$) | .16(.66) |
| Free and reduced lunch ($\gamma_{06}$) | -.89(1.78) |
| Black ($\gamma_{10}$) | -.18(.66) |
| Hispanic ($\gamma_{20}$) | 1.52(.69)* |
| All other ethnicities ($\gamma_{30}$) | 1.64(.64)* |
| Female ($\gamma_{40}$) | .44(.34) |
| 4th grade ($\gamma_{80}$) | 4.24(.73)*** |
| 6th grade ($\gamma_{90}$) | 2.88(.68)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 53.64 |
| LVL 2 Variance Intercept ($\tau_{00}$) | 3.32*** |

*** p<.001,** p<.01,* p<.05,~<.10

Table 9.12 presents the variance decomposition for SES. Unlike mathematics, average district SES (22% free and reduced lunch[21]) does not explain variation in average student

[21] This sample includes different teachers and districts than the mathematics sample which changes the district free and reduced lunch percentage by 1%.

reading gains across teachers.

**Table 9.12** Variance in Average Reading Gains Explained by SES

| Outcome: Read diff | Model 6 unconditional | Model 7 SES |
|---|---|---|
| Intercept ($\gamma_{00}$) | .18(.65) | .16(.66) |
| SES ($\gamma_{01}$) | | -.89(1.78) |
| LVL 1 Variance ($\sigma 2$) | 53.65 | 53.64 |
| LVL 2 Variance ($\tau_{00}$) | 3.25*** | 3.32*** |
| % variance explained | | 0% |

*** p<.001,** p<.01,* p<.05,~<.10

One standard deviation in increase in free and reduced lunch reduces average student reading gains across teachers by nearly one point over the course of one year. Although it does not vary across teachers, SES is retained in the teacher effects analysis as a control.

**Teacher Effects on Reading Achievement**

Using the same equations as presented for mathematics (EQ 8.18), *variable x* is replaced with teacher level variables for teacher practices, charactersitics, and professional development one at a time. This is the same process outlined for the mathematics decompositions in Chapter 8. One classroom practice (computer activities), one teacher characteristic (English preparation), and one professional development variable (professional development frequency) explained variance in average student scores across teachers. Results are presented in Table 9.13.

**Table 9.13** Variance Decomposition Table for Average Reading Gains Across Teachers in Connecticut

| Outcome: Read diff | Unconditional | SES | Computer Activities | English Preparation | Professional development frequency |
|---|---|---|---|---|---|
| Intercept ($\gamma_{00}$) | .18(.65) | .16(.66) | .13(.66) | .58(.63) | .10(.64) |
| SES ($\gamma_{01}$) | | -.89(1.78) | -2.47(1.91) | -.89(1.65)* | -.93(1.67) |
| ($\gamma_{02}$) | | | -.70(.29)* | -.60(.28) | -.53(.38) |
| LVL 1 Variance ($\sigma2$) | 53.65 | 53.64 | 53.62 | 53.66 | 53.65 |
| LVL 2 Variance ($\tau_{00}$) | 3.25*** | 3.32*** | 2.98*** | 3.01*** | 3.11*** |
| % variance explained | | 0% | 8.3% | 7.3% | 4.3% |

*** p<.001,** p<.01,* p<.05,~<.10

Computer activities explain 8.3% of the variance in student gains across teachers controlling for SES. A one standard deviation increase in computer activities decreases student gains by -.70 points. English preparation, like mathematics preparation, is negatively related to student gains. For every one standard deviation increase in teacher English preparation, -.60 points in reading gains are lost across teachers over the course of one year. English preparation explains 7.3% of the variation in average student gains across teachers. Professional development frequency explains 4.3% of the variation in student gains across teachers. About half a point is lost for every one standard deviation increase in professional development frequency.

<u>Summary of the effects of teacher variables on reading gains in Connecticut</u>. It is striking that the results from the reading analysis are so different than the mathematics results. Average gains by gender and ethnicity do not vary across teachers. SES does not explain any of the variance in average student reading gains across teachers. Computer activities, English Preparation and, professional development frequency account for variance across teachers in Connecticut but they are negatively related to average student

gains across teachers. One standard deviation increase in them decreases student gains by half to three quarters of a point over the course of one year.

**III. IHE Teacher Preparation, Teachers, and Student Achievement.**

Section III section addresses questions about university teacher preparation programs and their connections to teacher practices, characteristics, professional development, and subsequent student achievement in reading. In other words, do Connecticut teacher preparation programs vary in the average levels of teacher practice, characteristics, and professional development produced by their programs and do teacher preparation programs predict the teacher variables associated with average student achievement across teachers?

Effects of teacher preparation program on student reading achievement. The analysis in this section uses a smaller sample of teachers who indicated that they attended a teacher preparation program in Connecticut. There are 1285 students nested within 65 reading teachers within 13 identified university teacher preparation programs in Connecticut.

The unconditional three level analysis shows no significant variation in average student reading gains within teachers across teacher preparation programs. No further investigation into variance explained in student gains across teachers by teacher preparation programs was possible for this data.

Effects of teacher preparation institutions on teachers. This section addresses questions about Connecticut teacher preparation programs and their connections to teacher practice, characteristics, professional development. After finding relationships among student outcomes and certain teacher practices (computer activities), characteristics (English preparation), and professional development (professional development frequency), the next logical step is to explore the connections between teacher preparation institutions and those particular teacher variables.

A two level model (Model 8) with Connecticut teacher preparation programs at level two predicts one of three teacher variables at a time on level one (computer activities, English preparation, and professional development frequency). Results for the variance explained in average teacher variables across teacher preparation programs are presented in Table 9.22. This analysis included 67 reading/literacy teachers nested within 13 teacher preparation programs.[22]

***Model 8**. Base model for teacher variables across teacher preparation institutions.*

Level 1 - Teacher

$$variable\ x_{ij} = B_0 + e_{ij} \tag{9.17}$$

Level 2 - IHE teacher preparation program

$$\beta_{0j} = \gamma_{00} + r_{0j} \tag{9.18}$$

The coefficients for the intercepts and variances for computer activities, English preparation, and professional development frequency can be found in Table 9.14.

---

[22] Fixed effects are unbaised although standard errors are likely too small in small samples (Raudenbush, p. 281).

**Table 9.14** Unconditional Models for Levels of Teacher Variables Across Teacher Preparation Institutions

| Unconditional Models (outcomes) | Computer Activities | English Preparation | Professional Development Frequency |
|---|---|---|---|
| Intercept ($\gamma_{00}$) | 8.04(.32)*** | 3.75(.37)*** | 18.57(.35)*** |
| LVL 1 Variance ($\sigma 2$) | 6.94 | 3.93 | 4.35 |
| LVL 2 Variance ($\tau_{00}$) | .01 | .71* | .53 |

*** p<.001,** p<.01,* p<.05,~ p<.10

Of the teacher variables that were related to variation in student reading gains across teachers, only average English preparation varies significantly (p<.05) across teacher preparation institutions. 15% of the variance is located at the teacher preparation level and 85% of the variance is between teachers.

In Model 9, qualities of teacher preparation programs predict teacher English preparation. Variables available at the university level are faculty characteristics, field experiences, and program quality as reported by program graduates as well as rank on the Initial State Report Card based on teacher certification examinations.

***Model 9.*** *Base model to predict English preparation by teacher preparation program characteristics*

Level 1 - Teacher

$$EngPrep_{ij} = B_0 + e_{ij} \tag{9.19}$$

Level 2 - IHE teacher preparation program

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, program\ characteristic\ y + r_{0j} \tag{9.20}$$

where the *program characteristic y* is faculty characteristics, program quality, field experiences, or program rank.

**Table 9.15** Connecticut's Average English Preparation by Teacher Preparation Institutions

| English preparation | unconditional | field experiences | faculty characteristics | Program Quality | Rank on State Report Card |
|---|---|---|---|---|---|
| Intercept ($\gamma_{00}$) | 3.75(.37)*** | 3.70(.35)*** | 3.73(.35)*** | 3.86(.38)*** | 3.80(.40)*** |
| ($\gamma_{10}$) | | .44(.33) | -.40(.27) | -.45(.35) | -.05(.09) |
| LVL 1 Variance ($\sigma 2$) | 3.93 | 3.93 | 3.88 | 3.88 | 3.89 |
| LVL 2 Variance ($\tau_{00}$) | .71* | .59* | .79* | .80* | .91* |
| Variance explained | | 16.9% | 0% | 0% | 0% |

*** p<.001,** p<.01,* p<.05,~<.10

Field experience at the university level explains 16.9% of the variance in English preparation across teacher preparation institutions. For every one standard deviation increase in field experiences, teacher English preparation increases by .44. This is roughly equivalent to half of a major in English, English education, or half a course in English, writing, or literature.

Effects of specific teacher preparation institutions on reported teacher practices. A two level analysis (Model 9) uses dummies for teacher preparation institutions on level two to predict English preparation by program.

**Model 9.** *13 teacher preparation institutions predicting English preparation*

Level 1- Teachers

$$\textit{teacher variable } x_{ij} = B_0 + e_{ij} \tag{9.21}$$

Level 2- University teacher preparation programs

$$\begin{aligned} B_{00} = {} & \gamma_{00} + \gamma_{01}\, Univ1 + \gamma_{02}\, Univ2 + \gamma_{03}\, Univ3 + \gamma_{04} Univ4 \\ & + \gamma_{05} Univ5 + \gamma_{06} Univ6 + \gamma_{07} Univ7 + \gamma_{08} Univ8 \\ & + \gamma_{09} Univ9 + \gamma_{10} AllOthUniv \end{aligned} \tag{9.22}$$

The university teacher preparation coefficients predicting English preparation are presented in Table 9.16 followed by program rankings in Table 9.17.

**Table 9.16** Regression Coefficients for Teacher English Preparation by Teacher Preparation Institutions

| | Unconditional | English preparation |
|---|---|---|
| Intercept ($\gamma_{00}$) | 3.75(.37)*** | 3.99(.85) |
| University 1 ($\gamma_{01}$) | | -.63(3.99) |
| University 2 ($\gamma_{02}$) | | -2.23(3.98) |
| University 3 ($\gamma_{03}$) | | -2.18(3.97) |
| University 4 ($\gamma_{04}$) | | -1.43(3.94) |
| University 5 ($\gamma_{05}$) | | .95(4.02) |
| University 6 ($\gamma_{06}$) | | -3.30(4.15) |
| University 7 ($\gamma_{07}$) | | .20(4.01) |
| University 8 ($\gamma_{08}$) | | 1.80(3.95) |
| University 9 ($\gamma_{09}$) | | 1.20(4.37) |
| Three Other Universities ($\gamma_{30}$) | | .26(3.61) |
| LVL 1 Variance ($\sigma 2$) | 3.93 | 3.79 |
| LVL 2 Variance ($\tau_{00}$) | .71* | 7.26* |
| Variance explained at Lvl 2 | | 0% |

*** p<.001, ** p<.01, * p<.05, ~ p<.10

**Table 9.17** Teacher English Preparation by 13 Connecticut Teacher Preparation Programs (low to high)

| | |
|---|---|
| University 6 | 0.69 |
| University 2 | 1.76 |
| University 3 | 1.81 |
| University 8 | 2.19 |
| University 4 | 2.56 |
| University 1 | 3.36 |
| Three Other Universities | 3.73 |
| University 0 | 3.99 |
| University 7 | 4.19 |
| University 5 | 4.94 |
| University 9 | 5.19 |

The maximum score for this scale was 12. Interpretation is complicated. An average program score of less that one point means roughly that reading teachers from this program have degrees in majors other than English or English education with no coursework in English, writing, and literature. A score of 1-2 means roughly that universities whose graduates score in this range average up to an undergraduate minor in English or English education or up to 5 courses in English, writing, and literatures or an equivalent combination. 3-4 points is roughly equivalent to an undergraduate major with additional coursework up to a master's degree or an equivalent amount of coursework. Three universities scored over 4 points indicates at least Masters degree in English or English Education with up to two additional courses in English writing or literature or an undergraduate major with at least 3-5 additional courses.

Although it is possible to rank programs on the amount of English preparation their teacher graduates have had (fixed effects), the addition of 13 programs to the model does not explain variance across teacher preparation institutions and coefficient significance could not be calculated by HLM.

***Summary of the HLM results for reading gains in Connecticut.***

Once the contextual effects for gains by ethnic and gender groups controlling for grade level were estimated at level one, it is possible to answer the main research questions regarding how those relationships changed when teacher practices, characteristics, and professional development were added. Then the results for student outcomes and teacher variables predicted by teacher preparation program in Connecticut are presented. The findings for reading gains in Connecticut can be summarized as follows:

## Results Summary

### Section I Results Summary. Contextual effects of gender, ethnicity, and district SES on reading gains in Connecticut

**1.** Fixed effects at the student level show significant gains by ethnicity and grade level. Hispanic, and All other students are predicted to gain more than White and Black students over the course of one year and this gain is significantly different from zero. The categories for Hispanic and All other are also the ones most likely to contain English as second language students. Gender does not play a significant part in gains although fixed effects predict that girls tend to score slightly higher in reading than boys. There are no significant interactions between ethnicity and gender.

**2.** After grade level is introduced to the model, the estimated gains by demographic group pattern remains the same with Hispanic and All other students outpacing their White and Black counterparts. Grade level, however, is the largest predictor of reading gains in Connecticut and highly significant ($p<.000$). All students are predicted to gain 4 points from the beginning to the end of the 4$^{th}$ grade year. During 6$^{th}$ grade, students add nearly 3 points to their reading gains. Black and White students gain less than one point in reading during 8$^{th}$ grade. Hispanic and All other students gain an additional +1.5 points in reading during 8$^{th}$ grade.

**3.** Most variation in Connecticut reading classrooms is between students (90%). Only about 10% of the variation is located across teachers.

**4.** Average district SES measured at the teacher level did not vary significantly across teachers when gender, ethnicity, and grade level were controlled.

**Section II Results Summary. Teacher practice, characteristics, and professional development effects on reading gains in Connecticut**

**5**. Average reading gains did not vary significantly by demographic groups or gender across teachers in Connecticut when gender, ethnicity, and grade level were controlled. Consequently, there are no differential effects to report.

**a.** Of the *teacher practices* variables introduced at level two, only the practice of computer activities explained average reading gain variance across teachers (8.3%). For every one standard deviation increase in the use of computer activities, students are expected to lose about three quarters of a point over the year.

**b.** Of the *teacher characteristics* variables introduced at level 2, only English preparation accounted for average reading gain variance across teachers. English preparation explained only 7.3% of the variance in average reading gains across teachers. For every one standard deviation increase in teacher efficacy, average student gains dropped by -.60 points over the course of the year.

**c.** Of the teacher *professional development* variables, only professional development frequency explained average student reading gain variance across teachers (4.3%). For every one standard deviation increase in professional development frequency, average student gains dropped by half a point.

**Section III Results Summary. University teacher preparation program effects on student reading gains and teacher variables**

**9.** Average 4th-8th grade student *reading gains* produced by graduates of specific teacher preparation programs did not vary across Connecticut teacher preparation programs in a three level model.

**10.** Only one teacher preparation variable that explained variance in average student scores across teachers also varies across teacher preparation institutions, English preparation. 15% of the variance in teacher English preparation is attributable to teacher preparation programs while 85%of the variance is between teachers. Increases in English preparation are related to decreases in student scores.

**11.** Of the institutional variables available in the BTPS data, faculty characteristics, program quality, and field experiences, only field experience accounted for significant variation in English preparation (16.9%). Note that the program ranking from the Initial State Report Card did not predict levels of English preparation, the only variable shown to vary across Connecticut institutions.

**12.** It is possible to produce fixed effects rankings of teacher preparation programs based on levels of teacher English preparation (see Table 9.17); however, knowing specific teacher preparation programs does not account for the variance in average English preparation across institutions.

# 10. RESULTS OF THE HLM ANALYSIS FOR 1996 NAEP 4TH GRADE MATHEMATICS

The HLM analysis using NAEP data is intended to address the research questions about the relationships among student achievement and

- Teacher characteristics and practices
- Characteristics of professional development

The analysis also is intended to address the ways in which these relationships vary by student demographic group, including gender and ethnicity, across the United States and to see if the national mathematics results reflect trends similar to those found in the Connecticut BTPS data, a more detailed data set. Analyses concerning the relationships between teacher preparation institutions and teachers/students are not possible for this data since the names of teacher preparation programs are not collected in NAEP. The HLM analyses are weighted (origwt) and use five plausible values (mrpcm1-5) for the student outcome. There are 4,409 students and 584 teachers in the 1996 mathematics sample.[23]

To set the context for addressing the questions, results are presented first which describe the differences in student achievement (achievement status in NAEP data) across student demographic groups. The analyses yield results in terms of regression coefficients. To aid in understanding what the coefficients mean, predicted achievement (achievement status score) is computed for selected combinations of student

---

[23] The *n* varies from that in the structural equation models due to the number of variables included in maximum likelihood imputation vs. listwise deletion when creating the .mdm file in HLM.

characteristics.

The full set of mathematics variables for NAEP 1996 is given in Table 10.1. The outcomes and most teacher variables are the same as those used for the earlier SEM models.

**Table 10.1** Description of Variables Used in the NAEP 1996 4$^{th}$ Grade Mathematics Analyses

| Type | Variable Abbreviation | Variable Description |
|---|---|---|
| | | **Level 1 - students** |
| Continuous (Outcomes) | mrpcm1 | 1$^{st}$ of 5 multiple plausible values for the achievement of each student (mean=222.38, sd=30.53) |
| | mrpcm2 | 2$^{nd}$ of 5 multiple plausible values for the achievement of each student (mean=222.51, sd=30.56) |
| | mrpcm3 | 3$^{rd}$ of 5 multiple plausible values for the achievement of each student (mean=222.78, sd=30.33) |
| | mrpcm4 | 4$^{th}$ of 5 multiple plausible values for the achievement of each student (mean=222.18, sd=30.67) |
| | mrpcm5 | 5$^{th}$ of 5 multiple plausible values for the achievement of each student (mean=222.64, sd=30.06) |
| Dichotomous | dblack | Black (Black=1, else=0; mean=.18, sd=.38) |
| | dhispanic | Hispanic (Hispanic=1, else=0; mean=.16, sd=.36) |
| | dallother | All other ethnicities (all other=1, else=0; mean=.07, sd=.25) |
| | dfem | Female (female=1, else=0; mean=.51, sd=.50) |
| | SES | Average free and reduced lunch (eligible=1, else=0 mean=, sd=) |
| Effects coded | egenxbl | Effect of Black female; interaction of eBlack x egender where Black =1, White=-1, otherwise=0 multiplied by female=1, male=-1 (mean=.01, sd=.88) |
| | egenxhisp | Effect of Hispanic female; interaction of eHispanic x egender where Hispanic=1, White=-1, otherwise=0 multiplied by female=1, male=-1 (mean=-.02, sd=.87) |
| | egenxoth | Effect of All other female; interaction of eAllother x egender where All other =1, White=-1, otherwise=0 multiplied by female=1, male=-1 (mean=-.01, sd=.81) |
| Continuous | Origwt | Original weight (correction for oversampling) |

**Table 10.1** (cont'd).

| | | |
|---|---|---|
| | | **Level 2 - teachers** |
| Continuous | | Grand mean centered |
| Teacher Practices | gpa | Frequency of group and project activities (mean=11.97, sd=2.66) |
| | gpaasses | Emphasize group and project assessment (mean=7.56, sd=2.41) |
| | sensemaking | Frequency of sensemaking activities (mean=10.47, sd=2.71) |
| Teacher Characteristics | moref | Teacher's reform orientation (mean=3.02,sd=1.60) |
| | mathprep | Major/coursework/license in mathematics, mathematics education (mean=.59, sd=1.00) |
| | NCTM | Knowledge of NCTM standards (mean=1.88, sd=.90) |
| | mathknow | Math knowledge (mean=7.46, sd=.91) |
| | mathinst | Preparation for mathematics instruction, pedagogical coursework (mean=3.34, sd=1.01) |
| | divprep | Preparation for diversity (mean=1.65, sd=1.11) |
| Professional Development | pdf | Professional development in the last year (2.84, sd=1.23) |
| Continuous | avSES | Average percent free and reduced lunch per teacher/ classroom aggregated from level 1 (mean=.39,sd=.49) |

**I. Setting the Stage: Variation in Mathematics Achievement Gains by Student Demographic Groups and SES** (Results at the Student Level)

In this section, sets of background variables are added to look at changes in mathematics gains in the United States. Model 1 predicts mean student gains across the sample and partitions the variance between students and teachers where students are nested within teachers. Models 2-5 predict mean gains across gender, ethnicity, and SES (eligibility for free and reduced lunch).

Mean Gains. To set the context for answering the questions about connections between student achievement and teacher variables, results are presented first for differences in student achievement across student demographic groups. The level one and level two equations for the model are given followed by their estimated regression coefficients in Tables 10.2.

***Model 1.*** *Mean achievement*

Level 1 - Students

$$Y_{ij} = \beta_{0j} + e_{ij} \tag{10.1}$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

where $Y_{ij}$ is the mathematics achievement for student i in teacher j's classroom, $\beta_{0j}$ is the intercept or mean students' mathematics achievement within teacher j's classroom, $e_{ij}$ is the level 1 error, $\gamma_{00}$ is the mean value of classroom achievement for all teachers, and $r_{0j}$ is the level 2 error.

**Table 10.2** NAEP 1996 Average Mathematics Achievement

| | Model 1 |
|---|---|
| Intercept ($\gamma_{00}$) | 223.13(1.04)*** |
| LVL 1 Variance ($\sigma^2$) | 574.32 |
| LVL 2 Variance ($\tau_{00}$) | 381.34*** |

*** p<.001,** p<.01,* p<.05

The predicted average achievement score for 4th grade mathematics students for this sample of US students is 223.13 points ($\gamma_{00}$=223.13, sd=1.04) when no other variables are included in the model.

$\sigma^2$ and $\tau_{00}$ are used to compare variation in the errors at the teacher level with variation at the student level using the formulas:

$$\frac{\sigma^2}{\sigma^2 + \tau_{00}} = \textit{level one variance} \qquad \frac{\tau^2{}_{00}}{\sigma^2 + \tau_{00}} = \textit{level two variance} \qquad (10.2)$$

60% of the variance in achievement for Model 1 is within teachers (between students) leaving 40% of the variance in achievement between teachers. That is to say, there is 1.5 times more variation within teachers than between teachers.

Mean Achievement by Demographic Group. Next, demographic variables are added to form Models 2, 3 and 4. Model 5 introduces the control variable for SES. The equations for Model 2, ethnic group, are presented first (Eq.10.3) To interpret the coefficients, the estimated score for ethnicity can be computed using the appropriate set of 1s and 0s for the dummy variables. As an example, Equations 10.4 and 10.5 demonstrate the computation for obtaining Black student achievement. Estimated achievement by ethnicity are presented in Table 10.3.[24]

***Model 2.*** *Ethnic groups*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + e_{ij};$$

Level 2 - Teachers (10.3)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

---

[24] Note that the category for All other includes Asian and those students who did not report ethnicity. Often these smaller groups are folded into the intercept since they are not significantly different from the intercept. In this case, the coefficient for all other was significant at p<.05. Additionally, it was important to maintain the meaning of the intercept as a specific group, in this case white students, before analyzing the differential effects of teacher variables.

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

The regression coefficients for predicting student achievement from ethnic background are found in Table 10.3.

**Table 10.3** Regression Coefficients for NAEP 4th Grade Mathematics 1996 by Ethnicity

| | Model 2 |
|---|---|
| Intercept ($\gamma_{00}$) | 229.23(.98)*** |
| Black ($\gamma_{01}$) | -22.14(1.77)*** |
| Hispanic ($\gamma_{02}$) | -18.30(1.63)*** |
| All other ($\gamma_{03}$) | -4.64(1.92)* |
| LVL 1 Variance ($\sigma^2$) | 547.99 |
| LVL 2 Variance ($\tau_{00}$) | 254.23*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by ethnicity group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[25] For example, to determine the predicted achievement for a black student who would have ones for the intercept and black the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} \quad (10.4)$$

[25] A table showing ones and zeros for all dummy variables is provided in the Appendix.

Level 2 - Teachers

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 229.23$$

$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -22.14$$

When the slope values, the betas from Equation 10.4, are substituted in the level one equation (Eq.10.4), the result for black achievement is produced (Eq. 10.5).

$$\hat{Y}_{ij} = (229.23) + (-22.14)*1 \qquad (10.5)$$

$$= 207.09$$

Predicted mathematics achievement for all demographic groups are presented in Table 10.4. This table shows that the model predicts that White students score higher in mathematics than the other ethnic groups. The coefficients for ethnicity are significantly different from zero.

**Table 10.4** Predicted Mathematics Achievement Student Demographic Group

| Ethnic Group | Achievement |
|---|---|
| White | 229.23 |
| Black | 207.09 |
| Hispanic | 210.93 |
| All other | 224.59 |

In the next model, gender predicts average mathematics achievement.

***Model 3.*** *Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{4j}\, dfem_{ij} + e_{ij};$$

Level 2 - Teachers (10.6)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{4j} = \gamma_{10}$$

The regression coefficients for predicting student achievement from gender are found in

Table 8.5.

**Table 10.5** Regression Coefficients for NAEP Mathematics Achievement 1996 by Gender

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | 224.76(1.22)*** |
| Female ($\gamma_{04}$) | -3.27(1.02)** |
| | |
| LVL 1 Variance ($\sigma^2$ ) | 571.75 |
| LVL 2 Variance ($\tau_{00}$) | 381.73*** |

*** p<.001,** p<.01,* p<.05

These coefficients indicate that male mean achievement in mathematics is predicted to be 224.76 points compared to 221.49 points for females. The coefficient for female is significantly different from zero (p<.01).

***Model 4****. Combined gender and ethnicity*

Model 4 tests whether or not there is an interaction between gender and ethnicity in the 1996 NAEP 4th grade mathematics data. Equation 10.7 computes coefficients for combinations of gender and ethnicity.

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + \beta_{4j}\,dfem_{ij} + \beta_{5j}egenxbl_{ij} + \beta_{6j}egenxhisp_{ij} + \beta_{7j}egenxoth_{ij} + e_{ij}; \qquad (10.7)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background and gender are found in Table 10.6.

**Table 10.6** Regression Coefficients for NAEP 4th Grade Mathematics Achievement 1996 by Gender and Ethnicity

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | 231.02(1.18)*** |
| Black ($\gamma_{01}$) | -22.07(1.77)*** |
| Hispanic ($\gamma_{02}$) | -18.38(1.62)*** |
| All other ($\gamma_{03}$) | -4.59(1.90)* |
| Female ($\gamma_{04}$) | -3.61(1.27)** |
| Effect Black x female ($\gamma_{05}$) | 1.33(1.23) |
| Effect Hispanic x female ($\gamma_{06}$) | -.71(1.19) |
| Effect All other x female ($\gamma_{07}$) | -.82(1.59) |
| LVL 1 Variance ($\sigma^2$) | 545.33 |
| LVL 2 Variance ($\tau_{00}$) | 255.02*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by demographic group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[26] For example, to determine the predicted achievement for a black

[26] A table showing ones and zeros for all dummy variables is provided in the Appendix.

female student who would have ones for the intercept, black, female and black female, the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j}dblack_{ij} + \hat{B}_{4j}dfem_{ij} + \hat{B}_{5j}eblackxefem_{ij}$$

Level 2 - Teachers (10.8)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 231.02$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -22.07$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = -3.61$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = 1.33$$

When the level two betas from Equation 10.8 are substituted in the level one equation (Eq. 10.8), the result for black female achievement is produced (Eq. 10.9).

$$\hat{Y}_{ij} = (231.02) + (-22.07)*1 + (-3.61)*1 + (1.33)*1 \quad (10.9)$$
$$= 206.67$$

Predicted mathematics achievement for all subsets of students by demographic groups are presented in Table 10.6. The table shows that the model predicts female students from all ethnic groups have lower scores than their male counterparts.

**Table 10.7** Predicted Mathematics Achievement by Student Demographic Group with Differences by Gender

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 231.02 | 227.41 | - 3.61 |
| Black | 208.95 | 206.67 | - 2.28 |
| Hispanic | 212.64 | 208.32 | - 4.32 |
| All other | 226.42 | 222.00 | - 4.42 |

Score differences between males and females of the same ethnicity are predicted to range from 2.28 to 4.42 points. Coefficients for gender times ethnicity effects are not significantly different from zero over the course of one year.

Another known source of variation in student achievement is SES as measured by eligibility for free and reduced lunch. In NAEP, this variable is measured at the student level. Model 5 predicts mathematics achievement by gender, ethnicity, and SES.

***Model 5.*** **Gender, Ethnicity, and SES**

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} \quad (10.10)$$

$$+ \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij}$$

$$+ \beta_{8j} SES_{ij} + e_{ij};$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background, gender, and SES are found in Table 10.8.

**Table 10.8** Regression Coefficients for NAEP 4th Mathematics Achievement 1996 by Gender, Ethnicity, and SES

| | Model 5 |
|---|---|
| Intercept ($\gamma_{00}$) | 234.18(1.22)*** |
| Black ($\gamma_{01}$) | -18.46(1.74)*** |
| Hispanic ($\gamma_{02}$) | -16.48(1.57)*** |
| All other ($\gamma_{03}$) | -2.87(1.87) |
| Female ($\gamma_{04}$) | -3.26(1.23)** |
| Effect Black x female ($\gamma_{05}$) | 1.12(1.22) |
| Effect Hispanic x female ($\gamma_{06}$) | -.62(1.18) |
| Effect All other x female ($\gamma_{07}$) | -.59(1.55) |
| SES ($\gamma_{08}$) | -12.65(1.34)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 533.28 |
| LVL 2 Variance ($\tau_{00}$) | 211.30*** |

*** p<.001,** p<.01,* p<.05

As demonstrated earlier, the computational example to determine the predicted achievement for a black female student eligible for free and reduced lunch who would have ones for the intercept, black, female, black female, and free and reduced lunch, would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j}dblack_{ij} + \hat{B}_{4j}dfem_{ij} + \hat{B}_{5j}egenxbl_{ij} + \hat{B}_{9j}SES_{ij}$$

Level 2 - Teachers (10.11)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 234.18$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -18.46$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = -\ 3.26$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = \ 1.12$$
$$\hat{\beta}_{8j} = \hat{\gamma}_{80} = -12.65$$

When the betas from Level 2, Equation 8.11, are substituted in the Level 1 equation, the result for black female achievement is produced (Eq. 8.12).

$$\hat{Y}_{ij} = (234.18) + (-18.46)*1 + (-3.26)*1 + (1.12)*1 + (-12.65)*1 \quad (10.12)$$

$$= 200.93$$

Predicted mathematics achievement for all subsets of students by demographic group and grade are presented in Table 10.9. The estimated point loss associated with free and reduced lunch eligibility is -12.65 points.

**Table 10.9** Predicted Mathematics Achievement by Student Demographic Group and SES

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 221.53 | 218.27 | - 3.26 |
| Black | 203.07 | 200.93 | - 2.14 |
| Hispanic | 205.05 | 201.17 | - 3.88 |
| All other | 218.66 | 214.81 | - 4.85 |

The table also indicates that the model predicts that all non-white student groups have lower mathematics scores than white students once ethnicity, gender, and SES are taken into consideration. Girls are predicted to score lower than boys. Hispanic students show the largest differences by gender with girls predicted to lose -3.88 points. The point gap between males and females in Table 10.9 is somewhat less than the point gap in Table 10.7 which did not include SES.

*Summary of Section 1.* Developing the level one equation for students. As the level one base equation for 1996 NAEP 4$^{th}$ grade mathematics developed, average yearly achievement was first predicted with a simple unconditional model. This model indicated that achievement variance was 1.5 times larger within teachers (between students) than across teachers at level 2. Variables known to affect mean achievement (ethnicity, gender, ethnicity x gender, and SES) were added to the level one equation revealing that while differences in mathematics achievement by ethnicity and gender exist within teachers, the interaction terms for gender times ethnicity were not significant predictors of student achievement.

**II. Differential Effects of Teacher Characteristics, and Teacher Practices on Student Achievement** (Results at the teacher level)

This section addresses the questions:

1. Does mathematics achievement for demographic groups vary across teachers?
2. How much variation in student mathematics achievement do variables about teacher practices, characteristics, and professional development explain?
3. Is the amount of variance explained by teacher practices, characteristics, and professional development different for different student demographic groups?

The analyses in this section first test if average achievement for the various demographic groups varies across teachers. The unit of analysis shifts to teachers. Nonsignificant error terms for variance ($r_{00}$, $r_{10,}$ ...$r_{x0}$) indicate that group mean achievement does not vary across teachers so the error terms predicting mean achievement for those demographic groups are dropped from the model. Although the error terms are dropped, the fixed level two equations for non-significant variation by demographic groups are retained in the analysis. Analyses indicated that average achievement for Hispanic and All other males and females across teachers did not vary significantly from the intercept ($\tau_{..0}$=0) so the error terms have been eliminated. Their variance is then captured as part of the error term for the intercept ($r_{00}$).

Mathematics achievement varies across teachers for three demographic groups, females, Blacks, and Black females. The $\chi^2$ tests indicate that the variance of the intercept and the average mathematics achievement of females, Black, and Black female are significant ($\tau_{..0} \neq 0$). Average scores for each group are addressed in a series of analyses which include teacher practices, teacher characteristics, and teacher professional development.[27] The results for average achievement by gender across teachers is presented first followed by average achievement by Black students across teachers. Finally, results for average achievement by Black female students across teachers conclude the chapter.

**Effects of teacher level variables on mean gains by gender**

While the fixed effects model for gender, Model 3, predicted that average achievement varied significantly by gender, a random slopes and intercepts model, Model 6, is necessary to test whether or not the effects of teachers vary when predicting the average achievement for males and females. If they vary, do teacher practices, characteristics, or professional development have differential effects on the average mathematics achievement for one gender as opposed to the other? If variance is significant across teachers, it is then possible to determine how much of the variation can be explained by a specific level two variables.

---

[27] Average student SES is nearly significant ($p<.08$).

***Model 6.*** *Unconditional model for Teacher effects on Male/Female average achievement*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + \beta_{4j}\,dfem_{ij}$$

$$+\beta_{5j}SES_{ij} + e_{ij} \tag{10.13}$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \tag{10.14}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + r_{4j}$$

$$\beta_{5j} = \gamma_{50}$$

where

$\gamma_{00j}$ = intercept, average male achievement across teachers all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{40j}$ = intercept, average female achievement across teachers all other predictors held constant

$r_{4j}$ = the error term or unique teacher effects

**Table 10.10** Regression Coefficients for NAEP Mathematics Achievement 1996 by Gender

| | Model 6 |
|---|---|
| Intercept ($\gamma_{00}$) | 234.12(1.20)*** |
| Black ($\gamma_{10}$) | -18.41(1.75)*** |
| Hispanic ($\gamma_{20}$) | -16.42(1.56)*** |
| All other ($\gamma_{30}$) | -2.73(1.88) |
| Dfem ($\gamma_{40}$) | -3.09(1.00)** |
| SES ($\gamma_{50}$) | -12.67(1.33)*** |
| | |
| LVL 1 Variance ($\sigma 2$) | 518.07 |
| LVL 2 Intercept ($\tau_{00}$)~ | 253.12*** |
| Female ($\tau_{40}$)~ | 61.84** |

*** p<.001,** p<.01,* p<.05

The effect of female varies significantly across teachers ($\tau_{40}$=61.84, p<.01) as does the effect of male ($\tau_{00}$=253.12, p<.001). Selected coefficients from this model serve as the base for the variance decomposition presented in Tables 10.13-10.16. Variance decompositions systematically demonstrate the amount of variance explained by each teacher variable included at level two.

***Model 7.*** *SES and Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} SES_{ij} + e_{ij} \quad (10.15)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + r_{0j} \quad (10.16)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + \gamma_{41}\, avSES + r_{4j}$$

$$\beta_{5j} = \gamma_{50}$$

where

$\gamma_{00j}$ = intercept, average male achievement across teachers all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{40j}$ = intercept, average female achievement across teachers all other predictors held constant

$r_{4j}$ = the error term or unique teacher effects

**Table 10.11** Regression Coefficients for NAEP Mathematics Achievement 1996 by Gender and Average Classroom SES

| | Model 7 |
|---|---|
| Intercept ($\gamma_{00}$) | 233.35(1.22)*** |
| AvSES ($\gamma_{01}$) | -5.87(2.23)** |
| Black ($\gamma_{10}$) | -17.82(1.74)*** |
| Hispanic ($\gamma_{20}$) | -16.15(1.56)*** |
| All other ($\gamma_{30}$) | -2.65(1.89) |
| Dfem ($\gamma_{40}$) | -3.12(.97)** |
| AvSES ($\gamma_{41}$) | -.51(2.16) |
| SES ($\gamma_{50}$) | -11.61(1.39)*** |
| | |
| LVL 1 Variance ($\sigma 2$) | 516.45 |
| LVL 2 Intercept ($\tau_{00}$)~ | 252.63*** |
| Female ($\tau_{40}$)~ | 63.57** |

*** p<.001,** p<.01,* p<.05

How much variation does average classroom SES explain in the average mathematics achievement of males and females across teachers? To answer this question, the following equation computes the variance in female student achievement explained by average SES.

$$\frac{\tau_{40}(\text{model 6}) - \tau_{40}(\text{model 7})}{\tau_{40}(\text{model 6})} = \frac{61.84 - 63.57}{61.84} = -.28 = 0\% \qquad (10.17)$$

This means that 0%[28] of the variance in predicted average achievement for female students is explained by average SES across teachers (classrooms). Using the same process for the remaining tau, less than 1% of the variance in male average achievement is attributable to SES across teachers. SES does not explain variation in average achievement by gender.

**Gender and Teacher variables**

Now that the context is set for the analyses at Level Two, it is possible to track the variance explained by each teacher level variable controlling for classroom SES and to determine teacher effects on average mathematics achievement for male and female students across teachers. One variable at a time is added to the level two equation, EQ 10.14. The term "*variable x*" is exchanged for a single Mathematics Practice (math sensemaking activities, group and project activities, or group and project assessment activities), Teacher Characteristic (mathematics preparation, preparation for mathematics

---

[28] Negative estimates for variance are puzzling. Sometimes variability in a data set is large enough to produce a negative estimate, even though the true value of the variance component is positive. NAEP average classroom SES data is aggregated from student level records so it is not an independent measure - students may not only share teachers but schools and districts. It is common practice to treat negative variance components as if they are zero (SAS Institute Inc., 1999).

instruction, mathematics knowledge, reform orientation, diversity preparation), or Professional Development (professional development frequency in the last year).

***Model 8***. Template for Level 2 equations

Level 1 remains constant (Eq. 10.10)

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + \gamma_{02..}\, variable\ x + r_{0j} \tag{10.18}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40} + \gamma_{41}\, avSES + \gamma_{42}\, variable\ x + r_{4j}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{02..j}$ = change in average in average male student achievement associated with a one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

and

$\gamma_{42..j}$ = change in average female student achievement associated with one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

**Gender and Teacher Practices**. A variance decomposition for teacher practices can be found in Table 10.13. The table indicates that almost no additional variance in average achievement for female students is explained by teacher practices across teachers when SES is controlled. Small amounts are explained by teacher practices for male students.

**Table 10.12** Variance Decomposition for Male/Female Controlling SES - Teacher Practices

| | | avSES | Practice | | |
|---|---|---|---|---|---|
| Variables | Uncond | Free/Red Lu | Sensemaking | GPA assessment | Group&Project Activities |
| Intercept ($\gamma_{00}$) | 234.12(1.20)*** | 233.35(1.22)*** | 233.43(1.22)*** | 233.44(1.21)*** | 233.52(1.20)*** |
| ($\gamma_{01}$) | | | .96(.43)* | .73(.46) | 1.35(.36)*** |
| avSES($\gamma_{02}$) | | -5.87(2.23)** | -5.48(2.21)* | -5.97(2.23)** | -5.95(2.21)** |
| Female ($\gamma_{10}$) | -3.09(1.00)** | -3.12(.97)** | -3.18(.99)** | -3.22(.98)*** | -3.13(.98)** |
| ($\gamma_{11}$) | | | -.54(.42) | -.79(.41)~ | .42(.36) |
| avSES($\gamma_{12}$) | | -.51(2.16) | -.77(2.12) | -.40(2.16) | -.48(2.16) |
| Intrcpt Var($\tau_{00}$) | 253.12*** | 252.63*** | 243.66*** | 248.96*** | 239.40*** |
| Fem Var ($\tau_{40}$) | 61.84** | 63.57** | 61.27* | 61.32** | 61.93** |
| ($\sigma^2$) | 518.07 | 516.45 | 516.92 | 516.41 | 516.60 |
| % variance reduction intercept | | <1% | 3.7% | 3.5% | 5.4% |
| % variance reduction fem slope | | 0% | <1% | <1% | 0% |

For every one standard deviation increase in mathematics sensemaking and group and project assessment activities, male scores are predicted to increase by up to .96 points while female scores decrease by up to -.79 points. This is an indication of differential teacher practice effects in 1996 NAEP 4th grade mathematics data. For every one standard deviation increase in group and project assessment activities, male scores increase by .73 points while female scores decrease by -.79 points, another indicator of differential effects. A one standard deviation increase in group and project activities increases average achievement for both boys and girls.

**Gender, Teacher Characteristics, and Professional Development.** The variance decomposition for reform orientation follows in Table 10.14. Most of the teacher characteristic and professional development variables explain less than one percent of the variance in the scores of both genders. The one variable which explains variance across teachers for NAEP 1996 mathematics is reform orientation when SES is controlled.

**Table 10.13** Variance Decomposition for Male/Female Controlling for SES - Teacher Characteristics

| | | SES | Teacher Characteristics |
|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Unconditional | Free/Red Lunch | Reform Orientation |
| Intercept ($\gamma_{00}$) | 234.12(1.20)*** | 233.35(1.22)*** | 233.551.21)*** |
| ($\gamma_{01}$) | | | 1.65(.44)*** |
| avSES($\gamma_{02}$) | | -5.87(2.23)** | -5.47(2.22)* |
| Female ($\gamma_{10}$) | -3.09(1.00)** | -3.12(.97)** | -3.15(.98)** |
| ($\gamma_{11}$) | | | -.65(.45) |
| avSES($\gamma_{12}$) | | -.51(2.16) | -.69(2.15) |
| Intercept Var($\tau_{00}$) | 253.12*** | 252.63*** | 238.14*** |
| Female Var ($\tau_{40}$) | 61.84** | 63.57** | 60.57* |
| ($\sigma^2$) | 518.07 | 516.45 | 517.15 |
| % variance reduction intercept | | <1% | 5.9% |
| % variance reduction female slope | | 0% | 2.1% |

A one standard deviation increase in teacher reform orientation differentially increases the average male score by 1.65 points while decreasing the average female score by -.65 points.

*Summary.* Although the average scores for males and females vary across teachers, the effects of teacher practices, characteristics and professional development measured in this data do not explain much of the variance in scores for boys and girls. With the exception of group and project activities, it appears that these small effects are differential (i.e., increases in the teacher variables increase scores for boys while decreasing scores for girls on average). Increases in group and project activities increases average scores for both groups.

**Effects of teacher level variables on mean achievement by Black and White**

The fixed effects model for ethnicity, Model 2, predicts that average achievement varies significantly by ethnicity but a random slopes and intercepts model, Model 9, tests whether or not the average achievement for white and black students varies significantly across teachers. In other words, do teacher practices, characteristics, or professional development variables have differential effects on the average mathematics achievement of Black and/or White 4th grade students in the U.S.? If variance is significant across teachers, it is then possible to determine how much of the variation in each groups' achievement can be explained by single level two teacher variable and if that effect is different for each group.

The unconditional Model 9 establishes the maximum amount of variance that can be explained by teacher level variables for this data.

***Model 9.*** *Unconditional model for 4th grade Black/White achievement*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} SES_{ij} + e_{ij} \quad (10.19)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \quad (10.20)$$

$$\beta_{1j} = \gamma_{10} + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{70}$$

$$\beta_{5j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement/achievement in teacher j's classroom all other predictors held constant

$\gamma_{01j}$ = the change in average achievement associated with mean free and reduced lunch, all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, average black student achievement across teachers all other predictors held constant

$\gamma_{11j}$ = the change in average black achievement associated with one unit change in average free and reduced lunch

$r_{1j}$ = the error term or unique teacher effects

The estimated level two regression coefficients for Model 6 are given in Table 10.15.

**Table 10.14** Regression Coefficients for 1996 NAEP 4$^{th}$ Mathematics Achievement by Ethnicity

| | Model 9 |
|---|---|
| Intercept ($\gamma_{00}$) | 234.16(1.20)*** |
| Black ($\gamma_{10}$) | -19.09(1.78)*** |
| Hispanic ($\gamma_{20}$) | -16.47(1.56)*** |
| All other ($\gamma_{30}$) | -2.79(1.89) |
| Dfem ($\gamma_{40}$) | -3.06(.98)** |
| SES ($\gamma_{50}$) | -12.68(1.33)*** |
| | |
| LVL 1 Variance ($\sigma 2$) | 530.45 |
| LVL 2 Intercept ($\tau_{00}$) | 215.41*** |
| Black ($\tau_{10}$) | 57.10* |

*** p<.001,** p<.01,* p<.05

The effect of Black varies significantly across teachers ($\tau_{10}$=57.10, p<.05) as does the effect of White ($\tau_{00}$=215.41, p<.000). Selected coefficients from this analysis serve as the base for the variance decompositions found in Tables 10.12-10.15. Next district SES is added to the model to determine its effect on the average achievement for Black and White students across teachers. The variable for student free and reduced lunch has been aggregated to the teacher level as a control for classroom SES.

***Model 10.*** 4th grade *Black and White achievement controlling for SES*

Level 1- Students - Same as unconditional Equation 10.10

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + r_{0j} \tag{10.21}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, avSES + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement in teacher j's classroom holding all other predictors constant

$\gamma_{01j}$ = the change in average achievement associated with mean free and reduced lunch, all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, black student achievement in teacher j's classroom all other predictors held constant

$\gamma_{11j}$ = the change in average black student achievement associated with one unit change in average free and reduced lunch

$r_{1j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for Model 10 are given in Table 10.15.

**Table 10.15** Regression Coefficients for Black/White Achievement Predicted by Classroom SES

| | Model 10 |
|---|---|
| Intercept ($\gamma_{00}$) | 233.52(1.23)*** |
| avSES ($\gamma_{01}$) | -5.16(1.97)** |
| Black ($\gamma_{10}$) | -17.85(1.80)*** |
| avSES ($\gamma_{11}$) | -5.22(3.37) |
| Hispanic ($\gamma_{20}$) | -16.33(1.55)*** |
| All other ($\gamma_{30}$) | -2.80(1.89) |
| Dfem ($\gamma_{80}$) | -3.05(.98)** |
| SES ($\gamma_{90}$) | -11.69(1.38)*** |
| LVL 1 Variance ($\sigma^2$) | 529.49 |
| LVL 2 Variance ($\tau_{00}$) | 215.04*** |
| ($\tau_{10}$) | 48.48* |

*** p<.001,** p<.01,* p<.05

How much variation does average SES explain in the average mathematics achievement for White and Black students across teachers? To answer this question the following equation computes the variance in black achievement explained by average SES. The estimated $\tau_{10}$ values from Tables 10.10 and 10.11 would be substituted in the following equation:

$$\frac{\tau_{10}(\text{model } 9) - \tau_{10}(\text{model } 10)}{\tau_{10}(\text{model } 9)} = \frac{57.10 - 48.48}{57.10} = .151 = 15.1\% \qquad (10.22)$$

This means that 15.1% of the variance in predicted average achievement for black students is explained by average SES across teachers (classrooms). Using the same process for the remaining tau, less than 1% of the variance in White average achievement is attributable to SES across teachers. SES is a strong predictor of Black student achievement.

**Table 10.16** Variance Decomposition for Black/White Achievement by SES

| | | SES |
|---|---|---|
| Variables | Unconditional | Free/Red Lunch |
| Intercept ($\gamma_{00}$) | 234.16(1.20)*** | 233.52(1.23)*** |
| avSES($\gamma_{02}$) | | -5.16(1.97)** |
| Black ($\gamma_{10}$) | -19.09(1.78)*** | -17.85(1.80)*** |
| avSES($\gamma_{12}$) | | -5.22(3.37) |
| Intercept Variance ($\tau_{00}$) | 215.41*** | 215.04*** |
| Black Variance ($\tau_{10}$) | 57.10* | 48.48* |
| Level 1 variance ($\sigma 2$) | 530.45 | 529.49 |
| % variance reduction intercept | | <1% |
| % variance reduction black | | 11.7% |

*** p<.001,** p<.01,* p<.05

A one standard deviation increase in average classroom free and reduced lunch across teachers reduces average Black achievement by -5.12 points and decreases average White student achievement by -4.83 points. The variance decomposition from the table indicates that average classroom SES explains 11.7% of the variance in average Black achievement across teachers and almost none of the variance in average White achievement. SES, therefore, has a differential effect on Black and White student (non-Black) achievement.

**The Effects of Black, SES, and Teacher Variables**

Now that the context is set for the analyses at Level Two, it is possible to track the variance explained by each teacher level variable controlling for SES and to determine its effect on average mathematics achievement for Black and White students across teachers.

One variable at a time is added to the level two equation, EQ 10.14. The term "*variable x*" is exchanged for a single Mathematics Practice (math sensemaking activities, group and project activities, or group and project assessment activities), Teacher Characteristic (mathematics preparation, preparation for mathematics instruction, mathematics knowledge, reform orientation, diversity preparation, knowledge of NCTM standards), or Professional Development (professional development frequency in the last year).

**Model 8**. Template for Level 2 equations

Level 1 remains constant (Eq. 10.10)

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + \gamma_{02..}\, variable\ x + r_{0j} \qquad (10.23)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, avSES + \gamma_{12..}\, variable\ x + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{02..j}$ = change in average in average student achievement associated with a one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

and

$\gamma_{12..j}$ = the change in average black student achievement associated with one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

**Ethnicity and Teacher Practices**. A variance decomposition for teacher practices can be found in Table 10.17. The table indicates that little additional variance in average achievement for Black students is explained by teacher practices across teachers once classroom SES is controlled. For instance, average SES explains 15.1% of the variance in average Black achievement compared to average SES plus sensemaking which explains only 10.1% of the variance. Average SES by itself does a better job of explaining variation in Black achievement. Small amounts of additional variance are explained for White students.

**Table 10.17** Teacher Practices - Variance Decomposition for Average White/Black Achievement Across Teachers Controlling SES

| | | avSES | Practice | | |
|---|---|---|---|---|---|
| Variables | Unconditional | Free/Red Lun | Sensemaking | GPA assessment | Group&Project Activities |
| Intercept ($\gamma_{00}$) | 234.16(1.20)*** | 233.52(1.23)*** | 233.60(1.23)*** | 233.62(1.22)*** | 233.71(1.21)*** |
| ($\gamma_{01}$) | | | .72(.35)* | .43(.39) | 1.14(.33)** |
| avSES($\gamma_{02}$) | | -5.16(1.97)** | -4.84(1.98)* | -5.14(1.97)* | -5.19(1.97)** |
| Black ($\gamma_{10}$) | -19.09(1.78)*** | -17.85(1.80)*** | -17.97(1.80)*** | -17.91(1.80)*** | -18.17(1.84)*** |
| ($\gamma_{11}$) | | | -.11(.49) | -.44(.59) | .02(.56) |
| avSES($\gamma_{12}$) | | -5.22(3.37) | -5.40(3.38) | -5.15(3.45) | -5.34(3.35) |
| InterceptVar($\tau_{00}$) | 215.41*** | 215.04*** | 210.71*** | 213.49*** | 205.86*** |
| Black Var ($\tau_{40}$) | 57.10* | 48.48* | 51.31* | 50.35* | 53.55* |
| ($\sigma^2$) | 530.45 | 529.49 | 529.73 | 529.64 | 529.40 |
| % variance reduction intercept | | <1% | 2.2% | <1% | 4.4% |
| % variance reduction black | | 15.1% | 10.1% | 11.8% | 6.2% |

For every one standard deviation increase in sensemaking and group and project assessment activities, White scores increase by up to .72 points while Black scores decrease by up to -.44 points. This is an indication that there are differential teacher practice effects present in 1996 NAEP 4$^{th}$ grade mathematics data. For every one standard deviation increase in group and project activities, White scores increase by 1.14 points while Black scores remain nearly level. None of the practice variables explain more variation in average Black student scores across teachers than SES alone; however, the variables do explain small amounts of variance in White average achievement across teachers (up to 4.4%) after controlling for SES.

**The Effects of Teacher Characteristics on Average Black Achievement.** The variance decomposition for teacher characteristics follows in Table 10.19. There are four variables which explain more variance in average achievement across teachers than classroom SES alone - mathematics preparation, knowledge of NCTM standards, reform orientation, and mathematics knowledge. Mathematics preparation explains additional variation in achievement for both Black and White Students across teachers while mathematics knowledge explains additional variance in achievement for Black students only. Reform orientation explains a little additional variation in White average achievement controlling for SES as does knowledge of NCTM standards. A more detailed explanation follows the table.

**Table 10.18** Variance Decomposition for Black/White Controlling for SES - Teacher Characteristics

| | | SES | Teacher Characteristics | |
|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | MthPrep | NCTM |
| Intercept ($\gamma_{00}$) | 234.16(1.20)*** | 233.52(1.23)*** | 233.48(1.23)*** | 233.56(1.21)*** |
| ($\gamma_{01}$) | | | 1.48(1.00) | 2.35(1.05)* |
| avSES($\gamma_{02}$) | | -5.16(1.97)** | -4.92(1.96)* | -4.96(1.95)* |
| Black ($\gamma_{10}$) | -19.09(1.78)*** | -17.85(1.80)*** | -17.77(1.77)*** | .17.91(1.80)*** |
| ($\gamma_{11}$) | | | -1.38(1.49) | -2.17(1.67) |
| avSES($\gamma_{12}$) | | -5.22(3.37) | -5.33(3.36) | -5.25(3.36) |
| Intercept Var ($\tau_{00}$) | 215.41*** | 215.04*** | 212.77*** | 210.26*** |
| Black Variance ($\tau_{40}$) | 57.10* | 48.48* | 47.01* | 50.55* |
| Level 1 Variance ($\sigma^2$) | 530.45 | 529.49 | 529.69 | 529.53 |
| % variance reduction intercept | | <1% | 1.3% | 2.4% |
| % variance reduction black | | 15.1% | 17.7% | 11.4% |

| | | SES | Teacher Characteristics | |
|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | Reform Or | MthKnow |
| Intercept ($\gamma_{00}$) | 234.16(1.20)*** | 233.52(1.23)*** | 233.72(1.22)*** | 233.54(1.23)*** |
| ($\gamma_{01}$) | | | 1.29(.41)** | .27(.92) |
| avSES($\gamma_{02}$) | | -5.16(1.97)** | -4.82(1.96)* | -5.10(1.97)* |
| Black ($\gamma_{10}$) | -19.09(1.78)*** | -17.85(1.80)*** | -18.19(1.78)*** | -18.34(1.85)*** |
| ($\gamma_{11}$) | | | .28(.65) | 2.96(1.944) |
| avSES($\gamma_{12}$) | | -5.22(3.37) | -5.67(3.38) | -5.11(3.34) |
| Intercept Var ($\tau_{00}$) | 215.41*** | 215.04*** | 206.38*** | 215.58*** |
| Black Variance ($\tau_{40}$) | 57.10* | 48.48* | 53.17* | 46.94* |
| Level 1 Variance ($\sigma^2$) | 530.45 | 529.49 | 529.73 | 529.42 |
| % variance reduction intercept | | <1% | 4.2% | 0% |
| % variance reduction black slope | | 15.1% | 6.9% | 17.8% |

For every standard deviation increase in mathematics preparation (degrees/licenses), average White scores increase by 1.48 points and Black scores drop by -1.38 points. Mathematics preparation explains 17.7% of the variation in average Black achievement across teachers controlling for classroom SES but less than 1% for average White achievement.

Reform orientation controlling for SES across teachers explains 4.2% of the variation in white average achievement over the course of one year while explaining no additional variance in the average achievement for Black students across teachers controlling for SES. For every one standard deviation increase in teachers' reform orientation, black students are expected to gain .28 points over the course of one year. White students are expected to gain 1.29 points.

Teachers' math knowledge controlling for SES explains 17.8% of the variation in average Black achievement across teachers but nothing for average White achievement. A one standard deviation increase in mathematics knowledge increases Black achievement by 2.96 points while increasing White achievement by .27 points.

SES is again the most important single predictor of Black achievement across teachers but SES explains little about White achievement across teachers. Two teacher characteristic variables, mathematics preparation and mathematics knowledge, explained more variance in average Black achievement across teachers once SES was controlled than SES alone. The effect of mathematics preparation was differential. Increases in mathematics knowledge and reform orientation increased the average scores for both student groups across teachers.

Ethnicity and Teacher Professional Development. Professional development frequency does not explain variance for either group of students across teachers.

**Effects of teacher level variables on the mean achievement of Black Females**

Using the same process and equations demonstrated earlier for female and Black average achievement, the resulting variance decomposition for Black female is presented below. A more detailed explanation follows the table.

**Table 10.19** Variance Decomposition for Black Female Controlling for SES - Teacher Practice, Characteristics, and Professional Development

| | | SES | Practice | |
|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Free/Red Lu | GPA | GPA Assess |
| Intercept ($\gamma_{00}$) | 234.02(1.23)*** | 233.17(1.26)*** | 233.40(1.24)*** | 233.34(1.25)*** |
| ($\gamma_{01}$) | | | 1.14(.31)** | .33(.37) |
| avSES($\gamma_{02}$) | | -6.18(1.83)** | -6.29(1.82)** | -6.24(1.83)** |
| Black Fem ($\gamma_{10}$) | 1.42(1.22 | 1.56(1.23) | 1.48(1.21) | 1.36(1.22) |
| ($\gamma_{11}$) | | | .43(.20)* | .58(.23)* |
| avSES($\gamma_{12}$) | | -1.41(1.15) | -1.39(1.14) | -1.37(1.14) |
| Intercept Var ($\tau_{00}$) | 211.04*** | 209.04*** | 199.34*** | 207.28*** |
| Black Fem Var ($\tau_{40}$) | 20.25* | 19.99* | 18.86* | 17.81* |
| Level 1 Variance ($\sigma^2$) | 517.45 | 516.28 | 516.29 | 516.82 |
| % variance reduction intercept | | <1% | 5.5% | 1.8% |
| % variance reduction Black Female | | 1.3% | 6.9% | 12.0% |

| | | Practice | Teacher Characteristics | |
|---|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | Sensemaking | MthPrep | Reform Or |
| Intercept ($\gamma_{00}$) | 234.16(1.20)*** | 233.26(1.25)*** | 233.13(1.26)*** | 233.45(1.25)*** |
| ($\gamma_{01}$) | | .69(.33)* | 1.31(.92) | 1.32(.37)** |
| avSES($\gamma_{02}$) | | -5.95(1.83)** | -5.99(1.82)** | -5.89(1.81)** |
| Black Fem ($\gamma_{10}$) | -19.09(1.78)*** | 1.56(1.23) | 1.55(1.23) | 1.44(1.23) |
| ($\gamma_{11}$) | | .30(.24) | .24(.47) | .62(.25)* |
| avSES($\gamma_{12}$) | | -1.26(1.13) | -1.39(1.15) | -1.19(1.15) |
| Intercept Var ($\tau_{00}$) | 211.04*** | 204.36*** | 207.62*** | 199.48*** |
| Black Fem Var ($\tau_{40}$) | 20.25* | 19.54* | 20.33* | 19.31* |
| Level 1 Variance ($\sigma^2$) | 517.45 | 516.56 | 529.69 | 516.09 |
| % variance reduction intercept | | 3.2% | 1.6% | 4.2% |
| % variance reduction black slope | | 3.5% | 0% | 6.9% |

**Table 10.19** (cont'd).

| | | Characteristics | Professional Dev |
|---|---|---|---|
| Variables ($\gamma_{01}$ & $\gamma_{11}$) | Uncond | NCTM knowledge | PDF |
| Intercept ($\gamma_{00}$) | 234.02(1.23)*** | 233.23(1.25)*** | 233.17(1.25)*** |
| ($\gamma_{01}$) | | 2.07(.99)* | -.32(.68) |
| avSES($\gamma_{02}$) | | -5.93(1.82)** | -6.19(1.83)** |
| Black Fem ($\gamma_{10}$) | 1.42(1.22 | 1.54(1.23) | 1.55(1.23) |
| ($\gamma_{11}$) | | .84(.62) | .61(.45) |
| avSES($\gamma_{12}$) | | -1.32(1.14) | -1.36(1.15) |
| Intercept Var ($\tau_{00}$) | 211.04*** | 205.58*** | 209.39*** |
| Black Fem Var ($\tau_{40}$) | 20.25* | 19.68* | 19.63* |
| Level 1 Variance ($\sigma^2$) | 517.45 | 516.29 | 516.30 |
| % variance reduction intercept | | 2.6% | <1% |
| % variance reduction black female | | 2.8% | 3.1% |

The effect of the included teacher variables on Black female average achievement across teachers is positive. Group and project activities and group and project assessment activities explain the most variance[29] in black female achievement across teachers after SES is controlled. Note that SES accounts for little variation in this analysis, similar to the results from the male/female variance decomposition. Group and project activities explains nearly 7% of the variance in black female scores across teachers and group/project assessment activities explain 12% of the variance. A one standard deviation increase in these two teacher variables results in an increase of about half a point for Black females.

The only differential effect present is that for professional development frequency which is very small. For every one standard deviation increase in professional

[29] Variance explained in Black female achievement across teachers should be interpreted with caution since it is the effect of level two teacher variables on the level one interaction of black and female.

development frequency, Black female average achievement increases by .61 points while average achievement for white male students decreases by -.32 points.

# 11. RESULTS OF THE HLM ANALYSIS FOR 2000 NAEP 4TH GRADE MATHEMATICS

The HLM analysis using NAEP data is intended to address the research questions about the relationships among student achievement and

- Teacher characteristics and practices
- Characteristics of professional development

The analysis also is intended to address the ways in which these relationships vary by student demographic group, including gender, ethnicity, and SES across the United States and to see if the national mathematics results reflect trends similar to those found in the Connecticut BTPS data, a more detailed data set. Analyses concerning the relationships between teacher preparation institutions and teachers/students are not possible for this data since the names of teacher preparation programs are not collected in NAEP. The HLM analyses are weighted (origwt) and use five plausible values (mrpcm1-5) for the student outcome. There are 6,205 students and 1,245 teachers in the 2000 mathematics sample.

To set the context for addressing the questions, results are presented first which describe the differences in student achievement (achievement status in NAEP data) across student demographic groups. The analyses yield results in terms of regression coefficients. To aid in understanding what the coefficients mean, predicted achievement (achievement status score) is computed for selected combinations of student characteristics.

The full set of mathematics variables for NAEP is given in Table 11.1. The outcomes and most teacher variables are the same as those used for the earlier SEM models. In certain cases, variables that were only available in this NAEP data set were included as well.

**Table 11.1** Description of Variables Used in the 2000 NAEP 4th Grade Mathematics Analyses

| Type | Variable Abbreviation | Variable Description |
|---|---|---|
| | | **Level 1 - students** |
| Continuous (Outcomes) | avepv | The average of the 5 plausible values for mathematics achievement (mean 224.16, sd=29.75) |
| | mrpcm1 | 1st of 5 multiple plausible values for the achievement of each student (mean=224.01, sd=31.18) |
| | mrpcm2 | 2nd of 5 multiple plausible values for the achievement of each student (mean=224.35, sd=30.77) |
| | mrpcm3 | 3rd of 5 multiple plausible values for the achievement of each student (mean=224.09, sd=31.28) |
| | mrpcm4 | 4th of 5 multiple plausible values for the achievement of each student (mean=224.17, sd=31.00) |
| | mrpcm5 | 5th of 5 multiple plausible values for the achievement of each student (mean=224.16, sd=30.90) |
| Dichotomous | dblack | Black (Black=1, else=0) |
| | dhispanic | Hispanic (Hispanic=1, else=0) |
| | dallother | All other ethnicities (all other=1, else=0) |
| | dfem | Female (female=1, else=0) |
| | SES | Average free and reduced lunch (eligible=1, else=0) |
| Effects coded | egenxbl | Effect of Black female; interaction of eBlack x egender where Black =1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| | egenxhisp | Effect of Hispanic female; interaction of eHispanic x egender where Hispanic=1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| | egenxoth | Effect of All other female; interaction of eAllother x egender where All other =1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| Continuous | Origwt | Original weight (correction for oversampling) |

**Table 11.1** (cont'd).

| | | |
|---|---|---|
| | | **Level 2 - teachers** |
| Continuous | | Grand mean centered |
| Teacher Practices | gpa | Frequency of group and project activities (mean=12.52, sd=2.61) |
| | dla | Frequency of drill and lecture activities (mean=8.62, sd=.80) |
| | sensemaking | Frequency of sensemaking activities (mean=11.20, sd=2.61) |
| Teacher Characteristics | NCTMstds | Knowledge of NCTM standards (mean=1.96,sd=.89) |
| | statestds | Knowledge of state standards (mean=3.38,sd=.60) |
| | mathknow | Knowledge of mathematics (mean=27.93, sd=3.31) |
| | assessprep | Preparation for assessment (mean=11.36, sd=2.24) |
| | mathinst | Preparation for mathematics instruction, pedagogical coursework (mean=27.95, sd=3.39) |
| | motrad | Traditional orientation, emphasize subject (mean=8.42, sd=1.63) |
| | moref | Reform orientation, emphasize reasoning (mean=7.64, sd=1.30) |
| Professional Development | pdf | Professional development in the last year (2.84, sd=1.22) |
| Continuous | avSES | Average percent free and reduced lunch per teacher/ classroom aggregated from level 1 (mean=.39, sd=.49) |

**I. Setting the Stage: Variation in Mathematics Achievement by Student Demographic Groups and SES** (Results at the Student Level)

In this section, sets of background variables are added to look at changes in mathematics achievement in the United States. Model 1 predicts mean student achievement across the sample and partitions the variance between students and teachers where students are nested within teachers. Models 2-5 predict differences in mean achievement across gender, ethnicity, and SES (eligibility for free and reduced lunch).

Mean Achievement. To set the context for answering the questions about connections between student achievement and teacher variables, results are presented first for differences in student achievement across student demographic groups. The level one and level two equations for the model are given followed by their estimated regression coefficients in Tables 8.2.

***Model 1.*** *Mean achievement*

Level 1 - Students

$$Y_{ij} = \beta_{0j} + e_{ij} \tag{11.1}$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

where $Y_{ij}$ is the mathematics achievement for student i in teacher j's classroom, $\beta_{0j}$ is the intercept or mean students' mathematics achievement within teacher j's classroom, $e_{ij}$ is the level 1 error, $\gamma_{00}$ is the mean value of classroom achievement for all teachers, and $r_{0j}$ is the level 2 error.

**Table 11.2** 2000 NAEP 4th Grade Average Mathematics Achievement

| | Model 1 |
|---|---|
| Intercept ($\gamma_{00}$) | 225.57(.80)*** |
| LVL 1 Variance ($\sigma^2$) | 681.44 |
| LVL 2 Variance ($\tau_{00}$) | 362.09*** |

*** p<.001,** p<.01,* p<.05

The predicted average achievement score for 4th grade mathematics students for this sample of US students is 225.57 points ($\gamma_{00}$=225.57, sd=.80) when no other variables are included in the model.

$\sigma^2$ and $\tau_{00}$ are used to compare variation in the errors at the teacher level with variation at the student level using the formulas:

$$\frac{\sigma^2}{\sigma^2 + \tau_{00}} = \textit{level one variance} \qquad \frac{\tau^2{}_{00}}{\sigma^2 + \tau_{00}} = \textit{level two variance} \quad (11.2)$$

65% of the variance in achievement for Model 1 is within teachers (between students) leaving 35% of the variance in achievement between teachers. That is to say, there is almost 2 times more variation within teachers (at the student level) than between teachers.

Mean Achievement by Demographic Group. Next, demographic variables are added to form Models 2, 3 and 4. Model 5 introduces the level one control variable for SES. The equations for Model 2, ethnic group, are presented first (Eq.11.3) To interpret the coefficients, the estimated score for ethnicity can be computed using the appropriate set of 1s and 0s for the dummy variables. As an example, Equations 11.4 and 11.5 demonstrate the computation for obtaining Black student achievement. Estimated achievement by ethnicity are presented in Table 11.3.[30]

***Model 2.*** *Ethnic groups*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + e_{ij};$$

Level 2 - Teachers (11.3)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

---

[30] Note that the category for All other includes Asian and those students who did not report ethnicity. Often these smaller groups are folded into the intercept since they are not significantly different from the intercept. In this case, the coefficient for all other was significant at $p<.05$. Additionally, it was important to maintain the meaning of the intercept as a specific group, in this case white students, before analyzing the differential effects of teacher variables.

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

The regression coefficients for predicting student achievement from ethnic background are found in Table 11.3.

**Table 11.3** Regression Coefficients for 2000 NAEP 4th Grade Mathematics by Ethnicity

| | Model 2 |
|---|---|
| Intercept ($\gamma_{00}$) | 233.43(.84)*** |
| Black ($\gamma_{01}$) | -25.86(1.39)*** |
| Hispanic ($\gamma_{02}$) | -20.42(1.29)*** |
| All other ($\gamma_{03}$) | -5.59(2.53)* |
| | |
| LVL 1 Variance ($\sigma^2$ ) | 644.18 |
| LVL 2 Variance ($\tau_{00}$) | 232.97*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by ethnicity group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[31] For example, to determine the predicted achievement for a black student who would have ones for the intercept and black the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} \tag{11.4}$$

Level 2 - Teachers

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 233.43$$

$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -25.86$$

---

31 A table showing ones and zeros for all dummy variables is provided in the Appendix.

When the slope values, the betas from Equation 8.4, are substituted in the level one equation (Eq. 8.4), the result for black achievement is produced (Eq. 8.5).

$$\hat{Y}_{ij} = (233.43) + (-25.86)*1 \qquad (11.5)$$

$$= 207.57$$

Predicted mathematics achievement for all demographic groups are presented in Table 11.4. This table shows that the model predicts that White students score higher in mathematics than the other ethnic groups. The coefficients for ethnicity are significantly different from zero.

**Table 11.4** Predicted Mathematics Achievement Student Demographic Group

| Ethnic Group | Achievement |
|---|---|
| White | 233.43 |
| Black | 207.57 |
| Hispanic | 213.01 |
| All other | 227.84 |

In the next model, gender predicts average mathematics achievement.

***Model 3.*** *Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{4j}\, dfem_{ij} + e_{ij};$$

Level 2 - Teachers (11.6)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{4j} = \gamma_{10}$$

The regression coefficients for predicting student achievement from gender are found in Table 11.5.

**Table 11.5** Regression Coefficients for 2000 NAEP Mathematics Achievement by Gender

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | 226.55(.91)*** |
| Female ($\gamma_{04}$) | -2.00(.84)* |
| LVL 1 Variance ($\sigma^2$) | 680.53 |
| LVL 2 Variance ($\tau_{00}$) | 362.03*** |

*** p<.001,** p<.01,* p<.05

These coefficients indicate that male mean achievement in mathematics is predicted to be 226.55 points compared to 224.56 points for females. The coefficient for female is significantly different from zero.

***Model 4****. Combined gender and ethnicity*

Model 4 tests whether or not there is an interaction between gender and ethnicity in the NAEP 4th grade mathematics data. Equation 11.7 computes coefficients for combinations of gender and ethnicity.

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij} + e_{ij}; \qquad (11.7)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background and gender are found in Table 11.6.

**Table 11.6** Regression Coefficients for 2000 NAEP 4th Grade Mathematics Achievement by Gender and Ethnicity

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | 233.51(.95)*** |
| Black ($\gamma_{01}$) | -25.75(1.39)*** |
| Hispanic ($\gamma_{02}$) | -20.22(1.29)*** |
| All other ($\gamma_{03}$) | -5.58(2.52)* |
| Female ($\gamma_{04}$) | -.24(1.18) |
| Effect Black x female ($\gamma_{05}$) | 2.19(.82)** |
| Effect Hispanic x female ($\gamma_{06}$) | -.91(.82) |
| Effect All other x female ($\gamma_{07}$) | .54(1.38) |
| | |
| LVL 1 Variance ($\sigma^2$) | 641.50 |
| LVL 2 Variance ($\tau_{00}$) | 232.27*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by demographic group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[32] For example, to determine the predicted achievement for a black female student who would have ones for the intercept, black, female and black female,

[32] A table showing ones and zeros for all dummy variables is provided in the Appendix.

the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} eblackxefem_{ij}$$

Level 2 - Teachers (11.8)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 233.51$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -25.75$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = -\ .24$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = 2.19$$

When the slope values, the level two betas from Equation 11.8, are substituted in the level one equation (Eq. 11.8), the result for Black female achievement is produced (Eq. 11.9).

$$\hat{Y}_{ij} = (233.51) + (-25.75)*1 + (-.24)*1 + (2.19)*1 \quad (11.9)$$
$$= 209.71$$

Predicted mathematics achievement for all subsets of students by demographic groups are presented in Table 11.7. The table shows that the model predicts that White males are predicted to have the highest scores followed closely by white females.

**Table 11.7** Predicted Mathematics Achievement by Student Demographic Group with Differences by Gender

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 233.51 | 233.27 | - .24 |
| Black | 207.76 | 209.71 | +1.95 |
| Hispanic | 213.29 | 212.14 | - 1.15 |
| All other | 227.93 | 228.23 | + .30 |

Score differences between males and females of the same ethnicity are predicted to be less than 2 points. Both Black and Other females are expected to do slightly better than their male counterparts. The coefficient for the effect of female x Black is significantly different from zero ($p<.05$).

### *Gender, Ethnicity, and SES*

Another known source of variation in student achievement is SES as measured by eligibility for free and reduced lunch. In NAEP, this variable is measured at the student level.

***Model 5.*** **Gender, Ethnicity, and SES**

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij} + \beta_{8j} SES_{ij} + e_{ij}; \tag{11.10}$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background, gender, and SES are found in Table 11.8.

**Table 11.8** Regression Coefficients for 2000 NAEP 4th Mathematics Achievement by Gender, Ethnicity, and SES

| | Model 5 |
|---|---|
| Intercept ($\gamma_{00}$) | 236.94(1.00)*** |
| Black ($\gamma_{01}$) | -20.42(1.43)*** |
| Hispanic ($\gamma_{02}$) | -16.40(1.27)*** |
| All other ($\gamma_{03}$) | -3.64(2.46) |
| Female ($\gamma_{04}$) | .09(1.15) |
| Effect Black x female ($\gamma_{05}$) | 2.08(.80)* |
| Effect Hispanic x female ($\gamma_{06}$) | -1.13(.81) |
| Effect All other x female ($\gamma_{07}$) | 1.04(1.37) |
| SES ($\gamma_{08}$) | -14.25(1.17)*** |
| LVL 1 Variance ($\sigma^2$) | 628.74 |
| LVL 2 Variance ($\tau_{00}$) | 185.41*** |

*** p<.001,** p<.01,* p<.05

As demonstrated in prior sections, the computational example to determine the predicted achievement for a black female student eligible for free and reduced lunch who would have ones for the intercept, black, female, black female, and free and reduced lunch, would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} egenxbl_{ij} + \hat{B}_{9j} SES_{ij}$$

Level 2 - Teachers (11.11)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 236.94$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -20.42$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = .09$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = 2.08$$
$$\hat{\beta}_{8j} = \hat{\gamma}_{80} = -14.25$$

When the slope values, the betas from Level 2, Equation 8.11, are substituted in the Level 1 equation, the result for black female achievement is produced (Eq. 8.12).

$$\hat{Y}_{ij} = (236.94) + (-20.42)*1 + (.09)*1 + (2.08)*1 + (-14.25*1) \quad (11.12)$$
$$= 204.44$$

Predicted mathematics achievement for all subsets of students by demographic group and grade are presented in Table 11.9. The estimated point loss associated with free and reduced lunch eligibility is -14.25 points and its coefficient is significantly different from zero (p<.000).

**Table 11.9** Predicted Mathematics Achievement by Student Demographic Group and SES

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 222.69 | 222.78 | + .09 |
| Black | 202.27 | 204.44 | + 2.17 |
| Hispanic | 206.29 | 205.25 | - 1.04 |
| All other | 219.05 | 220.18 | + 1.13 |

This table indicates that the model predicts that all non-white student groups achieve lower mathematics scores than white students once ethnicity, gender, and SES are taken into consideration. Black students show the largest score difference by gender with girls predicted to score 2.17 points higher than boys. Only Hispanic girls score lower than their male counterparts (-1.04 points).

*Summary of Section 1.* Developing the level one equation for students. As the level one base equation for 2000 NAEP 4$^{th}$ grade mathematics developed, average yearly achievement was first predicted with a simple unconditional model. This model indicated that achievement variance was nearly two times larger within teachers (between students) than between teachers at level 2. Variables known to affect mean achievement (ethnicity, gender, ethnicity x gender, and SES) were added to the level one equation revealing that the coefficients for the intercept, Black, Hispanic, Black female, and SES were significantly different from zero. Being eligible for free and reduced lunch predicted a large scores drop for all students and White, Black, and All other girls have higher scores than their male counterparts when SES is controlled. In other words, the scores of boys from these groups drop farther than girls when SES is controlled.

**II. Differential Effects of Teacher Characteristics, and Teacher Practices on Student Achievement** (Results at the teacher level)

This section addresses the questions:

1. Does mathematics achievement for demographic groups vary across teachers?
2. How much variation in student mathematics achievement do variables about teacher practices, characteristics, and professional development explain?
3. Is the amount of variance explained by teacher practices, characteristics, and professional development different for different student demographic

groups?

The analyses in this section first test if average achievement for the various demographic groups varies across teachers. Nonsignificant error terms for variance ($r_{00}$, $r_{10}$, ...$r_{x0}$) indicate that group mean achievement does not vary across teachers so the error terms predicting mean achievement for those demographic groups are dropped from the model. Although the error terms are dropped, the fixed level two equations for non-significant achievement by demographic groups are retained in the analysis. Analyses indicated that average achievement for female, the female times ethnicity interactions, Black, and All other did not vary significantly across teachers ($\tau_{..0}=0$) so their error term have been eliminated. Their variance is then captured as part of the error term for the intercept ($r_{00}$). The $\chi^2$ tests for the intercept[33] and the average mathematics achievement for Hispanic students across teachers are significant ($\tau_{..0}\neq 0$).

2000 NAEP mathematics achievement varies across teachers for White and Hispanic. The $\chi^2$ tests indicate that the variances of the intercept and the average mathematics achievement of Hispanics across teachers are significant ($\tau_{..0}\neq 0$). The effects teacher practices, teacher characteristics, and teacher professional development on mean achievement for White and Hispanic is addressed below.[34]

**Effects of teacher level variables on mean Hispanic achievement**

While the fixed effects model for ethnicity, Model 2, predicts that average

---

[33] When main effects variables are dummy coded, not only does $\beta_0$ represent the mean for white male achievement, but the variance of $\beta_0$, $\tau_{00}$, is the variance of white male achievement (see Raudenbush, p.34).

[34] Average free and reduced lunch varies significantly across teachers (classrooms). The level two control variable for average SES is introduced in Model 7.

achievement varies significantly by ethnicity, a random slopes and intercepts model, Model 6, tests whether or not the average achievement for White and Black students varies significantly across teachers. In other words, do teacher practices, characteristics, or professional development variables have differential effects on the average mathematics achievement of Hispanic and/or White 4th grade students in the U.S.? If variance is significant across teachers, it is then possible to determine how much of the variation in each groups' achievement can be explained by single level two teacher variable and if that effect is different for each group.

The unconditional Model 6 establishes the maximum amount of variance that can be explained by teacher level variables for this data.

***Model 6.*** *Unconditional model for 4th grade Hispanic/White achievement*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + \beta_{4j}\,dfem_{ij} + \beta_{5j}SES_{ij} + e_{ij} \quad (11.10)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \quad (11.11)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20} + r_{2j}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{70}$$

$$\beta_{5j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement/achievement across teachers all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, average Hispanic student achievement/achievement across teachers all other predictors held constant

$r_{1j}$ = the error term or unique teacher effects

The estimated level two regression coefficients for Model 6 are given in Table 11.10.

**Table 11.10** Regression Coefficients for 2000 NAEP 4th Mathematics Achievement by Ethnicity

| | Model 6 |
|---|---|
| Intercept ($\gamma_{00}$) | 238.07(.94)*** |
| Black ($\gamma_{10}$) | -20.57(1.44)*** |
| Hispanic ($\gamma_{20}$) | -16.21(1.29)*** |
| All other ($\gamma_{30}$) | -3.86(2.46) |
| Dfem ($\gamma_{40}$) | -2.08(.82)* |
| SES ($\gamma_{50}$) | -14.26(1.18)*** |
| LVL 1 Variance ($\sigma2$) | 622.71 |
| LVL 2 Intercept ($\tau_{00}$)~ | 182.54*** |
| Hispanic ($\tau_{20}$)~ | 59.40* |

*** p<.001,** p<.01,* p<.05

The effect of Hispanic varies significantly across teachers ($\tau_{20}$=59.40, p<.05) as does the effect of White ($\tau_{00}$=182.54, p<.000). Selected coefficients from this analysis serve as the base for the variance decompositions found in Tables 11.12-11.15. Next classroom SES is added to the model to determine its effect on the average achievement for Hispanic and White students across teachers. The variable for student free and reduced lunch has been aggregated to the teacher level as control for classroom SES.

***Model 7.*** *$4^{th}$ grade Hispanic, White achievement and SES*

Level 1- Students - Same as unconditional Equation 11.10

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + r_{0j} \tag{11.12}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}\, avSES + r_{2j}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement/achievement across teachers all other predictors held constant

$\gamma_{01j}$ = the change in average achievement/achievement associated with one unit change in average free and reduced lunch

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{20j}$ = intercept, Hispanic achievement across teachers all other predictors held constant

$\gamma_{21j}$ = the change in average Hispanic student achievement associated with one unit change in average free and reduced lunch

$r_{1j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for Model 7 are given in Table 11.11.

**Table 11.11** Regression Coefficients for Black/White Achievement Predicted by Average Classroom SES

| | Model 7 |
|---|---|
| Intercept ($\gamma_{00}$) | 236.43(.97)*** |
| avSES ($\gamma_{01}$) | -7.29(1.47)*** |
| Black ($\gamma_{10}$) | -19.39(1.44)*** |
| Hispanic ($\gamma_{20}$) | -15.60(1.31)*** |
| avSES ($\gamma_{21}$) | 3.59(2.59) |
| All other ($\gamma_{30}$) | -3.45(2.45) |
| Dfem ($\gamma_{80}$) | -2.05(.82)* |
| SES ($\gamma_{90}$) | -12.56(1.25)*** |
| LVL 1 Variance ($\sigma^2$) | 620.34 |
| LVL 2 Variance ($\tau_{00}$) | 180.04*** |
| Hispanic ($\tau_{20}$) | 55.90* |

*** p<.001,** p<.01,* p<.05

How much variation does average SES explain in the average mathematics achievement for White and Black students across teachers? To answer this question the following equation computes the variance in black achievement explained by average SES. The estimated $\tau_{10}$ values from Tables 11.10 and 11.11 would be substituted in the following equation:

$$\frac{\tau_{20}(\text{model 6}) - \tau_{20}(\text{model 7})}{\tau_{20}(\text{model 6})} = \frac{59.40 - 55.90}{59.40} = .059 = 5.9\% \qquad (11.13)$$

This means that 5.9% of the variance in predicted average achievement for Hispanic students is explained by average SES across classrooms (teachers). Using the same process for the remaining tau, 1.2% of the variance in White average achievement is attributable to SES across teachers.

**Table 11.12** Variance Decomposition for Hispanic/White Achievement by SES

| | | SES |
|---|---|---|
| Variables | Uncond | Free/Red Lunch |
| Intercept ($\gamma_{00}$) | 238.07(.94)*** | 236.43(.97)*** |
| avSES($\gamma_{02}$) | | -7.29(1.47)*** |
| Hispanic ($\gamma_{10}$) | -16.21(1.29)*** | -15.60(1.31)*** |
| avSES($\gamma_{12}$) | | 3.59(2.59) |
| Intercept Variance ($\tau_{00}$) | 182.54*** | 180.04*** |
| Hispanic Variance ($\tau_{40}$) | 59.40* | 55.90* |
| Level 1 variance ($\sigma^2$) | 622.71 | 620.34 |
| % variance reduction intercept | | 5.9% |
| % variance reduction Hispanic | | 1.2% |

SES explains variance in the average mathematics achievement of White and Hispanic students across teachers in this data. For every one standard deviation increase in free and reduced lunch, White students (which includes all non-Hispanic students) lose -7.29 points while Hispanic students gain 3.59 points. SES explains 5.9% of the variance in average White scores but only 1.2% of the variance in average Hispanic scores across teachers.

**Ethnicity and Teacher variables**

Now that the context is set for the analyses at Level Two, it is possible to track the variance explained by each teacher level variable and to determine its effect on average mathematics achievement for Hispanic and White students across teachers. One variable at a time is added to the level two equation, EQ 11.14. The term "*variable x*" is added to "*avSES*" in Eq. 11.12 and represents a single Mathematics Practice (math sensemaking activities, group and project activities, or group and project assessment activities), Teacher Characteristic (mathematics preparation, preparation for mathematics instruction, mathematics knowledge, reform orientation, diversity preparation), or Professional

Development (professional development frequency in the last year).

**Ethnicity and Teacher Practices**. A variance decomposition for teacher practices can be found in Table 11.13. The table indicates that only Group and project activities explains additional variance in average achievement for White students across teachers once SES is controlled (2.5%).

**Table 11.13** Variance Decomposition for Hispanic/White Controlling SES - Teacher Practices

| | | SES | Practices | | |
|---|---|---|---|---|---|
| Variables | Uncond | Free/Red Lunch | Sensemaking | Group&Project Activities | DLA |
| Intercept ($\gamma_{00}$) | 238.07(.94)*** | 236.43(.97)*** | 236.45(.97)*** | 236.45(.97)*** | 236.45(.97)*** |
| ($\gamma_{01}$) | | | .29(.27) | .68(.25)** | -1.12(1.01) |
| avSES($\gamma_{02}$) | | -7.29(1.47)*** | -7.19(1.47)*** | -7.29(1.46)*** | -7.32(1.47)*** |
| Hispanic ($\gamma_{10}$) | -16.21(1.29)*** | -15.60(1.31)*** | -15.71(1.32)*** | -20.69(1.43)*** | -15.59(1.32)*** |
| ($\gamma_{11}$) | | | -.39(.47) | .06(.49) | -.42(1.79) |
| avSES($\gamma_{12}$) | | 3.59(2.59) | 3.55(2.59) | 3.52(2.59) | 3.60(2.60) |
| Intercept Var ($\tau_{00}$) | 182.54*** | 180.23*** | 180.23*** | 178.17*** | 179.15*** |
| Hispanic Var ($\tau_{40}$) | 59.40* | 55.90* | 56.72* | 57.94* | 57.06* |
| Level 1 Vari ($\sigma^2$) | 622.71 | 620.34 | 620.12 | 619.83 | 620.49 |
| % variance reduction intercept | | 1.3% | 1.3% | 2.5% | 1.9% |
| % variance reduction Hispanic | | 5.9% | 4.5% | 2.5% | 3.9% |

For every one standard deviation increase in sensemaking activities, White scores increase by up .29 points while Hispanic scores decrease by up to -.39 points. For every one standard deviation increase in group and project activities, White scores increase by .68 points while Hispanic scores remain nearly level. Both Hispanic and White scores decrease withe a one standard deviation increase in Drill and lecture activities. Teacher practices explain very little about the variance of Hispanic and White mathematics achievement across teachers.

The results for teacher characteristics and professional development variables are similar to the results for teacher practices. They explain little variance in average scores of Hispanics and Whites across teachers once SES is taken into account.

# 12. RESULTS OF THE HLM ANALYSIS FOR 1998 NAEP 4TH GRADE READING

The HLM analysis using NAEP data is intended to address the research questions about the relationships between student achievement and

- Teacher characteristics and practices
- Characteristics of professional development

The analysis also is intended to address the ways in which these relationships vary by student demographic group, including gender, ethnicity, and SES across the United States and to see if the national reading results reflect trends similar to those found in the Connecticut BTPS data, a more detailed data set. Analyses concerning the relationships between teacher preparation institutions and teachers/students are not possible for this data since the names of teacher preparation programs are not collected in NAEP. The HLM analyses are weighted (origwt) and use five plausible values (rrpcm1-5) for the student outcome. There are 4,303 students and 828 teachers in the 1998 reading sample.

To set the context for addressing the questions, results are presented first which describe the differences in student achievement (achievement status in NAEP data) across student demographic groups. The analyses yield results in terms of regression coefficients. To aid in understanding what the coefficients mean, predicted achievement (achievement status score) is computed for selected combinations of student characteristics.

The full set of reading variables for NAEP is given in Table 12.1. The outcomes and most teacher variables are the same as those used for the earlier SEM models. In certain cases, variables that were only available in this NAEP data set were included as well.

**Table 12.1** Description of Variables Used in the 1998 NAEP 4th Grade Reading Analyses

| Type | Variable Abbreviation | Variable Description |
|---|---|---|
| | | **Level 1 - students** |
| Continuous (Outcomes) | rrpcm1 | 1st plausible value for reading achievement (mean 213.36, sd=37.50) |
| | rrpcm2 | 2nd plausible value for reading achievement (mean 212.96, sd=37.78) |
| | rrpcm3 | 3rd plausible value for reading achievement (mean 212.64, sd=38.05) |
| | rrpcm4 | 4rth plausible value for reading achievement (mean 213.04, sd=37.55) |
| | rrpcm5 | 5th plausible value for reading achievement (mean 213.32, sd=37.43) |
| Dichotomous | dblack | Black (Black=1, else=0) |
| | dhispanic | Hispanic (Hispanic=1, else=0) |
| | dallother | All other ethnicities (all other=1, else=0) |
| | dfem | Female (female=1, else=0) |
| | SES | Average free and reduced lunch (eligible=1, else=0) |
| Effects coded | egenxbl | Effect of Black female; interaction of eBlack x egender where Black =1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| | egenxhisp | Effect of Hispanic female; interaction of eHispanic x egender where Hispanic=1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| | egenxoth | Effect of All other female; interaction of eAllother x egender where All other =1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| Continuous | Origwt | Original weight (correction for oversampling) |
| | | **Level 2 - teachers** |
| Continuous | | Grand mean centered |
| Teacher Practices | reformact | Frequency of reform activities (mean=38.18, sd=6.86) |
| | tradact | Frequency of traditional activities (mean=37.69, sd=3.91) |
| | compact | Frequency of computer activities (mean=6.00, sd=2.54) |

**Table 12.1** (cont'd).

| | | |
|---|---|---|
| Teacher Characteristics | readprep | Major/coursework/license in reading/English/writing, reading/English education (mean=15.49, sd=7.48) |
| | blito | Balanced literacy orientation (mean=14.93, sd=3.27) |
| | techprep | Technology preparation (mean=5.67,sd=1.59) |
| | divprep | Preparation for diversity(mean=3.95, sd=1.24) |
| | assessprep | Preparation for assessment (mean=28.95, sd=5.48) |
| | readinst | Preparation for reading instruction, pedagogical coursework (mean=23.42, sd=3.32) |
| Professional Development | pdf | Professional development in the last year (mean=3.15, sd=1.15) |
| Continuous | avSES | Average percent free and reduced lunch per teacher/ classroom aggregated from level 1 (mean=.46,sd=.50) |

**I. Setting the Stage: Variation in Reading Achievement Gains by Student Demographic Groups and SES** (Results at the Student Level)

In this section, sets of background variables are added to look at changes in reading gains in the United States. Model 1 predicts mean student gains across the sample and partitions the variance between students and teachers where students are nested within teachers. Models 2-5 predict differences in mean gains across gender, ethnicity, and SES (eligibility for free and reduced lunch).

Mean Gains. To set the context for answering the questions about connections between student achievement and teacher variables, results are presented first for differences in student achievement across student demographic groups. The level one and level two equations for the model are given followed by their estimated regression coefficients in Tables 8.2.

***Model 1.*** *Mean achievement*

Level 1 - Students

$$Y_{ij} = \beta_{0j} + e_{ij} \qquad (12.1)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

where $Y_{ij}$ is the reading achievement for student i in teacher j's classroom, $\beta_{0j}$ is the intercept or mean students' reading achievement within teacher j's classroom, $e_{ij}$ is the level 1 error, $\gamma_{00}$ is the mean value of classroom achievement for all teachers, and $r_{0j}$ is the level 2 error.

**Table 12.2** 1998 NAEP 4th Grade Average Reading Achievement

| | Model 1 |
|---|---|
| Intercept ($\gamma_{00}$) | 216.12(1.08)*** |
| LVL 1 Variance ($\sigma^2$) | 1025.62 |
| LVL 2 Variance ($\tau_{00}$) | 451.13*** |

*** $p<.001$,** $p<.01$,* $p<.05$

The predicted average achievement score for 4th grade reading students for this sample of US students is 223.41 points ($\gamma_{00}$=214.66, sd=1.07) when no other variables are included in the model.

$\sigma^2$ and $\tau_{00}$ are used to compare variation in the errors at the teacher level with variation at the student level using the formulas:

$$\frac{\sigma^2}{\sigma^2 + \tau_{00}} = \textit{level one variance} \qquad \frac{\tau^2{}_{00}}{\sigma^2 + \tau_{00}} = \textit{level two variance} \qquad (12.2)$$

70% of the variance in achievement for Model 1 is within teachers (between students) leaving 30% of the variance in achievement between teachers. That is to say, there is 2.3 times more variation within teachers than between teachers.

Mean Achievement by Demographic Group. Next, demographic variables are added to form Models 2, 3 and 4. Model 5 introduces the control variables for SES. The equations for Model 2, ethnic group, are presented first (Eq.12.3) To interpret the coefficients, the estimated score for ethnicity can be computed using the appropriate set of 1s and 0s for the dummy variables. As an example, Equations 12.4 and 12.5 demonstrate the computation for obtaining Black student achievement. Estimated achievement by ethnicity are presented in Table 12.3.[35]

***Model 2.*** *Ethnic groups*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + e_{ij};$$

Level 2 - Teachers (12.3)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

The regression coefficients for predicting student achievement from ethnic background are found in Table 12.3.

---

[35] Note that the category for All other includes Asian and those students who did not report ethnicity. Often these smaller groups are folded into the intercept since they are not significantly different from the intercept. In this case, the coefficient for all other was significant at $p<.05$. Additionally, it was important to maintain the meaning of the intercept as a specific group, in this case white students, before analyzing the differential effects of teacher variables.

**Table 12.3** Regression Coefficients for 1998 NAEP 4th Grade Reading by Ethnicity

| | Model 2 |
|---|---|
| Intercept ($\gamma_{00}$) | 224.40(1.04)*** |
| Black ($\gamma_{01}$) | -28.29(2.00)*** |
| Hispanic ($\gamma_{02}$) | -25.24(2.08)*** |
| All other ($\gamma_{03}$) | -9.78(3.20)** |
| LVL 1 Variance ($\sigma^2$) | 988.20 |
| LVL 2 Variance ($\tau_{00}$) | 276.13*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by ethnicity group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[36] For example, to determine the predicted achievement for a black student who would have ones for the intercept and black the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} \tag{12.4}$$

Level 2 - Teachers

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 224.40$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -28.29$$

When the slope values, the betas from Equation 8.4, are substituted in the level one equation (Eq. 8.4), the result for black achievement is produced (Eq. 8.5).

[36] A table showing ones and zeros for all dummy variables is provided in the Appendix.

$$\hat{Y}_{ij} = (224.40) + (-28.29)*1 \quad (12.5)$$
$$= 196.11$$

Predicted reading achievement for all demographic groups are presented in Table 12.4. This table shows that the model predicts that White students score higher in reading than the other ethnic groups. The coefficients for ethnicity are significantly different from zero.

**Table 12.4** Predicted Reading Achievement Student Demographic Group

| Ethnic Group | Achievement |
|---|---|
| White | 224.42 |
| Black | 196.11 |
| Hispanic | 199.16 |
| All other | 214.62 |

In the next model, gender predicts average reading achievement.

***Model 3.*** *Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{4j}\, dfem_{ij} + e_{ij};$$

Level 2 - Teachers (12.6)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{4j} = \gamma_{10}$$

The regression coefficients for predicting student achievement from gender are found in Table 12.5.

**Table 12.5** Regression Coefficients for 1998 NAEP Reading Achievement by Gender

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | 213.69(1.35)*** |
| Female ($\gamma_{04}$) | 4.80(1.37)*** |
| LVL 1 Variance ($\sigma^2$) | 1019.61 |
| LVL 2 Variance ($\tau_{00}$) | 451.34*** |

*** p<.001,** p<.01,* p<.05

These coefficients indicate that male mean achievement in reading is predicted to be 213.69 points compared to 218.49 points for females. On average, girls are expected to score nearly 5 points higher than boys on NAEP 4th Grade Reading. The coefficient for female is significantly different from zero.

***Model 4**. Combined gender and ethnicity*

Model 4 tests whether or not there is an interaction between gender and ethnicity in the NAEP 4th grade reading data. Equation 12.7 computes coefficients for combinations of gender and ethnicity.

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j}dblack_{ij} + \beta_{2j}dhispanic_{ij} + \beta_{3j}dallother_{ij} + \beta_{4j}\,dfem_{ij} + \beta_{5j}egenxbl_{ij} + \beta_{6j}egenxhisp_{ij} + \beta_{7j}egenxoth_{ij} + e_{ij}; \qquad (12.7)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background and gender are found in Table 12.6.

**Table 12.6** Regression Coefficients for 1998 NAEP 4th Grade Reading Achievement by Gender and Ethnicity

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | 220.94(1.41)*** |
| Black ($\gamma_{01}$) | -28.62(2.04)*** |
| Hispanic ($\gamma_{02}$) | -25.15(2.08)*** |
| All other ($\gamma_{03}$) | -10.13(3.14)** |
| Female ($\gamma_{04}$) | 6.94(2.01)** |
| Effect Black x female ($\gamma_{05}$) | -.37(1.31) |
| Effect Hispanic x female ($\gamma_{06}$) | -.83(1.40) |
| Effect All other x female ($\gamma_{07}$) | 2.30(2.07) |
| LVL 1 Variance ($\sigma^2$) | 980.99 |
| LVL 2 Variance ($\tau_{00}$) | 276.24*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by demographic group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[37] For example, to determine the predicted achievement for a black

[37] A table showing ones and zeros for all dummy variables is provided in the Appendix.

female student who would have ones for the intercept, black, female and black female, the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} eblackxefem_{ij}$$

Level 2 - Teachers (12.8)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 220.94$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -28.62$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = 6.94$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = -.37$$

When the slope values, the level two betas from Equation 12.8, are substituted in the level one equation (Eq. 12.8), the result for Black female achievement is produced (Eq. 12.9).

$$\hat{Y}_{ij} = (220.94) + (-28.62)*1 + (6.94)*1 + (-.37)*1 \quad (12.9)$$
$$= 198.89$$

Predicted reading achievement for all subsets of students by demographic groups are presented in Table 12.6. The table shows that the model predicts all female students are predicted to have higher scores than their male counterparts by at least six points. Coefficients for gender times ethnicity effects are not significantly different from zero.

**Table 12.7** Predicted Reading Achievement by Student Demographic Group with Differences by Gender

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 220.94 | 227.88 | + 6.94 |
| Black | 192.32 | 198.89 | + 6.57 |
| Hispanic | 195.79 | 201.90 | + 6.11 |
| All other | 210.81 | 220.05 | + 9.24 |

Black students are predicted to score lower on 4th Grade NAEP Reading than the other demographic groups, including those students who likely have English as a second language.

***Gender, Ethnicity, and SES***

Another known source of variation in student achievement is SES as measured by eligibility for free and reduced lunch. In NAEP, this variable is measured at the student level.

***Model 5.* Gender, Ethnicity, and Grade Level**

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} \qquad (12.10)$$
$$+ \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij}$$
$$+ \beta_{8j} SES_{ij} + e_{ij};$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$
$$\beta_{4j} = \gamma_{40}$$
$$\beta_{5j} = \gamma_{50}$$
$$\beta_{6j} = \gamma_{60}$$
$$\beta_{7j} = \gamma_{70}$$
$$\beta_{8j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background, gender, and SES are found in Table 12.8.

**Table 12.8** Regression Coefficients for 1998 NAEP 4th Reading Achievement by Gender, Ethnicity, and SES

| | Model 5 |
|---|---|
| Intercept ($\gamma_{00}$) | 225.10(1.40)*** |
| Black ($\gamma_{01}$) | -22.65(2.13)*** |
| Hispanic ($\gamma_{02}$) | -20.92(2.03)*** |
| All other ($\gamma_{03}$) | -8.37(3.09)** |
| Female ($\gamma_{04}$) | 7.23(1.98)** |
| Effect Black x female ($\gamma_{05}$) | -.38(1.30) |
| Effect Hispanic x female ($\gamma_{06}$) | -.69(1.30) |
| Effect All other x female ($\gamma_{07}$) | 2.39(1.39) |
| SES ($\gamma_{08}$) | -15.94(1.64)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 962.97 |
| LVL 2 Variance ($\tau_{00}$) | 220.26*** |

*** p<.001,** p<.01,* p<.05

As demonstrated in prior sections, the computational example to determine the predicted achievement for a black female student eligible for free and reduced lunch who would have ones for the intercept, black, female, black female, and free and reduced lunch, would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} egenxbl_{ij} + \hat{B}_{9j} SES_{ij}$$

Level 2 - Teachers (12.11)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 225.10$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -22.65$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = 7.23$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = -.38$$
$$\hat{\beta}_{8j} = \hat{\gamma}_{80} = -15.94$$

When the slope values, the betas from Level 2, Equation 12.11, are substituted in the Level 1 equation, the result for black female achievement is produced (Eq. 12.12).

$$\hat{Y}_{ij} = (225.10) + (-22.65)*1 + (7.23)*1 + (-.38)*1 + (-15.94*1) \quad (12.12)$$

$$= 193.36$$

Predicted reading achievement for all subsets of students by demographic group and grade are presented in Table 12.9. The estimated point loss associated with free and reduced lunch eligibility is -12.31 points.

**Table 12.9** Predicted Reading Achievement by Student Demographic Group and SES

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 209.16 | 216.39 | + 7.23 |
| Black | 186.51 | 193.36 | + 6.85 |
| Hispanic | 188.24 | 194.78 | + 6.54 |
| All other | 200.79 | 210.41 | + 9.62 |

This table also indicates that the model predicts that all non-white student groups achieve lower reading scores than white students once ethnicity, gender, and SES are taken into consideration. Hispanic students show the largest differences by gender with girls predicted to lose -4.23 points. The point gap between males and females in Table 12.9 is nearly the same as the point gap in Table 12.7 which did not include SES.

*Summary of Section 1.* Developing the level one equation for students. As the level one base equation for 1998 NAEP $4^{th}$ grade reading developed, average yearly achievement was first predicted with a simple unconditional model. This model indicated that achievement variance was 1.5 times larger within teachers (between students) than across teachers at level 2. Variables known to affect mean achievement (ethnicity, gender, ethnicity x gender, and SES) were added to the level one equation revealing that while differences in achievement exist within teachers, the interaction terms for gender time ethnicity were not significant predictors of student achievement.

**II. Differential Effects of Teacher Characteristics, and Teacher Practices on Student Achievement** (Results at the teacher level)

This section addresses the questions:

1. Does reading achievement for demographic groups vary across teachers?
2. How much variation in student reading achievement do variables about teacher practices, characteristics, and professional development explain?
3. Is the amount of variance explained by teacher practices, characteristics, and professional development different for different student demographic groups?

The analyses in this section first test if average achievement for the various demographic groups varies across teachers. Nonsignificant error terms for variance ($r_{00}$, $r_{10}$, ...$r_{x0}$) indicate that group mean achievement does not vary across teachers so the error terms predicting mean achievement for those demographic groups are dropped from the model. Although the error terms are dropped, the fixed level two equations for non-significant achievement by demographic groups are retained in the analysis. Analyses indicated that average achievement for female, Hispanic and All other across teachers did not vary significantly from the intercept ($\tau_{..0}$=0) so the error terms have been eliminated. Their variance is then captured as part of the error term for the intercept ($r_{00}$). The $\chi^2$

tests for the intercept, average reading achievement of White students,[38] and for the average reading achievement of Black students across teachers are significant ($\tau_{..0} \neq 0$) so they are addressed in a series of analyses which include teacher practices, teacher characteristics, and teacher professional development.

**Effects of teacher level variables on mean achievement by Black and White**

While the fixed effects model for ethnicity, Model 2, predicts that average achievement varies significantly by ethnicity, a random slopes and intercepts model, Model 6, tests whether or not the average achievement for white and black students varies significantly across teachers. In other words, do teacher practices, characteristics, or professional development variables have differential effects on the average reading achievement of Black and/or White 4th grade students in the U.S.? If variance is significant across teachers, it is then possible to determine how much of the variation in each groups' achievement can be explained by single level two teacher variable and if that effect is different for each group.

The unconditional Model 6 establishes the maximum amount of variance that can be explained by teacher level variables for this data.

***Model 6.*** *Unconditional model for 4th grade Black/White achievement*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} SES_{ij} + e_{ij} \quad (12.10)$$

[38] When main effects variables are dummy coded, not only does $\beta_0$ represent the mean for white male achievement, but the variance of $\beta_0$, $\tau_{00}$, is the variance of white male achievement (see Raudenbush, p.34).

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \tag{12.11}$$

$$\beta_{1j} = \gamma_{10} + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{70}$$

$$\beta_{5j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement/achievement in teacher j's classroom all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, average black student achievement/achievement across teachers all other predictors held constant

$r_{4j}$ = the error term or unique teacher effects

The estimated level two regression coefficients for Model 6 are given in Table 12.10.

**Table 12.10** Regression Coefficients for 1998 NAEP 4th Reading Achievement by Ethnicity

| | Model 6 |
|---|---|
| Intercept ($\gamma_{00}$) | 226.19(1.28)*** |
| Black ($\gamma_{10}$) | -22.50(2.09)*** |
| Hispanic ($\gamma_{20}$) | -21.20(2.02)*** |
| All other ($\gamma_{30}$) | -8.34(3.13) |
| Dfem ($\gamma_{40}$) | 5.36(1.29) |
| SES ($\gamma_{50}$) | -16.05(1.62)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 956.25 |
| LVL 2 Intercept ($\tau_{00}$)~ | 201.52*** |
| Black ($\tau_{10}$)~ | 124.27* |

*** $p \leq .001$, ** $p < .01$, * $p < .05$
~ $X^2$ tests from analysis using average of mrpcm1-5, no weights

The effect of Black varies significantly across teachers ($\tau_{10}$=63.36, p<.05) as does the effect of White ($\tau_{00}$=213.02, p<.000). Selected coefficients from this analysis serve as the base for the variance decompositions found in Tables 12.12-12.15. Next district SES is added to the model to determine its effect on the average achievement for Black and White students across teachers. The variable for student free and reduced lunch has been aggregated to the teacher level as control for classroom SES.

***Model 7.*** 4th grade *Black, White achievement and SES*

Level 1- Students - Same as unconditional Equation 12.10

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + r_{0j} \qquad (12.12)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, avSES + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement/achievement in teacher j's classroom holding all other predictors constant

$\gamma_{01j}$ = slope, the change in average achievement/achievement associated with mean free and reduced lunch, all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

$\gamma_{10j}$ = intercept, black student achievement/achievement in teacher j's classroom all other predictors held constant

$\gamma_{11j}$ = slope, the change in average black student achievement/achievement associated with one unit change in average free and reduced lunch

$r_{1j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for Model 7 are given in Table 12.11.

**Table 12.11** Regression Coefficients for Black/White Achievement Predicted by Average District SES

| | Model 7 |
|---|---|
| Intercept ($\gamma_{00}$) | 224.41(1.32)*** |
| avSES ($\gamma_{01}$) | -7.93(2.06)*** |
| Black ($\gamma_{10}$) | -21.67(2.08)*** |
| avSES ($\gamma_{11}$) | -2.09(3.42) |
| Hispanic ($\gamma_{20}$) | -20.06(2.05)*** |
| All other ($\gamma_{30}$) | -7.81(3.14)* |
| Dfem ($\gamma_{80}$) | -3.11(.98)*** |
| SES ($\gamma_{90}$) | -14.31(1.74)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 952.93 |
| LVL 2 Variance ($\tau_{00}$) | 198.92*** |
| ($\tau_{10}$) | 116.26* |

*** p<.001,** p<.01,* p<.05

How much variation does average SES explain in the average reading achievement for White and Black students across teachers? To answer this question the following equation computes the variance in black achievement explained by average SES. The estimated $\tau_{10}$ values from Tables 12.10 and 12.11 would be substituted in the following equation:

$$\frac{\tau_{10}(\text{model 6}) - \tau_{10}(\text{model 7})}{\tau_{10}(\text{model 6})} = \frac{63.36 - 55.93}{63.36} = .117 = 11.7\% \qquad (12.13)$$

This means that 11.7% of the variance in predicted average achievement for black students is explained by average SES across teachers (classrooms). Using the same process for the remaining tau, less than 1% of the variance in White average achievement is attributable to SES across teachers.

**Table 12.12** Variance Decomposition for Black/White Achievement by SES

| | | SES |
|---|---|---|
| Variables | Uncond | Free/Red Lunch |
| Intercept ($\gamma_{00}$) | 226.19(1.28)*** | 224.41(1.32)*** |
| avSES($\gamma_{02}$) | | -7.93(2.06)*** |
| Black ($\gamma_{10}$) | -22.50(2.09)*** | -21.67(2.08)*** |
| avSES($\gamma_{12}$) | | -2.09(3.42) |
| Intercept Variance ($\tau_{00}$) | 201.52*** | 198.92*** |
| Black Variance ($\tau_{10}$) | 124.27* | 116.26* |
| Level 1 variance ($\sigma 2$) | 956.25 | 952.93 |
| % variance reduction intercept | | 1.6% |
| % variance reduction black slope | | 6.4% |

A one standard deviation increase in average free and reduced lunch across teachers reduces average Black achievement by -5.12 points and decreases average White student achievement by -4.83 points. The variance decomposition from the table indicates that average SES explains 11.7% of the variance in average Black achievement across teachers and almost none of the variance in average White achievement. SES, therefore, has a differential effect on Black and White student (non-Black) achievement.

**Ethnicity, SES, and Teacher variables**

Now that the context is set for the analyses at Level Two, it is possible to track the variance explained by each teacher level variable controlling for SES and to determine its effect on average reading achievement for Black and White students across teachers. One variable at a time is added to the level two equation, EQ 12.14. The term "*variable x*" is exchanged for a single Reading Practice (read sensemaking activities, group and project activities, or group and project assessment activities), Teacher Characteristic (reading preparation, preparation for reading instruction, reading knowledge, reform orientation, diversity preparation), or Professional Development (professional development frequency in the last year).

***Model 8***. Template for Level 2 equations

Level 1 remains constant (Eq. 12.10)

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + \gamma_{02..}\, variable\ x + r_{0j} \qquad (12.14)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, avSES + \gamma_{12..}\, variable\ x + r_{4j}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{02..j}$ = change in average in average student achievement associated with a one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

and

$\gamma_{12..j}$ = the change in average black student achievement associated with one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

**Ethnicity and Teacher Practices**. A variance decomposition for teacher practices can be found in Table 12.13. The table indicates that no additional variance in average achievement for Black students is explained by teacher practices across teachers when SES is controlled. Small amounts are explained for White students.

**Table 12.13** Variance Decomposition for White/Black Controlling SES - Teacher Practices

| | | avSES | Practice | | |
|---|---|---|---|---|---|
| Variables | Uncond | Free/Red Lu | Reform activities | Traditional Activities | Computer Activities |
| Intercept ($\gamma_{00}$) | 226.19(1.28)*** | 224.41(1.32)*** | 224.39(1.34)*** | 233.87(1.21)*** | 233.94(1.21)*** |
| ($\gamma_{01}$) | | | .02(.14) | .46(.39) | 1.11(.33)** |
| avSES($\gamma_{02}$) | | -7.93(2.06)*** | -7.99(2.07)*** | -4.82(1.96)* | -4.93(1.97)* |
| Black ($\gamma_{10}$) | -22.50(2.09)*** | -21.67(2.08)*** | -21.75(2.08)*** | -18.42(1.86)*** | -18.68(1.90)*** |
| ($\gamma_{11}$) | | | .40(.28) | -.48(.59) | .03(.56) |
| avSES($\gamma_{12}$) | | -2.09(3.42) | 1.56(3.40) | -5.04(3.53) | -5.21(3.44)** |
| Intrcpt Var($\tau_{00}$) | 201.52*** | 198.92*** | 199.52*** | 211.15*** | 204.45*** |
| Black Var ($\tau_{40}$) | 124.27* | 116.26* | 112.15* | 57.61* | 60.71* |
| ($\sigma^2$) | 956.25 | 952.93 | 952.77 | 530.08 | 529.84 |
| % variance reduction intercept | | 1.6% | >1% | 0% | 0% |
| % var reduction black slope | | 6.4% | 9.8% | 53.6% | 49.9% |

For every one standard deviation increase in sensemaking and group and project assessment activities, White scores increase by up to .74 points while Black scores decrease by up to -.48 points. This is an indication that there may be differential teacher practice effects present in 1998 NAEP 4th grade reading data. For every one standard deviation increase in group and project activities, White scores increase by 1.11 points while Black scores remain nearly level. None of the variables explain more variation in average Black student scores across teachers than SES alone; however, the variables do explain small amounts of variance in White average achievement across teachers (up to 4%) after controlling for SES.

**Ethnicity and Teacher Characteristics.** The variance decomposition for teacher characteristics follows in Table 12.14. There are only two variables which explain more variance in achievement across teachers than SES alone - reading preparation and reform orientation. Reading preparation explains additional variation in achievement for both

Black and White Students across teachers while reform orientation only explains additional variance in achievement for Black students. There is twice as much variation in the average achievement for Black students across teachers when compared to White students. There is 5 times as much variance for black achievement across teachers explained by reform orientation. A more detailed explanation follows the table.

**Table 12.14** Variance Decomposition for Black/White Controlling for SES - Teacher Characteristics

| | | SES | Teacher Characteristics | |
|---|---|---|---|---|
| Variables | Uncond | Free/Red Lu | Reading Prep | Assess Prep |
| Intercept ($\gamma_{00}$) | 226.19(1.28)*** | 224.41(1.32)*** | 224.35(1.32)*** | 224.19(1.32)*** |
| ($\gamma_{01}$) | | | -.21(.12) | -.28(.17) |
| avSES($\gamma_{02}$) | | -7.93(2.06)*** | 7.76(2.06)** | 7.84(2.06)** |
| Black ($\gamma_{10}$) | -22.50(2.09)*** | -21.67(2.08)*** | 21.65(2.08)*** | 21.46(2.08)*** |
| ($\gamma_{11}$) | | | .14(.24) | .37(.30) |
| avSES($\gamma_{12}$) | | -2.09(3.42) | 1.89(3.43) | 2.01(3.45) |
| Intercept Var($\tau_{00}$) | 201.52*** | 198.92*** | 197.99*** | 197.83*** |
| Black Var ($\tau_{40}$) | 124.27* | 116.26* | 114.98* | 114.29* |
| ($\sigma^2$) | 956.25 | 952.93 | 952.70 | 952.71 |
| % variance reduction intercept | | 1.6% | 1.8% | 1.8% |
| % variance reduction black slope | | 6.4% | 7.5% | 8.0% |

Note that for every standard deviation increase in reading preparation (degrees/ licenses), White students scores increase by 1.49 points and Black scores drop by 1.59 points. The negative relationship between student achievement and teacher preparation persists for Black students and it accounts for 16% of the variation in average Black achievement across teachers in this data. White students gain points but teacher reading preparation only account for a little more than 1% of the variance in average White achievement across teacher - a differential effect. Across the US, low SES Black students are likely to have teachers with less reading preparation than low SES White students. It is puzzling that this is not the case in Connecticut where the explanation for the negative relationship between reading preparation and all student achievement controlling for SES appears related to the types of degrees that early childhood educators often receive.

Reform orientation controlling for SES across teachers explains 18.4% of the variation in Black average achievement over the course of one year while explaining little variance in the average achievement for White students. For every one standard deviation increase in teachers' reform orientation, black students are expected to lose an additional -1.81 points over the course of one year. White students are expected to gain 1.02 points but it explains little of the variation in average White student scores across teachers.

Teachers' read knowledge explains 13.8% of the variation in average Black achievement across teachers but nothing for average White achievement when SES is controlled. A one standard deviation increase in reading knowledge increases Black achievement by 2.81 points while leaving White achievement almost level.

SES continues to be the most important single predictor of Black achievement across teachers but SES explains little about white achievement. Three teacher characteristic variables increased the variance explained in average Black achievement across teachers once SES was controlled, read preparation, reform orientation, and reading knowledge. The effect of these variables was differential since they only explained variance in average Black achievement. Reform orientation and reading preparation reduces Black achievement while increasing White achievement. One exception is teachers' reading knowledge which increases average Black achievement while leaving average White achievement virtually unchanged..

**Ethnicity and Teacher Professional Development**. This analysis focuses on the differential effects of teacher professional development as measured by professional development impact, mentoring frequency, and professional development frequency. All professional development variables explain additional variance in reading achievement achievement across teachers for non-blacks after SES is controlled. Only professional development frequency explains additional variance for black achievement across teachers once SES is controlled. The variance decomposition for professional development follows in Table 8.16.

**Table 12.15** Variance Decomposition for Black/White Controlling for SES - Teacher Professional Development

| Variables | Uncond | SES | PD Frq |
|---|---|---|---|
| Intercept ($\gamma_{00}$) | 226.19(1.28)*** | 224.41(1.32)*** | 224.27(1.34)*** |
| ($\gamma_{01}$) | | | -.68(.86) |
| avSES($\gamma_{02}$) | | -7.93(2.06)*** | 7.93(2.07)*** |
| Black ($\gamma_{10}$) | -22.50(2.09)*** | -21.67(2.08)*** | -21.54(2.08)*** |
| ($\gamma_{11}$) | | | 3.38(1.58)* |
| avSES($\gamma_{12}$) | | -2.09(3.42) | -1.72(3.35) |
| Intercept Var($\tau_{00}$) | 201.52*** | 198.92*** | 200.72*** |
| Black Var ($\tau_{40}$) | 124.27* | 116.26* | 105.09* |
| ($\sigma^2$) | 956.25 | 952.93 | 951.69 |
| % variance reduction intercept | | 1.6% | >1% |
| % variance reduction black slope | | 6.4% | 15.4% |

For every standard deviation increase in professional development frequency, black students are predicted to gain about .17 points compared to White students who lose about half a point. Professional development frequrncy controlling for SES districts nothing more than SES alone for either group.

**<u>Summary of differential effects by ethnicity in 1998 NAEP 4th grade reading</u>**

- Average achievement for Black and White students varies across teachers for 1998 NAEP 4th grade reading data. SES, group and project assessment activities (teacher practice), reading preparation, reform orientation and reading knowledge (teacher characteristics) explain up to 18.4% of the variance in average Black achievement across teachers. Reading sensemaking activities, group and project activities, and group and project assessment activities (teacher practices), reading preparation and reform orientation (teacher charactersitics) explain very small amounts of the variation in average White student achievement (up to 4%). Professional development frequency does not explain variation for either group.

Of the variables, predominantly teacher practices, which account for variation in average White scores when SES is controlled, a one standard deviation increase in those variables increases White scores while simultaneously decreasing Black scores. They are examples of differential effects. Of the variables which account for variance in average Black achievement controlling for SES, predominantly teacher characteristics, a one standard deviation increase teachers read knowledge increases Black scores (2.81 points) while leaving White achievement level. The rest of the variables differentially decrease Black scores while increasing White scores by up to 1.8 points.

# 13. RESULTS OF THE HLM ANALYSIS FOR 2000 NAEP 4$^{TH}$ GRADE READING

The HLM analysis using 2000 NAEP 4$^{th}$ grade reading data is intended to address the research questions about the relationships between student achievement and

- Teacher characteristics and practices
- Characteristics of professional development

The analysis also is intended to address the ways in which these relationships vary by student demographic group, including gender, ethnicity, and SES across the United States and to see if the national reading results reflect trends similar to those found in the Connecticut BTPS data, a more detailed data set. Analyses concerning the relationships between teacher preparation institutions and teachers/students are not possible since the names of teacher preparation programs are not collected in NAEP. The HLM analyses are weighted (origwt) and use five plausible values (rrpcm1-5) for the student outcome.There are 4,548 students and 918 teachers in the 1998 reading sample.

To set the context for addressing the questions, results are presented first which describe the differences in student reading achievement (achievement status in NAEP data) across student demographic groups. The analyses yield results in terms of regression coefficients. To aid in understanding what the coefficients mean, predicted achievement (achievement status score) is computed for selected combinations of student characteristics.

The full set of reading variables for the NAEP analysis is given in Table 13.1.The outcomes and most teacher variables are the same as those used for the earlier SEM models. In certain cases, variables that were only available in this NAEP data set were included as well.

**Table 13.1** Description of Variables Used in the 2000 NAEP $4^{th}$ Grade Reading Analyses

| Type | Variable Abbreviation | Variable Description |
|---|---|---|
| | | **Level 1 - students** |
| Continuous (Outcomes) | rrpcm1 | $1^{st}$ plausible value for reading achievement (mean 212.98, sd=39.61) |
| | rrpcm2 | 2nd plausible value for reading achievement (mean 213.45, sd=39.19) |
| | rrpcm3 | 3rd plausible value for reading achievement (mean 213.45, sd=39.11) |
| | rrpcm4 | $1^{rth}$ plausible value for reading achievement (mean 213.16, sd=39.31) |
| | rrpcm5 | 5th plausible value for reading achievement (mean 213.53, sd=39.37) |
| Dichotomous | dblack | Black (Black=1, else=0) |
| | dhispanic | Hispanic (Hispanic=1, else=0) |
| | dallother | All other ethnicities (all other=1, else=0) |
| | dfem | Female (female=1, else=0) |
| | SES | Average free and reduced lunch (eligible=1, else=0) |
| Effects coded | egenxbl | Effect of Black female; interaction of eBlack x egender where Black =1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| | egenxhisp | Effect of Hispanic female; interaction of eHispanic x egender where Hispanic=1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| | egenxoth | Effect of All other female; interaction of eAllother x egender where All other =1, White=-1, otherwise=0 multiplied by female=1, male=-1. |
| Continuous | Origwt | Original weight (correction for oversampling) |
| | | **Level 2 - teachers** |
| Continuous | | Grand mean centered |
| Teacher Practices | refrmact | Frequency of reform activities (mean=15.03, sd=2.79) |
| | tradact | Frequency of traditional activities (mean=24.36, sd=2.32) |
| | compact | Frequency of computer activities (mean=7.31, sd=2.94) |

**Table 13.1** (cont'd).

| | | |
|---|---|---|
| Teacher Characteristics | readprep | Major/coursework/license in reading/English/writing, reading/English education (mean=.62, sd=1.27) |
| | techprep | Technology preparation (mean=7.85,sd=2.13) |
| | divprep | Preparation for diversity(mean=6.27, sd=1.68) |
| | assessprep | Preparation for assessment (mean=16.56, sd=3.20) |
| | readinst | Preparation for reading instruction, pedagogical coursework (mean=21.12, sd=3.01) |
| Professional Development | pdf | Professional development in the last year (mean=, sd=) |
| Continuous | avSES | Average percent free and reduced lunch per teacher/ classroom aggregated from level 1 (mean=.45, sd=.38) |

**I. Setting the Stage: Variation in NAEP 2000 Reading Achievement Status by Student DemographicGroups and SES** (Results at the Student Level)

In this section, sets of background variables are added to look at changes in reading gains in the United States. Model 1 predicts mean student gains across the sample and partitions the variance between students and teachers where students are nested within teachers. Models 2-5 predict differences in mean gains across gender, ethnicity, and SES (eligibility for free and reduced lunch).

Mean Gains. To set the context for answering the questions about connections between student achievement and teacher variables, results are presented first for differences in student achievement across student demographic groups. The level one and level two equations for the model are given followed by their estimated regression coefficients in Tables 8.2.

***Model 1.** Mean achievement*

Level 1 - Students

$$Y_{ij} = \beta_{0j} + e_{ij} \qquad (13.1)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

where $Y_{ij}$ is the reading achievement for student i in teacher j's classroom, $\beta_{0j}$ is the intercept or mean students' reading achievement within teacher j's classroom, $e_{ij}$ is the level 1 error, $\gamma_{00}$ is the mean value of classroom achievement for all teachers, and $r_{0j}$ is the level 2 error.

**Table 13.2** 2000 NAEP 4th Grade Average Reading Achievement

| | Model 1 |
|---|---|
| Intercept ($\gamma_{00}$) | 214.66(1.07)*** |
| LVL 1 Variance ($\sigma^2$) | 1183.40 |
| LVL 2 Variance ($\tau_{00}$) | 479.40*** |

*** p<.001,** p<.01,* p<.05

The predicted average achievement score for 4th grade reading students for this sample of US students is 214.66 points ($\gamma_{00}$=214.66, sd=1.07) when no other variables are included in the model.

$\sigma^2$ and $\tau_{00}$ are used to compare variation in the errors at the teacher level with variation at the student level using the formulas:

$$\frac{\sigma^2}{\sigma^2 + \tau_{00}} = \textit{level one variance} \qquad \frac{\tau^2{}_{00}}{\sigma^2 + \tau_{00}} = \textit{level two variance} \qquad (13.2)$$

A little over 70% of the variance in 2000 reading achievement for Model 1 is within teachers (between students) leaving a little less than 30% of the variance in achievement between teachers. That is to say, there is 2.5 times more variation within teachers than between teachers, very similar to proportion found in 1998 reading.

Mean Achievement by Demographic Group. Next, demographic variables are added to form Models 2, 3 and 4. Model 5 introduces the control variable for SES. The equations for Model 2, ethnic group, are presented first (Eq.13.3) To interpret the coefficients, the estimated score for ethnicity can be computed using the appropriate set of 1s and 0s for the dummy variables. As an example, Equations 13.4 and 13.5 demonstrate the computation for obtaining Black student achievement. Estimated achievement by ethnicity are presented in Table 13.3.[39]

***Model 2.*** *Ethnic groups*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij}\ e_{ij};$$

Level 2 - Teachers (13.3)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

The regression coefficients for predicting student achievement from ethnic background are found in Table 13.3.

[39] Note that the category for All other includes Asian and those students who did not report ethnicity. Often these smaller groups are folded into the intercept since they are not significantly different from the intercept. Additionally, it was important to maintain the meaning of the intercept as a specific group, in this case white students, before analyzing the differential effects of teacher variables.

**Table 13.3** Regression Coefficients for 2000 NAEP 4th Grade Reading by Ethnicity

| | Model 2 |
|---|---|
| Intercept ($\gamma_{00}$) | 222.77(1.07)*** |
| Black ($\gamma_{01}$) | -25.01(2.20)*** |
| Hispanic ($\gamma_{02}$) | -23.06(1.81)*** |
| All other ($\gamma_{03}$) | -5.66(3.64) |
| | |
| LVL 1 Variance ($\sigma^2$) | 1163.26 |
| LVL 2 Variance ($\tau_{00}$) | 311.95*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by ethnicity group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[40] For example, to determine the predicted achievement for a black student who would have ones for the intercept and black the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} \tag{13.4}$$

Level 2 - Teachers

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 222.77$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -25.01$$

When the slope values, the betas from Equation 8.4, are substituted in the level one equation (Eq. 8.4), the result for black achievement is produced (Eq. 8.5).

$$\hat{Y}_{ij} = (222.77) + (-25.01)*1 \tag{13.5}$$

[40] A table showing ones and zeros for all dummy variables is provided in the Appendix.

$= 197.76$

Predicted reading achievement for all demographic groups are presented in Table 13.4. This table shows that the model predicts that White students score higher in reading than the other ethnic groups. The coefficients for the intercept, Black, and Hispanic are significantly different from zero. Hispanic and All other students, groups most likely to contain English as a second language students, are expected to outperform Black students on NAEP 2000 reading.

**Table 13.4** Predicted Reading Achievement Student Demographic Group

| Ethnic Group | Achievement |
|---|---|
| White | 222.77 |
| Black | 197.76 |
| Hispanic | 199.71 |
| All other | 217.11 |

In the next model, gender predicts average reading achievement.

***Model 3.*** *Gender*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{4j}\, dfem_{ij} + e_{ij};$$

Level 2 - Teachers (13.6)

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{4j} = \gamma_{10}$$

The regression coefficients for predicting student achievement from gender are found in Table 13.5.

**Table 13.5** Regression Coefficients for 2000 NAEP Reading Achievement by Gender

| | Model 3 |
|---|---|
| Intercept ($\gamma_{00}$) | 209.38(1.37)*** |
| Female ($\gamma_{04}$) | 10.34(1.62)*** |
| | |
| LVL 1 Variance ($\sigma^2$ ) | 1155.36 |
| LVL 2 Variance ($\tau_{00}$) | 479.40*** |

*** p<.001,** p<.01,* p<.05

These coefficients indicate that male mean achievement in reading is predicted to be 209.38 points compared to 219.72 points for females. The coefficient for female is significantly different from zero. On average, girls are expected to outscore boys by more than 10 points.

***Model 4****. Combined gender and ethnicity*

Model 4 tests whether or not there is an interaction between gender and ethnicity in the NAEP 4th grade reading data. Equation 13.7 computes coefficients for combinations of gender and ethnicity.

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij}$$
$$+ \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij} + e_{ij}; \qquad (13.7)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background and gender are found in Table 13.6.

**Table 13.6** Regression Coefficients for 2000 NAEP 4th Grade Reading Achievement by Gender and Ethnicity

| | Model 4 |
|---|---|
| Intercept ($\gamma_{00}$) | 216.29(1.41)*** |
| Black ($\gamma_{01}$) | -25.39(2.19)*** |
| Hispanic ($\gamma_{02}$) | -22.82(1.80)*** |
| All other ($\gamma_{03}$) | -5.08(3.64) |
| Female ($\gamma_{04}$) | 12.71(1.86)*** |
| Effect Black x female ($\gamma_{05}$) | .74(1.58) |
| Effect Hispanic x female ($\gamma_{06}$) | -1.05(1.31) |
| Effect All other x female ($\gamma_{07}$) | 2.09(2.34) |
| | |
| LVL 1 Variance ($\sigma^2$) | 1134.42 |
| LVL 2 Variance ($\tau_{00}$) | 310.92*** |

*** p<.001,** p<.01,* p<.05

To make it easier to interpret the results from the achievement by demographic group analysis, these coefficients can be used to compute the achievement for a student with a given set of demographic characteristics. Because the demographic characteristics are coded as 1 or 0, the computation involves summing the terms which are coded 1 for a particular student.[41] For example, to determine the predicted achievement for a black female student who would have ones for the intercept, black, female and black female,

[41] A table showing ones and zeros for all dummy variables is provided in the Appendix.

the resulting equation would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} eblackxefem_{ij}$$

Level 2 - Teachers (13.8)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 216.29$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -25.39$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = 12.71$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = .75$$

When the slope values, the level two betas from Equation 13.8, are substituted in the level one equation (Eq. 13.8), the result for black female achievement is produced (Eq. 13.9).

$$\hat{Y}_{ij} = (216.29) + (-25.39)*1 + (12.71)*1 + (.75)*1 \quad (13.9)$$
$$= 204.36$$

Predicted reading achievement for all subsets of students by demographic groups are presented in Table 13.7. The table shows that the model predicts female students from all groups are predicted to have higher average achievement than their male counterparts.

**Table 13.7** Predicted Reading Achievement by Student Demographic Group with Differences by Gender

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 216.29 | 229.00 | + 12.71 |
| Black | 190.90 | 204.36 | + 13.46 |
| Hispanic | 193.47 | 205.13 | + 11.66 |
| All other | 211.21 | 226.01 | + 14.80 |

Score differences between males and females of the same ethnicity are predicted to range from 11.66 to 13.46 points. Coefficients for gender times ethnicity effects are not significantly different from zero over the course of one year. Black average achievement is less than the average reading achievement for other groups for both boys and girls.

***Gender, Ethnicity, and SES***

Another known source of variation in student achievement is SES as measured by eligibility for free and reduced lunch. In NAEP, this variable is measured at the student level.

***Model 5.* Gender, Ethnicity, and Grade Level**

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} \quad (13.10)$$
$$+ \beta_{5j} egenxbl_{ij} + \beta_{6j} egenxhisp_{ij} + \beta_{7j} egenxoth_{ij}$$
$$+ \beta_{8j} SES_{ij} + e_{ij};$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$
$$\beta_{4j} = \gamma_{40}$$
$$\beta_{5j} = \gamma_{50}$$
$$\beta_{6j} = \gamma_{60}$$
$$\beta_{7j} = \gamma_{70}$$
$$\beta_{8j} = \gamma_{70}$$

The regression coefficients for predicting student achievement from ethnic background, gender, and SES are found in Table 13.8.

**Table 13.8** Regression Coefficients for 2000 NAEP 4th Grade Reading Achievement by Gender, Ethnicity, and SES

| | Model 5 |
|---|---|
| Intercept ($\gamma_{00}$) | 221.35(1.41)*** |
| Black ($\gamma_{01}$) | -17.50(1.80)*** |
| Hispanic ($\gamma_{02}$) | -17.47(1.57)*** |
| All other ($\gamma_{03}$) | -2.72(3.55) |
| Female ($\gamma_{04}$) | 13.48(1.79)*** |
| Effect Black x female ($\gamma_{05}$) | .99(1.58) |
| Effect Hispanic x female ($\gamma_{06}$) | -1.43(1.27) |
| Effect All other x female ($\gamma_{07}$) | 2.17(2.29) |
| SES ($\gamma_{08}$) | -20.43(1.66)*** |
| LVL 1 Variance ($\sigma^2$) | 238.75*** |
| LVL 2 Variance ($\tau_{00}$) | 1099.54 |

*** p<.001,** p<.01,* p<.05

As demonstrated in prior sections, the computational example to determine the predicted achievement for a black female student eligible for free and reduced lunch who would have ones for the intercept, black, female, black female, and free and reduced lunch, would be:

Level 1 - Students

$$\hat{Y}_{ij} = \hat{\beta}_{0j} + \hat{B}_{1j} dblack_{ij} + \hat{B}_{4j} dfem_{ij} + \hat{B}_{5j} egenxbl_{ij} + \hat{B}_{8j} SES_{ij}$$

Level 2 - Teachers (13.11)

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} = 221.35$$
$$\hat{\beta}_{1j} = \hat{\gamma}_{10} = -17.50$$
$$\hat{\beta}_{4j} = \hat{\gamma}_{40} = 13.48$$
$$\hat{\beta}_{5j} = \hat{\gamma}_{50} = .99$$
$$\hat{\beta}_{8j} = \hat{\gamma}_{80} = -20.43$$

When the slope values, the betas from Level 2, Equation 8.11, are substituted in the Level 1 equation, the result for black female achievement is produced (Eq. 8.12).

$$\hat{Y}_{ij} = (221.35) + (-17.50)*1 + (13.48)*1 + (.99)*1 + (-20.43*1) \quad (13.12)$$
$$= 197.89$$

Predicted reading achievement for all subsets of students by demographic group and grade are presented in Table 13.9. The estimated point loss associated with free and reduced lunch eligibility is -20.43 points.

**Table 13.9** Predicted Reading Achievement by Student Demographic Group and SES

| | Males | Females | Difference in Achievement by Gender |
|---|---|---|---|
| White | 200.92 | 214.40 | + 13.48 |
| Black | 183.42 | 197.89 | + 14.47 |
| Hispanic | 183.45 | 195.50 | + 12.05 |
| All other | 198.20 | 213.85 | + 15.65 |

This table indicates that females score higher than their male counterparts by more than 12 points. Average Black and Hispanic reading achievement controlling for SES is expected to be similar, around 183 points for males and about 197 points for females. The same is true for average White and All other achievement, around 200 points for males and 214 points for females.

*Summary of Section 1.* Developing the level one equation for students. As the level one base equation for 2000 NAEP 4th grade reading developed, average yearly achievement was first predicted with a simple unconditional model. This model indicated that achievement variance was 2.5 times larger within teachers (between students) than across teachers at level two. Variables known to affect mean achievement (ethnicity, gender, ethnicity x gender, and SES) were added to the level one equation revealing that while differences in achievement by gender and ethnicity exist within teachers, the interaction terms for gender time ethnicity were not significant predictors of student achievement.

**II. Differential Effects of Teacher Characteristics, and Teacher Practices on Student Achievement** (Results at the teacher level)

This section addresses the questions:

1. Does reading achievement for demographic groups vary across teachers?
2. How much variation in student reading achievement do variables about teacher practices, characteristics, and professional development explain?
3. Is the amount of variance explained by teacher practices, characteristics, and professional development different for different student demographic groups?

The analyses in this section first test if average achievement for the various demographic groups varies across teachers. Nonsignificant error terms for variance ($r_{00}$, $r_{10,}$ ...$r_{x0}$) indicate that group mean achievement does not vary across teachers so the error terms predicting mean achievement for those demographic groups are dropped from the model. Although the error terms are dropped, the fixed level two equations for non-significant achievement by demographic groups are retained in the analysis.

Average achievement did not vary across teachers for any of the demographic groups in NAEP 2000 4th grade reading. Consequently, there are no differential effects of teacher practices evident in this data. It is possible, however, to look at the amounts of variation that teacher practice variables explain for average achievement for all students across teachers.

**Effects of teacher level variables on mean achievement**

The unconditional Model 6 establishes the maximum amount of variance that can be explained by teacher level variables for this data. Note that the interaction terms for gender times ethnicity have been dropped as they were in preceding sections about the effects of teacher practices, characteristics, and professional development.

***Model 6.*** *Unconditional model for 4th grade achievement*

Level 1- Students

$$Y_{ij} = \beta_{0j} + \beta_{1j} dblack_{ij} + \beta_{2j} dhispanic_{ij} + \beta_{3j} dallother_{ij} + \beta_{4j}\, dfem_{ij} + \beta_{5j} SES_{ij} + e_{ij} \quad (13.10)$$

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + r_{0j} \quad (13.11)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{70}$$

$$\beta_{5j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average in student achievement/achievement in teacher j's classroom all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

The estimated level two regression coefficients for Model 6 are given in Table 13.10.

**Table 13.10** Regression Coefficients for 2000 NAEP 4th Grade Reading Achievement by Ethnicity

| | Model 6 |
|---|---|
| Intercept ($\gamma_{00}$) | 222.45(1.28)*** |
| Black ($\gamma_{10}$) | -17.35(2.30)*** |
| Hispanic ($\gamma_{20}$) | -17.46(1.85)*** |
| All other ($\gamma_{30}$) | -2.94(3.58) |
| Dfem ($\gamma_{40}$) | 11.25(1.54)*** |
| SES ($\gamma_{50}$) | -20.40(1.67)*** |
| | |
| LVL 1 Variance ($\sigma 2$) | 1100.32 |
| LVL 2 Intercept ($\tau_{00}$)~ | 238.75*** |

*** p<.001,** p<.01,* p<.05

Selected coefficients from this analysis serve as the base for the variance decompositions found in Tables 13.12-13.15. Next district SES is added to the model to determine its effect on the average achievement for Black and White students across teachers. The variable for student free and reduced lunch has been aggregated to the teacher level as control for classroom SES.

***Model 7.*** 4th grade *average achievement and SES*

Level 1- Students - Same as unconditional Equation 13.10

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + r_{0j} \tag{13.12}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{00j}$ = intercept, average student achievement across teachers holding all other predictors constant

$\gamma_{01j}$ = change in average achievement associated with a one standard deviation increase in mean free and reduced lunch all other predictors held constant

$r_{0j}$ = the error term or unique teacher effects

The estimated level two regression coefficients and variance components for Model 7 are given in Table 13.11.

**Table 13.11** Regression Coefficients Mean Achievement Predicted by Average District SES

| | Model 7 |
|---|---|
| Intercept ($\gamma_{00}$) | 218.44(1.46)*** |
| avSES ($\gamma_{01}$) | -18.25(3.29)*** |
| Black ($\gamma_{10}$) | -14.85(2.40)*** |
| Hispanic ($\gamma_{20}$) | -15.96(1.88)*** |
| All other ($\gamma_{30}$) | -2.65(3.55) |
| Dfem ($\gamma_{80}$) | 11.08(1.53)*** |
| SES ($\gamma_{90}$) | -15.21(1.95)*** |
| | |
| LVL 1 Variance ($\sigma^2$) | 1093.81 |
| LVL 2 Variance ($\tau_{00}$)~ | 225.51*** |

*** p<.001** p<.01,* p<.05

How much variation does average SES explain in the average reading achievement across teachers? To answer this question the following equation computes the variance in mean achievement explained by average SES. The estimated $\tau_{00}$ values from Tables 13.10 and 13.11 would be substituted in the following equation:

$$\frac{\tau_{00}(\text{model 6}) - \tau_{00}(\text{model 7})}{\tau_{00}(\text{model 6})} = \frac{238.76 - 225.51}{238.75} = .055 = 5.5\% \qquad (13.13)$$

This means that 5.5% of the variance in predicted average achievement is explained by average SES across teachers (classrooms).

**Table 13.12** Variance Decomposition for Mean Achievement by SES

| | | SES |
|---|---|---|
| Variables | Uncond | Free/Red Lunch |
| Intercept ($\gamma_{00}$) | 222.45(1.28)*** | 218.44(1.46)*** |
| avSES($\gamma_{02}$) | | -18.25(3.29)*** |
| Intercept Variance ($\tau_{00}$) | 238.75*** | 225.51*** |
| Level 1 variance ($\sigma 2$) | 1100.32 | 1093.81 |
| % variance reduction intercept | | 5.5% |

A one standard deviation increase in average free and reduced lunch across teachers reduces average achievement across teachers by -18.25 points.

**Teacher Effects on Average Achievement**

Now that the context is set for the analyses at Level Two, it is possible to track the variance explained by each teacher level variable controlling for SES and to determine its effect on average reading achievement across teachers. One variable at a time is added to the level two equation, EQ 13.14. The term "*variable x*" is exchanged for a single Reading Practice (reform activities, traditional activities, and computer activities), Teacher Characteristic (reading preparation, preparation for reading instruction, technology preparation, assessment preparation, and diversity preparation), or

Professional Development (professional development frequency in the last year).

**Model 8**. Template for Level 2 equations

Level 1 remains constant (Eq. 13.10)

Level 2 - Teachers

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, avSES + \gamma_{02..}\, variable\ x + r_{0j} \qquad (13.14)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{8j} = \gamma_{70}$$

$$\beta_{9j} = \gamma_{70}$$

where

$\gamma_{02..j}$ = change in average in average student achievement associated with a one unit change in "*variable x*" holding SES constant at the classroom (teacher) average

The analysis process followed the same pattern for all variables as that presented in prior sections. One teacher practice variable, computer activities, and one teacher characteristic, preparation for reading instruction, explained variation across teachers in excess of that explained by SES. Results are presented in Table 13.13.

**Table 13.13** Teacher Effects on Mean Achievement Controlling SES

| | | avSES | Practice | Characteristic |
|---|---|---|---|---|
| Variables | Uncond | Free/Red Lu | Computer activities | Reading Inst |
| Intercept ($\gamma_{00}$) | 222.45(1.28)*** | 218.44(1.46)*** | 218.39(1.46)*** | 218.49(1.45)*** |
| ($\gamma_{01}$) | | | -.38(.28) | .74(.27)** |
| avSES($\gamma_{02}$) | | -18.25(3.29)*** | -18.03(3.29)*** | -18.01(3.29)*** |
| Intercept Var($\tau_{00}$) | 238.75*** | 225.51*** | 224.54*** | 224.07*** |
| ($\sigma^2$) | 1100.32 | 1093.81 | 1093.90 | 1092.27 |
| % variance reduction intercept | | 5.5% | 5.9% | 6.1% |

For every one standard deviation increase in computer activities, average achievement is expected to decrease by -.38 points. Computer activities explain 5.9% of the variance in average achievement controlling for SES. Preparation for reading instruction explains 6.1% of the variation average achievement across teachers. A one standard deviation increase in preparation for reading instruction increase average achievement by .74 points.

## 14. COMPARISON OF HLM RESULTS FOR MATHEMATICS AND READING ACROSS DATA SETS

The HLM analyses addressed questions about the relationships among student achievement and

- Teacher characteristics and practices
- Characteristics of professional development
- Characteristics of teacher preparation

The clearest way to summarize the HLM results across data sets is to use a series of tables. Table 14.1 summarizes the demographic distribution of students by gender, ethnicity, SES, student and teacher n, as well as the variation present at both student and teacher levels for all data sets.

**Table 14.1** Comparison of Variation Across Data Sets

| | **Mathematics** | | | **Reading** | | |
|---|---|---|---|---|---|---|
| Demographics | CT | NAEP 1996 | NAEP 2000 | CT | NAEP 1998 | NAEP 2000 |
| White | 68% | 59% | 54% | 67% | 55% | 54% |
| Black | 14% | 18% | 18% | 16% | 19% | 19% |
| Hispanic | 9% | 16% | 21% | 9% | 21% | 22% |
| Other | 9% | 7% | 7% | 8% | 5% | 6% |
| Male | 48% | 49% | 50% | 47% | 50% | 51% |
| Female | 52% | 51% | 50% | 53% | 50% | 49% |
| Free/Red Lunch | 20% | 36% | 43% | 22% | 43% | 42% |
| Student n | 1550 | 4409 | 6205 | 1776 | 4303 | 4548 |
| Teacher n | 81 | 584 | 1245 | 90 | 828 | 918 |
| Variation at Lvl 1 | 91% | 60% | 65% | 90% | 70% | 70% |
| Variation at Lvl 2 | 9% | 40% | 35% | 10% | 30% | 30% |

Connecticut has fewer minorities and fewer free and reduced lunch eligible students than the national sample. A large proportion of the variation in student achievement in Connecticut is between students within classrooms - about 90%. This percentage captures variation specific to each student in a particular classroom such as individual variation due to parental education, home environment, peer influence, etc. At level two, about 10% of the variation in student scores is attributable to teachers. It is this variation that can be impacted systematically by state policy, teacher education, and professional development.

In NAEP, there are more minority students than in the Connecticut data. It is interesting to note that by data collection in 2000 the national Hispanic student population had grown larger than the Black population, which remained relatively stable across administrations. Less variation is located at the student level and more initial variation across teachers, over 20% more, than found in Connecticut.

SES, as measured by free and reduced lunch, plays a large role in student achievement by demographic group. Although SES was not a focus of this study, it is important to acknowledge its relationship to student achievement and control for it. NAEP samples have higher proportions of free and reduced lunch students than the Connecticut samples.

Student gain scores in mathematics and reading in Connecticut show a different picture of achievement than do achievement status scores in NAEP. In Connecticut, minority and female students generally gain more over the course of the year than do White and male students. As in NAEP, White and male students tend to score higher overall than minority and female students. This is to say that minorities and females have not yet achieved the same level of academic attainment as their White and male counterparts although they gain more over the course of a single year.

To tackle questions about the relationships between student achievement and teacher practices, characteristics, and professional development, it is first necessary to determine which group means vary across teachers. If group means do not vary across teachers, it is only possible to talk about the relationships of average achievement to teacher practices, characteristics, and professional development and the variation attributable to teachers for all students. If group means vary by gender and/or ethnicity across teachers, then differential relationships are present depending on group membership. Table 14.2 indicates the group means which varied significantly across teachers (marked with an x).

**Table 14.2** Group Means Which Varied Across Teachers (x) for All Data Sets

| | **Mathematics** | | | **Reading** | | |
|---|---|---|---|---|---|---|
| Demographics | CT | NAEP 1996 | NAEP 2000 | CT | NAEP 1998 | NAEP 2000 |
| White (Intercept) | x | x | x | x | x | x |
| Black | x | x | | x | x | |
| Hispanic | | | x | | | |
| Other | | | | | | |
| Male (Intercept) | x | x | | | | |
| Female | x | x | | | | |
| Black female[42] | | x | | | | |
| Student n | 1550 | 4409 | 6205 | 1776 | 4303 | 4548 |
| Teacher n | 81 | 584 | 1245 | 90 | 828 | 918 |

Analyses using average gain scores by gender and ethnicity in Connecticut were more sensitive to variation across teachers than achievement status in NAEP. Teacher

[42] It's difficult to interpret three way interactions among black x female x teacher variables.

mathematics variables were more likely to be associated with mathematics achievement than teacher reading variables were likely to be associated with reading achievement across data sets. Group means for at least one group varied across teachers in all data sets except NAEP 2000 Reading.

Black mean achievement varied across teachers in four of the six data sets, two for mathematics (CT mathematics, NAEP 1996) and two for reading (CT reading, NAEP 1998). Group means for Black (CT Reading) varied across teachers but teacher variables did not account for any of the variation. Female mean mathematics achievement varied across teachers in two data sets. Hispanic mean mathematics (NAEP 2000) and Black female mean mathematics achievement (NAEP 1996) varied across teachers in one data set each.

Table 14.3 summarizes the variance explained by teacher variables on mean achievement by ethnicity and Table 14.4 presents the variance explained by teacher variables on Male/Female achievement. Only one teacher variable at a time controlling for SES was entered into an analysis.

**Ethnicity**

**Table 14.3** Variance in Mean Achievement by Ethnicity Explained by Teacher Variables Across All Data Sets

| | **Mathematics** | | | **Reading** | | |
|---|---|---|---|---|---|---|
| Demographics | CT | NAEP 1996 | NAEP 2000 | CT | NAEP 1998 | NAEP 2000 |
| Intercept | White | White | White | All** | White | All |
| SES | 2.2% | <1% | 1.3% | 0% | 1.6% | 5.5% |
| Practice-Sensemaking | 0%* | 2.2% | 1.3%* | x | x | x |
| GPA | x | 4.4% | 2.5% | x | x | x |

**Table 14.3** (cont'd).

| | | | | | | |
|---|---|---|---|---|---|---|
| Reform activities | x | x | x | x | <1%* | x |
| Traditional activities | x | x | x | x | 0%* | x |
| Computer activities | x | x | x | 8.3% | 0%* | 5.9% |
| Characteristics-Math prep | 25.6% | 1.3% | x | 7.3% | X | x |
| Math knowledge | x | 0% | x | x | x | x |
| Reform orientation | 0%* | 4.2% | x | x | x | x |
| Read Prep | x | x | x | x | 1.8% | 6.1% |
| Assess Prep | x | x | x | x | 1.8% | x |
| Professional Dev- PD Impact | 9.9% | x | x | x | x | x |
| Ment Frq | 17.5% | x | x | x | x | x |
| PD frq | 44.4% | x | x | 4.3% | <1%* | x |
| Group | Black | Black | Hispanic | All | Black | All |
| SES | 45.9% | 15.1% | 5.9% | x | 6.4% | X |
| Practice-Sensemaking | 47.5% | 10.1%* | 4.5%* | x | x | x |
| GPA | x | 6.2%* | 2.5%* | x | x | x |
| Reform activities | x | x | x | x | 9.8% | x |
| Traditional activities | x | x | x | x | 53.6% | x |
| Computer activities | x | x | x | x | 49.9% | x |
| Characteristics-Math prep | 57.9% | 17.7% | x | x | x | x |
| Math knowledge | x | 17.8% | x | x | x | x |
| Reform orientation | 55.5% | 6.9%* | x | x | x | x |
| Read Prep | x | x | x | x | 7.5% | x |
| Assess Prep | x | x | x | x | 8.0% | x |
| Professional Dev- PD Impact | 43%* | x | x | x | x | x |
| Ment Frq | 43.7%* | x | x | x | x | x |
| PD frq | 52.5% | x | x | x | 15.4% | x |

* Less than SES alone. Likely due to multicollinearity.

** Black mean achievement varied but selected teacher variables did not explain variation. The variance decomposition is presented for the intercept only (all students).

Different proportions of variance in average White/Black achievement are explained across teachers in four data sets. Since teacher variables did not explain variation in Black achievement across teachers in Connecticut Reading, the results are not included here Results for the HLM analyses across data sets are summarized below.

**Controlling for SES**. SES explains a large proportion of the variation in average Black mathematics achievement (15%-45%) across teachers in CT and NAEP 1996 but only a small proportion of the variance in average White achievement (1%-2%). Low SES has a large, negative effect on average Black achievement in Connecticut (-19 points in average SES districts). NAEP 1996 shows a loss of about -5 points for both groups in average SES classrooms.

While SES plays a part in NAEP 2000 White/Hispanic mathematics achievement and NAEP1998 Black/White reading achievement, it only explains 1-2% of the variance in White scores and approximately 6% of the variance in Black and Hispanic scores across teachers. Average White scores decrease as a function of increased free and reduced lunch but Hispanic students increase average mathematic achievement by +3.5 points in NAEP 2000. When the effect of teachers on Hispanic achievement is compared to the effect on White achievement, it is important to realize that the intercept also includes all non-Hispanic groups. No teacher level variables explain more variation in average Hispanic mathematics achievement across teachers than SES alone.

**Relationships between teacher practice and Black/White achievement.** Positive effects of teacher classroom practices on average White/Black student achievement controlling for SES are evident when teachers use math sensemaking activities in CT. Teacher use of math sensemaking activities explains none of the variance in White average achievement across teachers in CT but it explains 47.5% of the variation in Black average achievement. For every one standard deviation increase in sensemaking activities, Black achievement increases by +1.25 points. White achievement increases by +.18 points. In NAEP 1996, less variance is explained across teachers but average White achievement increases by +.72 points while Black average achievement decreases by -.11 points after controlling for SES.

Teacher practice variables explain small amounts of variance in average White reading achievement across teachers in NAEP 1998. Average Black reading achievement varies from 10%-50% across teachers controlling for SES depending on the specific teacher practice. A one standard deviation increase in reform activities leaves average achievement for White students virtually unchanged but increases Black achievement by +.4 points. A one standard deviation increase in traditional activities increases average White achievement by about half a point while simultaneously decreasing Black average achievement by half a point. No similar effects were detected in CT reading or NAEP Reading 2000. A one standard deviation increase in computer activities increases average White reading achievement by over one point while leaving Black average achievement unchanged..

**Relationships between teacher characteristics and Black/White achievement.** Teacher preparation in mathematics, i.e. major/credentials, explains 25.6% of the variation in average White gains and 57.9% of the variation in average Black gains controlling for SES in CT. A one standard deviation increase in teacher mathematics preparation signals a decrease in average White gains of -1 point per year and a decrease in average Black gains of -2 points per year. A slightly different phenomenon occurs in NAEP mathematics 1996. Teacher preparation in mathematics explains 1.3% of the variation in White average achievement and 17.7% of the variation in average Black achievement controlling for SES. For every one standard deviation increase in teacher preparation, White scores increase by +1.5 points while Black average achievement decreases by -1.4 points. The counterintuitive relationship between mathematics achievement and teacher preparation has been noted by Valli and Reckase (2000) and Rowan, Correnti, and Miller (2004). Possible explanations include that elementary mathematics teachers often major and get licenses in subjects other than mathematics.

Reform orientation in mathematics explains less than 4% of the variation in White achievement across teachers controlling for SES while it explains 55% of the variation in average Black achievement in CT and 7% of the variation in NAEP mathematics 1996. A one standard deviation increase in teacher reform orientation raises Black gains in CT by +1.33 points and raises achievement in NAEP 1998 by +.28 points. White achievement stays nearly level in CT and increases by +1.3 points in NAEP 1996.

Teacher reading preparation and assessment preparation explain approximately 8% of the variation in Black average reading achievement in NAEP 1998 and less than 2% for White average achievement. Increases in reading preparation and assessment preparation increase Black average achievement by nearly +.3 points while decreasing White average achievement by similar amounts.

**Relationship between teacher professional development and Black/White achievement.** Teacher professional development variables do not explain variation in Black/White average mathematics achievement at the national level but they do explain variance in CT. The variable that is most consistent across data sets and that has the largest effect on Black/White achievement is professional development frequency. It explains 44.4% of the variance in average White gains and 52.5% of the variance in average Black gains across teachers once SES is controlled. A one standard deviation increase in professional development frequency has a differential effect on students. Average White achievement decreases by -1 point while average Black mathematics achievement increases by nearly +1 point.

Teacher professional development frequency explains 15.4% of the variance in average Black reading achievement across teachers in NAEP 1998. A one standard deviation increase in professional development frequency increases Black average achievement by +3.38 points while it decreases White achievement by -.68 points.

**Gender**

The variance explained results for Connecticut and NAEP 1996 mathematics for male and female are presented in Table 14.4. Teacher effects did not vary by gender for reading achievement or NAEP Mathematics 2000. Average achievement for Black female in NAEP 1996 mathematics varies across teachers but the three way interactions among black, female, and teacher variables are difficult to interpret since the main groups for Black and female also vary across teachers. Results for Black female are not discussed again here.

**Table 14.4** Variance in Mean Achievement by Gender Explained by Teacher Variables

| | Mathematics | | | |
|---|---|---|---|---|
| Demographics | CT | | NAEP 1996 | |
| Intercept | Male | Female | Male | Female |
| SES | 6.1% | 23.2% | <1% | 0% |
| Practice-Sensemaking | 8.5% | 29.3% | 3.7% | <1% |
| GPA | 8.6% | 37.5% | 5.4% | 0%* |
| GPA assessment | x | x | 3.5% | <1% |
| DLA | 5.7%* | 33.8% | x | x |
| Characteristics-Math prep | 5.4%* | 13.6%* | x | x |
| Math knowledge | 2.8%* | 22%* | x | x |
| Reform orientation | 6.4% | 23.6% | 5.9% | 2.1% |
| Assess Prep | 6.6% | 30.1% | x | x |
| Efficacy | 7.9% | 41.7% | x | x |
| Professional Dev- PD Impact | 5.0%* | 15.5%* | x | x |
| Ment Frq | 1.8%* | 27.8% | x | x |
| PD frq | 17.6% | 15.5% | x | x |

* Less than SES alone. Likely due to multicollinearity.

**Controlling for SES.** SES explains 23.2% of the variation in average female mathematics achievement across teachers in CT but none of the variation in NAEP mathematics. For every one standard deviation increase in district free and reduced lunch in CT, boys lose more than -6 points. Girls gain them back (+8 points).

Teacher variables explain variance in average mathematics achievement for males and females across teachers in both data sets. Variance explained in female average achievement is very large when compared to the variance explained in male average achievement across teachers in CT. In NAEP 1996, more variance is explained in the achievement of males than females.

**Relationships between teacher practice and male/female achievement.** One teacher practice variable, math sensemaking, explains variance in male/female achievement across data sets. A one standard deviation increase in sensemaking activities increases male average achievement across teachers by .70 points in CT and .96 points in NAEP 1996. Sensemaking decreases female achievement by -.54 in CT and -.98 in NAEP 1996. Reform orientation has a similar pattern. A one standard deviation increase in teacher reform orientation increase male achievement by +.58 points in CT and +1.65 points in NAEP 1996. Female achievement decreases by -.70 and -.65 respectively.

The one teacher practice that appears to increase female mathematics gains controlling for SES in CT is drill and lecture activities. It explains 33.8% of the variance in female achievement but only 5.7% of the variance in male gains across teachers. A one standard deviation increase in drill and lecture activities increases female achievement by +.60 points while male gains decrease by -.30 points. This effect was not present in the national sample.

**Relationships between teacher characteristics and male/female achievement.** One teacher characteristic, reform orientation, explains variances in male/female achievement across both data sets. In CT, reform orientation explains 6.4% of the variance in male achievement and 23.6% of the variation in female achievement controlling for SES. A one standard deviation increase in teacher reform orientation increases male scores by .58 points and decreases female achievement by -.70 points. In NAEP 1996, a one standard deviation increase in reform orientation, increases male mean achievement by 1.65 points and decreases female achievement by -.65 points.

**Relationship between teacher professional development and male/female achievement.** While professional development frequency does not explain variance in NAEP 1996 mathematics achievement for male/female, it does explain male/female achievement variance across teachers in CT. For average male achievement, it accounts for 17.6% of the variance and for 15.5% of the variance in female achievement. A one standard deviation increase in professional development frequency decreases male gains by -.12 points while female gains remain level.

Teacher practices, characteristics, and professional development do not have the same effects on achievement for all students across subject matter and data sets. Generally, teacher variables explain more variance in female and Black average achievement across teachers than for male and White achievement. Relationships among achievement and teacher variables are stronger for mathematics than for reading. Teacher variables for mathematics sensemaking, reform orientation, and professional development frequency (for mathematics and reading) have the strongest relationships with student achievement across multiple data sets but the results showing who ultimately benefits are mixed. In general, gain scores offer more insight into which students are benefitting from teacher practice, characteristics, and professional development.

**Relationship between teacher preparation and student achievement**. Connecticut student data could be nested within teachers and teachers nested within teacher preparation programs. Average student gains produced by teacher graduates of specific programs did not vary by teacher preparation program in CT for either reading or mathematics.

Two level analyses were possible with teacher practices and characteristics as the outcomes. 38% of the variance in teacher mathematics knowledge was explained by teacher preparation program and 15% of the variation in teacher English preparation was explained by university teacher preparation. In the two level student/teacher analyses for CT, a one standard deviation in mathematics knowledge was related to an increase in female gains of .89 points and +.41 point increase in Black/White gains. Male scores remained level. This implied relationship should be interpreted with caution since there was not enough variation in teacher preparation institutions to support a three-level model using student gains over one year as the outcome.

# 15. DISCUSSION AND IMPLICATIONS

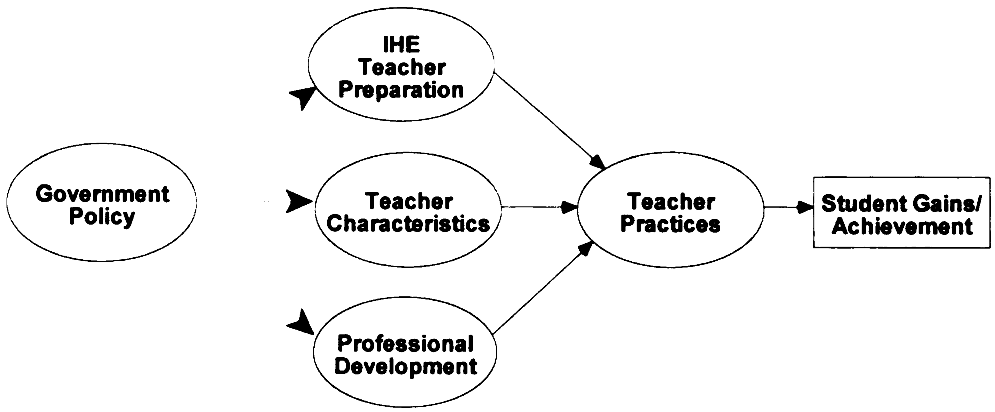
Using results from the analysis of six state and national data sets, this study provides evidence on the strength of the relationships and relative contributions of four important elements in the education system to student achievement in mathematics and reading: teacher preparation, teacher characteristics, professional development, and teacher practice. Results indicate that teacher practices, teacher characteristics, and professional development across teachers generally have moderate to large direct effects on teacher practice and small but significant indirect effects on student achievement in mathematics and reading. Sometimes those effects differentially impact student achievement by ethnic and gender groups. Findings vary based on curriculum and policy context. While there are many interesting results related to specific data sets throughout the analyses, this discussion will concentrate on the larger picture by identifying common threads across data sets and analyses.

The discussion begins with a review of the conceptual model which also addresses design and methodological strengths and weaknesses, Next, the discussion moves to the relative importance and contributions of teacher preparation, teacher characteristics, and professional development to teacher practice and student achievement. The discussion of differential effects on student achievement by gender and ethnicity follows. Implications for the present and future conclude.

**The Conceptual Model.**

For the purposes of framing the discussion about the strength of relationships and relative contributions of variables to student achievement, a quick review of the conceptual model is in order.

**Figure 14.1.** Conceptual Model (repeated)




The model graphically illustrates study research questions concerning the relationships among variables. In the model, state and/or federal government education policy is directed at teacher preparation, teacher characteristics, and professional development which, in turn, are expected to influence teacher practices and subsequently increase student achievement. Prior research results supported the connections but no single study looked at all the connections simultaneously. This study used SEM to look at the relationships among variables simultaneously and to estimate standardized direct and indirect effects. HLM analyzed the effects of individual teacher variables on student achievement as well as differential teacher effects on achievement by gender and ethnicity. Additionally, the study presented effects of teacher preparation programs on teacher characteristics.

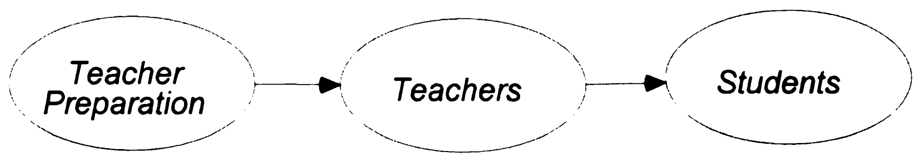
Methodological choices which may have affected the interpretations for both SEM and HLM analyses include large variations in sample size and variations in the specific variables available across data sets. Sample size varied from 168 teachers and 1550 students in Connecticut to over 1000 teachers and 6000 students in NAEP. Larger sample sizes have more stability and accuracy than smaller ones. To improve interpretation across differing sample sizes, analysis proceeded with the smallest data set serving as the base analysis with results confirmed in successively larger data sets. Variation in available variables across data sets was unavoidable. Reasonable care was taken to ensure similarities of scales and variable contributions to constructs.

**SEM Analysis.** The conceptual model served well as the structural model for SEM analysis across all data sets. The unit of analysis for SEM was teachers with average classroom achievement standardized by grade level as the outcome. Each latent variable (other than government policy) had a series of descriptor variables that formed separate measurement models. The measurement models successfully defined the latent variables for teacher preparation, teacher characteristics, professional development, and teacher practice. When the measurement models were combined with the structural model, the final hybrid model had good fit (NFI>.984, p<.001, RMSEA=.05 to .116) and worked in similar ways across data sets. Variations that occurred appeared related to subject matter and government policy context.

Methodological choices which may have affected the interpretation of SEM results include the use of correlation matrices as input for the SEM and potential naming fallacy. Although covariance matrices are the input of choice for SEM, the use of correlation matrices was not possible since student data from Tennessee was only available as part of a correlation matrix. In order to treat all data sets in similar fashion, weighted correlation matrices were constructed for NAEP data to ensure appropriate standard errors. There is always a chance that the name chosen for a construct is not appropriate (otherwise called "naming fallacy"). Correlations suggest that teacher preparation and teacher characteristics may not be distinctly different constructs and consequently one or both may be misnamed.

**HLM Analysis.** HLM analysis was based on the conceptual model although nesting requirements necessitated slight modifications. The model took the following form:

**Figure 14.2.** Three Level HLM model




Students were nested within teachers who were nested within teacher preparation programs. The three level models did not have enough variation present at level three for meaningful analysis across teacher preparation programs. Therefore, two level models with students nested within teachers or teachers nested within teacher preparation programs were employed to look at the effects of teachers on students and the effects of teacher preparation programs on teachers. In the first case, students were the unit of analysis. In the second case, teachers were the unit of analysis. There was evidence of differential effects of teacher characteristics, professional development, and practices on achievement by gender and ethnic group. It was also possible to rank teacher preparation programs based on the levels of certain characteristics found in beginning teachers who attended that program. While interesting, results were not significantly different from zero.

Methodological issues which may affect the interpretation of HLM results include a non-random sample of teachers, variations in the measure of SES included across data sets and the use of gains for state analyses versus achievement status in NAEP. Students are randomly sampled in NAEP and teacher variables are included as a part of the student record. It is not a random sample of teachers. District SES was added to the Connecticut data as part of the teacher data since no measure of SES was available in BTPS data. SES is measured at the student level in NAEP and aggregated to the teacher level.

**Gains Versus Achievement Status.** In general, gains are more informative of progress than achievement status. They more accurately portray the variance associated with teachers by reducing variation attributable to factors outside of the classroom. When achievement status is the outcome, white students have higher overall scores than other demographic groups and boys have higher average scores than girls. In the HLM analyses, gain scores pointed to a narrowing of the achievement gap over the course of the year with minorities and girls gaining the most points. It is very important not to confuse narrowing the achievement gap with attaining equal levels of achievement. When either pre- or post-test scores in Connecticut were analyzed separately, results revealed the same patterns as achievement status -- White and male students had the highest overall scores.

**State Versus National Data Sets.** State data used gain scores as the student outcome while the national data used achievement status. In general, state samples where students varied in similar ways based on the education policy environment were more informative than the national samples which tended to average out variance found at the state level. In other words, a random sample of students from all states averages out indicators that might be important at the state level, masking important trends. High achievement in one state is balanced by low achievement in another. Specifically, a highly centralized state might have strong positive connections from policy to student achievement while a decentralized state might have strong negative connections. At the national level this would show up as no connection.

**Relationships Across Models**

What do the SEM analyses say about the structure of the education system? When looking at the absolute size of effects, there are strong similarities in structural equation modeling results across data sets. Analyses based on the conceptual model indicate relatively stable direct and indirect effect sizes for teacher preparation, teacher characteristics, professional development, and teacher practice across policy contexts. Most surprising is the persistent, small indirect effect ($\delta \approx |.05|$) for teacher preparation, teacher characteristics, and professional development through teacher practice to student outcomes regardless of whether the outcome is gains or achievement status in mathematics or reading.

While some might argue that an effect of this size is too small to be worth pursuing, a small effect can be very important (Camilli et al., 2003; Reckase, 2002; Valli et al., 2003; Wayne & Youngs, 2003). Sanders et al. (1999) have shown that teacher effects are additive across years and can be of great importance to future student achievement. If policy makers plan to evaluate the education system based on student achievement outcomes, knowing the size of the effect that can be reasonably traced back to teacher preparation, teacher characteristics, and professional development is key.

Teacher preparation, teacher characteristics, and professional development produce measurable direct effects on teacher practice. Generally, teacher preparation has medium direct effects on teacher practice ($\delta \approx |.30|$)[43]. On average, teacher characteristics have large direct effects on teacher practice ($\delta > |.40|$). The direct effects of professional development on teacher practice are mostly small ($\delta \approx |.10|$ to $|.20|$). Teacher practice has small direct effects on student achievement although the effects are larger than the other constructs ($\delta \approx |.01|$ to $|.30|$).

In other words, teacher characteristics have the largest effects on teacher practice followed by teacher preparation and then professional development. The size of indirect effects on student achievement are in the range of $\delta \approx |.05|$ for all three. The direct effects of teacher practice on student achievement, as might be expected, are larger than any of the indirect effects of teacher preparation program, teacher characteristics, or professional development. Predictably, 68-85% of the variance in teacher mathematics practice is yet to be explained by the SEM models. Generally, 50-75% of the variation in teacher reading practices is left to be explained.[44] The unexplained variance represents unmeasured variables that may be related to a teachers' personal backgrounds, individual styles, peer influences, school climate, and other factors not included in the models but which appear to have great influence on the practices teachers use.

**Contexts**

This section discusses the potential roles that policy and subject matter context play in differences found across data sets. While effect size is important to understanding the strength of relationships among variables, the sign of the effect indicates differences

---

[43]Data is only available for Connecticut and Tennessee.

[44]Atypically, NAEP 1998 model only has 18% of the variation in reading practices left to be explained.

by subject matter and policy context.

**Government Policy**. Since there were no empirical measures of government policy present in the data, this study could only review policies at the national and state levels that were believed to impact teacher preparation, teacher characteristics, and professional development. Prior to 1998, most federal policy did not intervene in the rights of states to regulate education except in the area of equality, a fundamental and federally protected right for all citizens. Since 1998, two major pieces of legislation, the Amendments to the Elementary and Secondary School Act (ESEA, Title I) of 1965 and HEA (Higher Education Act, Title II) have had far-reaching effects into state regulation of education. These acts include No Child Left Behind (NCLB) and the Title II State Report Card on Teacher Preparation which tie student achievement and teacher preparation to federal funding. Neither was in force during the data collection years for this study. Connecticut BTPS data was augmented with results from the Initial State Report Card on Teacher Preparation (1998) but there were no detectable effects of IHE teacher preparation program certification pass rates on teacher practices or student achievement.

A review of education policies from Connecticut and Tennessee indicated that the states took different approaches to regulating education. Two ways in which the state policy environments varied during 1999-2000 were in the choice of method for policy implementation (top-down centralized-CT vs. bottom-up decentralized-TN) and testing purposes (low-stakes-CT vs. high-stakes-TN). Given the similarities of analysis results by state and the lack of within-state policy variation by data set, any conclusions about the influence of policy would be circumstantial at best.

**Subject matter context.** Positive values for teacher characteristics and professional development translate into positive values for teacher practice and student achievement across all data sets. In the two data sets that have a measure of teacher preparation program (Connecticut and Tennessee), teacher preparation program effects are negative on both teacher mathematics practice and student mathematics achievement. Valli et al. (2003) suggest that this negative relationship may be related to the level of subject matter preparation required for teaching elementary mathematics. There were also indications that teacher preparation program as measured in this data was not a distinctly different construct from teacher characteristics.

With the exception of Connecticut, effects for reading are mixed across data sets. Positive effects at one level of the system sometimes translate into negative effects at the next level and vice versa. This suggests that education system is not aligned to support reading achievement in the same way it is to support mathematics achievement. It also suggests that Connecticut is doing something right when it comes to reading achievement. Lessons from Connecticut could be valuable as other states move toward comprehensive reading policy in order to reach NCLB reading achievement requirements.

## Differential Effects

The preceding discussion describes a model of education that is connected in reasonable ways. The system is linked to mathematics achievement and less well linked to reading achievement. Teacher preparation programs, teacher characteristics, and professional development have strong effects on teacher practices that in turn have small but significant effects on average student achievement. What does this mean for various subgroups of students who come from a variety of ethnic and gender groups?

A large portion of this study was devoted to the potential differential impact of teacher characteristics and practices on student achievement in mathematics and reading. Several earlier studies indicated that subjects such as reading are culturally embedded and may be more effectively addressed with explicit rather than implicit instructional strategies (Delpit, 1988; Hirsch, 1996; Heath, 1983). Mathematics also embodies both computational and conceptual components which can be approached in both explicit and implicit ways (NCTM 2000, Ma 1999). Does a particular approach disadvantage some students while advantaging others? HLM was the appropriate method for addressing questions regarding the differential effects of teacher variables on student achievement by gender and ethnicity. Tennessee was excluded from the analyses since it did not have individual student scores nested within teachers.

**Effects of teachers on student achievement.** Achievement by gender and ethnicity varied significantly across teachers in some but not all data sets. The following is a list of groups whose achievement varied and the data sets in which they are found.

1. Black mathematics and reading achievement varied across teachers in four data sets: Connecticut and NAEP 1996 mathematics, Connecticut and NAEP 1998 reading.
2. Hispanic mathematics achievement varied across teachers in NAEP 2000.
3. Female mathematics achievement varied across teachers in two data sets: Connecticut and NAEP 1996.
4. Black female mathematics achievement varied across teachers in NAEP1996.

For the most part, teacher effects which explained variance in average achievement by demographic groups were not consistent across data sets. Results for average Black achievement across teachers are discussed below followed by results for average female achievement. Although average Hispanic achievement and average Black female achievement varied across teachers in certain data sets, teacher level variables explained little variance in their average achievement so they are not discussed further.

**Teacher effects on average Black mathematics and reading achievement**. Teacher mathematics preparation, an indicator of teacher quality, explains variance in achievement for both Black and White students across teachers in Connecticut and NAEP 1996. Controlling for SES, variance explained for average Black achievement (57.9%) is more than twice that for White (25.6%). This suggests that while teacher mathematics preparation varies considerably for all students, Black students are more than twice as likely to have mathematics teachers with less disciplinary course work.

Teachers' frequency of sensemaking activities, which are facilitated by subject matter preparation, account for nearly 48% of the variance in Black average gains and none of the variance in White gains across teachers in Connecticut. If teachers increased the frequency of sensemaking activities by 1 SD, Black average gains would increase by an +1.25 points over the year. White students would also increase their scores slightly by about +.18. In NAEP 1996, sensemaking activities accounted for 10% of the variation in Black average achievement and only 2% of the variation in White achievement. Every one SD increase of sensemaking activities explained a -.11 difference in overall Black achievement status compared to White achievement status. In other words, the frequency of sensemaking activities was less for Black students than for white students and resulted in a -.11 drop in average achievement across teachers.

By NAEP 2000, Black mathematics achievement did not vary across teachers at the national level. It is worth noting that the Connecticut data was also collected in 2000. The lack in variance in NAEP is likely due to the balancing out state level effects in the national sample.

Other differential teacher effects that explain variation in Black average mathematics achievement across teachers in Connecticut include reform orientation and professional development frequency but these effects are not present in multiple data sets. These results are discussed in the HLM section but not repeated here.

Average Black reading achievement varied in two data sets but only the teacher variables in NAEP 1998 explained variance. NAEP variables related to teacher quality - teacher reading preparation and assessment preparation - explained 7-8% of the variance in Black achievement across teachers and none of the variance in White achievement controlling for SES. All three teacher reading practices explained variance in Black achievement (10-50%) across teachers. None of the practice variables explained variance in White achievement. As in mathematics, the effect disappeared at the national level in NAEP 2000. The pattern of differential teacher effects on Black reading and mathematics achievement across subject matter and policy contexts is worth further exploration.

**Teacher effects on average female mathematics achievement**. Differential effects of teachers on average female mathematics achievement were also present in Connecticut and NAEP 1996. Differential effects attributable to teachers were very large in Connecticut and very small in NAEP 2000. In general, teacher variables explained three to five times more variation in girls' average mathematics gains compared to boys. District SES also explained four times more variance in average gains for girls across teachers. Interpretation, however, is key. Boys are disproportionately disadvantaged by lower SES than girls in this sample. For every one standard deviation increase in free and reduced lunch, boys lose more than -6 points in mathematics gains over the year while girls gain +8 points. For each standard deviation increase in teacher variables, boys tend to gain less than a point while girls tend to lose less than a point.

In NAEP 1996, a similar pattern emerges where a one standard deviation increase in free and reduced lunch (SES) signals an achievement status drop for boys of -6 to -7 points while girls lose less than one point. As in Connecticut, each standard deviation increase in teacher variables signals a gain in average achievement status for boys of less than a point while girls tend to lose less than a point. One exception is the effect of teacher reform orientation where boys increase average achievement status by +1.65 points across teachers while girls decrease by -.65 points.

**Implications for the Present and Future**

The education system in the United States is connected and can be measured from policy to student achievement. Teacher preparation programs, teacher characteristics, and professional development have reasonably strong effects on teacher practices that in turn have small but significant effects on average student achievement in mathematics and reading. As a whole, the system supports the development of mathematics teachers in more predictable ways than it supports the development of reading teachers but the ultimate effect on student achievement for both subjects is in the range of $\delta \approx .05$. There is room for tightening the system and increasing the size of those effects.

It is good news for government that it is possible to track the contributions of teacher preparation, teacher characteristics, and professional development, to teacher practice and student achievement. This makes it possible to measure improvements in the connections among the variables as well as predict future student achievement. Using data from this study, one could predict that a Black student in Connecticut with gains at the average on the state mathematics achievement test would grow from +6.68 points to +33.4 points from 2000 to 2005. Adding a one standard deviation increase in teacher sensemaking activities (+1.25 points) would result in a further increase from 33.4 points to almost 40 points over that same time period.

The bad news is that this effect may not be enough to attain the NCLB goal of proficiency for all students by the 2014-15 school year. For the sake of argument, take a White student who has a mathematics CAPT score of 199 (below basic in CT) and project this score until 2005 using the +4.05 average increase until 2005. The anticipated outcome score would be just under 220 by 2005, just short of the proficient level (223-260). By 2010, the projection would reach 240; by 2015, the student would reach the Connecticut Goal of 261, just above proficient, the level required for AYP. Of course, this presumes a perfect world with no intervening circumstances and steady progress at the predicted average. In spite of average increases over time, $\delta \approx .05$ may not be enough to make AYP requirements in Connecticut or anywhere else. Increases in certain teacher practices, such as mathematics sensemaking, could add additional point increments; however, in the case of differential effects, one disaggregated group may improve at the expense of another.

Improvement in academic achievement does not exist in a vacuum. Students move in and out of districts and states frequently. Immigration patterns guarantee that many schools start over with new students every year, new students who will struggle to acclimate to their new school, curriculum, and testing. Although it is possible to improve the achievement of students currently in the system, the constant flow of new students means that the system will always have an unending supply of students who will need to be brought up to speed.

It is good news that minority groups and women in Connecticut are closing achievement gaps in mathematics and reading when gain scores are the outcome. This bodes well for reporting improvement in subgroup scores for NCLB. It is bad news, however, that achievement status from NAEP still indicates that minorities have yet to equal or outpace the overall scores of their majority peers.

The biggest hurdle for all students is still low SES. The effect of one standard deviation increase in free and reduced lunch at the district level lowers the average gains of Black students by around -19 points for mathematics in Connecticut. In NAEP 2000 Reading, the effect of one standard deviation increase in free and reduced lunch at the classroom level lowers student scores by -18 points. While it is clear that the effects of certain teachers variables on average achievement by ethnicity and gender is differential, no teacher level variables can overcome the overwhelming effects associated with SES. As long as socioeconomic inequality exists in the extremes, adequate yearly progress may be impossible.

As always, more research needs to be done with data that spans government policy to student achievement which means appropriate databases and/or ways to bridge data sets need to be developed. Stronger connections need to be made throughout the education system, especially in reading, so that all elements of the system are working in concert. More evidence of teacher practices that improve student achievement for all students needs to be gathered. And in all this, it can't be forgotten that human beings are not perfectly predictable so education systems will never be.

# APPENDICES

**APPENDIX A:** Correlation Tables

**Table A.1** Connecticut 2000 4th-8th Mathematics (n=168)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 MMSDF | 1.000 | | | | | | | | | | | |
| 2 ASSINST | -0.077 | 1.000 | | | | | | | | | | |
| 3 MATHINST | -0.006 | 0.578 | 1.000 | | | | | | | | | |
| 4 MATHPREP | -0.327 | 0.053 | 0.163 | 1.000 | | | | | | | | |
| 5 SCHQUAL | -0.231 | 0.194 | 0.135 | -0.008 | 1.000 | | | | | | | |
| 6 MTHSM | 0.117 | 0.290 | 0.210 | -0.159 | 0.205 | 1.000 | | | | | | |
| 7 MTHDLA | 0.077 | 0.078 | 0.185 | 0.071 | 0.013 | 0.176 | 1.000 | | | | | |
| 8 MTHGPA | 0.006 | 0.333 | 0.256 | 0.162 | 0.198 | 0.507 | 0.289 | 1.000 | | | | |
| 9 MTHKNW | 0.032 | 0.275 | 0.418 | 0.273 | 0.218 | 0.062 | 0.101 | 0.247 | 1.000 | | | |
| 10 MOTRAD | 0.074 | 0.039 | 0.207 | 0.116 | 0.015 | -0.010 | 0.318 | 0.012 | 0.044 | 1.000 | | |
| 11 MOREF | 0.053 | -0.002 | 0.044 | -0.050 | 0.024 | 0.182 | -0.089 | 0.254 | 0.069 | -0.214 | 1.000 | |
| 12 PRFDEVI | -0.105 | 0.133 | 0.152 | 0.026 | 0.348 | 0.035 | 0.136 | 0.163 | -0.058 | 0.183 | 0.068 | 1.000 |
| 13 PRFDEVF | -0.396 | 0.094 | -0.049 | -0.015 | 0.049 | -0.034 | 0.044 | 0.195 | -0.083 | 0.086 | -0.021 | 0.480 |
| 14 EFFICACY | -0.007 | 0.204 | 0.308 | 0.057 | 0.214 | 0.336 | -0.031 | 0.191 | 0.009 | -0.101 | 0.086 | 0.016 |
| 15 DIVPREP | 0.120 | 0.487 | 0.421 | -0.079 | 0.283 | 0.187 | 0.036 | 0.116 | 0.159 | 0.012 | 0.038 | 0.299 |
| 16 PROGQUAL | -0.032 | 0.488 | 0.572 | 0.005 | 0.193 | 0.177 | 0.052 | 0.115 | 0.217 | -0.007 | 0.064 | 0.186 |
| 17 FIELDEXP | -0.018 | 0.305 | 0.353 | -0.094 | 0.266 | 0.215 | 0.117 | 0.164 | 0.080 | 0.172 | -0.088 | 0.289 |
| 18 FACCHAR | -0.104 | 0.455 | 0.295 | 0.069 | 0.295 | 0.096 | 0.095 | 0.163 | 0.031 | -0.042 | -0.036 | 0.211 |
| 19 MENTFRQ | -0.237 | -0.026 | 0.036 | 0.170 | 0.208 | -0.095 | 0.024 | -0.023 | -0.097 | 0.146 | -0.036 | 0.526 |
| 20 MENTWRTH | -0.240 | 0.001 | 0.060 | 0.093 | 0.292 | -0.033 | -0.018 | -0.085 | -0.082 | 0.089 | 0.041 | 0.532 |
| mean | 0.132 | 30.037 | 47.802 | 2.690 | 40.230 | 30.202 | 25.162 | 22.296 | 25.946 | 12.723 | 15.622 | 46.351 |
| stddev | 0.032 | 0.506 | 0.803 | 0.192 | 0.660 | 0.276 | 0.278 | 0.402 | 0.313 | 0.298 | 0.198 | 0.874 |

**Table A.1** (continued) Connecticut 2000 4th-8th Mathematics (n=168)

| Variable | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|
| 13 PRFDEVF | 1.000 | | | | | | | |
| 14 EFFICACY | -0.038 | 1.000 | | | | | | |
| 15 DIVPREP | 0.255 | 0.195 | 1.000 | | | | | |
| 16 PROGQUAL | -0.051 | 0.241 | 0.419 | 1.000 | | | | |
| 17 FIELDEXP | 0.086 | 0.112 | 0.255 | 0.295 | 1.000 | | | |
| 18 FACCHAR | 0.141 | 0.137 | 0.228 | 0.464 | 0.251 | 1.000 | | |
| 19 MENTFRQ | 0.361 | 0.051 | 0.101 | 0.070 | 0.124 | 0.126 | 1.000 | |
| 20 MENTWRTH | 0.385 | 0.169 | 0.121 | 0.091 | 0.171 | 0.193 | 0.876 | 1.000 |
| mean | 18.778 | 29.918 | 21.445 | 18.593 | 16.391 | 23.609 | 31.105 | 22.054 |
| stddev | 0.905 | 0.302 | 0.472 | 0.286 | 0.254 | 0.350 | 1.096 | 0.707 |

**Table A.2** Connecticut 2000 4th-8th Reading (n=168)

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 READINST | 1 | | | | | | | | | | | |
| 2 LITMAT | 0.018 | 1 | | | | | | | | | | |
| 3 DPRDIF | 0.039 | 0.001 | 1 | | | | | | | | | |
| 4 LITCOMP | 0.145 | -0.01 | -0.202 | 1 | | | | | | | | |
| 5 LITTRAD | 0.175 | 0.086 | 0.209 | 0.205 | 1 | | | | | | | |
| 6 LITREFM | 0.151 | -0.016 | 0.163 | 0.267 | 0.398 | 1 | | | | | | |
| 7 MOTPREP | 0.651 | 0.099 | 0.132 | 0.11 | 0.101 | 0.224 | 1 | | | | | |
| 8 ASSINST | 0.765 | 0.009 | 0.021 | 0.112 | 0.051 | 0.14 | 0.631 | 1 | | | | |
| 9 LITKNOW | 0.528 | -0.021 | -0.015 | 0.18 | 0.175 | 0.249 | 0.347 | 0.477 | 1 | | | |
| 10 SCHQUAL | 0.272 | -0.164 | -0.151 | 0.343 | 0.121 | 0.238 | 0.291 | 0.194 | 0.33 | 1 | | |
| 11 PRFDEVI | 0.271 | 0.215 | 0.031 | 0.359 | 0.157 | 0.176 | 0.219 | 0.154 | 0.192 | 0.369 | 1 | |
| 12 PRFDEVF | -0.025 | 0.137 | 0.026 | 0.277 | -0.437 | 0.035 | 0.018 | 0.052 | 0.106 | -0.022 | 0.291 | 1 |
| 13 EFFICACY | 0.219 | -0.051 | 0.114 | 0.106 | 0.131 | 0.253 | 0.382 | 0.205 | 0.185 | 0.201 | 0.028 | -0.329 |
| 14 DIVPREP | 0.533 | 0.105 | 0.115 | 0.123 | 0.174 | 0.204 | 0.606 | 0.486 | 0.331 | 0.247 | 0.284 | 0.177 |
| 15 PROGQUAL | 0.517 | 0.008 | 0.095 | 0.012 | 0.102 | 0.146 | 0.434 | 0.455 | 0.246 | 0.182 | 0.19 | -0.083 |
| 16 FIELDEXP | 0.383 | 0.109 | 0.077 | 0.149 | 0.056 | 0.234 | 0.405 | 0.309 | 0.241 | 0.25 | 0.287 | -0.124 |
| 17 FACCHAR | 0.434 | -0.032 | 0.034 | 0.042 | 0.093 | 0.158 | 0.407 | 0.463 | 0.174 | 0.262 | 0.233 | -0.057 |
| 18 BLITO | 0.104 | -0.21 | 0.068 | 0.014 | -0.07 | 0.037 | 0.149 | 0.167 | 0.196 | 0.213 | 0.056 | 0.068 |
| 19 MENTWRTH | 0.231 | 0.309 | 0.215 | 0.029 | -0.002 | 0.175 | 0.071 | 0.103 | 0.103 | 0.311 | 0.479 | 0.206 |
| mean | 44.757 | 7.083 | 0.288 | 8.136 | 31.77 | 33.125 | 25.108 | 30.028 | 39.887 | 40.249 | 46.566 | 21.013 |
| stddev | 0.741 | 0.233 | 0.288 | 0.237 | 0.239 | 0.317 | 0.391 | 0.51 | 0.462 | 0.653 | 0.876 | 0.928 |

**Table A.2** (continued) Connecticut 2000 4th-8th Reading (n=168)

| Variables | 13 | 14 | 14 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|
| 13 EFFICACY | 1 | | | | | | |
| 14 DIVPREP | 0.168 | 1 | | | | | |
| 15 PROGQUAL | 0.226 | 0.445 | 1 | | | | |
| 16 FIELDEXP | 0.099 | 0.255 | 0.292 | 1 | | | |
| 17 FACCHAR | 0.113 | 0.222 | 0.453 | 0.255 | 1 | | |
| 18 BLITO | 0.181 | -0.031 | 0.147 | 0.041 | 0.215 | 1 | |
| 19 MENTWRTH | 0.056 | 0.138 | 0.174 | 0.19 | 0.237 | 0.164 | 1 |
| mean | 29.89 | 21.446 | 18.583 | 16.396 | 23.636 | 35.013 | 21.964 |
| stddev | 0.299 | 0.473 | 0.285 | 0.254 | 0.351 | 0.254 | 0.782 |

**Table A.3** Tennessee 2000 4th-8th Mathematics and Reading Correlations (n=442)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 FACCHAR | 1.000 | | | | | | | | | | | | | | |
| 2 FIELDEXP | 0.439 | 1.000 | | | | | | | | | | | | | |
| 3 MTHKNW | 0.016 | 0.072 | 1.000 | | | | | | | | | | | | |
| 4 MATHPREP | -0.036 | -0.064 | 0.192 | 1.000 | | | | | | | | | | | |
| 5 PROGQUAL | 0.616 | 0.417 | 0.236 | -0.012 | 1.000 | | | | | | | | | | |
| 6 MENTWRTH | 0.186 | 0.204 | 0.011 | 0.010 | 0.231 | 1.000 | | | | | | | | | |
| 7 MENTFRQ | 0.199 | 0.333 | 0.027 | -0.017 | 0.287 | 0.866 | 1.000 | | | | | | | | |
| 8 PRFDEVF | 0.140 | 0.152 | -0.080 | 0.076 | 0.139 | 0.236 | 0.296 | 1.000 | | | | | | | |
| 9 PRFDEVQ | 0.219 | 0.241 | 0.203 | -0.007 | 0.407 | 0.129 | 0.234 | 0.384 | 1.000 | | | | | | |
| 10 MOTRAD | -0.018 | 0.057 | -0.066 | -0.069 | 0.055 | 0.184 | 0.094 | 0.263 | 0.061 | 1.000 | | | | | |
| 11 LITKNOW | 0.257 | 0.201 | 0.398 | 0.032 | 0.462 | 0.079 | 0.076 | 0.158 | 0.381 | 0.010 | 1.000 | | | | |
| 12 EFFICACY | 0.275 | 0.204 | -0.018 | -0.012 | 0.132 | 0.053 | 0.089 | 0.075 | 0.228 | -0.064 | 0.135 | 1.000 | | | |
| 13 MATHINST | 0.444 | 0.349 | 0.276 | 0.073 | 0.640 | 0.207 | 0.246 | 0.191 | 0.388 | -0.006 | 0.403 | 0.137 | 1.000 | | |
| 14 PRFDEVI | 0.290 | 0.316 | 0.065 | -0.064 | 0.335 | 0.483 | 0.501 | 0.454 | 0.530 | 0.113 | 0.287 | 0.194 | 0.327 | 1.000 | |
| 15 MOREF | 0.079 | 0.096 | 0.185 | 0.090 | 0.147 | 0.135 | 0.078 | 0.150 | 0.165 | -0.020 | 0.276 | 0.118 | 0.223 | 0.088 | 1.000 |
| 16 LITMAT | 0.152 | 0.155 | -0.010 | 0.012 | 0.152 | 0.219 | 0.190 | 0.186 | 0.177 | 0.322 | 0.166 | 0.123 | 0.180 | 0.142 | 0.006 |
| 17 MTHGPA | 0.111 | 0.123 | 0.238 | 0.081 | 0.143 | -0.035 | -0.052 | 0.309 | 0.217 | 0.013 | 0.249 | 0.096 | 0.251 | 0.283 | 0.299 |
| 18 MTHDLA | 0.024 | 0.048 | 0.148 | 0.030 | 0.162 | 0.154 | 0.080 | 0.189 | 0.173 | 0.234 | 0.084 | 0.021 | 0.057 | 0.019 | 0.153 |
| 19 BLITO | 0.123 | 0.088 | 0.147 | -0.007 | 0.207 | 0.197 | 0.161 | 0.057 | 0.099 | 0.196 | 0.195 | -0.012 | 0.123 | 0.145 | 0.184 |
| 20 MTHSM | 0.233 | 0.190 | 0.317 | 0.144 | 0.345 | 0.080 | 0.090 | 0.186 | 0.300 | -0.080 | 0.424 | 0.262 | 0.410 | 0.284 | 0.359 |
| 21 LITREFM | 0.179 | 0.121 | 0.136 | 0.075 | 0.248 | 0.081 | 0.076 | 0.369 | 0.288 | 0.035 | 0.360 | 0.179 | 0.194 | 0.287 | 0.280 |
| 22 LITTRAD | 0.129 | 0.061 | 0.119 | 0.087 | 0.259 | 0.133 | 0.160 | 0.235 | 0.197 | 0.001 | 0.190 | 0.104 | 0.102 | 0.088 | 0.103 |
| 23 ASSINST | 0.474 | 0.366 | 0.198 | -0.035 | 0.624 | 0.215 | 0.262 | 0.216 | 0.383 | 0.120 | 0.408 | 0.161 | 0.669 | 0.351 | 0.136 |
| 24 DIVPREP | 0.436 | 0.384 | 0.106 | -0.082 | 0.477 | 0.142 | 0.203 | 0.278 | 0.360 | 0.054 | 0.311 | 0.190 | 0.519 | 0.341 | 0.177 |
| 25 READINST | 0.432 | 0.384 | 0.167 | -0.083 | 0.626 | 0.244 | 0.312 | 0.242 | 0.381 | 0.125 | 0.506 | 0.163 | 0.622 | 0.355 | 0.152 |
| 25 MATH | 0.029 | 0.003 | 0.202 | -0.003 | -0.076 | -0.091 | -0.009 | 0.126 | 0.120 | -0.114 | -0.008 | 0.178 | -0.085 | -0.058 | 0.091 |
| 27 READING | 0.012 | -0.031 | -0.034 | -0.023 | -0.054 | -0.139 | -0.069 | 0.024 | -0.001 | 0.005 | 0.022 | 0.073 | -0.108 | -0.003 | -0.120 |
| mean | 21.715 | 14.715 | 23.731 | 2.204 | 17.337 | 22.223 | 31.441 | 19.896 | 19.626 | 13.778 | 36.176 | 27.818 | 44.339 | 46.426 | 15.167 |
| stddev | 4.915 | 3.848 | 4.270 | 1.418 | 4.279 | 7.643 | 13.811 | 8.577 | 4.875 | 2.926 | 7.230 | 4.609 | 11.413 | 12.312 | 2.183 |

**Table A.3** (continued) Tennessee 2000 4th-8th Mathematics and Reading Correlations (n=442)

| Variable | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 litmat | 1.000 | | | | | | | | | | | |
| 17 mthgpa | 0.035 | 1.000 | | | | | | | | | | |
| 18 mthdla | 0.267 | 0.074 | 1.000 | | | | | | | | | |
| 19 blito | 0.033 | 0.019 | 0.106 | 1.000 | | | | | | | | |
| 20 mthsm | 0.154 | 0.504 | 0.278 | 0.117 | 1.000 | | | | | | | |
| 21 litrefm | 0.115 | 0.530 | 0.158 | 0.080 | 0.509 | 1.000 | | | | | | |
| 22 littrad | 0.188 | 0.221 | 0.405 | 0.035 | 0.341 | 0.444 | 1.000 | | | | | |
| 23 assinst | 0.169 | 0.168 | 0.130 | 0.145 | 0.244 | 0.198 | 0.103 | 1.000 | | | | |
| 24 divprep | 0.175 | 0.250 | 0.063 | 0.102 | 0.353 | 0.222 | 0.081 | 0.554 | 1.000 | | | |
| 25 readinst | 0.178 | 0.107 | 0.131 | 0.167 | 0.225 | 0.178 | 0.092 | 0.716 | 0.525 | 1.000 | | |
| 25 math | 0.002 | 0.122 | -0.004 | -0.057 | 0.169 | 0.120 | 0.213 | -0.058 | 0.002 | -0.162 | 1.000 | |
| 27 reading | 0.054 | -0.044 | -0.014 | -0.062 | 0.026 | -0.046 | 0.054 | -0.125 | 0.032 | -0.058 | 0.000 | 1.000 |
| mean | 11.141 | 19.180 | 26.099 | 34.431 | 28.743 | 29.648 | 31.636 | 27.158 | 19.053 | 40.523 | -0.474 | -0.436 |
| stddev | 3.765 | 5.064 | 2.843 | 3.391 | 3.831 | 5.074 | 3.329 | 6.742 | 6.079 | 9.977 | 7.489 | 4.113 |

**Table A.4** 1996 4th Grade NAEP Mathematics Correlations (n=584 teachers)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 MRPCM1-PV1 | 1.000 | | | | | | | | | | | | | |
| 2 MRPCM2-PV2 | 0.965 | 1.000 | | | | | | | | | | | | |
| 3 MRPCM3-PV3 | 0.958 | 0.958 | 1.000 | | | | | | | | | | | |
| 4 MRPCM4-PV4 | 0.959 | 0.957 | 0.955 | 1.000 | | | | | | | | | | |
| 5 MRPCM5-PV5 | 0.961 | 0.955 | 0.952 | 0.955 | 1.000 | | | | | | | | | |
| 6 Professional Dev Freq | -0.065 | -0.064 | -0.051 | -0.050 | -0.053 | 1.000 | | | | | | | | |
| 7 Group & Project Activities | 0.104 | 0.096 | 0.115 | 0.105 | 0.110 | 0.296 | 1.000 | | | | | | | |
| 8 Sensemaking activities | 0.088 | 0.087 | 0.097 | 0.091 | 0.103 | 0.364 | 0.503 | 1.000 | | | | | | |
| 9 Mathematics Knowledge | 0.062 | 0.051 | 0.057 | 0.070 | 0.058 | 0.060 | 0.126 | 0.176 | 1.000 | | | | | |
| 10 GPA assessment | -0.046 | -0.044 | -0.022 | -0.030 | -0.026 | 0.374 | 0.483 | 0.629 | 0.086 | 1.000 | | | | |
| 11 Reform Orientation | 0.067 | 0.072 | 0.074 | 0.076 | 0.064 | 0.148 | 0.211 | 0.191 | 0.065 | 0.203 | 1.000 | | | |
| 12 Preparation for math instruction | 0.008 | 0.006 | 0.006 | -0.003 | -0.004 | 0.136 | 0.225 | 0.230 | 0.177 | 0.217 | 0.259 | 1.000 | | |
| 13 Preparation for diversity | -0.026 | -0.037 | -0.026 | -0.029 | -0.035 | 0.214 | 0.300 | 0.282 | 0.189 | 0.313 | 0.284 | 0.482 | 1.000 | |
| 14 Math preparation | 0.098 | 0.091 | 0.110 | 0.104 | 0.094 | 0.182 | 0.150 | 0.226 | 0.170 | 0.123 | 0.129 | 0.154 | 0.189 | 1.000 |
| Mean | 223.738 | 224.133 | 223.747 | 223.402 | 223.979 | 2.784 | 11.887 | 10.471 | 7.448 | 7.475 | 3.017 | 3.298 | 1.620 | 0.626 |
| SD | 22.894 | 22.513 | 22.467 | 23.481 | 22.288 | 1.212 | 2.688 | 2.726 | 0.916 | 2.406 | 1.582 | 1.035 | 1.113 | 1.018 |

**Table A.5** NAEP 2000 4th Mathematics Correlations (n=1029)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 PV1 | 1.000 | | | | | | | | | | | | |
| 2 PV2 | 0.951 | 1.000 | | | | | | | | | | | |
| 3 PV3 | 0.946 | 0.945 | 1.000 | | | | | | | | | | |
| 4 PV4 | 0.943 | 0.941 | 0.950 | 1.000 | | | | | | | | | |
| 5 PV5 | 0.946 | 0.947 | 0.953 | 0.951 | 1.000 | | | | | | | | |
| 6 Professional dev frq | 0.012 | -0.004 | 0.020 | 0.017 | 0.031 | 1.000 | | | | | | | |
| 7 Knowledge of NCTM stds | 0.125 | 0.118 | 0.109 | 0.092 | 0.092 | 0.223 | 1.000 | | | | | | |
| 8 Knowledge of state stds | 0.078 | 0.067 | 0.074 | 0.052 | 0.047 | 0.150 | 0.308 | 1.000 | | | | | |
| 9 Mathematics instruction prep | 0.075 | 0.075 | 0.074 | 0.045 | 0.064 | 0.169 | 0.318 | 0.378 | 1.000 | | | | |
| 10 DLA activities | -0.044 | -0.025 | -0.022 | -0.038 | -0.045 | -0.107 | -0.009 | 0.056 | 0.094 | 1.000 | | | |
| 11 Preparation for mathematics | -0.010 | -0.013 | -0.007 | -0.005 | 0.015 | 0.079 | 0.078 | 0.048 | 0.160 | -0.012 | 1.000 | | |
| 12 Sensemaking activities | 0.078 | 0.060 | 0.068 | 0.061 | 0.061 | 0.308 | 0.281 | 0.181 | 0.223 | -0.099 | 0.023 | 1.000 | |
| 13 GPA | 0.092 | 0.083 | 0.095 | 0.089 | 0.103 | 0.194 | 0.150 | 0.186 | 0.242 | 0.044 | 0.036 | 0.486 | 1.000 |
| mean | 227.163 | 227.412 | 226.858 | 226.760 | 226.713 | 2.439 | 1.969 | 3.363 | 27.850 | 8.643 | 0.350 | 13.718 | 9.444 |
| stddev | 24.151 | 23.890 | 24.289 | 24.007 | 23.735 | 1.156 | 0.907 | 0.618 | 3.447 | 0.769 | 0.764 | 3.087 | 2.217 |

**Table A.6** NAEP 1998 Reading (n=757)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 PV1 | 1.000 | | | | | | | | | | | | |
| 2 PV2 | 0.906 | 1.000 | | | | | | | | | | | |
| 3 PV3 | 0.908 | 0.912 | 1.000 | | | | | | | | | | |
| 4 PV4 | 0.911 | 0.903 | 0.910 | 1.000 | | | | | | | | | |
| 5 PV5 | 0.914 | 0.904 | 0.905 | 0.904 | 1.000 | | | | | | | | |
| 6 PD - reading & writing in last year | -0.140 | -0.086 | -0.130 | -0.103 | -0.134 | 1.000 | | | | | | | |
| 7 Reading Instruction Prep | -0.004 | -0.001 | 0.006 | 0.027 | -0.003 | 0.225 | 1.000 | | | | | | |
| 8 Technology preparation | -0.082 | -0.083 | -0.071 | -0.064 | -0.066 | 0.112 | 0.321 | 1.000 | | | | | |
| 9 Diversity preparation | -0.197 | -0.180 | -0.170 | -0.175 | -0.189 | 0.283 | 0.359 | 0.206 | 1.000 | | | | |
| 10 Balanced literacy instruction | -0.113 | -0.099 | -0.124 | -0.115 | -0.104 | 0.056 | 0.163 | 0.159 | 0.133 | 1.000 | | | |
| 11 Reform activities | -0.051 | -0.031 | -0.031 | -0.025 | -0.035 | 0.234 | 0.365 | 0.322 | 0.316 | 0.388 | 1.000 | | |
| 12 Traditional activities | -0.114 | -0.091 | -0.098 | -0.101 | -0.100 | 0.197 | 0.307 | 0.275 | 0.302 | 0.253 | 0.634 | 1.000 | |
| 13 Computer activities | -0.096 | -0.112 | -0.091 | -0.092 | -0.090 | 0.118 | 0.134 | 0.313 | 0.060 | 0.169 | 0.245 | 0.234 | 1.000 |
| 14 Assessment variety frq | -0.135 | -0.131 | -0.129 | -0.121 | -0.128 | 0.165 | 0.253 | 0.176 | 0.197 | 0.289 | 0.502 | 0.452 | 0.286 |
| 15 Reading preparation | -0.081 | -0.086 | -0.099 | -0.100 | -0.077 | 0.093 | 0.024 | 0.063 | 0.000 | 0.140 | 0.103 | 0.031 | 0.147 |
| mean | 216.721 | 216.334 | 215.595 | 216.280 | 216.803 | 3.013 | 23.386 | 5.632 | 3.777 | 14.811 | 37.810 | 37.545 | 5.836 |
| stddev | 27.183 | 26.636 | 27.014 | 26.977 | 27.028 | 1.143 | 3.367 | 1.586 | 1.200 | 3.160 | 6.697 | 3.732 | 2.496 |

**Table A.6** (continued) NAEP 1998 Reading (n=757)

| Variable | 14 | 15 |
|---|---|---|
| 14 Assessment variety frq | 1.000 | |
| 15 Reading preparation | 0.097 | 1.000 |
| mean | 28.469 | 15.131 |
| stddev | 5.321 | 7.239 |

**Table A.7** NAEP Reading 2000 (n=867)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 PD in last year | 1 | | | | | | | | | | | |
| 2 DIV PREP | 0.094 | 1 | | | | | | | | | | |
| 3 ASSESS | 0.21 | 0.063 | 1 | | | | | | | | | |
| 4 TECH PREP | 0.151 | 0.211 | 0.204 | 1 | | | | | | | | |
| 5 DIV PREP | 0.2 | 0.285 | 0.251 | 0.321 | 1 | | | | | | | |
| 6 INST PREP | 0.208 | 0.238 | 0.277 | 0.337 | 0.448 | 1 | | | | | | |
| 7 READ PREP | 0.025 | 0.024 | 0.076 | 0.064 | 0.029 | 0.162 | 1 | | | | | |
| 8 REFORM ACTIVITIES | 0.245 | 0.075 | 0.49 | 0.227 | 0.308 | 0.313 | 0.061 | 1 | | | | |
| 9 TRAD ACTIVITIES | 0.175 | 0.077 | 0.402 | 0.154 | 0.259 | 0.193 | 0.072 | 0.553 | 1 | | | |
| 10 COMPUTER ACTIVITIES | 0.003 | 0.063 | 0.194 | 0.25 | 0.126 | 0.132 | 0.025 | 0.209 | 0.154 | 1 | | |
| 11 RRPCM1_1 | -0.021 | 0.098 | -0.045 | 0.019 | -0.078 | 0.138 | 0.096 | 0.006 | -0.029 | -0.075 | 1 | |
| 12 RRPCM2_1 | -0.034 | 0.101 | -0.06 | 0.007 | -0.077 | 0.116 | 0.094 | 0.005 | -0.029 | -0.06 | 0.908 | 1 |
| 13 RRPCM3_1 | -0.005 | 0.112 | -0.055 | 0.034 | -0.073 | 0.138 | 0.093 | 0.013 | -0.018 | -0.067 | 0.907 | 0.922 |
| 14 RRPCM4_1 | -0.038 | 0.084 | -0.061 | 0.004 | -0.073 | 0.125 | 0.086 | 0.001 | -0.026 | -0.066 | 0.914 | 0.92 |
| 15 RRPCM5_1 | -0.028 | 0.102 | -0.07 | 0.003 | -0.074 | 0.123 | 0.099 | -0.005 | -0.029 | -0.081 | 0.91 | 0.925 |
| mean | 3.117 | 2.016 | 16.414 | 7.765 | 6.112 | 21.029 | 0.624 | 15.006 | 24.277 | 7.205 | 213.279 | 214.027 |
| stddev | 1.142 | 0.694 | 3.137 | 2.127 | 1.664 | 3.001 | 1.246 | 2.815 | 2.289 | 2.939 | 29.122 | 29.327 |

**Table A.7** (continued) NAEP Reading 2000 (n=867)

| Variable | 13 | 14 | 15 |
|---|---|---|---|
| 13 RRPCM3_1 | 1 | | |
| 14 RRPCM4_1 | 0.909 | 1 | |
| 15 RRPCM5_1 | 0.921 | 0.918 | 1 |
| mean | 214.427 | 213.565 | 214.531 |
| stddev | 29.101 | 28.762 | 29.519 |

**APPENDIX B:** Miscellaneous Tables

**Table B.1** Student Demographics - Dummy Coding

| | Intercept | Black | Hispanic | Other | Female | Black female | Hispanic female | Other female |
|---|---|---|---|---|---|---|---|---|
| White male | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Black male | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Hispanic male | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Other male | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Female | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Black female | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| Hispanic female | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| Other female | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |

# REFERENCES


Amrein, A. L., & Berliner, D. C. (2002). High stakes testing, uncertainty and student learning. *Education Policy Analysis Archives, 10*(18).

Angus, D. L. (2001). Professionalism and the public good: A brief history of teacher certification. In J. Mirel (Ed.). Washington, CD: Thomas B. Fordham Foundation.

Arbuckle, J. L. (1994-1999). AMOS (Version 4.01 (Build 344)). Chicago, IL: SmallWaters, Corp.

Archer, J. (1999). Sanders 101. *Education Week.*

Ball, D. L. (1997). Developing mathematics reform: What we don't know about teacher learning -- but would make a good working hypothesis. In S. N. Friel & G. W. Bright (Eds.), *Reflecting on our work: NSF Teacher enhancement in K-6 mathematics* (pp. 77-111). Latham, MD: University Press of America.

Barton, R., & Oja, S. (1999). Predicting beginning teachers' feelings of efficacy. *Educational Reseach Quarterly.*

Borko, H., & Putnam, R. (1996). Learning to teach. In R. C. Calfee & D. C. Berliner (Eds.), *Handbook of Educational Psychology* (pp. 673-708). New York: Macmillan.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Camilli, G., Vargas, S., & Yurecko, M. (2003). Teaching children to read: The fragile link between science and federal education policy. *Education Policy Analysis Archives, 11*(15).

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis, 12*(3), 327-345.

Cohen, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform works.* New Haven: Yale University Press.

Cohen, D. K., & Spillane, J. (1993). Policy and practice: The relations between governance and instruction. In *Designing coherent educational instruction* (pp. 35-95). San Franciso: Jossey-Bass.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cohen, J. (2002). AM statistical software (Version 0.06.01): American Institutes for Research.

Cusick, P. A. (1992). *The educational system: its nature and logic*. New York: McGraw-Hill.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1), 1-56.

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Reseacher, 31*(9), 13-25.

Delpit, L. D. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review, 58*, 280-298.

Eberts, R. W., & Stone, J. A. (1984). *Unions and public schools: The effect of collective bargaining on American education*. Lexington, MA: Heath.

Education Counts Research Center. (2002). *Education Counts*, 2004, from http://www.edweek.org/rc/edcounts/

Ferguson, R. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal of Legislation, 28*(Summer 1991), 465-498.

Floden, R. E., & Buchmann, M. (1993). *Detachment and concern: Conversations in the philosophy of teaching and teacher education*. New York: Teachers College Press.

Foster, M., & Peele, T. (1999). Teaching and learning in the contexts of African American English culture. *Education and Urban Society, 31*(2), 177-189.

Fuhrman, S. H. (2001). Introduction. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 81-98). Chicago: The Unviersity of Chicago Press.

Goertz, M. E., Duffy, M. C., & with Le Floch, K. C. (2001). *Assessment and accountability systems in the 50 states: 1999-2000* (No. RR-046). Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Education Research, 66*(Fall 1966), 361-396.

Grossman, P. L. (1989). Learning to teach without teacher education. *Teachers College Record, 91*(2), 192-208.

Heath, S. B. (1983). *Ways with words: language, life, and work in communities and classrooms*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Hirsch, E. D. (1996). *The schools we need and why we don't have them* (1st ed.). New York: Doubleday.

Kennedy, M. M. (1998). *The effects of inservice programs on student achievement.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego.

Kennedy, M. M. (1999). The problem of evidence in teacher education. In R. Roth (Ed.), *The Role of the University in the Preparation of Teachers* (pp. 87-107). London: Falmer Press.

Kingdon, J. (1995). *Agendas, alternatives, and public policies*. Boston: Little, Brown and Co.

Labaree, D. F. (1997). *Public goods, private goods: The American struggle over educational goals*.

Linn, R. L. (2001). *The design and evaluation of educational assessment and accountability systems* (No. CSE Technical Report 539). Los Angeles: Center for the Study of Evaluation.

Linn, R. L., & Baker, E. L. (1998). School quality: Some missing pieces. *The CRESST Line: Newsletter of the National Center for Research on Evaluation, Standards, and Student Testing, 2-3*(7).

Linn, R. L., Baker, E. L., & Betebenner, D., W. (2002). Accountability systems, Implications of requirements of the No Child Left Behind Act. *Educational Researcher, 31*(6), 3-16.

Lipsky, M. (1980). *Street-level bureaucracy : dilemmas of the individual in public services*. New York: Russell Sage Foundation.

Lortie, D. C. (1975). *Schoolteacher: a sociological study*. Chicago,: University of Chicago Press.

McCartney, K., & Dearing, E. (2002). Evaluating effect sizes in the policy arena. *The Evaluation Exchange Newsletter, 8*(1), 4 & 7.

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*, 173-180.

McDonald, L. M., & Elmore, R. F. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis, 9*(2), 133-152.

Muncey, D. E., & McQuillan, P. J. (1996). *Reform and resistance in schools and classrooms*. New Haven: Yale University Press.

National Center for Research on Teaching and Learning. (1993). *Findings on learning to teach*. East Lansing: Michigan State Univerisity.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA.

National Reading Panel. (2000). *Teaching Children to Read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: NICHD.

No Child Left Behind Act of 2001, (2002).

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Education Policy Analysis Archives, 26*(3), 217.

Paige, R. (2002). *Meeting the highly qualified teachers challenge: The Secretary's annual report on teacher quality*. Washington, DC: U. S. Department of Education.

Paige, R. (2003). *Meeting the highly qualified teachers challenge: The Secretary's annual report on teacher quality*. Washington, DC: U. S. Department of Education.

Paige, R. (2004). *Meeting the highly qualified teachers challenge: The Secretary's annual report on teacher quality*. Washington, DC: U. S. Department of Education.

Pearson, P. D. (1996). Reclaiming the center. In *The First R: Every Child's Right to Read*: Teachers College Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Second edition ed. Vol. 1). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2005). HLM for Windows (Version 6.01h). Chicago, IL: Scientific Sotware International.

Raudenbush, S. W., Kasim, R., M., Eamsukkawat, S., & Miyazaki, Y. (1996). Social origins, schooling, adult literacy, and employment: Results from the National Adult Literacy Survey.

Ravitch, D. (2002, March 2002). *A brief history of teacher professionalism.* Paper presented at the White House Conference on Preparing Tomorrow's Teachers, Washington, DC.

Reckase, M. (2002, May 2002). *What factors in the preparation of teachers influence how much students learn?* Paper presented at the First International Education Conference, Famagusta, North Cyprus.

Riley, R. W. (2000). *The initial report of the Secretary on the quality of teacher preparation.* Washington, DC: U.S. Department of Education.

Rowan, B., Chiang, F.-S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 70*(October), 256-284.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of elementary schools. *Teachers College Record, 104*(8), 1525-1567.

Sanders, W., & Horn, S. (1998). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. In *Models for Teacher Evaluation.*

Sanders, W., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Sheff v O'Neill (Connecticut Supreme Court 1996).

Shulman, L. S. (1987). *Knowledge and teaching: Foundations of the new reform.*

Spellings, M. (2005). *Meeting the highly qualified teachers challenge: The Secretary's annual report on teacher quality.* Washington, DC: U. S. Department of Education.

Spillane, J. (2005). Standards deviation: How schools musunderstand education policy. *CPRE Policy Briefs, RB-43*(June 2005).

Supovitz, J. (2001). Translating teaching practice into improved student achievement. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 81-98). Chicago: The Unviersity of Chicago Press.

Tennessee Small School Systems v. McWherter (Supreme Court of Tennessee 1993).

Traub, J. (2002, April 7, 2002). The test mess. *The New York Times on the Web.*

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia : a century of public school reform.* Cambridge, Mass.: Harvard University Press.

Amendments to the Higher Education Act, (1998a).

Reauthorization of the Elementary and Secondary Education Act, (1998b).

Valli, L. (2000). Beginning Teacher Survey. In National Partnership for Education and Accountability in Teaching (Ed.). Columbia, PA: National Computer Systems, Inc.

Valli, L., Raths, J., Reckase, M., & Rennert-Ariev, P. (2001). *Beginning Teacher Survey Study: Progress report & preliminary analysis of findings.* College Park, MD: University of Maryland.

Valli, L., Raths, J., & Rennert-Ariev, P. (2001). *The Beginning Teacher Survey Study:.* Paper presented at the American Educational Research Association, Seattle, WA.

Valli, L., Reckase, M., & Raths, J. (2003). *Teacher education, program outcomes, teaching practice, and pupil achievement.* Paper presented at the Annual meeting of the American Educational Research Association, Chicago, IL.

Valli, L., Reckase, M., Raths, J., & Rennert-Ariev, P. (2001). *Beginning Teacher Survey Study final report.* College Park, MD: University of Maryland.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122.

What Works Clearinghouse. (2002). Overview. http://www.whatworks.ed.gov/whatwedo/overview.html.

Wilson, S., Darling-Hammond, L., & Berry, B. (2000). *Teaching policy: Connecticut's long-term efforts to improve teaching and learning.*

Wilson, S. M., Floden, R. E., & Ferrini-Mundy. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations*: Prepared for US Department of Education by the Center for the Study of Teaching and Policy in collaboration with Michigan State University.

Youngs, P., & King, M. B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly, 38*(5), 643-670.