

LEARNING TRAJECTORIES RELATED TO BIVARIATE DATA
IN CONTEMPORARY HIGH SCHOOL MATHEMATICS TEXTBOOK SERIES
IN THE UNITED STATES

A Dissertation
presented to
the Faculty of the Graduate School
University of Missouri – Columbia

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy

by
DUNG TRAN

Dr. James E. Tarr, Dissertation Supervisor

JULY 2013

© Copyright by Dung Tran 2013

All Rights Reserved

The undersigned, appointed by the Dean of the Graduate School, have examined the
dissertation entitled

LEARNING TRAJECTORIES RELATED TO BIVARIATE DATA
IN CONTEMPORARY HIGH SCHOOL MATHEMATICS TEXTBOOK SERIES
IN THE UNITED STATES

presented by Dung Tran,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that in their opinion it is worthy of acceptance.

Dr. James E. Tarr

Dr. Barbara Dougherty

Dr. Barbara J. Reys

Dr. Kathryn Chval

Dr. Lori Thombs

ACKNOWLEDGEMENTS

It takes a whole village to raise a child (African proverb)

I, a child in the profession of mathematics education, have been carefully raised for four years in my Ph.D. program in the U.S. I wish to fully express my thankfulness to all the people in the village who have encouraged, assisted, guided, and challenged me directly or indirectly.

To Drs. James Tarr, Barbara Dougherty, Barbara Reys, Kathryn Chval, and Lori Thombs, thanks for serving on my dissertation committee. Above all, I would like to express my deepest appreciation to my advisor, Dr. James Tarr, for your excellent feedback and caring. I could not count have many times I have got your encouragements, supports, and critical challenges to grow in the field. Specifically, it must be patience and sympathy for you to be a guide in my long journey from brainstorming research ideas to the final stage completing the dissertation. I could not say how hard the journey is without such supports.

To Dr. Barbara Reys, I am so grateful to work with you for two years at MU. I have learned, from you, how to think hard and think big, how to take care of other colleagues, and how to become a great mathematics educator and researcher. I appreciate your caring for me, especially arranging the internship at Western Michigan University (WMU) on my research interest.

To Drs. Robert Reys, Kathryn Chval, Oscar Chavez, Barbara Dougherty, Douglas Grouws, and John Lannin, I want to express my gratitude to all of you. You, mathematics education faculties, all have been setting examples while I am a student in your classes or through informal setting. I really appreciate all your time available for conversation about

scholarly as well as personal issues and your tolerance for having me accidentally stop by without appointment. That means to me a lot.

To Drs. Christine Spinka and Lori Thombs at the Statistics Department, thanks for serving on my committees and stimulating my interest in statistics. To Dr. Christian Hirsch at WMU, it has been an invaluable experience working with you on curriculum development. To Dr. Christine Franklin at the University of Georgia, thanks for your timely responses at the early stage of my dissertation and throughout my study helping clarifying the GAISE Framework.

To my colleagues at MU, I have been learned with you and from you. I might not name each and all of you, but I am thankful for making my time in MU enjoyable and meaningful. Ruthmae Sears, you always encouraged me to be confident in myself, shared your experience with me, and are also one of my closest collaboration partners. Special thanks go to Joann I and Victor Soria for your commitment to working on establishing the reliability of coding in my dissertation. I know that my dissertation has not been completed without such help.

To Carolyn Magnuson, I appreciate your helping me write from scratch. Your comments as an outsider always helped me to carefully consider how to make the writing understandable to native speakers. I also appreciate your unconditional support, encouragement and showing me the out-of-school culture around Columbia. I know that I will help others as one way to pay back the people who have supported me.

To the people of Columbia, MO and the U.S., thank you for your friendliness, I have always learned something new when working and talking with you. Special appreciation goes to Steve Rasmussen, who helped initiate bridging my education

experience from Vietnam to the U.S. It has been the most valuable experience I have ever had in my life.

To *Vietnam International Education Development* (VIED), thanks for supporting me the first two years in my Ph.D. program. The support comes from the Vietnamese people as a whole. I am always thinking of ways to create similar educational opportunities for my fellow citizens.

To VSA@MU, you have been established a great environment to help new students here. That is invaluable and we should be continued to do that. I owe special thanks to my roommate, Cao Tiến Đạt. You have been so sympathetic and helpful with me when I was so weak and could not afford to walk by myself. I would have suffered greatly without your help when I was far away from my family.

To my family: my parents, my siblings, nieces, and nephews — you have always been the great motivation to better myself, to be a great person, and to be a helpful human. It is indirect support, but its value is immeasurable. To my wife, Trần Nguyễn Tân Khoa, thanks for your patience and being with me in the critical time. It was tough, but we overcame together.

I am mature enough to become a villager, helping to raise other children from now!!!

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	xi
LIST OF FIGURES.....	xiii
ABSTRACT.....	xvii
CHAPTER 1.....	1
Rationale for the Study	1
Statement of the Problem	4
Purpose of the Study	5
Research Questions	6
Conceptual Perspectives.....	7
Content Analysis.....	7
Learning Trajectories.....	9
Instructional Tasks.....	10
GAISE Framework.....	11
Task-Technique-Theory Framework.....	12
Purpose and Utility Framework.....	14
Definitions.....	15
Learning Trajectories.....	15
Bivariate Relationships.....	16
Significance of the Study.....	16
Summary.....	18

CHAPTER 2: LITERATURE REVIEW.....	19
Learning Trajectories.....	19
Multiple Definitions of Learning Trajectories.....	20
Graphical Representations of Learning Trajectories	25
Examples of Forming and Using Learning Trajectories.....	27
Bivariate Relationships (Covariation) – Association and Correlation.....	32
Representing Covariation in Graphs.....	33
Interpret Covariation.....	36
Research involves categorical data.....	37
Research involves numerical data.....	44
Descriptions of the Progressions within the GAISE Framework.....	49
Two categorical variables.....	49
One categorical and one numerical variable.....	50
Two numerical variables.....	50
Textbook Analysis.....	53
Horizontal Approach.....	53
Vertical Approach.....	55
Summary.....	60
CHAPTER 3: METHODOLOGY.....	62
Selection of Textbook Sample.....	63
Criterion 1: The Series Elicits the Potential for In-depth Analyses of	
Differences in Sequencing and Organizing Bivariate Data Content....	65
Holt McDougal Larson	65

The University of Chicago School Mathematics Project.....	65
Core-Plus Mathematics Project.....	66
Criterion 2: The Series Explicitly Displays Learning Goals to Support Teachers Implementing the Curriculum.....	67
Holt McDougal Larson	67
The University of Chicago School Mathematics Project.....	67
Core-Plus Mathematics Project.....	68
Criterion 3: The Series is Currently Used in U.S. Educational Systems..	68
Holt McDougal Larson	68
The University of Chicago School Mathematics Project.....	68
Core-Plus Mathematics Project.....	68
Analyses of the Textbooks.....	69
Unit of Analysis.....	69
Coding.....	72
Learning trajectories.....	72
Task features.....	77
Summary of coding the coding scheme for bivariate data.....	80
Sample Application of the coding scheme.....	82
Learning trajectories.....	82
Task features.....	83
Data Analysis.....	85
Question 1	86
Question 2	89

Variability and Reliability.....	91
Summary.....	94
CHAPTER 4: ANALYSIS OF DATA AND RESULTS.....	96
Learning Trajectories	96
Distribution of Instances of Bivariate Data across Three	
Textbook Series.....	97
Descriptions of Learning Trajectories within Each Series.....	103
Holt McDougal Larson	104
The University of Chicago School Mathematics Project.....	121
Core-Plus Mathematics Project.....	140
Comparison of Textbook Series Learning Trajectories with Common	
Core State Standards for Mathematics Learning Expectations.....	162
Holt McDougal Larson	165
The University of Chicago School Mathematics Project.....	166
Core-Plus Mathematics Project.....	167
Comparison of Learning Trajectories with the GAISE Framework... ..	169
Holt McDougal Larson	169
The University of Chicago School Mathematics Project.....	172
Core-Plus Mathematics Project.....	175
Summary of Learning Trajectories for Bivariate Data.....	178
Task Features.....	179
Level of Mathematical Complexity.....	179

Distribution of levels of mathematical complexity across the three series.....	179
Examples of instances for the three levels of mathematical complexity.....	180
The GAISE Framework Coding.....	188
Four components of doing statistics.....	188
Developmental levels of the GAISE Framework.....	192
Association Between Levels of Mathematical Complexity and the GAISE Frameworks.....	195
Purpose and Utility Framework Coding.....	197
The Common Core State Standards-Mathematics: Mathematical Practices Coding.....	202
Summary of Task Features for Bivariate Data.....	207
CHAPTER 5: DISCUSSION, SUMMARY, AND RECOMMENDATIONS.....	209
Summary of the Study, Findings and Discussions.....	209
Purpose of the Study.....	210
Methodology.....	214
Results of the Study and Discussions.....	212
Learning trajectories.....	212
Task features.....	222
Limitations of the Study.....	227
Implications and Recommendations for Future Research.....	228
Curriculum Development.....	228

Content Analysis.....	230
Teacher Education.....	231
Future Research.....	232
Summary.....	234
References.....	235
APPENDIX.....	247
VITA.....	250

LIST OF TABLES

Table 1	<i>Factors to consider in content analysis of mathematics curricular materials...</i>	8
Table 2	<i>Combinations of two variables in bivariate relationships.....</i>	33
Table 3	<i>Characteristics of tasks at different levels of mathematical complexity</i>	77
Table 4	<i>Example of description of SMP and its coding framework</i>	80
Table 5	<i>Framework for coding bivariate data</i>	80
Table 6	<i>CCSSM learning expectation related to bivariate data at the high school</i>	88
Table 7	<i>The Mapping of Data Sources to Research Questions</i>	91
Table 8	<i>Reliability measures</i>	94
Table 9	<i>Instances related to bivariate data</i>	97
Table 10	Summary of learning trajectories for two numerical variables in the HML series	119
Table 11	Summary of learning trajectories for two numerical variables in the UCSMP series.....	137
Table 12	Summary of learning trajectories for two numerical variables in the CPMP series	160
Table 13	Frequency of instances addressing CCSSM LEs related to bivariate data across the three series.....	163
Table 14	<i>List of Learning Expectations found across the three series but not included in CCSSM... ..</i>	168
Table 15	<i>Means (SDs) of mathematical complexity across the three series.....</i>	180
Table 16	<i>Percentages of instances addressing the components of doing statistics in each of the three series</i>	188
Table 17	<i>Percentages of instances addressing the components of doing statistics in each of the three series.....</i>	189
Table 18	<i>Means (SDs) of Analyze Data across the three series</i>	193
Table 19	<i>Means (SDs) of Interpret Results across the three series.....</i>	194
Table 20	<i>Association between Analyze Data and Interpret Results.....</i>	195
Table 21	<i>Association between mathematical complexity and Analyze Data.....</i>	197
Table 22	<i>Association between mathematical complexity and Interpret Results.....</i>	197
Table 23	<i>Number of instances address across purpose and utility features in HML.....</i>	199

Table 23 <i>Number of instances address across purpose and utility features in UCSMP</i>	199
Table 25 Number of instances address across purpose and utility features in CPMP.....	200
Table 26 <i>Percentages of instances address zero to eight MPs for each series</i>	203

LIST OF FIGURES

<i>Figure 1.</i> GAISE Framework	13
<i>Figure 2.</i> Mathematics teaching cycle	20
<i>Figure 3.</i> Ladder-like linear sequences, branching tree diagram, and genetic inheritance tree models of learning trajectories.....	26
<i>Figure 4.</i> Big ideas tree model	27
<i>Figure 5.</i> An instance containing more than one task.....	70
<i>Figure 6.</i> An instance of super-problems with related sub-problems.....	70
<i>Figure 7.</i> An instance of problems without sub-problems.....	71
<i>Figure 8.</i> Instances of author examples and author texts	71
<i>Figure 9.</i> An instance that was not coded for the CC association	73
<i>Figure 10.</i> An instance that was not coded for the CN association	75
<i>Figure 11.</i> An instance that was not coded for the NN association	76
<i>Figure 12.</i> Sample coding for one instance	82
<i>Figure 13.</i> Number of instances related to bivariate data across the textbooks in each series.....	98
<i>Figure 14.</i> Distribution of the three combinations of bivariate data instances by chapters in the HML series.....	100
<i>Figure 15.</i> Distribution of the three combinations of bivariate data instances by chapter in the UCSMP series	100
<i>Figure 16.</i> Distribution of the three combinations of bivariate data instances by chapter in the CPMP series	101
<i>Figure 17.</i> Percentages of instances in the three combinations of bivariate data across three series	101
<i>Figure 18.</i> Types of functional models in the three series	103
<i>Figure 19.</i> Illustrating frequencies in a two-way table	105
<i>Figure 20.</i> Reading/constructing two-way tables.....	105
<i>Figure 21.</i> Probabilities of independent and dependent events.....	106
<i>Figure 22.</i> Learning trajectory for the CC association in the HML series	107
<i>Figure 23.</i> Comparing two distributions using back-to-back stem-and-leaf plot.....	108
<i>Figure 24.</i> Comparing distributions using box-and-whisker plots	109
<i>Figure 25.</i> Distinguishing among three types of studies.....	112

<i>Figure 26.</i> Learning trajectories for the CN association in the HML series.....	113
<i>Figure 27.</i> Using linear regression to make a prediction.....	116
<i>Figure 28.</i> Learning trajectories for the NN association in the HML series	121
<i>Figure 29.</i> Simulation of the likelihood for a Chi-square value with technology	123
<i>Figure 30.</i> Simpson's Paradox.....	125
<i>Figure 31.</i> Learning trajectories for the CC association in the UCSMP series	126
<i>Figure 32.</i> Comparing two distributions	127
<i>Figure 33.</i> Learning trajectories for the CN association in the UCSMP series	128
<i>Figure 34.</i> Examining of linear fitting with a moveable line	130
<i>Figure 35.</i> Two approaches of exponential fitting	132
<i>Figure 36.</i> Properties of the regression line.....	135
<i>Figure 37.</i> Learning trajectories for the NN association in the UCSMP series	139
<i>Figure 38.</i> Property of conditional probability of independent events.....	141
<i>Figure 39.</i> Using a randomization test to examine the difference in proportions	142
<i>Figure 40.</i> Comparing distributions.....	144
<i>Figure 41.</i> Experimental studies	147
<i>Figure 42.</i> Learning trajectories for the CC and CN associations in the CPMP series..	149
<i>Figure 43.</i> Pattern of chance.....	151
<i>Figure 44.</i> Spearman's and Kendall's rank correlations	155
<i>Figure 45.</i> Different types of bivariate relationships	156
<i>Figure 46.</i> Linearization of data.....	159
<i>Figure 47.</i> Learning trajectories for the NN association in the CPMP series	162
<i>Figure 48.</i> Learning trajectory of the CC association in terms of the GAISE Framework in the HML series	170
<i>Figure 49.</i> Learning trajectory of the CN association in terms of the GAISE Framework in the HML series.....	171
<i>Figure 50.</i> Learning trajectory of the NN association in terms of the GAISE Framework in the HML series.....	172
<i>Figure 51.</i> Learning trajectory of the CC association in terms of the GAISE Framework in the UCSMP series.....	173

<i>Figure 52. Learning trajectory of the CN association in terms of the GAISE Framework in the UCSMP series.....</i>	173
<i>Figure 53. Learning trajectory of the NN association in terms of the GAISE Framework in the UCSMP series.....</i>	174
<i>Figure 54. Learning trajectory of the CC association in terms of the GAISE Framework in the CPMP series</i>	175
<i>Figure 55. Learning trajectory of the CN association in terms of the GAISE Framework in the CPMP series</i>	176
<i>Figure 56. Learning trajectory of the NN association in terms of the GAISE Framework in the CPMP series</i>	178
<i>Figure 57. Proportions of levels of mathematical complexity across the three series...</i>	180
<i>Figure 58. Low-level mathematical complexity instance</i>	181
<i>Figure 59. Low-level mathematical complexity instance.....</i>	182
<i>Figure 60. Low-level mathematical complexity instances</i>	182
<i>Figure 61. Moderate-level mathematical complexity instance</i>	183
<i>Figure 62. Moderate-level mathematical complexity instance</i>	183
<i>Figure 63. Moderate-level mathematical complexity instance.....</i>	184
<i>Figure 64. High-level mathematical complexity instance</i>	185
<i>Figure 65. High-level mathematical complexity instance</i>	186
<i>Figure 66. High-level mathematical complexity instance.....</i>	187
<i>Figure 67. Questions for data analysis</i>	190
<i>Figure 68. A full statistical investigation process</i>	191
<i>Figure 69. Percentages of Analyze Data instances across the three series</i>	192
<i>Figure 70. Percentages of Interpret Results across the three series</i>	193
<i>Figure 71. Association between mathematical complexity and Analyze Data</i>	196
<i>Figure 72. Association between mathematical complexity and Interpret Results</i>	196
<i>Figure 73. Percentages of instances that address the purpose feature across the three series</i>	198
<i>Figure 74. Percentages of instances that address the utility feature across the three series</i>	199
<i>Figure 75. An instance that addresses the purpose but not utility feature</i>	201
<i>Figure 76. An instance that addressed neither the purpose nor the utility feature</i>	202

<i>Figure 77.</i> Percentages of instances that address the SMP in each of the three series	204
<i>Figure 78.</i> An instance that addressed MP1, MP2, MP3, MP4, MP6, and MP7	205
<i>Figure 79.</i> An instance that addressed MP5	206
<i>Figure 80.</i> An instance that addressed MP8	207

ABSTRACT

Bivariate relationships play a critical role in school statistics, and textbooks are significant in determining student learning. In recent years, researchers have emphasized the importance of learning trajectories (LTs) in mathematics education. In this study, I examined LTs for bivariate data in relation to the development of covariational reasoning in three high school textbooks series: Holt McDougal Larson (HML), The University of Chicago School of Mathematics Project (UCSMP), and Core-Plus Mathematics Project (CPMP). The LTs were generated by coding for the presence of variable combinations, learning goals, and techniques and theories. Task features were analyzed in relation to the GAISE Framework, NAEP mathematical complexity, purpose and utility, and the CCSSM Standards for Mathematical Practice.

The LTs varied by the presence, development, and emphases of bivariate content and alignment with the GAISE Framework and CCSSM. Across three series, about 80% to 90% of the 582 bivariate instances addressed two numerical variables. The CPMP series followed the GAISE's developmental progression for all combinations whereas UCSMP deviated for two categorical variables. All CCSSM learning expectations were found in HML and CPMP but not in UCSMP. At the same time, several bivariate learning expectations present in textbooks were not found in CCSSM. For the task features, few instances were at a high level of mathematical complexity and rarely included a *Collect Data* component. Analyses revealed the accordance of the GAISE and mathematical complexity frameworks. Research findings provide implications for curriculum development, content analysis, and teacher education, and challenge the notion of CCSSM-aligned curricula.

CHAPTER 1: INTRODUCTION AND RATIONALE FOR STUDY

Rationale for the Study

In today's modern world, people are surrounded by data. Television news, newspapers, and labels for grocery products contain data in some form; all require appropriate interpretation. Students need to be data-literate when graduating from high school in order to be critical consumers of information (Franklin, Kader, Mewborn, Moreno, Peck, Perry, & Scheaffer, 2007). Societal needs are a force shaping priorities in school mathematics curricula and, hence, there is a push to place a greater emphasis on statistical literacy, reasoning and thinking (Ben-Zvi & Garfield, 2004; Shaughnessy, 2007). Statistical literacy is especially important because "statistics has some claim to being a fundamental method of inquiry, a general way of thinking that is more important than any of the specific techniques that make up the discipline" (Moore, 1990, p. 134). Thus, in order to be productive citizens, it is essential for students to develop abilities for engaging in statistical inquiry.

In addition to societal needs, professional organizations are another force shaping priorities in school mathematics curricula. The National Council of Teachers of Mathematics' (NCTM) report of *Curriculum and Evaluation Standards for School Mathematics* (1989) and *Principles and Standards for School Mathematics* (PSSM) (2000) elevated probability and statistics from enrichment content to equal standing with other traditional domains such as Number Sense, Algebra, Geometry, and Measurement as the foundation for school mathematics. Specifically, NCTM (1989, 2000) recommended the introduction of probability and statistics concepts throughout the K–12 school years. In the years since the publication of the NCTM *Standards*, probability and statistics have become more important in the school mathematics curriculum of the U.S.

In line with aiming to promote statistical literacy, in 2007, the American Statistical Association proposed the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework (2007)*. The *GAISE Report* described a cohesive and coherent framework for statistical education at Grades PreK–12 in the U.S. It was the first professional report that put an exclusive focus on statistics education and provided important connections to mathematics education.

Yet another force influencing priorities in mathematics education was the recent joint efforts of the National Governors’ Association (NGA) and the Council of Chief State School Officers (CCSSO). In 2010, the NGA Center for Best Practices and the CCSSO released the *Common Core State Standards for Mathematics (CCSSM)*, which ostensibly serves as a de facto K–12 mathematics ‘national curriculum’ for the U.S. (Confrey & Krupa, 2012; Reys, Thomas, Tran, Kasmer, Newton & Teuscher, 2013). The primary purpose of the CCSSM is to provide a blueprint for the mathematics that students should learn in school (Heck, Weiss, & Pasley, 2011). The CCSSM identifies statistics as one of the major mathematics content areas that all students should have the opportunity to learn in school. Statistics is included in the Measurement and Data domain of the CCSSM at the elementary school level, in the Statistics and Probability domain at the middle school level, and in the Statistics and Probability conceptual category at the high school level. In addition, at the high school level, Statistics and Probability are closely connected to the conceptual categories of Functions and Modeling. Hence, according to the CCSSM, statistics holds a prominent position in the school curriculum for mathematics in the U.S.

Despite the important role that probability and statistics play in students' lives, there are inherent challenges in learning and teaching statistics (Ben-Zvi & Garfield, 2004). Given that textbooks are a strong determinant of what students experience in the classroom (Senk & Thompson, 2003; Stein, Remillard, & Smith, 2007), it is essential to examine how the learning of statistics is organized in textbook materials, yet there is a paucity of research examining statistical content in mathematics textbooks.

Curriculum standards, textbooks, and assessments collectively provide insights into what is considered important for students to learn. Curriculum standards represent an *intended curriculum* but do not constitute a *written curriculum*, *textbook curriculum* or *implemented curriculum*. Along with the intended curriculum, textbooks serve as another indication of what is important in school mathematics (Senk & Thompson, 2003; Stein et al., 2007). In conjunction with the *assessed curriculum*, textbooks serve accountability and control functions (Woodward, 1994). Of all the factors influencing the teaching and learning of mathematics, textbooks arguably play the most crucial role in helping students learn mathematics.

Although textbooks are not solely accountable for students' achievement, they do play an important role in student learning (Stein et al., 2007; Willoughby, 2010). Students draw upon textbooks as one of the primary resources of examples and tasks for learning (Reys, Reys, & Chavez, 2004; Stein et al., 2007). It is essential to examine the opportunities in textbooks for students to learn mathematics. However, there is limited formal research on curriculum analysis (Mesa, 2004).

Embedded in textbooks are lists of topics, a sequence of the topics, and instructional strategies, which serve as a resource for teachers' instructional decision

making (Grouws & Smith, 2000; Tarr, Chavez, Reys, & Reys, 2006). In textbooks, teachers find guidance about important topics and the sequencing of those topics. Examining the introduction and development of content evident in textbooks can identify the opportunities for students to learn. However, few studies have undertaken the systematic evaluation of statistical content in secondary mathematics textbooks.

The construct of learning trajectories appeared in recent research literature as a way to conceptualize the progression of student thinking in specific content (e.g., Clements & Sarama, 2004; Confrey, Maloney, Nguyen, Mojica, & Myers, 2009; Daro, Mosher, & Corcoran, 2011). Researchers have emphasized the role of learning trajectories in different aspects of mathematics education including curriculum development, instruction, and assessment (Battista, 2004; Clements & Sarama, 2004; Confrey et al., 2009; Daro et al., 2011). Research on learning trajectories has primarily been conducted for elementary and middle school mathematics (Battista, 2004; Clements & Sarama, 2004; Confrey et al., 2009). In contrast, there has been limited research about LTs for secondary mathematics. Moreover, there have been no formal studies regarding LTs for bivariate data, which is arguably one of the most important topics in high school statistics.

Statement of the Problem

The importance of textbooks and the lack of systematic curriculum analysis in regard to textbook content related to statistics and probability indicate a need for content analysis based on rigorous methodological standards (NRC, 2004). Such an analysis has the potential to provide information and guidance for textbook selection as well as curriculum design. This study addresses the lack of available information regarding

analysis of statistical content as presented in textbooks currently used in mathematics courses in the U.S.

There is also a need to examine existing textbook curriculum in terms of alignment with the CCSSM (Heck et al., 2011); such an analysis may provide curriculum developers' and teachers' perspective about the focus of the core content and what alignment with the CCSSM means. "Monitoring the extent and nature of alignment of a variety of curriculum materials with the CCSSM...will be revealing regarding the potential reach of the standards" (Heck et al. 2011, p.20). An analysis of existing curricular materials has the potential to inform the revision process of textbook materials.

Bivariate relationships are among the important statistical concepts in school curricula (CCSSI, 2010; Jones, Langrall, Mooney, & Thornton, 2004; NCTM, 2000). In particular, the CCSSM (2010) addresses the concept of *covariation* as one important content area at the high school level. Further, many problems in statistics relate to multivariate situations that involve covariational reasoning (Shaughnessy, 2007). This research examines constructs of bivariate relationships in secondary textbooks through learning trajectories and nature of tasks related to bivariate data.

Purpose of the Study

The purpose of the study is three-fold: (1) to analyze current secondary mathematics textbooks' portrayal of topics in bivariate relationships, (2) to assess the alignment of the textbooks' bivariate relationships with the CCSSM and the GAISE Framework as found in teacher's guide of the textbooks, and (3) to examine the nature of tasks related to bivariate data.

Research Questions

Given the importance of textbooks in determining the scope and sequence of bivariate data content, this study investigates the following research questions:

1. What learning trajectories for topics in bivariate data are represented in high school mathematics textbooks in the U.S.? How are those learning trajectories similar to and different from those embedded within the *Common Core State Standards for Mathematics* (CCSSM)? To what degree are the learning trajectories found in textbooks aligned with the developmental levels described in the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) *Report*?
 - a. What topics in bivariate data are addressed in the textbooks and what learning trajectories for bivariate data are evident in the textbooks?
 - b. How are connections made between the topics in bivariate data and topics in univariate data in the textbooks? How are connections made between bivariate data and other conceptual categories in mathematics at the high school level?
 - c. To what extent are the CCSSM standards for bivariate data at the high school level evident in the textbook materials, and to what extent does the approach to, and sequence of, content in the materials reflect the developmental progressions of the topics described by the GAISE Framework?

2. With respect to bivariate data, what is the nature of the instructional tasks presented in textbooks? Do the tasks provide opportunities for students to access the CCSSM's Standards for Mathematical Practice (SMP)?
 - a. What levels of mathematical complexity are required by the tasks related to bivariate data?
 - b. To what degree are the GAISE developmental levels reflected in the tasks related to bivariate data?
 - c. What is the quality level of the tasks in terms of purpose and utility?
 - d. How do the tasks provide opportunities for students to access the CCSSM's Standards for Mathematical Practice?

Conceptual Perspectives

Content Analysis

Rigorous and comprehensive curriculum analysis is crucial to meaningful evaluation of curricular effectiveness. In 2004, the National Research Council (NRC) provided recommendations about curricular evaluation, hereafter referenced as the NRC Report. The NRC Report identified three approaches to measuring curricular effectiveness: content analyses, comparative studies, and case studies. Situated at the heart of curricular evaluation, content analysis plays a pivotal role. Ensuring “comprehensiveness, completeness, and accuracy of topic” and considering “if the sequencing forms a coherent, logical, and age-appropriate progression” (p. 43) provides insightful information about content. The authors of the NRC Report recommended three dimensions to address in content analysis:

1. Clarity, comprehensiveness, accuracy, depth of mathematical inquiry and mathematical reasoning, organization, and balance (disciplinary perspectives).
2. Engagement, timeliness, and support for diversity, and assessment (learner-oriented perspectives).
3. Pedagogy, resources, and professional development (teacher- and resource-oriented perspectives). (p. 93)

While the learner-oriented and teacher- and resource-oriented perspectives of curricula are crucial in content analysis, the focus of this study is on the disciplinary component as presented in textbooks. Specifically, I consider “the importance, quality, and sequencing of mathematics content” (p. 40). The NRC Report further identified a number of factors to consider in content analysis that I have slightly modified for use in this study. The factors I considered include: (textbook) emphasis on context and modeling activities, the type and extent of explanations provided, problem solving, the use and emphasis on reasoning and sense making, the relationships among the mathematical strands, and the focus on calculation, symbolic manipulations, and conceptual development (Table 1).

Table 1

Factors to consider in content analysis of mathematics curricular materials (Adapted from NRC, 2004)

- | |
|--|
| <ul style="list-style-type: none"> • Listing of topics • Sequence of topics • Clarity, accuracy, and appropriateness of topic representation • Pace, depth, and emphasis of topics • Overall structure: integrated, interdisciplinary, or sequential • Types of tasks and activities, purposes, and level of engagement • Use of prior knowledge, attention to (mis)conceptions, and student strategies • Focus on conceptual ideas and algorithmic fluency • Emphasis on analytic/symbolic, visual, or numeric approaches • Types and levels of reasoning, communication, and reflection • Type and use of explanation • Form of practice • Approach to formalization • Use of contextual problems and/or elements of quantitative literacy |
|--|

Learning Trajectories

Learning comes about and accumulates over time; effective instruction is often based on information that came before and will come after the current learning goal. The construct of *learning progressions* in general and *learning trajectories* (LT) in mathematics education offer pathways for the development of student learning in mathematics. According to Daro et al. (2011), both learning progressions and learning trajectories are rooted in Simon's (1995) use of the term *hypothetical learning trajectory* (HLT) and consist of: "the learning goals that define the direction, the learning activities, and the hypothetical learning process— a prediction of how the students' thinking and understanding will evolve in the context of the learning activities" (p. 136).

Simply put, LTs are paths that lead to students' learning and understanding of increasingly complex mathematical concepts. They show key waypoints along the path in which students' knowledge and skills are likely to grow and develop in school subjects (Corcoran, Mosher, & Rogat, 2009). They "involve both the order and nature of the steps in the growth of students' mathematical understanding, and about the nature of the instructional experiences that might support them in moving step by step toward the goals of school mathematics" (Daro et al., 2011, p. 12).

Historically, LTs were *scope and sequence* charts that developers used when designing curricular materials (Daro et al., 2011). The scope and sequence of learning formed the backbone of curriculum and assessment. They were normally based on the logic of mathematics and on the conventional wisdom of practice that the profession thinks school mathematics should be. Learning trajectories go further to include testable hypotheses related to their practicality in classroom with support from empirical data on

student learning. The hypotheses are validated when the learning of the content is accomplished. In this sense, LTs are “empirical choices of when to teach what to whom” (Daro et al., 2011, p. 12)

Within this study, the focus is on hypothetical learning trajectories, not on empirical validation. I seek to provide insight as to how curriculum developers and textbook authors choose to include topics related to bivariate data, how the topics are sequenced, and how the topics build upon each other. In addition, I examine the connections among different strands in mathematics (e.g., functions and statistics) and in what way curriculum developers link these strands.

Instructional Tasks

The construct of instructional tasks has been the focus of several studies. “An instructional task is an activity engaged in by teachers and students during classroom instruction that is oriented toward the development of a particular skill, concept, or idea” (Stein & Lane, 1996). In the Mathematical Task Framework, Stein and Lane provided a series of task variables and associated factors. In particular, instructional tasks pass through three phases: (1) as represented in curriculum/instructional materials, (2) as set up by the teacher in the classroom, and (3) as implemented by students during the lesson. For the scope of this study, I focus on instructional tasks that are represented in textbooks with the assumption that there will be transformations of the tasks in subsequent phases.

Furthermore, instructional tasks (as presented in textbooks) are examined in terms of *cognitive demands* (Stein & Lane, 1996). The cognitive demands facet refers to the kinds of thinking processes that are involved in solving each task. It is related to the level and kind of cognitive effort that students make when engaging in the tasks. The cognitive

demands facet of tasks is approached somewhat differently than that of Stein and Lane (1996) by the National Center for Education Statistics (NCES, 2007). They provided a framework that identifies three levels of mathematical complexity: low complexity, moderate complexity, and high complexity. These levels specify the cognitive demands of the items appearing on the National Assessment of Educational Progress (NAEP). The specific characteristics of each level are subsequently described; NCES framework for mathematical complexity was used to code the levels of cognitive demands in each task presented in the textbooks analyzed.

Along with standards for content, the CCSSM contains eight mathematical practice standards that are included in students' learning process. The eight practices describe expectations that students:

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precisions.
7. Look for and make use of structures.
8. Look for and express regularity in repeated reasoning. (pp. 6–8)

In this study, I carefully examine the features of tasks in order to gauge whether they offer the opportunity for students to learn the SMPs.

GAISE Framework

The GAISE Framework, a 6 x 3 matrix, lists four key components of the

statistical process and provides three developmental levels for each. More specifically, the first four rows identify the four process components when learning statistics including: formulate question, collect data, analyze data, and, interpret results; the fifth and the sixth rows relate to the nature of variability and the focus of variability. The columns represent three developmental levels: students begin awareness of the statistics question distinction, students increase awareness of the statistics question distinction, and students can make the statistics question distinction (for detailed descriptions of the framework, see Figure 1). Using the GAISE Framework, I examined each learning task and identified which components of statistics were addressed and the task's developmental level. Further, I used the progression for bivariate data suggested in the GAISE Framework to compare with those found in the textbooks.

Task – Technique – Theory Framework (TTT)

Artigue (2000, 2002) based on Chevallard's theory of practices, built a model consisting of *task*, *technique*, and *theory*. In general, *task* refers to something that needs to be done, to be undertaken; specifically, the task often has mathematical objects such as numbers and relations embedded within it. *Technique* is "a complex assembly of reasoning and routine work" that is used to solve the tasks (Artigue, 2002, p. 248). *Technique* is a method to carry out the process required of the tasks. Hence, the technique can be a simple routine process or a complex reasoning process when the solution is not readily available. *Theory* serves as a discourse explaining or justifying the techniques. It might be a cognitive structure constructed to serve as a warrant for validating techniques. Using this framework, I identify techniques that are likely to be employed when solving the tasks related to bivariate data suggested by the curriculum/textbook authors or on the

Process Component	Level A	Level B	Level C
I. Formulate Question	Beginning awareness of the statistics question distinction Teachers pose questions of interest Questions restricted to the classroom	Increased awareness of the statistics question distinction Students begin to pose their own questions of interest Questions not restricted to the classroom	Students can make the statistics question distinction Students pose their own questions of interest Questions seek generalization
II. Collect Data	Do not yet design for differences Census of classroom Simple experiment	Beginning awareness of design for differences Sample surveys; begin to use random selection Comparative experiment; begin to use random allocation	Students make design for differences Sampling designs with random selection Experimental designs with randomization
III. Analyze Data	Use particular properties of distributions in the context of a specific example Display variability within a group Compare individual to individual Compare individual to group Beginning awareness of group to group Observe association between two variables	Learn to use particular properties of distributions as tools of analysis Quantify variability within a group Compare group to group in displays Acknowledge sampling error Some quantification of association; simple models for association	Understand and use distributions in analysis as a global concept Measure variability within a group; measure variability between groups Compare group to group using displays and measures of variability Describe and quantify sampling error Quantification of association; fitting of models for association

Process Component	Level A	Level B	Level C
IV. Interpret Results	Students do not look beyond the data No generalization beyond the classroom Note difference between two individuals with different conditions Observe association in displays	Students acknowledge that looking beyond the data is feasible Acknowledge that a sample may or may not be representative of the larger population Note the difference between two groups with different conditions Aware of distinction between observational study and experiment Note differences in strength of association Basic interpretation of models for association Aware of the distinction between association and cause and effect	Students are able to look beyond the data in some contexts Generalize from sample to population Aware of the effect of randomization on the results of experiments Understand the difference between observational studies and experiments Interpret measures of strength of association Interpret models of association Distinguish between conclusions from association studies and experiments
Nature of Variability	Measurement variability Natural variability Induced variability	Sampling variability	Chance variability
Focus on Variability	Variability within a group	Variability within a group and variability between groups Covariability	Variability in model fitting

Figure 1. GAISE Framework. From *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report – A Pre-K–12 Curriculum Framework* by C. Franklin, G. Kader, D. Mewborn, J. Moreno, R. Peck, M. Perry, & R. Scheaffer, 2007, Alexandria, VA: American Statistical Association, pp. 14-15. Copyright 2007 by the American Statistical Association. Reprinted with permission.

logic of mathematics. The data to be gained from analyzing tasks, techniques and theories provided information about learning goals, learning activities, and the hypothetical learning processes for the learning trajectories. In particular, analyzing the set of tasks related to bivariate relationships provides data describing the learning activities; the techniques and theories provide information about what is to be gained after solving the tasks and the process in which students engage in order to achieve the goals.

Purpose and Utility Framework

Ainley and colleagues (Ainley, Bills, & Wilson, 2003; Ainley & Patt, 2002; Ainley, Patt, & Hansen, 2006) provided a framework for the development of instructional tasks. The framework includes two aspects: the *purposeful* and *utility* features of tasks. The *purposeful* feature of tasks relates to the outcome of completing the task for the learner. A *purposeful* task provides challenge and relevance and creates the necessity for the learner to use and/or apply the target knowledge to create a meaningful outcome (Ainley et al. 2003; Ainley & Patt, 2002; Ainley et al., 2006). The *utility* feature of tasks relates to “knowing how, when, and why that mathematical idea is useful” (Ainley et al., 2003, p.195). Students know how the mathematical concepts are used as tools to solve a set of problems.

In this study, I use the framework developed by Ainley and colleagues to code the purpose and utility features of each task to evaluate how the bivariate data tasks are set up in the textbooks. In particular, I examine tasks to determine if, when solving the task, students are afforded opportunities to understand when, how, and why the mathematical ideas are used (utility) and if the task produces an end product (purpose). Data about the purpose and utility features of tasks provide another source of information to add to

mathematical complexity, statistical components, developmental levels, and mathematical practices when analyzing the multiple aspects of the opportunities textbooks provide for students to learn bivariate data content.

Definitions

In this section, I provide a list of terms that are used throughout this study. The terms are operationalized to fit the purpose of this study. They are rooted in the disciplines of mathematics and statistics as well as in mathematics education, and serve to facilitate communication and understanding in the field.

Learning Trajectories

Within the scope of this study, learning trajectories are considered at two levels: *macro level* and *micro level*. At the *macro level*, learning trajectories include big ideas (Baroody et al., 2004), an ordered network of constructs that students are likely to move through when learning particular content (bivariate data, in this study). At this level, they are extensions of the scope and sequence of topics “embedded within conceptual corridors” (Confrey et al., 2009, p. 346). At the *micro level* (lesson level), learning trajectories are:

descriptions of children’s learning in a specific mathematical domain, and a related conjectured route through a set of instructional tasks designed to engender those mental processes or actions hypothesized to move children through a developmental progression of levels of thinking, created with the intent of supporting children’s achievement of specific goals in that mathematical domain. (Clements & Sarama, 2004, p. 83)

Given that the order of units, lessons, and tasks in the textbooks reflect the flow of the development of a concept, examining the sequence reveals learning trajectories.

Bivariate Relationships (Covariation)

Bivariate data are pairs of linked variables. Bivariate relationships or covariation concerns correspondence of variations; that is, how two variables change together.

“*Statistical covariation* refers to the correspondence of variation of two statistical variables that vary along numerical scales” (Moritz, 2004, p. 228). It includes the *correlation* when dealing with numerical variables, and *association* for categorical variables. Reasoning about covariation is the process of translation among raw numerical data, graphical representations, and verbal statements about statistical covariation and causal association (Moritz, 2004). Another process might involve using mathematical tools to interpret the association of data or fitting the data to a functional equation (Moritz, 2004). These collectively serve as a framework to examine the concept of covariation in student learning.

Significance of the Study

Results from this study describe how the content of bivariate data is treated in U.S. secondary mathematics textbooks. The study has the potential to contribute to the understanding of different approaches to teaching bivariate data to foster student learning. An additional benefit is the information the study provides regarding the alignment between current curricular materials with the CCSSM; this information might help curriculum developers bridge the gap when reviewing and developing the next generation of high school mathematics textbooks.

A careful analysis of learning trajectories in mathematics textbooks can shed additional light on different approaches used when teaching and learning the topics. In doing so, the study serves as a means to evaluate curriculum coherence. The results might also provide insight for future researchers to compare multiple hypothetical learning trajectories. Such information has the potential to offer background for further research focusing on advantages and disadvantages of the trajectories, to consider whether the trajectories are suitable for the targeted students. In addition, results could provide useful guidance for developing a seamless set of curricular materials for student learning of topics in high school statistics.

It is predicted that the CCSSM will have a major impact on revision of curriculum materials and curriculum design projects, and in turn influence teaching and learning in the U.S. (Heck et al., 2011). To assess the influence of the standards in mathematics education, it is critical to address the alignment of curriculum materials to the intent of the standards. This study can help bridge the gap in the research about the meaning of alignment and, in turn, to examine its influence on mathematics education. In addition, the results of this study can provide to curriculum developers with the suggestion about the development of bivariate data the CCSSM addresses and how to present the content in order to support meaningful learning.

As the NRC (2004) noted, when evaluating a curriculum it is not valid to draw conclusions without first analyzing the mathematical content. This analysis can address the content portion in program evaluation. This research can provide information about content storylines in bivariate data when researching the impact of the curriculum on student learning.

Summary

There are numerous high school textbooks series available in the U.S. The quality of content, sequence, and presentation is important in helping teachers make instructional decisions and foster student learning. Consequently, there is an immediate need to conduct textbook analysis particularly in key topics in school mathematics such as bivariate data. However, there has been a scarcity of research on this topic. Therefore, this study addresses an existing gap in the research literature and has the potential to inform curriculum developers as they revise and design new materials. Curriculum evaluators will find the study useful when addressing fidelity of implementation of the curriculum. Moreover, this study can provide insight into the coherence and rigor of the content. In turn, it can offer teachers a perspective in the interpretation of the CCSSM, and what alignment with the CCSSM means in practice. Finally, it adds to the research on content analyses in mathematics education using the construct of learning trajectory as the primary focus of this study.

CHAPTER 2: LITERATURE REVIEW

A wide range of literature informed this study of learning trajectories for bivariate data in high school textbooks. In particular, research related to learning trajectories helped formalize the components of learning trajectories and informed the selection of data sources collected from curricular materials. Furthermore, research about bivariate relationships, covariation, helped to set the boundary for the inclusion and exclusion of tasks to be coded and served as a repertoire of learning expectations, techniques, and strategies involved in solving bivariate data tasks. In addition, the learning trajectories related to bivariate data suggested in the GAISE Framework are summarized. Finally, the literature related to textbook analysis helped inform the research design of this study.

In this chapter, I summarize how different researchers conceptualize the construct of learning trajectories (LTs) and use examples to illustrate their views. I also discuss how researchers link LT with many aspects including curriculum development, instruction, and assessment. Next, I examine research related to students' covariational reasoning in different fields. Finally, I conclude the chapter with a summary of the research methods employed in the analyses of curriculum materials.

Learning Trajectories

In this section, I describe the construct of learning trajectories (LTs). In addition, I discuss definitions of LTs, their origins, prominent models of the construct, and the ways the construct is utilized in conjunction with other aspects of mathematics education that are aimed toward improving student learning.

Multiple Definitions of Learning Trajectories

The construct of LTs is rooted in Simon's (1995) notion of *hypothetical learning trajectory* (HLT). Simon placed the construct at the core of the *mathematics teaching cycle*, using it to frame mathematics pedagogy from the constructivist perspective, which emphasizes learning over teaching. In the cycle, the reflective relationship between teachers' design of tasks and consideration of students' thinking and learning becomes apparent. According to Simon, the teacher uses knowledge of mathematics and his or her hypotheses about students' understanding to frame the trajectory; the assessment of students' knowledge helps modify the trajectory as learning progresses (Figure 2).

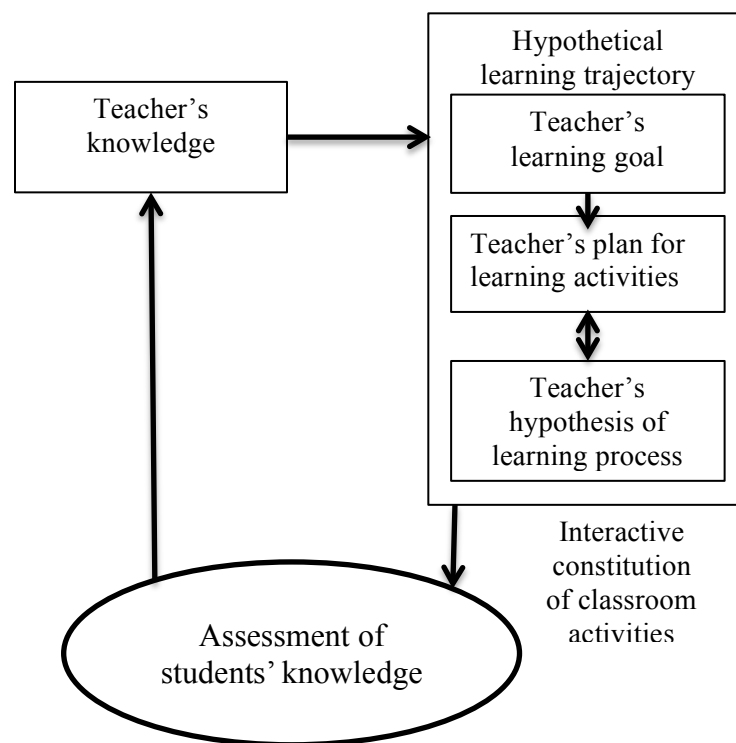


Figure 2. Mathematics teaching cycle. Adapted from “Reconstructing mathematics pedagogy from a constructivist perspective,” by M. A. Simon, 1995, *Journal for Research in Mathematics Education*, 26, p. 136. Copyright 1995 by Copyright Clearance Center, Inc.

According to Simon (1995), an HLT consists of three components: “the learning goal that defines the direction, the learning activities and the hypothetical learning

process – a prediction of how the students’ thinking and understanding will evolve in the context of learning activities” (p. 136). From the learning goal, the teacher plans the instructional activities or sequence to address the goal and determines the student learning process accordingly. An HLT “refers to the teacher’s prediction as to the path by which learning might proceed” (p. 135). However, because the teaching may not go as predicted, the teacher might change the initial goal and plan based on an assessment of students’ actual learning.

To develop a meaningful HLT, it is necessary to specify the goals in terms of knowledge and skills the students will apply. It is not that the teacher always pursues one goal or one trajectory at a time, but it underscores the importance of having a learning goal and a rationale for instructional decision-making, and for developing a hypothetical trajectory of students’ thinking processes. Teachers might pose a task, but the students’ experiences when they engage in the task determine their actual learning. Teachers plan for classroom learning activities; however, interactions with students may create a learning experience that is different from the planned experience. There is a modification of the teacher’s ideas and knowledge as he or she makes sense of what is actually occurring in the classroom. The assessment of student thinking brings about a further adaptation in the teacher’s knowledge and understanding that brings about the creation of a new or modified HLT in the mathematics teaching cycle.

Building on Simon’s (1995) HLT construct, Clements and Sarama (2004) defined learning trajectories as:

descriptions of children’s thinking and learning in a specific mathematical domain and a related, conjectured route through a set of instructional tasks designed to

engender those mental processes or actions hypothesized to move children through a developmental progression of levels of thinking, created with the intent of supporting children's achievement of specific goals in that mathematical domain. (p. 83)

According to Clements and Sarama (2004), a LT includes three components: learning goals, a developmental progression, and a sequence of instructional tasks. Starting from a learning goal, researchers constructed student-learning models to illustrate the developmental progressions through the levels of "sophistication, complexity, abstraction, power, and generality" (p. 83). From the model, key tasks are developed to promote students' targeted learning. By interacting with objects and actions inherent in the tasks, students can achieve the desired learning goal. The tasks are sequenced to correspond to the model developed and "constitute a particularly efficacious educational program" (p. 84). Clements and Sarama used LTs as paths the student thinking process is likely to follow and that have the likelihood of happening when learning specific mathematical concepts.

In line with Simon's (1995) conceptualization, Clements and Sarama (2004) asserted that the LT of specific topics is not completely known in advance; rather it is a conjectured path that provides important key points throughout the student learning process and an ongoing construction of students' model of learning when teachers/researchers interact with the students. From a constructivist perspective, the LT serves as a cognitive model for student learning that leads to building the instructional sequence. However, the LT might be viewed from a social perspective through a reconceptualization of the construct as "a sequence (or set) of (taken-as-shared)

classroom mathematics practices that emerges through interaction (especially through classroom discourse – with the proactive involvement of the teacher)” (p. 85). Thus, learning is created within the social context of the classroom.

Standing on an empirical perspective, Confrey et al., (2009) used the LT construct as a basis for their synthesis “to unpack complexity by revealing characteristics of gradual student learning over time” (p. 346). In that sense they defined learning trajectory as:

A researcher-conjectured, empirically supported description of the ordered network of constructs a student encounters through instruction (i.e. activities, tasks, tools, forms of interaction and methods of evaluation), in order to move from informal ideas, through successive refinements of representation, articulation, and reflection, towards increasingly complex concepts over time. (p. 346)

For this definition, they assume multiple possible paths and various obstacles that students typically encounter. They highlight the characteristics of LT: researcher-conjectured, empirically supported, and successively refined through instruction. Confrey and Maloney (2010) stated that the LT is not

a hard-and-fast order in which topics must be learned for students to be successful, they permit specification, at an appropriate and actionable level of detail, of ideas students need to know during the development and evolution of a given concept over time. (p. 1)

Studying LT within the context of classroom activities, Steffe (2004) described the LT as a co-product of teachers and students interacting with each other. That is, they

are both inquirers when participating in the same world. He asserted, “Through the construction of learning trajectories that are co-produced by children, it is possible to construct learning trajectories of children that include an account of one’s own ways and means of acting and operating as a teacher” (p. 130). According to Steffe, the LT involves intensive interaction with children and is abstracted from acts of teaching:

A learning trajectory of children includes a model of their initial concepts and operations, an account of the observable changes in those concepts and operations as a result of the children’s interactive mathematical activity in the situations of learning, and an account of the mathematical interactions that were involved in the changes. (p. 131)

Researchers from various cultures have conceptualized the construct of learning trajectories, but they use different terminologies. Gravemeijer (2004), in the tradition of the Realistic Mathematics Education (RME) group, provided the construct of *local instruction theories*, which is quite similar to the construct of learning trajectories. Consistent with Simon’s (1995) HLT, a *conjectured local instruction theory* is composed of three components:

(a) learning goals for students, (b) planned instructional activities and the tools that will be used, and (c) a conjectured learning process in which one anticipates how students’ thinking and understanding could evolve when the instructional activities are used in the classroom. (pp. 109–110)

The *local theory* serves as guidance for teachers to frame their instruction so that the activities help develop students’ current way of thinking into increasingly sophisticated ways of reasoning.

Graphical Representations of Learning Trajectories

In addition to narrative definitions and descriptions of LTs, several researchers have graphically conceptualized LTs. For example, Lesh and Yoon (2004) conceptualized learning trajectories, which they referred to as *model-eliciting activities*, as an inverted genetic inheritance tree model (Figure 3):

problem-solving situations in which goals include developing more powerful constructs or conceptual systems. Therefore, significant conceptual developments occur because students are challenged to repeatedly express, test, and revise their own current ways of thinking—not because they were guided along a narrow conceptual path toward (idealized versions of their teacher’s ways of thinking) (p. 205)

Lesh and Yoon (2004) criticized the ladder-like model of learning trajectories and pointed out the limitations of branch-tree models (Figure 3). The ladder-like approach – moving from informal to sophisticated understanding of specific constructs – is useful for curriculum developers whose goal is to guide students toward deeper understanding of important concepts in mathematics. However, the ladder-like approach has inherent problematic assumptions, such as the implication that development occurs linearly and bounded into a neat sequence. In contrast, branching tree diagrams have the advantage of fitting emerging ways of thinking; thus, allowing teachers to be flexible and adapt the activities. In addition, teachers may take both cognitive and social aspects into account. However, with the branching tree approach, it might be the case that “one perspective is treated as being more ‘politically correct’ than others” (p. 208). For example, there is the danger that some student thinking is rejected because the community does not favor it.

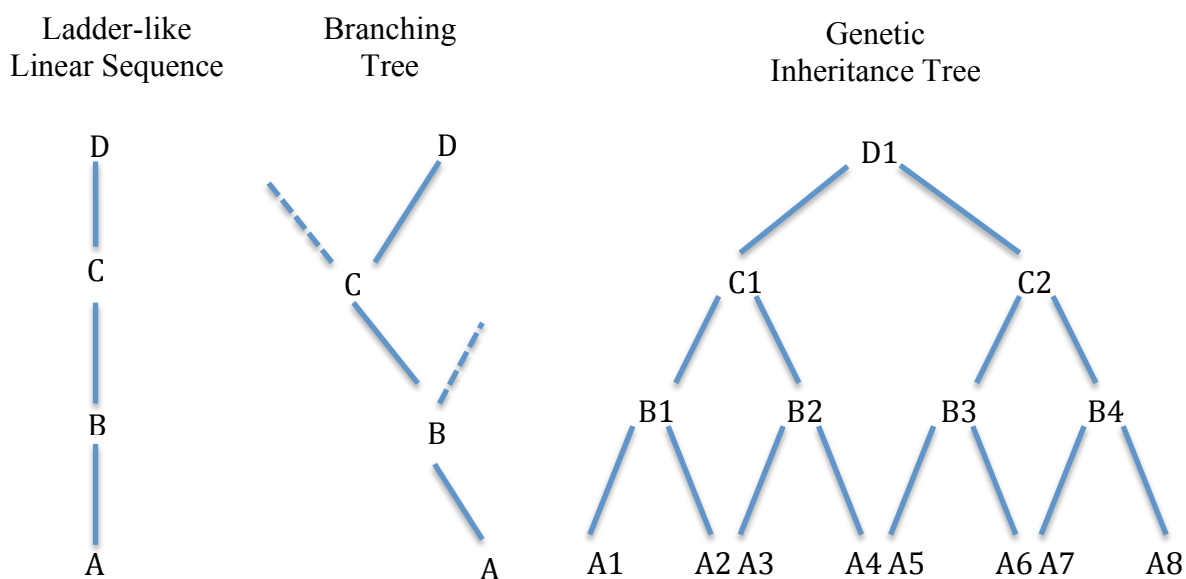


Figure 3. Ladder-like linear sequences, branching tree diagram, and genetic inheritance tree models of learning trajectories. Adapted from “Evolving communities of mind-in which development involves several interacting and simultaneously developing strands,” by R. Lesh & C. Yoon, 2004, *Mathematical Thinking and Learning*, 6, p. 225. Copyright 2004 by Lawrence Erlbaum Associates, Inc.

Along with attending to the role of teachers as guiding students along paths that lead to target goals, Lesh and Yoon suggested shifting the focus to activities that involve students in specifying learning trajectories. Hence, the LT resembles an inverted genetic inheritance tree (Figure 3), emphasizing the integration and differentiation of diverse ways of thinking.

Consistent with Lesh and Yoon’s (2004) perspective of the generic inheritance tree, Baroody, Cibulskis, Lai, and Li (2004) presented the concept of *big ideas* as guidance for hypothetical learning trajectories (HLTs). Baroody et al. suggested focusing on big ideas to make HLTs more practical. They modified the generic inheritance tree model proposed by Lesh and Yoon (2004) to create big ideas model of learning trajectory (Figure 4). This model links and integrates multiple branches of the inheritance tree. Thus, one large (overarching) idea can serve as the connector to several big ideas in

different content areas; in turn, each big idea can do the same for topics within the content.

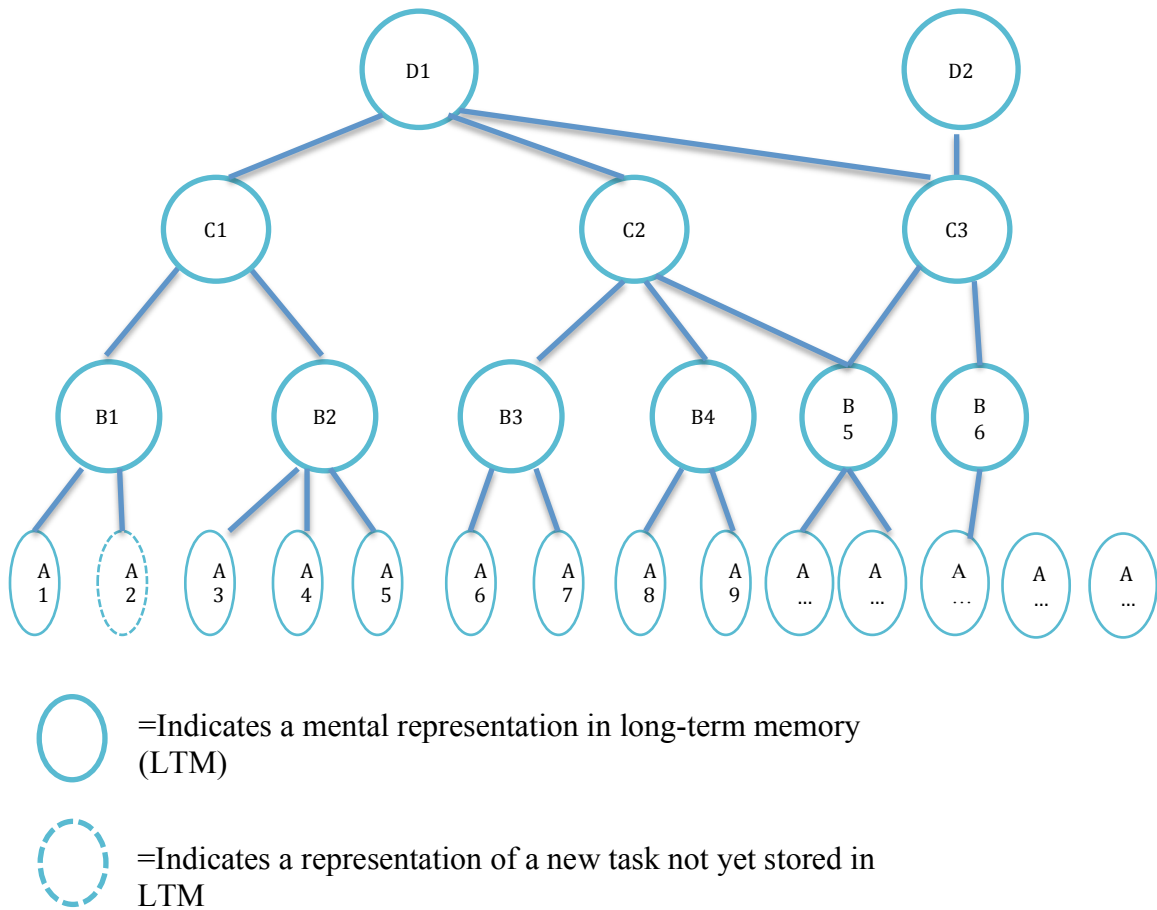


Figure 4. Big ideas tree model. Adapted from “Comments on the use of learning trajectories in curriculum development and research,” by A. J. Baroody, M. Cibulskis, M. –I. Lai, & X. Li, 2004, *Mathematical Thinking and Learning*, 6, p. 255. Copyright 2004 by Lawrence Erlbaum Associates, Inc.

Examples of Forming and Using Learning Trajectories

Many researchers use the construct of LT to operationalize student learning in early mathematics such as spatial sense (e.g., Clements, Wilson, & Sarama, 2004) and major topics in mathematics such as rational number reasoning (e.g., Confrey et al., 2009). Clements et al. (2004) operationalized the genesis of an HLT in young children’s learning of the composition of geometric shapes. They described the process of creating

an HLT and instrumentation to assess student learning along the developmental progression within the trajectory. Building on previous research on the geometric shapes topics, they developed a tentative developmental progression and relevant activities to guide student learning. Then they tested both developmental progression and instruments through a series of studies including formative case studies and summative studies of a larger number of participants. The process included both teaching and learning throughout the processes of designing, testing, recognizing variations, and further refinement of the initial HLT.

In another program of research, Confrey and colleagues (e.g., Confrey & Maloney, 2010; Confrey et al., 2009) described the sophisticated process leading to the genesis of *equipartitioning* learning trajectory. They started from an exhaustive research synthesis examining the theoretical claims and empirical evidence about equipartitioning, where they found areas of consensus in the field of mathematics education. Their synthesis of the research determined that the process of forming a LT passes through three phases: (a) using cases to build an initial learning trajectory; (b) iteratively developing the trajectory through the process of research synthesis and empirical investigation involving clinical and think-aloud interviews; and (c) designing assessment item response analysis and rubric development moving from small-scale to large-scale studies. Confrey et al. focus was on generating assessment tasks rather than developing curricular materials based on LTs research.

Focusing on curriculum design, the RME *local instruction theory* approach (e.g., Gravemeijer, 2004) is similar to the process of design experiments. Researchers formulate HLTs by identifying a possible learning route and specific means (learning

activities) to support that learning by using the historical development of mathematical concepts as a heuristic, and then taking students' informal solution strategies into account. A "best-case" instructional sequence is formed through several phases of refinement beginning with designing preliminary tasks and refining them in a series of processes of implementation in various classroom settings. This process leads to a local instruction theory that classroom teachers can use to construct trajectories that are sensitive to their student learning.

In the tradition of RME, Gravemeijer (2004) used an example of addition and subtraction up to 100 to illustrate characteristics of the local instruction theory. Instead of a ready-made plan or an instructional sequence that works, the research "offers teachers an empirically grounded theory on how a certain set of instructional activities can work" (p. 105). This process takes into account ongoing personal input from students about their learning; thus, placing student learning rather than curriculum materials as the guide for learning activities.

Previous studies report the process of forming LTs and use it in different phases of mathematics education. Indeed, learning trajectories have an essential role in improving learning and teaching of mathematics. Daro et al. (2011) stated, "We have concluded that learning trajectories hold great promise as tools for improving instruction in mathematics, and they hold promise for guiding the development of better curriculum and assessments as well" (p. 13).

Several researchers have provided models of using the LT to inform classroom instruction at the elementary level (Clement & Sarama, 2004; Gravemeijer, 2004; Simon, 1995). Learning trajectories are not deterministic; some LTs are more likely than others

to be followed and are more productive than others. Instruction that capitalizes on the “best case” learning trajectory will be the most effective in helping students and informing teachers about what to expect from students (Gravemeijer, 2004; Daro et al., 2011). A well-developed learning trajectory provides an empirical basis for the choice of what to teach, to whom, and at what time. They “offer a stronger basis for describing the interim goals that students should meet in if they are to reach the common core college and career-ready high school standards” (Daro et al., 2011, p. 12). With knowledge of learning trajectories, teachers are well positioned to understand the evidence of learning that can be derived from students’ work and to use such information to inform their classroom instruction.

In addition to the development and implementation of learning trajectories it is important to look at the role of LT in assessment. Battista (2004) provided the model of cognition-based assessment (CBA) for elementary school mathematics. Cognition-based assessment details the cognitive underpinnings of the progression of learning specific content, and emphasizes the developmental progression aspect of learning. Battista provided a conceptual framework based on levels of sophistication ranging from informal, pre-instructional reasoning to formal mathematical concepts. Under this model of assessment, teachers need to understand cognition-based research on students’ learning and use the results to determine and monitor the development of their own students’ reasoning. A CBA includes:

- (a) Descriptions of core mathematical ideas and reasoning processes,
- (b) Research-based descriptions of the cognitive constructions students must make in developing understanding of the idea, and

(c) Coherent sets of assessment items that enable educators to investigate students' cognitions and precisely locate students' positions in the constructive itineraries typically taken in acquiring competence with the idea. (pp. 188–189)

With CBA, teachers can identify students' position in the levels-of-sophistication framework to know what cognitive processes and conceptualizations students must acquire to make progress in learning/thinking. Using this information, teachers can make reasonable conjectures for instructional tasks that encourage development of higher levels of sophistication in the progression of student learning.

Researchers also emphasize roles of LTs in designing curricular materials including standards and textbooks (Clement & Sarama, 2004; Confrey et al. 2010; Daro et al., 2011). With the knowledge about LTs for content areas, standards can take advantage of the sequence of big ideas in school mathematics. Learning trajectories support coherence and rigor in standards and curriculum development by addressing cognitive development, mathematical coherence, and the pragmatics of instructional systems (Daro et al., 2011).

The common components of LTs, described by the previous researchers, helped inform data collection procedures for this study. That is, for the purpose of this study, LTs are identified as big ideas or a network of corridors (Baroody et al., 2004; Confrey et al., 2009). The learning trajectory is based on: (a) the learning goals (Clements & Sarama, 2004; Gravemeijer, 2004; Simon, 1995) the authors described in the curriculum materials, (b) the activities (Clements & Sarama, 2004) presented in the materials (such as the order of tasks), and (c) the link between the learning goals and learning activities and among activities. The student learning process (Clements & Sarama, 2004) while

engaged in the activities was interpreted accordingly based on the introduction and development of content evident in textbooks.

Bivariate Relationships (Covariation) – Association and Correlation

This section summarizes studies related to bivariate relationships. In particular, the conceptualization of multiple types of covariation and multiple aspects of covariational reasoning are included in the review. Furthermore, the learning trajectories related to bivariate data suggested in the GAISE Framework are summarized. Specific studies are described in the following paragraphs.

Covariation involves the association of two or more variables – the correspondence of variations. Moritz (2004) categorized covariations into types: *logical*, *numerical*, and *statistical covariation*. Logical covariation involves logical variables, which have values of either true or false as in $A = \text{NOT } (B)$. The value of A varies when the value of B changes from true to false to maintain the correct equation. Numerical covariation concerns functional relationships between two or more real variables. Common forms of functions, such as polynomial, rational and piecewise, are categorized in this type of covariation. The last type of covariation, “statistical covariation[,] refers to the correspondence of the variation of two statistical variables that vary along numerical scales” (p. 228). In most cases, the statistical association does not fit the deterministic models as in logical and numerical covariation. Instead, formal statistical tests are often involved to measure the degree of fit or variation of the models.

When considering the nature of two variables in a bivariate relationship, there are three combinations: (1) two numerical variables, (2) two categorical variables, and (3) one categorical variable and one numerical variable. As depicted in Table 2, the term

correlation is often used for linear relationships between two numerical variables and is involved in making scatterplots and calculating regressions; *statistical association* is generally used for two variables (e.g., expressed in frequency tables and chi-square tests for two categorical variables); and the *covariation* judgment is reasoning about the relationship between two variables.

Table 2

Combinations of two variables in bivariate relationships

	Numerical Data	Categorical Data
Numerical Data	Correlation and Regression	Comparing Groups
Categorical Data	Comparing Groups	Association – Two-way Tables

Moritz (2004) described the characteristics of reasoning about covariation – covariational reasoning – as follows:

Reasoning about covariation commonly involves translation processes among raw numerical data, graphical representations, and verbal statements about statistical covariation and causal association, ...calculating and interpreting statistical tests of association, mathematical modeling to fit data to a specific functional equation, and translating to and from symbolic expressions of algebraic functions. (p. 228)

Accordingly, I organize the literature on covariation into two sections: research related to graphing covariational relationship and studies that involved interpreting covariation.

Representing Covariation in Graphs

There has been substantial research involving point-wise construction and interpretation of graphs, such as, locating and plotting values (e.g., Bell & Janvier, 1993; Moritz, 2004). However, there has been a scarcity of research related to variation and qualitative graphs (Clement, 1989; Leinhardt, Zaslavsky, & Stein, 1990). Studies reported in this section involve examining student reasoning when transferring from

verbal descriptions related to covariation into graphical representations (Bell, Brekke, & Swan, 1987a, 1987b; Coulombe & Berenson, 2001; Krabbendam, 1982; Mavarech & Kramarsky, 1997; Moritz, 2000, 2002, 2004; Swan, 1985, 1988; Watson & Moritz, 1997).

Watson and Moritz (1997) provided tasks requiring students in grades 6 and 11 to produce graphs to describe a nearly perfect relationship between increased heart rate and motor vehicle use. They developed a three-tier framework: (1) a basic understanding of stochastic terminology, (2) an understanding of stochastic language and concepts embedded in context of wider social discussion, and (3) a questioning attitude to apply more sophisticated concepts to contradict claims for coding students' response. They found that most students were too willing to accept anything they read from the print related to claims about association and had difficulty graphing the association between the events. They also found that the students who performed well at the third level of the framework questioned whether there was an actual cause-and-effect relationship between the two events.

Moritz (2000) studied a similar problem, which asked students to graph increasing and curvilinear functions. Moritz used a three-tier coding scheme: *unsuccessful*, *partial*, and *complete* to categorize students' responses. Related to curvilinear function graphing, he found that most students in grades 4–6 responses were coded in the partial or complete tiers. Students in grade 4 had difficulty with multivariate association tasks, in particular about one-third of grade 4 students' responses were rated unsuccessful in multivariate association tasks. However, most grade 6 students' gave complete responses. The task of representing multiple variables posed difficulty for most students, and most students at

grade 4 tried to reduce complexity by attending to selected data or selected variables as a means to reduce the complexity.

In related studies, Moritz (2002, 2004) provided a similar task to students in grades 3, 5, 7, and 9, requiring them to graph positive and negative associations from a verbal description. Based on an analysis of student work, Moritz determined four tiers of responses: *nonstatistical*, *single aspect*, *inadequate covariation*, and *appropriate covariation*. He found that most students demonstrated at least tier-three responses while many older students performed at tier-four. More specifically, student responses were rated higher when graphing positive associations (one quantity tends to increase when the other increases) and when the association depicted by the graph was consistent with students' prior beliefs and knowledge about the relationship between the two quantities.

Krabbendam's (1982) research focused on 12- to 13-year-olds' ability to construct graphical representations of data. He concluded, "it appears to be rather difficult for children to keep an eye on two variables...[but that] time could play an important part in recording a relation" (p. 142). The findings support a view of attending to continuous variation rather than point-wise approach and utilization of time as an inherent independent variable when studying covariation.

Other researchers have examined students' ability to graph various kinds of association. Swan (1988) asked students to represent, for a fixed total cost of a bus, how the price per ticket varies with the size of the group. He found that more than a third of 192 of the 12- to 13-year-old students drew a graph that showed a decreasing trend. Mevarech and Kramarsky (1997) extended the research for different kinds of associations. They found that more than half of 92 grade 8 students appropriately labeled

two axis graphs to represent verbal statements of positive association (“the more she studies, the better her grades”), negative association, and no association, whereas most students had difficulty representing curvilinear association.

Some researchers have utilized contexts that are consistent with students’ prior beliefs of covariation for graph production. Researchers found that there is a natural mapping of time on the horizontal axis and another variable on the vertical axis (e.g., Ainley, 1995; Brasell & Rowe, 1993; Moritz, 2000, 2004). Such findings provide a rationale for using time as an implicit independent variable when developing the concept of covariation for students.

Graphically displaying covariation was shown to be important content to include in textbook analysis. In particular, the literature helps inform a focus for looking into curricula: examining opportunities for students to translate between verbal statements about association to graphs. In addition, types of association are of interest in the analysis because research has shown that students’ performance varies with types of association displayed in graphs (e.g., linear, curvilinear)

Interpreting Covariation

Many researchers have investigated *association* in a variety of disciplines such as in social psychology (e.g., Alloy & Tabachnik, 1984; Crocker, 1981; Ross & Cousins, 1993), science education (e.g., Bell & Janvier, 1981; Donnelly & Welford, 1989; Swatton, 1994; Swatton & Taylor, 1994), mathematics education (e.g., Brasell & Rowe, 1983; Carlson, Jacobs, Coe, Larsen, & Hsu, 2002; Clement, 1985), and statistics education (e.g., Batanero, Estepa, Godino, & Green, 1996; Batanero, Estepa, & Godino, 1997). The following section discusses research involving categorical data and numerical

data specifically as it relates to covariational reasoning.

Research involves categorical data.

In this section, I discuss the research literature related to categorical and numerical data and synthesize research from the fields of psychology, science education, mathematics education and statistics education.

Research in psychology. A substantial body of research related to people's judgments about categorical data, especially dichotomous data, has been conducted in the field of psychology. Generally, psychology researchers have been concerned with how people judge the covariation of events and what evidence they use to support their judgments (e.g., Crocker, 1981; Shaklee & Turker, 1980). Almost all researchers in this field have utilized one sample at one time-point to examine judgments related to covariation.

Research designs in this field have involved a traditional quantitative approach with large sample sizes of subjects ranging from young children to university students and adults. Most of the tasks included a two-way table, using dichotomous data. In only a few studies (e.g., Inhelder & Piaget, 1958; Lovell, 1961; Neimark, 1975; Smedslund, 1963) were students required to sort the data themselves. Most studies asked people if there was any relationship between the variables and how they justified their conclusions about the relationship with given data.

Data collected were often coded with preconceived schemes, which were similar to the following hierarchy:

	Event B	Not Event B
Event A	[a]	[b]
Not Event A	[c]	[d]

- Level 1: Subjects used data/information only from one cell; usually, cell [a]
- Level 2: Subjects compared [a] with [b] or [a] with [c]
- Level 3: Subjects compared [a] with [b] and [a] with [c]
- Level 4: Subjects used four cells, employing additive comparisons.
- Level 5: Subjects used four cells, employing multiplicative comparison (Perez Echevarria, 1990).

Most research studies have used randomized treatments and applied traditional quantitative analysis such as ANOVA to yield results.

Although covariational reasoning is important in everyday lives, many researchers from psychology have concluded that people are generally poor at assessing relationships within data. In particular, people more often base their interpretation on their prior beliefs rather than data at hand and use deterministic thinking when judging covariation (Alloy & Tabachnik, 1984; Crocker, 1982; Jennings, Amabile, & Ross, 1982; Kuhn, Amsel, & O'Loughlin, 1988; Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; Nisbett & Ross, 1980; Peterson, 1980; Smedslund, 1963; Snyder, 1981; Snyder & Swann, 1987; Ward & Jenkins, 1965; Wason & Johnson-Laird, 1972).

In addition to the above findings, researchers found that people often hold an illusory correlation when judging covariation. Chapman (1967) defined illusory correlation as when people form a correlation between two events that “(a) are not correlated, (b) are correlated to a lesser extent than reported, or (c) correlated in the opposite direction from that which is reported” (p. 151). People often form the correlation of events based on their memory rather than the existing data (such as a positive

relationship between high price and quality of the products even though the data might suggest the opposite). This finding was consistent with other research in the psychology literature (e.g., Chapman & Chapman, 1967, 1969; Crocker, 1981, Fiedler, 1991).

A second consistent finding regarding covariational reasoning, was that people tend to judge associations mostly based on the joint presence of the events (Level 1) and are less influenced by the joint absence of the events (Kao & Wasserman, 1993; Inhelder & Piaget, 1958; Levin, Wasserman & Kao, 1993; Lipe, 1990; Schustack & Sternberg, 1981; Smedslund, 1963; Wasserman, Dorner & Kao, 1990 or use the sum of two cells [a] and [d] (Shaklee & Tucker, 1980). People do not appear to understand that both the presence and absence of events have the same meaning concerning the association, even though they admitted that cell [d] in the two-way table was also related to the existence of the association. Complete strategies involved using the quantities either $R = \frac{(a+d)-(b+c)}{a+b+c+d}$ when confirming/disregarding the association (Inhelder & Piaget, 1958) or $\frac{a}{a+c} - \frac{b}{b+d}$, when marginal frequencies for the independent variable did not occur equally (Jenkins & Ward, 1965).

Findings from other studies indicated that people have difficulty with negative associations (e.g., Beyth-Marom, 1982; Erlick, 1966; Gray, 1968). Other researchers found that subjects' covariational judgment of the relationship between two variables tends to be smaller than the actual correlation presented in the data) (Jennings, Amabile, & Ross, 1982; Kuhn, 1989; Meyer, Taieb, & Flascher, 1997; Shaklee & Mims, 1981; Shaklee & Paszek, 1985) – a result that is in conflict with illusory correlation.

Yet another consistent finding was that people tend to form *causal relationships* based on the correlation between two events (Crocker, 1981; Heider, 1958; Inhelder &

Piaget, 1958; Kelley, 1967; Ross & Cousins, 1993; Smedslund, 1963; Shaklee & Tucker, 1980). In addition, people tend to dismiss an association in the data because there was no apparent causation or because other variables were more plausible causes (Crocker, 1981).

Other studies of covariational reasoning analyzed the conditions in which people tend to form more appropriate judgments about association. For example, people tend to present more accurate judgments about association when they are provided with detailed instruction addressing covariational reasoning (Allan & Jenkins, 1983; Ross and Cousins, 1993; Jerkins & Ward, 1965; Shaklee & Tucker, 1980). In addition, they formulate association judgments better when data are given in easy-to-process formats such as in table form (Ward & Jerkins, 1965) and low frequency of data/cases in the cells of two-way tables (Inhelder & Piaget, 1958). In other studies, if data are provided simultaneously people perform better than when given one case at a time (Seggie & Endersby, 1972; Smedslund, 1963).

Another issue in covariational judgment research concerns the role of theory in judging. Although people tend to use their prior beliefs to form covariational judgments, Alloy and Tabachnik (1984) found that it is better to start with a theory/belief about the relationship between the events in mind rather than to be blinded about the events. That is, theories and data at hand collectively influence covariational judgments.

Research in the field of psychology has not targeted the improvement of students' learning regarding covariational reasoning; however, it instructs us about common misconceptions and difficulties people might demonstrate when reasoning about covariation. Next, issues related to student learning are addressed in research in science

education and statistics education.

Research in science education. The research questions of interest in science education have been rather similar to those of psychologists. However, the focus of science educators has been how students used covariational reasoning in a specific educational context such as solving science-related problems. For instance, science education researchers asked students to judge if a correlation exists between events and how strong the relationship is (e.g., Adi, Karplus, Lawson, & Pulos, 1978). Echoing the claims about importance of covariational relationship, science educators assert that such reasoning is fundamental to scientific reasoning.

Research methodology in science education was similar to that in psychology, but often involved a smaller number of participants, without randomized assignment to treatments, and the employment of qualitative data analyses. Findings from science education are rather consistent with the field of psychology. For example, two findings that mirror those from psychology are that students tend to use sub-optimal solution strategies to determine whether relationships exist in the data and form a causal relationship from correlational analysis (e.g., Adi et al., 1978).

Research in statistics education. Several research studies in the field of statistics education considered (a) how technology impacts a student's ability to judge the degree of covariation between two variables, (b) if students respond differently to tasks if they hold prior beliefs about the covariation they expect to see, (c) how they utilize strategies when carrying out the tasks, and (d) in what ways instruction help build covariational reasoning in students (Batanero et al., 1997; Batanero et al., 1996; Batanero et al., 1998; Cobb, McClain, & Gravemeijer, 2003; Konold, 2002; Zieffler, 2006).

Similar to research in science education, investigations in statistics education have been often conducted when teaching bivariate data in secondary schools or at the college level. Statistics educators utilized in-depth interviews with smaller sample sizes or used responses to questionnaires with larger sample sizes. Researchers used emerging coding schemes in their analyses (e.g., Batanero et al., 1997; Batanero et al., 1996; Batanero et al., 1998). Research in this field tends to extend 2x2 tables to 2x3 and 3x3 contingency tables.

Batanero et al. (1996) found that three factors affect the performance of students on contingency tasks: the dimensions of tables, students' previous beliefs or theories about the context of the problem, and a lack of understanding of inverse association. In addition, their findings highlighted strategies that led to correct judgments including:

- (a) comparison of the conditional relative frequency distribution of one variable, when conditioned by the different values of the other variable, or comparison between these conditional distributions and the marginal relative frequency distribution, and
- (b) comparison of frequencies of cases in favor of and against each value of the dependent variable B or comparison of ratio of these frequencies in each value of the independent variable A. (p. 166)

In addition, Batanero and colleagues provided examples of incorrect strategies students demonstrated: (a) *determinist conception*, “considered that cells [b] and [c] in the [2x2] table ought to be null in order to assume association”; (b) *unidirectional conception*, “considered that the frequency in cell [d] ought to be greater than the frequency in cell [c] in order to admit the existence of association”; (c) and a *localist*

conception, “compared the relative frequencies in only one conditional distribution” (p. 165). Their findings also highlighted various combinations evident in students’ work: correct strategy and incorrect judgment, partially correct strategy and incorrect judgment, and incorrect strategy and correct judgment.

Concerned with the impact of technology on student learning, Batanero and colleagues (Batanero et al. 1997; Batanero et al. 1998) examined whether a computer-based teaching experiment would improve 17- and 18-year-old secondary and 20-year-old university students’ strategies of judging statistical association. Both studies asked students to assess the existence of correlation between two variables given to them in a 2x2 contingency table. In these studies, Batanero and colleagues found an overall improvement in students’ strategies, and a persistence of unidirectional and causal misconceptions in some students after an intensive teaching experiment including 21 sessions of 1.5 hours with one third of them working on statistical laboratory. In particular, students moved from comparing data points to the whole set of data, from absolute frequencies in the tables to relative frequencies, and overcame deterministic and localist conceptions of association. However, no improvement was observed in inverse association items and items in which correlation is due to concordance not causal influence.

Findings in judging the covariation of categorical data cited one main category to attend to in the coding process. The literature indicates that misconceptions when judging covariation are worthy of examination in curriculum analysis because it is likely that the misconceptions will be exhibited when students solve covariation tasks. In particular, findings related to students’ difficulty when confronting situations that were

contradictory to their general beliefs about relationship inform the need to focus on examination of the context provided in the tasks. Finally, it is essential to document strategies textbook authors offer for tasks related to judging covariation and if the difference between correlation and causation is addressed in curricular materials.

Research involves numerical data.

As is the case with covariation, researchers in the fields of psychology, science education, mathematics education, and statistics education have examined the concept of association. The problems of concern have involved how well people judge correlation based on data provided in table format or graph forms, how theories influence people's judgments, and what factors influence the judgment process.

Research in psychology. Researchers in psychology approach numerical data and categorical data similarly. They have utilized traditional empirical approaches with random treatments to compare large sample-sized group performances in their research designs. Traditional statistical analyses, such as t-test or ANOVA, have been employed to draw conclusions about people's covariational judgement (Bobko & Karren, 1979; Erlick & Mills, 1967; Cleveland, Diaconiss, & McGill, 1982; Konarski, 2005; Lane, Anderson, & Kellam, 1985; Ross & Cousins, 1993; Trolier & Hamilton, 1986; Wright & Murphy, 1984).

Consistent with students' reasoning about categorical variables, students often performed poorly on tasks requiring covariational reasoning, and often focus on individual points rather than the data as a whole when making judgments. Furthermore, students experienced more challenges when judging negative association and naïve students tended to bias toward positive correlation (Erlick & Mills, 1967). They tended to

carry their theories into judging covariation about current data (Ross & Cousins, 1993; Wright & Murphy, 1984) and form causal relationships when making judgments about covariation (Ross & Cousins, 1993). Moreover, when the data at hand confirmed their beliefs about the relationship, students felt more confident judging the covariation. However, when data at hand did not confirm their beliefs, accuracy in their judgment of covariation depended on the relative strength of the two sources of information, previous beliefs and data at hand (Wright & Murphy, 1984).

Several researchers explored the factors that affect correlational judgments, including representation forms, slope of the regressions, and deviations. In particular, students tended to make better judgments with graphs than tables, and their judgment was more influenced by the variance of error than variances of X or slope (Lane et al., 1985). Further, positive bias of estimates appeared when sequences included a few large and many small deviations versus all intermediate deviations (Erlick & Mills, 1967).

Across many studies (e.g., Beach & Scopp, 1966; Erlick & Mills, 1967, Jennings et al., 1982), one consistent finding was that people tend to perform better in covariational reasoning with numerical data than categorical data. In addition, it appears that the ability to make a valid judgment for one type of data (e.g., categorical data) is not transferred to the other type of data (e.g., numerical data) (Ross & Cousins, 1993). Furthermore, people tend to develop their covariational judgment when aging (Ross & Cousins, 1993).

Research in science education. Although a few studies related covariation of numerical data have been conducted in science education, most of them are related to research in the comprehension of graphs. For example, Wavering (1989) examined

students' ability when doing a variety of tasks related to graphs such as scaling the axes and found they performed poorly due to their deficiencies in logical reasoning. Wavering also linked the successful completion of different tasks to different levels of reasoning within the hierarchy of Piagetian framework related to covariational reasoning.

Research in mathematics education. Mathematics educators have approached covariational relationship somewhat differently, often looking at the covariational reasoning as related to functional thinking and rate of change (e.g., Carlson et al., 2002; Clement, 1989). Research in this field has used qualitative approaches to analyze students' reasoning while solving problems. With in-depth interviews, researchers have gained a deeper understanding of the thinking processes that were taking place.

Carlson et al. (2002) provided college students with the following task: “(i) imagine this bottle filling with water. Sketch a graph of the height as a function of the amount of water that is in the bottle (p. 360).” Researchers found that most students, even after training in Calculus had challenges addressing the rate of change. In addition, representation aids, such as aesthetics of graphs, kinesthetic or physical enactment of certain problems appeared to support the students in their ability to reason correctly about covariation (Carlson, 1998; Carlson, 2002; Carlson et al., 2002; Carlson, Larsen, & Jacobs, 2001). Their research suggested the need for students to learn the concept of function by focusing on the covariational aspect through real life events.

In line with other fields, mathematics educators have highlighted the importance of covariational reasoning specifically for learning algebra (Nemirovsky, 1996) and calculus (Carlson et al., 2002; Kaput, 1992; Thompson, 1994). Researchers found that the interpretation of covariation is slow to develop among students (e.g., Monk, 1992; Monk

& Nemirovsky, 1994; Nemirovsky, 1996), and students tend to look at graph of a function at pointwise rather than depicting covariation (Kaput, 1992; Thompson, 1994).

Research in statistics education. A rich body of research literature regarding numerical data comes from the field of statistics education. Some researchers have asked students to assess the correlation co-efficient from a scatterplot (e.g., Stockburger, 1982; Estepa & Batanero, 1996) and interpret correlation coefficient (e.g., Sotos, Vanhoof, Noortgate, & Onghena; Truran, 1995) while others have incorporated technology into teaching and learning covariational reasoning (e.g., Batanero et al., 1997, Batanero et al., 1998), and studied students' development processes when learning the concept of covariation (e.g., Dierdorp, Bakker, Eijkelhof, & Maanen, 2011; Zieffler, 2006). Most of these studies have applied qualitative methods with in-depth interviews or employed an experimental design with a small sample size of secondary or university students. Moreover, they have used qualitative coding schemes to categorize students' responses, applying coding schemes from their predecessors in psychology (Batanero, Estepa, & Godino, 1997; Batanero, Godino, & Estepa, 1998; Estepa & Batanero, 1996; Estepa & Cobo, 2001; Moritz, 2004; Stockburger, 1982).

Estepa and Batanero (1996) examined senior secondary school students' ability to read scatterplots and make correlation judgments. They found that students were accurate when provided with one coordinate to find the other coordinate using given graphs. They also found that the spread in the scatterplot influenced students' judgment of the correlation. Estepo and Batanero emphasized that strategies such as using increasing, decreasing or constant shape of a scatterplot and comparing the scatterplot with a familiar function helped students at least to be partially correct. Other strategies that hindered

students from making correct judgments included: determinist conception of correlation, localist conception of correlation, and causalistic conception of correlation.

More recently, research has explored the role of statistical technology tools in learning and teaching (e.g., Batanero et al., 1997; Batanero et al., 1998; Konold, 2002). Along with assessing students' reasoning with categorical data, researchers have examined students' learning of correlation for numerical data. Building a course of instruction for university students, Batanero et al. (1997, 1998) characterized the process. Working with statistical software that provided instant feedback, students showed improvement in their strategies, moving beyond merely choosing an algorithm to solve the tasks.

Some researchers have examined students' covariational reasoning in relation to the comprehension of graphs (e.g., Bell & Janvier, 1981; Brasell & Rowe, 1993). Bell and Janvier (1981), in their comparative teaching experiment, found that most secondary students holding a local perspective were weak in interpreting global graphical features in order to extract information about contextual situations. This result is consistent with that of Brasell and Rowe (1983) who found that high school students had more difficulty comprehending graphs with derived variables more than directly measurable variables, and students had difficulty linking graphs with verbal description.

Konold (2002) had a different view regarding people's judgment of covariation. He suggests that people are not poor at judgments, but have trouble decoding the ways in which the relationships are displayed when using TinkerPlots statistical software. When some features of TinkerPlots such as the color-superimposing gradient were used, middle school students performed well when judging the correlation. He posited that by

manipulating the software, students divided the process into two stages: (1) anticipating what will be seen, and (2) examining the new display.

Descriptions of the Progressions within the GAISE Framework

In the previous section, I describe research literature related to bivariate data. In this section, I summarize the recommendations for bivariate data learning trajectories in the GAISE Framework (Franklin et al., 2007). Specifically, I discuss the three levels of statistical maturity for bivariate data as described within the GAISE Framework.

Following is the summary of the progression for each combination of variables in the GAISE Framework.

Two categorical variables. The authors of the GAISE Framework identify Level B as an appropriate level to first address the relationship of two categorical variables. At Level B, problems concern the association between two events (categorical) such as: “Do students who like rock music tend to like or dislike rap music? (Franklin et al., -. 74). By making a two-way frequency table (or contingency table), students can observe trends of the association. Then, they can use and compare relative frequencies of the cells (such as percentage of students who like both and percentage of students who do not like either of them), which can be displayed by pie or bar graphs. To quantify the strength of the associations in a 2x2 table, the authors of *GAISE* suggest statistics called *Agreement-Disagreement Ratio* (ADR) or *Quadrant Count Ratio* (QCR) (see p. 97 of Franklin et al., 2007). Moving to Level C, the *GAISE* authors suggest using tests of significance or simulation to determine if the differences of the proportions are statistically significant and the Phi measure (see p. 98 of Franklin et al., 2007) to test the differences. The GAISE authors suggest addressing Simpson’s Paradox at Level C of the framework.

One categorical and one numerical variable. At Level A, students become aware of the difference between groups (usually two) and look for an association. Students might use a back-to-back stem plots or two dot plots to get a sense of shape (symmetric), and draw conclusions such as “16 year-old boys tend to have longer standing broad jumps than 16 year-old girls.” Students could then estimate or calculate the mean and range for each group to understand the difference. Moving to level B of the GAISE Framework, students might use boxplots to compare groups, interpret results based on the global characteristics of each distribution (center, spread, and shape and interquartile range [IQR]). At Level C, students’ concern might be to test for outliers and compare more than two groups using distribution sampling or statistical tests to examine whether the difference is statistically significant. For example, students might consider if the difference in means of the two groups of boys and girls is statistically significant and if there are any outliers in each distribution.

Two numerical variables. At Level A of the GAISE Framework, students consider if two variables are correlated. By making a scatterplot for the data, students look for trends and pattern such as “generally as one gets larger, so does the other. Based on these data, the organizers might feel comfortable ordering some complete outfits of sweatshirts and sweatpants based on size” (Franklin et al., 2007, p. 32). In addition, students might observe how the values of a numerical variable change over time at a time plot (chart the data using a line graph). Moving to Level B, students start to quantify the association between two variables, determine if one can be a predictor for the other, and whether a linear function is a good model for the data. At this level (B), along with scatterplots, students might use QCR, negative/positive correlation, and eyeballed lines

(mean-mean lines were suggested). For example, they consider the strength of the relationship between sweatshirts and sweatpants and what might be a linear model of fit for the set of data and ask if the correlation is positive, negative. Continuing to Level C of the framework, students might use regression to fit the data and residual plots and r to assess the goodness of fit. They might embrace the error (two standard deviations of the residuals) when predicting. For example, when using the linear regression to predict the size of sweatpants if they know the size of sweatshirts, how confident they can be with that prediction considering the error of prediction. They might assess the goodness of fit for the linear regression using residual plots to determine if it is an appropriate model.

The authors of the GAISE Framework also suggest three levels for learning about different types of scientific studies. At Level A, students engage in simple comparative experiments, e.g., conducting a census for the whole population (might be the classroom) without considering randomness. For example, they might divide the class into two groups and compare which group stacks pennies better. Moving to Level B of the framework, students conduct comparative experiments for two treatments, e.g., compare the effects of two or more treatments (including experimental conditions) on a specific response variable. Students start to design a strategy for collecting appropriate experimental data that involves randomness. For example, they might randomly assign students into two groups and randomly assign the treatments: stacking pennies with right hand or with left hand and examine which treatment yields a higher stack of pennies. At Level C of the framework, students distinguish observational studies and experiments, correlation vs. causation and link the understanding to drawing conclusions about causal relationships. Characteristics of well-designed studies are considered. For example,

students might realize that even though there is a high correlation between smoking and lung cancer, causation of lung cancer cannot be concluded from the observational study. An experiment would need to be conducted because observational studies do not establish causation.

There is limited research related to transferring from graphical representations to verbal statements about covariation. Moritz (2004) addressed this aspect of reasoning in students of grades 3, 5, 7, and 9 when he asked students to translate a scatterplot into a verbal statement. He gave students descriptions and graphs that portrayed a negative association (e.g., numbers of people in classroom vs. level of noise), asking students to explain to another who did not have access to the graph. Using a four-tier framework, *nonstatistical, single aspect, inadequate covariation, appropriate covariation*, to categorize student responses, he found the challenges students encountered included: focusing on isolated bivariate points only, focusing on a single variable rather than bivariate data, and handling negative covariations that were counter to their prior beliefs.

In addition to findings in research related to categorical data, studies in numerical data provide another main category to be coded in this study. Literature about judging numerical variables association advises this study to attend to multiple constructs including graph comprehension, interpretation of correlations, and the line of best fit. Taken together, the research studies related to bivariate data cited in the foregoing section, informed the coding framework that I describe in chapter 3.

There is limited research informing the development of learning trajectories for bivariate data. In particular, formal studies about how the constructs are treated in

secondary curriculum were not found. In the next section, I discuss how my study addresses existing gaps in the literature.

Textbook Analysis

This section summarizes research related to textbook analysis that informed this study. The focus of this section is more on the methodology that researchers used in their analyses because these had the potential to inform the research design of this study. In particular, textbook analysis is categorized into *horizontal* and *vertical* approaches, the terminologies that Charalombous, Delaney, Hsu, and Mesa (2010) used in their analysis. In general, horizontal analysis deals with the textbook as a whole, focusing “on general textbook characteristics (e.g., physical appearance, the organization of the content across the book)” (p. 119); whereas, vertical analysis examines the treatment of a single mathematical concept in the textbooks, which view textbooks as an “environment for construction of knowledge” (Herbst, 1995, p. 3).

Horizontal Approach

The research reported in this section provides background information about curriculum such as page size, numbers of pages, topics, and sequence of topics, new versus old, which provide information about whether textbooks were suitable for students, or teachers’ need (e.g., Flanders 1987, 1994a, 1994b; Stevenson & Bartsch, 1992; Valverde et al. 2002). The findings were often used for the purpose of informing educational policy, such as findings of Schmidt et al. (1997) who characterized the U.S. curriculum as “a mile wide, inch deep” (p. 122). However, a lack of details in the specific content and key questions were the way in which the concept was treated in the textbooks.

Flanders (1987, 1994a, 1994b) was concerned about the repetition of the topics in popular K–8 textbook series. Flanders (1987) used scope and sequence charts to specify numbers of pages devoted to content and if the content was new or old compared to content that appeared in the previous grade level. In his later studies, Flanders (1994a, 1994b) was interested in a similar topic when using items from the Second International Mathematics Study (SIMS) as an anchor to examine the presence of the items, and coded the content as old or new in six of the most commonly used grade-8 textbooks in the U.S. Such studies supported claims about the repetition of mathematics content in the U.S. curriculum.

Stevenson and Bartsch (1992) examined Japanese and American K–12 textbooks in terms of which topics appeared in each country and in which grade level. Their findings showed that some topics are introduced earlier in the American textbooks and others earlier in the Japanese textbooks and some topics appeared in the textbooks of one country and not in the other.

Valverde et al. (2002) were concerned about the structure and content of mathematics and science textbooks used in the schools of the students in the study, as a component of Trends in International Mathematics and Science Study (TIMSS). With a sample of 194 textbooks from 48 countries, they counted numbers of pages, calculated total page area of the text, identified the topics addressed in the text, and counted the numbers of times the text changed from one topic to another. They also analyzed the content of texts: primary nature of lessons (concrete and pictorial vs. textual and symbolic), other components of the lessons, and student performance expectations. They used a block, which “constituted narrative or graphical elements; exercise or question

sets, worked examples, or activities” (p. 141) as a unit of analysis. Such analytic framework highlights the general features of the textbooks as a measurement for comparing students’ achievement among the countries. For example, U.S. textbooks were among the upper extreme values of the sample in regard to frequently changing topics very often, “textbooks across all populations were mostly made up of exercises and question sets” (p. 143), and “the most common expectation for student performance was that they read and understand, recognize or recall or that they use individual mathematical notations, facts or objects. This is followed... by the use of routine mathematical procedures” (p. 128).

Vertical Approach

This section includes studies related to analysis in a single mathematical topic such as functions. Generally, researchers utilized frameworks closely related to the content as a lens to look into the curriculum. Some researchers used the mathematical task framework (Stein & Lane, 1986) as a common denominator to judge the cognitive demands of tasks and contingent opportunities for students to learn from textbooks.

Li (1999) conducted a comparative analysis of 12 textbooks from four countries: Hong Kong, Mainland China, Singapore and United States (non-NSF funded textbooks). The goal was to obtain information about students’ performance through an international study. His research focused on the ways textbooks present content of Algebra in relation to the achievement data provided by the TIMSS. In particular, he was concerned about content emphases and requirements, presentation and organization of content and desired competence via to-be-solved problems (TBS). He used a focus on function, relations, and equations from the TIMSS curriculum framework to decide what chapters to include. The

number of chapters and number of pages were used to measure relative emphasis of mathematics content. For the TBS problems, he considered the mathematics features of new/old, context of content of pure or real-life, performance requirements of explanation or no explanation, and cognitive requirements. His findings mirrored findings from previous studies highlighting the difference between the U.S. and Asian curricula, namely that, U.S. textbooks covered too many topics and not at a deep level. Furthermore, he found fine-grained results in terms of TBS problems, cognitive demands, the nature of tasks, and expectation of an explanation of response.

Mesa (2000, 2004) took another approach when examining the concept of function in a larger sample of textbooks from 18 countries (2000) and 15 countries (2004). She used the Balacheff's theory of conceptions and Biehler's notion of the prototypical domain of application of concepts to describe the practices associated with the notion of function. She further grouped countries in terms of similar textbook characteristics. Across countries, the textbooks fell into four clusters according to the predominant conceptions and uses of function: (1) rule-oriented, (2) abstract-oriented, (3) abstract-oriented with applications, and (4) applications-oriented. Moreover, her analysis yielded five different practices – symbolic rule, ordered pair, social data, physical phenomena, and controlling image – that were present both across and within the textbooks analyzed. In addition, the results suggested that there is no canonical trend for teaching function and organizing mathematics content on function in grade 8.

Mesa (2000) used ten items from the TIMSS achievement test for seventh- and eighth-grade students to code and compare the tasks in the textbook clusters. She also kept track of students' performance on the item in countries using textbooks promoting

the same conception. Two key findings were that (a) the test did not reflect any country's distribution of conceptions, and (b) using a textbook belonging to a particular cluster or promoting a certain conception did not provide an advantage to students. Thus, at the fine-grained level of textbook content, there is no evidence that one textbook organization yields higher student achievement than the other.

Ross (2011) also focused on the construct of function analyzing the contemporary curricula presented by three U.S. high school textbook series: the *Glencoe Mathematics* series, the *University of Chicago School Mathematics Project* series, and the *Core-Plus Mathematics Project* series. Taking another approach, he conducted an extensive literature review related to function in order to provide constructs for coding. In each series, functions were examined in four areas: language used in relation to functions, presence of functions, core features of function examples, and ancillary features of function examples.

Ross's (2011) findings regarding definitions, representations, and examples of functions in the textbook provide similar results to Mesa's (2000, 2004) conception. All textbook series included multiple representations of functions with examples. Many examples include verbal descriptions, while there were smaller proportions of numeric and graphic representations. In addition, he found other related characteristics such as the presence of examples, the context where examples appeared, and the utilization of technology in curriculum to aid student learning. For example, examples mostly appeared in homework exercises and in abstract settings, explicit recommendations for using technology with examples are relatively infrequent, and opportunities for students to generate function examples or explore non-examples are limited.

Jones (2004) traced probability content throughout four eras of mathematics education in the U.S.: the New Math (1957–72), Back to Basics (1973–83), a focus on Problem Solving (1984–93), and the advent of the NCTM Standards (1994–2004). Using popular and alternative middle grade mathematics textbook series from four eras, he was interested in the extent and nature of the treatment of probability and structure of probability lessons within these textbooks. He examined background information such as determining the number of pages that contained probability tasks and their placement in the textbooks, checking whether content in later books was new or repeated. He also tracked the introduction and development of probability concepts, determined the level of cognitive demands required by each probability task, and selected archetypal tasks to illustrate each era.

Jones (2004) highlighted the differences and similarities among the eras and provided comparable key suggestions in each era related to the probability topics. It was not surprising that Standards-era textbooks contained far more probability tasks than other eras and a wider variety of learning expectations. Consistent with findings in other areas, the majority of tasks in all series had low levels of cognitive demands and archetypal activities were remarkably similar across eras. Exceptions were the Standards alternate series and the Back to Basics alternate series, which contained more tasks with higher levels of cognitive demands.

Dingman (2007) studied the alignment of textbooks with ten states' grade-level learning expectations (LEs) related to fraction concepts. He used two popular elementary and two popular middle grade textbook series, those with statewide textbook adoption policies and those without such policies. He constructed a generalized set of LEs from a

sample of state frameworks as well as from research on rational numbers. He analyzed five types of instructional segments related to fraction concepts: lessons, pre-lessons, end-of-lesson extra features, end-of-chapter features, and games. For each instructional segment, he tracked the number of pages on which the concept occurred, the related learning objectives and the matching generalized LEs.

Dingman's findings along with other research (e.g., Reys, 2006) told the story about *de facto* national curriculum. Not surprisingly, with the diversity of U.S. educational system, he found considerable differences across states. There were many instructional segments in the textbooks that did not correspond to the states' LEs, although textbooks generally incorporated the majority of states' LEs related to fractions.

Other researchers have focused on proof and reasoning in their analyses (e.g., Johnson, Thompson, & Senk, 2010; Stylianides, 2009). They share similar approaches when using frameworks closely related to proof to code the textbooks. For example, Johnson et al. (2010) used NCTM Reasoning and Proof Standards to code non-geometry textbooks: *Algebra 1*, *Algebra 2*, and *Pre-calculus* from Glencoe, Holt, Rinehart, and Winston; Key Curriculum, Prentice Hall, *Core Plus Mathematics Project*, and the *University of Chicago School Mathematics Project*; whereas, Stylianides (2009) developed "an analytic framework" (p. 261) grounded in the research related to proof to code one middle school textbook series: *Connected Mathematics Program* (CMP).

Johnson et al. (2010) findings reveal what proof-related content in textbooks looks like and percentages of problems requiring reasoning. They found that the percentages of problems devoted to proof in the various series were quite low (less than 8%). Johnson et al. highlighted examples of different kinds of activities related to proof

in the textbooks. Stylianides (2009) provided more details about a proof framework; which can be utilized to see proof-related problems in other strands. In this framework, Stylianides found a higher percentage of sections in the textbook that provided students opportunities to engage in proof-related tasks.

Summary

Generally, researchers approach their analysis with one or a combination of (a) tracking background information such as page numbers, portions of a page related to specific content or proportions of old and new content, (b) gauging the alignment of textbooks to standards, and (c) focusing on a single content utilizing frameworks from the literature related to the topics to code textbooks. My study is not intended to be evaluative, but descriptive in its nature. Even though there is a rating component related to the tasks in the textbooks, my focus is on the description of learning trajectories related to bivariate data in the curricular materials. To that end, I developed a coding framework grounded in the literature related to the construct of covariation. In addition, inherent in this study is a component to monitor the alignment of curricular materials and two aspects of the CCSSM. First, the learning trajectories related to the construct of bivariate data in the materials are compared with the CCSSM learning expectations. Second, a detailed framework for coding the CCSSM mathematical practices (described in chapter 3) provides guidance when coding the tasks. This coding process highlights the opportunities the materials offered students to access the CCSSM practices standards.

In line with other research categorized in the vertical approach, I chose to do an in-depth analysis of one topic at the high school level as in Ross's (2011) study. In addition, I used the construct of LTs to document how the construct of bivariate data was

treated and highlight the opportunities for students to learn provided in the curricular materials. Admittedly, the horizontal approach provides a source of background information; however, the information yielded from one topic was not comprehensive enough to consider in this study.

In addition, following suggestions from NRC (2004), my research design has taken into consideration each of the following: selection of a set of standards with rationale for analysis, and components to consider in analysis. With training in advanced graduate courses in mathematics, a Ph.D. minor in statistics, and completing substantive coursework in mathematics education, I am well qualified to do the analysis. Furthermore, as I discuss in the next chapter, I selected frameworks to address the research questions, and specifically address the discipline perspectives suggested by NRC (2004).

The review of the research literature in the areas of learning trajectories, covariational reasoning, and textbook analyses revealed the existence of a gap and a need for this study. Heretofore, no studies have examined the treatment of covariation in high school textbooks. Using the construct of learning trajectory (LT) as guidance, I sought to understand how students' opportunities to learn covariational reasoning are treated in the textbooks. In addition, in the era of the CCSSM, it is crucial to compare the findings with recommendation from widely adopted standards. In the following chapter, I provide a detailed accounting of the research methods employed in this study.

CHAPTER 3: METHODOLOGY

In this chapter, I provide an overview of the methodology utilized in this study. First, I provide a rationale for the selection of the textbook sample, then describe the coding scheme that is grounded in the research literature, and finally offer potential limitations of the study. Through the process of data collection and analysis, described herein, I seek to answer the following research questions:

1. What learning trajectories for topics in bivariate data are represented in high school mathematics textbooks in the U.S.? How are those learning trajectories similar to and different from those embedded within the *Common Core State Standards for Mathematics* (CCSSM)? To what degree are the learning trajectories found in textbooks aligned with the developmental levels described in the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) *Report*?
 - a. What topics in bivariate data are addressed in the textbooks and what learning trajectories for bivariate data are evident in the textbooks?
 - b. How are connections made between the topics in bivariate data and topics in univariate data in the textbooks? How are connections made between bivariate data and other conceptual categories in mathematics at the high school level?
 - c. To what extent are the CCSSM standards for bivariate data at the high school level evident in the textbook materials, and to what extent does the approach to, and sequence of, content in the materials reflect the

developmental progressions of the topics described by the GAISE Framework?

2. With respect to bivariate data, what is the nature of the instructional tasks presented in textbooks? Do the tasks provide opportunities for students to access the CCSSM's Standards for Mathematical Practice (SMP)?
 - a. What levels of mathematical complexity are required by the tasks related to bivariate data?
 - b. To what degree are the GAISE developmental levels reflected in the tasks related to bivariate data?
 - c. What is the quality level of the tasks in terms of purpose and utility?
 - d. How do the tasks provide opportunities for students to access the CCSSM's Standards for Mathematical Practice?

Selection of Textbook Sample

A set of three criteria informed the selection of the mathematics textbooks series for analyses: The series (1) elicits the potential for in-depth analyses of differences in sequencing and organizing bivariate data content, (2) explicitly includes learning goals to support teachers implementing the curriculum, and (3) is currently used in U.S.

educational systems. The textbook series selected were:

Holt McDougal Larson series (HML):

- *Algebra 1* (Larson, R., Boswell, L., Kanold, T. D., & Stiff, L., 2012a)
- *Geometry* (Larson, R., Boswell, L., Kanold, T. D., & Stiff, L., 2012b)
- *Algebra 2* (Larson, R., Boswell, L., Kanold, T. D., & Stiff, L., 2012c)

The University of Chicago School Mathematics Project series (UCSMP):

- *Algebra* (Brown, S. A., Breunlin, R. J., Wiltjer, M. H., Degner, K. M., Eddins, S. K., Edwards, M. T., et al., 2008)
- *Geometry* (Benson, J., Klein, R., Miller, M. J., Capuzzi-Feuerstein, C., Fletcher, M., & Usiskin, Z., 2009)
- *Advanced Algebra* (Flanders, J., Lassak, M., Sech, J., Eggerding, M., Karafiol, P. J., McMullin, L., et al., 2010)
- *Functions, Statistics, and Trigonometry* (McConnell, J. W., Brown, S. A., Karafiol, P. J., Brouwer, S., Ives, M., Lassak, M., et al. 2010)
- *Precalculus and Discrete Mathematics* (Peressini, A. L., DeCraene, P. D., Rockstroh, M. A., Viktora, S. S., Canfield, W. E., Wiltjer, M. H., et al., 2010)

Core-Plus Mathematics Project series (CPMP):

- *Course 1* (Hirsch, C. R., Fey, J. T., Hart, E. W., Schoen, H. L., Watkins, A. E., Ritsema, B. E., et al., 2008a)
- *Course 2* (Hirsch, C. R., Fey, J. T., Hart, E. W., Schoen, H. L., Watkins, A. E., Ritsema, B. E., et al., 2008b).
- *Course 3* (Hirsch, C. R., Fey, J. T., Hart, E. W., Schoen, H. L., Watkins, A. E., Ritsema, B. E., et al., 2009).
- *Course 4* (Hirsch, C. R., Fey, J. T., Hart, E. W., Schoen, H. L., Watkins, A. E., Ritsema, B. E., et al., 2010)

For the purpose of this study, I primarily analyzed teacher's editions. However, I also examined other materials such as implementation guidelines in order to identify learning trajectories evident in the curricula.

The differences in organization and conceptualization of mathematical content among these textbook series creates a basis for comparing three different approaches to the teaching of bivariate data. In the next section, I describe how specific features of each textbook series address the selection criteria.

Criterion 1: The Series Elicits the Potential for In-depth Analyses of Differences in Sequencing and Organizing Bivariate Data Content

Holt McDougal Larson Series (HML), (Houghton Mifflin Harcourt Publishing Company, 2012). The HML series uses the conventional three-course sequence: *Algebra 1*, *Geometry*, and *Algebra 2*. Content is organized via many short examples and guided practices that include questions similar to those represented in the examples for students to learn specific skills and knowledge. Some lessons also include longer activities for students to relate content within a lesson. In addition, the authors emphasize that the textbooks align with the CCSSM and include processes/procedures that help students learn the CCSSM Standards for Mathematical Practices (SMP). This series offers the potential for a sequence of content that is aligned to the new standards.

The University of Chicago School Mathematics Project Series (UCSMP), (Wright Group/McGraw-Hill, 2008, 2009, 2010). The UCSMP series includes five textbooks with conventional titles: *Algebra*; *Geometry*; *Advanced Algebra* (AA); *Functions, Statistics and Trigonometry* (FST); and *Precalculus and Discrete Mathematics* (PD). Even though the names of individual textbooks are conventional, the authors integrate content from different strands in each of the textbooks while still focusing on the main content theme (e.g., geometry) of each textbook. Furthermore, the UCSMP textbook series includes substantial statistical content that offers the potential for

in-depth analyses. Each lesson includes different components ranging from short responses such as mental math, guided examples for students to fill in the blanks to complete the tasks, longer activities within the lesson, and projects at the end of chapters. The textbook authors argue that the series differs from other secondary mathematics textbooks by requiring more reading, greater use of technology, and a more deliberate sequencing of mathematical applications (Usiskin, 2007).

Core-Plus Mathematics Project Series (CPMP) (McGraw-Hill Companies, 2008, 2009, 2010). The CPMP textbook series utilizes an integrated approach. The content related to algebra, geometry, statistics, and discrete mathematics are integrated in each of four textbooks, *Course 1*, *Course 2*, *Course 3*, and *Course 4*. CPMP topics are sequenced over the series not only to build concepts, but also to provide what the authors considered the most important mathematics content in each course in the high school sequence. Lessons generally consist of a series of questions that lead students to investigate, reflect on, and summarize mathematical concepts. The authors described this is as an extended design experiment with principles focusing on: substantial reading, fewer but longer examples and exercises, and a greater emphasis on mathematical modeling of real-world situations including the extensive use of technology (Fey & Hirsch, 2007).

In contrast to HML, UCSMP and CPMP “offer an approach to mathematics teaching and learning that is qualitatively different from conventional practice in content, priorities, organization, and approaches” (Hirsch, 2007, p.1). In addition, these two textbook series were updated to include topics related to data analysis and probability, to focus on big ideas across grade levels with multiple representations connected among

ideas across mathematical strands and grade levels, and to emphasize depth over coverage to promote deeper understanding of important mathematical ideas (Hirsch, 2007).

The three textbook series offer fundamentally different organizations of the mathematical content and philosophy of curriculum design. Moreover, the UCSMP and CPMP series represent innovative, non-traditional approaches whereas the HML exemplifies a traditional approach while including many features that highlight the intentional alignment with the CCSSM. Collectively, these three series together elicit the potential for differences in sequencing and organizing bivariate contents.

Criterion 2: The Series Explicitly Displays Learning Goals to Support Teachers Implementing the Curriculum

Holt McDougal Larson Series (HML). The HML series includes *big ideas* (as called in the series) for each chapter and lesson. At the start of each lesson, the authors make reference to content that was learned *before* to the *now* content, and provide reasons *why* to make connections within and across the lessons. In addition, the authors map the contents in each lesson to the CCSSM standards to link with big ideas for each lesson.

The University of Chicago School Mathematics Project Series (UCSMP). The UCSMP series provides detailed information for teachers as they choose tasks to address learning goals. Specifically, for each lesson they include a list of objectives including skills, properties, uses, and representations. At the beginning of each chapter, the authors provide an overview of the chapters, and recommend teachers assign the final projects at the end of each chapter to target student learning through the chapter.

Core-Plus Mathematics Project Series (CPMP). The CPMP series is more overt in “conveying the mathematical goals at the unit, lesson, and sometimes problem level” (Fey & Hirsch, 2007, p. 139). This includes information about students’ level of proficiency expected when learning important topics. The authors also incorporate how students might approach the investigations and offer common misconceptions to support teachers’ understanding of the learning goals. In addition, scope and sequence materials, containing a list of core topics and how they are of *focus* and *connection* in each unit of the four courses, are provided to help teachers (and students) build connections among the content in investigations, lessons, units, and the textbook series as a whole.

Criterion 3: The Series is Currently Used in U.S. Educational Systems

Holt McDougal Larson Series (HML). The HML series was published by Holt McDougal in 2012, a subsidiary of by Houghton Mifflin Harcourt Publishing renamed after the merger of two companies Holt, Rinehart and Winston and McDougal Littell. This company, as a whole, publishes the textbooks that are most frequently used in U.S. schools (Dossey, 2012).

The University of Chicago School Mathematics Project Series (UCSMP). The UCSMP series is continually revised and new material is developed. In 2009, UCSMP made available online lessons to supplement the third edition of the curricular materials for grades 6-12. The series is also designed to follow *Everyday Mathematics* (McGraw-Hill Companies, 2007), which is the most popular elementary mathematics textbook series in the U.S. (Stanton, 2011).

Core-Plus Mathematics Project Series (CPMP). The CPMP is currently in its second edition. In addition, the same group of authors for CPMP is currently developing

the *Transition to College Mathematics and Statistics* as an alternative to Course 4 for non-calculus bound students to follow the first three courses, which implies that the series is currently used U.S. systems.

Analyses of the Textbooks

In this section, I describe the analyses of the textbooks. First, I define the unit of analysis for coding. Then, I discuss the coding scheme including learning trajectories and task features and use an example to illustrate the coding process. Finally, I describe the analysis of data to address the aforementioned research questions.

Unit of Analysis

The unit of analysis is defined as the primary entity to examine. For this study, I draw upon Olson (2010) and specify the unit of analysis as an *instance* that is either: (1) a super-problem includes one or more related tasks, (2) a problem with no sub-problems, (3) an example that the authors showed both prompt and solution, and (4) author text including definitions, explanation of content. The first two types of instance were referred to as to-be-solved (TBS) problems (Li, 1999) in which the solution was not readily available and students are required to solve the problems.

The method for defining an instance was (1) identifying prompts, or instructions for students to accomplish in each problem, (2) an example including questions, prompt or instruction and solution, or (3) portions consisted of the elaboration of concepts, definition of terminology, or presentation of content separate from author examples. For example, Figure 5, 6, and 7 show the first two types of instance; the prompts or instructions that appear at the beginning of each instance signify these types of super-problems or problem. In particular, Figure 5 shows an instance including a super-problem

with two tasks (as numbered in the HML *Algebra 1* textbook). Figure 6 depicts an instance of super-problems with related sub-problems including one task (as numbered in the textbook) from the CPMP *Course 1* textbook. Figure 7 shows an example of a problem for students to solve without sub-problems that appears in the UCSMP *Algebra*. Figure 8 illustrates two instances of two types, an author example and text taken from UCSMP series.

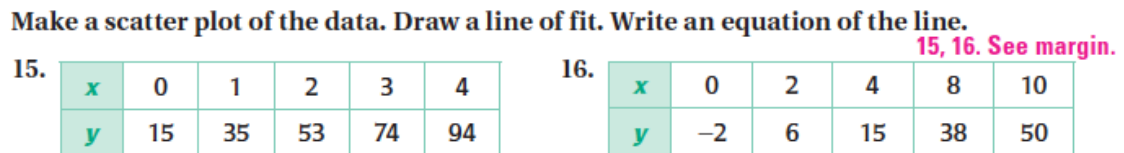


Figure 5. An instance containing more than one task. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson, L. Boswell, T. D. Kanold, & L. Stiff, 2012, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.291. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

1 Make a model bungee jump by attaching a weight to an elastic cord or to a chain of rubber bands.

a. Use your model to collect test data about bungee cord stretch for at least five weights. Record the data in a table and display it as a *scatterplot* on a graph.

Weight Attached (in grams)						
Length of Stretched Cord (in cm)						

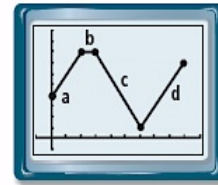
b. Use the pattern in your experimental data to predict length of the stretched bungee cord for weights different from those already tested. Then test the accuracy of your predictions.

c. Compare your results to those of others doing the same experiment. What might explain any differences in results?

Figure 6. An instance of super-problems with related sub-problems. From *Core-Plus Mathematics: Course 1 – Teacher’s Guide* by C. R. Hirsch, J. T. Fey, E. W. Hart, H. L.

Schoen, A. E. Watkins, B. E. Ritsema, R. K. Walker, S. Keller, R. Marcus, A. F. Coxford, G. Burrill, 2008a, Columbus: OH: Glencoe/McGraw-Hill, p.5. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

14. A bulldozer is working on a highway. The graph at the right shows the distance between the bulldozer and a fixed point on the highway measured over time. Use the graph to complete the table.



Section of Graph	Rate of Change (positive, negative, or zero)	Bulldozer's Movement (towards point, away from point, standing still)
a	positive ?	away from point
b	zero ?	? standing still
c	negative ?	? towards point
d	positive	? away from point

Figure 7. An instance of problems without sub-problems. From *The University of Chicago School Mathematics Project: Algebra – Teacher's Edition* by S. S. Brown, R. J. Breunlin, M. H. Wiltjer, K. M. Degner, S. K. Eddins, M. T. Edwards, N. A. Metcalf, N. Jakucyn, & Z. Usiskin R, 2008, Chicago: IL: Wright Group/McGraw-Hill, p.332. Copyright 2008 by Wright Group/McGraw-Hill. Reprinted with permission.

Example 1

The population of Chicago from 1830 to 2000 is shown in the table and graph on page 325. Find the rate of change of the population of Chicago (in people per year) during the given time period, and describe how the rate is pictured on the graph.

a. 1890 to 1900 b. 1970 to 1980

Solutions

a. the rate of change, in people per year, from 1890 to 1900:

$$\frac{\text{change in population}}{\text{change in years}} = \frac{1,698,575 - 1,099,850}{1900 - 1890}$$

$$= \frac{598,725 \text{ people}}{10 \text{ yr}}$$

$$= 59,872.5 \frac{\text{people}}{\text{yr}}$$

Between 1890 and 1900 the population increased, so the rate of change is positive. The graph slants upward as you read the graph from left to right.

b. the rate of change between 1970 and 1980:

$$\frac{\text{change in population}}{\text{change in years}} = \frac{3,005,072 - 3,369,357}{1980 - 1970}$$

$$= \frac{-364,285 \text{ people}}{10 \text{ yr}}$$

$$= -36,428.5 \frac{\text{people}}{\text{yr}}$$

Between 1970 and 1980 the population decreased, so the rate of change is negative. In that interval, the graph slants downward as you read the graph from left to right.

When you read graphs in algebra, read them from left to right just as you would read a line in a book. A *positive rate of change* indicates that the graph slants *upward*. A *negative rate of change* indicates that the graph slants *downward*.

Author
example

Author
text

Figure 8. Instances of author examples and author texts. From *The University of Chicago School Mathematics Project: Algebra – Teacher's Edition* by S. S. Brown et al., 2008,

Chicago: IL: Wright Group/McGraw-Hill, p.327. Copyright 2008 by Wright Group/McGraw-Hill. Reprinted with permission.

Coding

When relevant textbook content was located, the unit of analysis (i.e., each instance) was coded vis-à-vis *learning trajectories and task features*. All instances of “to-be-solved” (TBS) (Li, 2000) problems were coded for learning trajectories and task features. That is, instances without solutions provided or learning activities requiring students to work through several steps were coded for both learning trajectories and task features. This also includes other portions of tasks that students have to accomplish for homework, self-tests, or chapter review. However, instances including author text and author examples with solution at hand, in which students do not have to solve them, were coded only for learning trajectories. In short, the task feature was only coded for TBS instances.

Learning trajectories.

Textbooks were analyzed to determine the flow of topics related to bivariate data, in turn to discern learning trajectories. In order to conduct the curriculum analysis, I applied several codes using a well-defined coding scheme. In this section, I describe the data collected and the process used to assign codes.

Combination of two variables. For each instance, the data was codified as one of the three types: (1) two categorical variables (CC), (2) one categorical variable and one numerical variable (CN), or (3) two numerical variables (NN). After I identified the types of data, I assigned codes to indicate the form in which data appeared, size and form of representations involved, and types of association. More specifically, the process is described as follows:

- Two categorical variables: Form was coded as a *ready-made table* or as *raw data* that need to be sorted; the size was coded as dimension of the table, for example, 2x2, 2x3, 3x3; representation of data was coded in such ways as pie charts, bar charts, or numerical representation (e.g., proportions, probabilities); and types of association involved was coded as positive (one quantity increases when the other increases) or negative (one quantity increases when the other decreases). It should be noted that not all instances include two-way tables were coded. In particular, I excluded all instances related to reading and interpreting absolute frequencies of the cells in the table or those using the table to calculate conditional probabilities. Figure 9 is an example of such an instance taken from HML *Algebra 1* (p. 674). This instance might offer potentials for students to attend to the association between two variables, gender and position of living, but the authors did not explicitly use that context for the purpose.

EXAMPLE 3 Analyze a situation in a two-way table

The table shows where students at a university live.

	Live on Campus	Live off Campus	Total
Men	3216	4010	7226
Women	3824	3758	7582
Total	7040	7768	14,808

- Do more students live on campus or off campus?
- Is it also true that more women live off campus than on campus?

Solution

- Look at the marginal frequencies in the Total row. More students live off campus.
- No. Even though the marginal frequencies show that more students live off campus, looking at just the row for women, you can see that more women live on campus than off campus.

Figure 9. An instance that was not coded for the CC association. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.674. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

- One categorical variable and one numerical variable: form was coded as table or graph; the size was coded based on the number of data points in each distribution; representation of data was coded by types of graphs or other representations, such as back-to-back plots, histograms or box plots; and types of association involved were coded as positive (one quantity increases when the other increases) or negative (one quantity increases when the other decreases). Because tasks related to the association of one categorical variable and one numerical variable often involve comparing groups, all tasks related to comparing distributions were considered but were ultimately coded only if the goals addressed the *association* between the two variables as opposed to the simple *comparison* of two groups. Figure 10 is an instance from UCSMP *Functions, Statistics, and Trigonometry* (McConnell, 2010, p. 59) that involves *comparing* two distributions, length of bridges in the two countries. However, because it does not require students to identify the *association* between two variables, this instance was not coded.

- Two numerical variables: form was coded as table, graph, or both; the size was coded based on the number of data points required/provided in the task; and representation was coded as the type graphical representation of the data such as graphs, scatterplot, order-case value bars, scatterplot slices, superimposed color gradient graph or algebraic; types of associations were coded according to the types of functions used to fit the data such as linear, polynomial with degree 2 or more, exponential, power, logarithmic, trigonometric, rational. Situations such as fitting functions to data were considered examples of covariational reasoning. However, situations that asked students to plot points that form a perfect fitting function and situations that involve functional relationship were not coded. For example, the following instance from CPMP Course 1

(Hirsch et al., 2008a, p. 6) was not identified as bivariate data because it involves a perfect fit of linear function based on the data provided in the table (Figure 11).

7. The table below shows the length in meters of the 10 longest bridges in the United States and the 10 longest bridges outside of the United States. Write a descriptive paragraph comparing the length of the longest bridges in the United States to the longest bridges elsewhere in the world.

Longest Bridges	
U.S.	
Name	Length (m)
Lake Pontchartrain Causeway	38,422
Manchac Swamp Bridge	36,710
Atchafalaya Swamp Freeway	29,290
Chesapeake Bay Bridge Tunnel	24,140
Bonnet Carré Spillway	17,702
Jubilee Parkway	12,875
San Mateo-Hayward Bridge	11,265
Seven Mile Bridge	10,887
Sunshine Skyway Bridge	8,851
Twin Span	8,851
Outside U.S.	
Name	Length (m)
Bang Na Expressway	54,000
Yangcun Bridge of Beijing-Tianjin Intercity Rail	35,812
Hangzhou Bridge	35,673
Runyang Bridge	35,660
Donghai Bridge	32,500
King Fahd Causeway	26,000
The No. 1 Bridge of Tianjin Binhai Mass Transit	25,800
Liangshui River Bridge of Beijing-Tianjin Intercity Rail	21,563
Yongding River Bridge of Beijing-Tianjin Intercity Rail	21,133
6th October Bridge	20,500

Source: Bridgemeister



Are We There Yet?

The Donghai Bridge, in China, is 32.5 km long and links Shanghai with the Yangshan Deep Water Port.



Figure 10. An instance that was not coded for the CN association. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher's Edition* by J. W. Connell, S. S. Brown, P. J. Karfiol, S. Brouwer, M. Ives, M. Lassak, R. McCullagh, N. Jakucyn, & Z. Usiskin R, 2010, Chicago: IL: Wright Group/McGraw-Hill, p.59. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

- 3 The Five Star marketing staff did a survey of park visitors to find out the *number of customers* that could be expected each day for the bungee jump at various possible *price per jump* values. Their survey produced data that they rounded off and presented in this table.

Market Survey Data

Price per Jump (in dollars)	0	5	10	15	20	25	30
Likely Number of Customers	50	45	40	35	30	25	20

- Plot the (*price per jump*, *number of customers*) data on a coordinate graph. Then describe how the predicted *number of customers* changes as *price per jump* increases from \$0 to \$30.
- The Five Star data processing department proposed the rule $N = 50 - p$ for the relationship between *number of customers* N and *price per jump* p . Does this rule represent the pattern in the market research data? Explain your reasoning.

Figure 11. An instance that was not coded for the NN association. From *Core-Plus Mathematics: Course 1 – Teacher’s Guide* by C. R. Hirsch et al., 2008a, Columbus: OH: Glencoe/McGraw-Hill, p.6. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

Abstract/Contextual. The context of the data provided was coded as either a contextual or an abstract (decontextualized) setting. Types of associations were recorded for each instance (positive, negative). When contextual-based data were specified, the solving of tasks and the solution offered in the materials were used to determine if the type of association was consistent with or contradictory to general beliefs. That is, I considered whether the relationship between the two variables was positive or negative and how it compared to the data at hand.

Task – Technique – Theory Framework (Artique, 2000, 2002). This framework was utilized to code what techniques might be used to reach a solution. The teacher’s edition was used as a reference of *techniques* offered in the materials. Related *theory* (e.g., definition, properties provided in the textbooks) to explain the techniques used was also specified. Some constructs of bivariate data were referenced in this process such as when tasks offered opportunities for students to make links between association and

causation, tasks that involved interpreting, estimating, and using r , r^2 and tasks which required judgments about the magnitude of an association between two variables.

Learning Goals. Based on the *lesson objectives* or *big ideas* given for each lesson containing bivariate data, and the common techniques and theories for sets of instances related to bivariate data, I specify the learning goals for each lesson. Sometimes, the instances related to bivariate data were not included in lessons related to bivariate data but instead appeared for the purpose of review. In these cases, the codes of *technique* and *theory* provided information for learning goals.

Task features.

In this section, I describe how I coded the nature of tasks for the mathematical complexity, purpose and utility features, GAISE Framework, and CCSSM mathematical practices.

Mathematical complexity. The mathematical complexity measures the cognitive demands required by each instance in the materials. To code the mathematical complexity of each instance, I used a three-level framework of mathematical complexity (low, moderate, and high) based on the National Assessment of Education Project's (NAEP 2006) description of the features of each level (Table 3).

Table 3

Characteristics of tasks at different levels of mathematical complexity (National Assessment Governing Board, 2006, pp. 36-40)

<p>Characteristics of tasks at different levels of mathematical complexity</p> <p>Low Complexity</p> <p>This category relies heavily on the recall and recognition of previously learned concepts and principles. Items typically specify what the student is to do, which is often to carry out some procedure that can be performed mechanically. It is not left to the student to come up with an original method or solution. The following are some, but not all, of the demands that items in the low-complexity category might make:</p> <ul style="list-style-type: none"> • Recall or recognize a fact, term, or property. • Recognize an example of a concept. • Compute a sum, difference, product, or quotient.

- Recognize an equivalent representation.
- Perform a specified procedure.
- Evaluate an expression in an equation or formula for a given variable.
- Solve a one-step word problem.
- Draw or measure simple geometric figures.
- Retrieve information from a graph, table, or figure.

Moderate Complexity

Items in the moderate-complexity category involve more flexibility of thinking and choice among alternatives than do those in the low-complexity category. They require a response that goes beyond the habitual, is not specified, and ordinarily has more than a single step. The student is expected to decide what to do, using informal methods of reasoning and problem-solving strategies, and to bring together skill and knowledge from various domains. The following illustrate some of the demands that items of moderate complexity might make:

- Represent a situation mathematically in more than one way.
- Select and use different representations, depending on situation and purpose.
- Solve a word problem requiring multiple steps.
- Compare figures or statements.
- Provide a justification for steps in a solution process.
- Interpret a visual representation.
- Extend a pattern.
- Retrieve information from a graph, table, or figure and use it to solve a problem requiring multiple steps.
- Formulate a routine problem, given data and conditions.
- Interpret a simple argument.

High Complexity

High-complexity items make heavy demands on students, who must engage in more abstract reasoning, planning, analysis, judgment, and creative thought. A satisfactory response to the item requires that the student think in abstract and sophisticated ways. Items at the level of high complexity may ask the student to do any of the following:

- Describe how different representations can be used for different purposes.
- Perform a procedure having multiple steps and multiple decision points.
- Analyze similarities and differences between procedures and concepts.
- Generalize a pattern.
- Formulate an original problem, given a situation.
- Solve a novel problem.
- Solve a problem in more than one way.
- Explain and justify a solution to a problem.
- Describe, compare, and contrast solution methods.
- Formulate a mathematical model for a complex situation.
- Analyze the assumptions made in a mathematical model.
- Analyze or produce a deductive argument.
- Provide a mathematical justification.

GAISE Framework (Franklin et al., 2007). In addition to the NAEP framework for Mathematical Complexity, I applied the GAISE Framework to code the developmental levels of statistical components embedded in the tasks. The GAISE Framework consists of four statistical components: *formulate questions, collect data,*

analyze data, and interpret result. The framework also consists of the three developmental levels: *Level A, Level B, and Level C* (for a detailed description of the framework, see Figure 1).

Purpose and utility framework (Ainley et al. 2006). This framework was used to determine if the instances were purposeful (meaningful) and the mathematics ideas were useful for students: Were they introduced purposefully and did they lead to meaningful outcomes for the learner, either with an actual or virtual product? I also coded for the utility of the embedded mathematical/statistical concepts/objects when solving the problem at hand.

Common Core State Standards for Mathematics – Mathematical Practices (2010). The CCSSM includes eight Standards for Mathematical Practices (SMP) that students should know and be able to apply when doing mathematics. These practices are intended to be observed in the classroom during actual teaching and learning. However, in the absence of classroom observations, I identified references to SMP within the textbook instances and narrative in the teacher's editions. The performance expectation framework developed in TIMSS study (Schmidt et al., 1996, 1997) was used to provide examples for natures of the instances, which offers the opportunities to look for the practices. The eight mathematical practices are:

1. Make sense and persevere on solving problem.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.

6. Attend to precision.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

Table 4 provides example of one mathematical practice and descriptors of the practice that I use to code bivariate data instance. For full descriptions about SMPs and their coding framework see Appendix A.

Table 4

Example of description of SMP and its coding framework

CCSSM Statement	Descriptors of the coding framework
<i>Construct viable arguments and critique the reasoning of others</i> [S]tudents understand and use stated assumptions, definitions, and previously established results in constructing arguments. They make conjectures and build a logical progression of statements to explore the truth of their conjectures. They are able to analyze situations by breaking them into cases, and can recognize and use counterexamples. They justify their conclusions, communicate them to others, and respond to the arguments of others. They reason inductively about data, making plausible arguments that take into account the context from which the data arose. [S]tudents are also able to compare the effectiveness of two plausible arguments, distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is. ...Students learn to determine domains to which an argument applies. Students ... can listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments. (CCSSI, 2010, pp. 6-7)	The task provides opportunities for students to argue or critique. Look for the performance expectations from the task that ask students to (a) verify the computational correctness of a solution, or justify a step in the solution, (b) identify information relevant to verify or disprove a conjecture, (c) argue the truth of a conjecture or construct a plausible argument, (d) identify a contradiction (something that is never true), (e) critique a written or spoken mathematical idea, solution, result, or method for solving a problem and the efficiency of the method or similarly critique an algorithm and its efficiency.

Summary of the coding scheme for bivariate data.

All the coding is applied for individual instance related to bivariate data. Table 5 summarizes the framework for coding instances related to bivariate data.

Table 5

Framework for coding bivariate data

1. Learning Trajectory <ul style="list-style-type: none"> • Combination of two variables <ul style="list-style-type: none"> ▪ Two categorical variables <ul style="list-style-type: none"> ✓ Form of data provided (Raw data, tables, graphs, unspecified)
--

<ul style="list-style-type: none"> ✓ Size (Dimension of the two-way table such as 2x2, 2x3, 3x3, etc.) ✓ Types of representation involved (Tables, numeric/algebraic formulas, pie graphs, bar graphs, etc.) ✓ Types of association (Negative, positive, or unspecified) ▪ One categorical variable and one numerical variable <ul style="list-style-type: none"> ✓ Form of data provided (Raw data, tables, graphs, unspecified) ✓ Size (Number of data points for each sample/group) ✓ Types of representation involved (Tables, two boxplots, two histograms, back-to-back stem-and-leaf plots, etc.) ✓ Types of association (Negative, positive, or unspecified) ▪ Two numerical variables <ul style="list-style-type: none"> ✓ Form of data provided (Raw data, tables, graphs, unspecified) ✓ Size (Number of data points) ✓ Types of representation involved (Tables, scatterplots, order-case value bars, scatterplot slices, superimposed color gradient graph, algebraic, etc.) ✓ Types of association (Types of functional relationship)
<ul style="list-style-type: none"> • Contextualize/Decontextualized <ul style="list-style-type: none"> ▪ Context <ul style="list-style-type: none"> ✓ Consistent with theory ✓ Inconsistent with theory ✓ Unspecified ▪ Decontextualized • Task, technique, theory (TTT) framework <ul style="list-style-type: none"> ▪ (Estimating r, interpreting r, r^2, judgment of association, magnitude of association, association and causation) • Learning Goals
2. Task Features <ul style="list-style-type: none"> • Mathematical Complexity - Low, Moderate, and High • GAISE Framework <ul style="list-style-type: none"> ▪ Formulate questions – Level A, B, C ▪ Collect data – Level A, B, C ▪ Analyze data – Level A, B, C ▪ Interpret results – Level A, B, C • Purpose and Utility framework <ul style="list-style-type: none"> ▪ Purpose ▪ Utility • Standards for Mathematical Practice <ul style="list-style-type: none"> ▪ Make sense of problems and persevere in solving them (MP1) ▪ Reason abstractly and quantitatively (MP2) ▪ Construct viable arguments and critique reasoning of the others (MP3) ▪ Model with mathematics (MP4) ▪ Use appropriate tools strategically (MP5) ▪ Attend to precision (MP6) ▪ Look for and make use of structures (MP7) ▪ Look for and express regularity in repeated reasoning (MP8)

Sample Application of the Coding Scheme

To illustrate how the coding scheme was applied, I provide the following coding example from CPMP *Course 1*, Unit 1, Investigation 1 (Hirsch et al., 2008a, p. 5) relating the stretched length of a bungee cord to different weights that are attached (Figure 12).

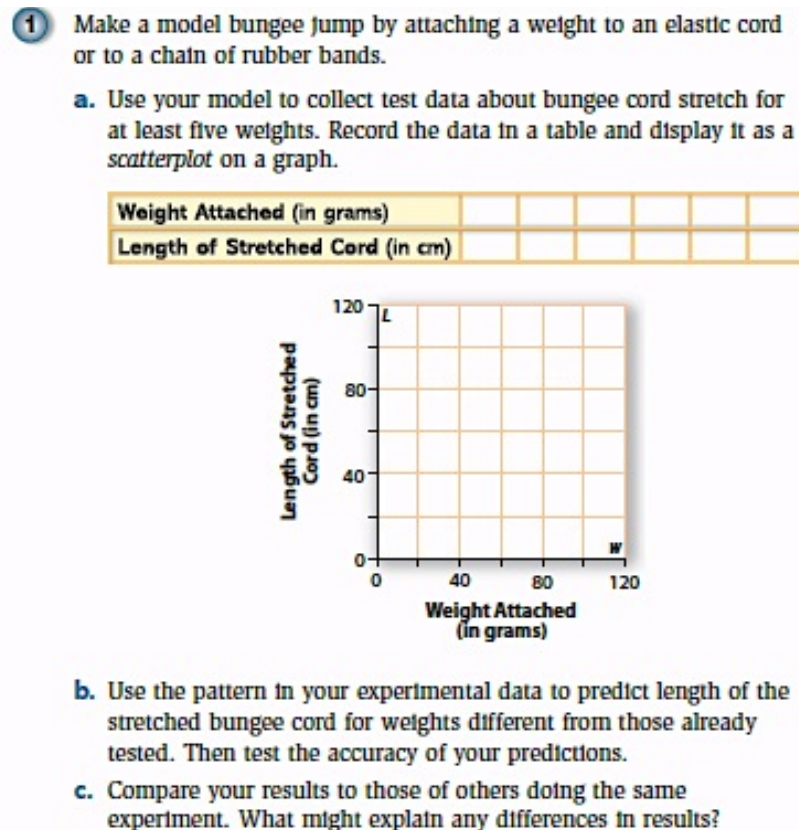


Figure 12. Sample coding for one instance. From *Core-Plus Mathematics: Course 1 – Teacher’s Guide* by C. R. Hirsch et al., 2008a, Columbus: OH: Glencoe/McGraw-Hill, p.5. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

1. Learning Trajectories

- **Combination of two variables**, namely two numerical variables
 - Form: No data are provided
 - Size: sample 6 data points
 - Type of association: positive linear

- **Contextualize/De-contextualized:** Data were provided in contextual setting and the type of association from the data is consistent with general belief (as the attached weight increases, the stretch of the cord also increases).
- **Task, technique, theory (TTT) framework:**
 - Task: Bungee Cord.
 - Techniques: Students carry out experiment to collect data, record data in a table and make a scatterplot. They observe the pattern of increasing/decreasing the scatterplot, draw a line passing through the data points and use the line to predict one variable given the other or write a linear equation and use it for prediction. Using the data, students compare the difference in the results, the error in measurement and the physical feature of the bungee cord that might account for the difference.
 - Theories: behind the action are: observation and functional relationship.
- **Learning goals:** “use table and graph, algebraic expression to express the relationship among variables” (p. 3).

2. Task Features

- **Mathematical Complexity:** Moderate complexity. The students carry out multiple steps that go beyond the habitual and need to decide what to do.
- **GAISE Framework (Level A, B or C)**
 - Formulate questions. This instance provides a context for a statistical investigation; however, students do not have to create a question, therefore it is coded as Level A.

- Collect data. Students have to do a simple experiment to collect data without being aware of design for differences, therefore the instance is coded as Level A.
- Analyze data. Students are supposed to use simple models for association (eyeballed line), therefore the instance was coded as Level B.
- Interpret results. Students make basic interpretation of models for association, using the equation or the eyeballed line to predict the stretch given a weight, therefore the instance was coded as Level B.
- **Utility and Purpose framework (Yes - No)**
 - Purpose: In this instance, students have a final product to use, which is a formula or a line to predict the value of stretch when given a weight, therefore this instance is coded YES for the purpose feature.
 - Utility: In this instance, students can understand why scatterplots are used and how they are used for eyeballing line and prediction, therefore this instance is coded YES for the utility feature.
- **Standards for Mathematical Practice (SMP)**
 - In this instance, students need to make sense of the weight and length and reason about the relationship between them (MP2: Reason abstractly and quantitatively); they need to “explain any differences in results” (MP3: Construct viable arguments and critique reasoning of others); they simplify to make an equation to model the relationship (MP4: Model with mathematics); they need to be precise when measuring, graphing, and calculating (MP6:

Attend to precision); and they need to look for a pattern to predict the length (MP7: Look for and make use of structures).

The coding scheme was applied for all instances related to bivariate data. Then, I used the collected data for further analyses, which are described in the next section.

Data Analysis

Question 1. Learning Trajectories for Bivariate Data in the Curricular Materials and the Comparison with those in the CCSSM and GAISE Framework

For each textbook series, I computed descriptive statistics of learning trajectories codes that are summarized in Table 5. Additional information related to the tasks including form of data provided, size, form of representations, and contextual vs. decontextualized codes are also reported as summary statistics. The total and percentages of tasks in each series are reported in various levels: by chapter/unit, by textbook, and the whole series. The distributions of instances in the three combinations of bivariate data are reported at the chapter level.

Qualitative analysis reveals the learning trajectories of these topics related to bivariate data. At the *lesson-level*, codes are specified for learning goals. The instructional activities related to the content and process of learning were examined using the technique and theory codes. The levels of the trajectories were specified: within the tasks, between the tasks, and between lessons. Information gained from the TTT framework analysis was used to further answer this research question by using the results coding the techniques and theories associated with each instance.

For each lesson that includes instances related to bivariate data, I identified the learning goals and the flow of content. Data from *learning goals* code were used for the

start of specifying the flow of content for the lesson. The flow of a lesson is primarily in the development of the content, which is referred as *development portion* (Jones, 2004), aimed for students to involve in classroom. The flow includes how textbook authors introduce and develop the content through definitions, explanations, and examples. In some textbook series (e.g., CPMP), the authors include investigations and a list of guiding questions for students to learn the content. By solving this task or referring to the suggested strategies offered in the teacher' edition, I specified the flow. In some cases, new learning expectations are introduced in the *assignment portion* (Jones, 2004) after the development of content is finished. When such cases are evident, these learning expectations are also reported. Learning goals and the flow of content for each lesson provide the learning trajectories at the *micro level*.

After lesson's learning goals are identified, I grouped related learning goals among lessons and chapters to form the learning trajectories for bivariate relationships for the whole textbook series. In addition, the prescribed scope and sequence was used to specify the connection among the topics. When the authors did not explicitly note the sequence in the teacher' edition, the order of topics as appeared in the textbooks series was used to specify the flow. The findings related to the connection among topics related to bivariate data provide the *macro level* for the learning trajectories. The reporting of learning trajectories are grouped into three combinations of two variables.

Along with this process, knowledge/skills from other mathematical strands that are needed to solve tasks related to bivariate data were identified. For example, to examine the association between one categorical and one numerical variable (e.g., gender vs. height), students need to compare distributions (in this case, distributions of male and

female heights). Therefore, students need to understand and use distributions of univariate data to examine the association for bivariate data. I also used the teacher's editions as guidance to specify connections among bivariate data, univariate data, and other CCSSM conceptual categories at the high school level.

In order to analyze the alignment of textbook series with the CCSSM regarding bivariate data, I examined the learning expectations (LEs) related to bivariate data present in the CCSSM. Eight standards related to bivariate data were found in the CCSSM. In some cases, the original standard is too broad because it comprises several outcomes. Hence, using this standard to identify if it was addressed in textbook series might sacrifice some level of precision. For example, consider the following standard S-ID.6:

Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models.

b. Informally assess the fit of a function by plotting and analyzing residuals.

c. Fit a linear function for a scatter plot that suggests a linear association. (p. 81)

This standard consists of eight LEs: (1) Represent data on two quantitative variables on a scatter plot, and describe how the variables are related; (2) Fit a function to the data (emphasize linear models); (3) Fit a function to the data (emphasize quadratic models); (4) Fit a function to the data (emphasize exponential models); (5) Use functions fitted to data to solve problems in the context of the data; (6) Use given functions or choose a function suggested by the context; (7) Informally assess the fit of a function by plotting;

and (8) Informally assess the fit of a function by analyzing residuals. More often, for any given instance, several of the eight LEs but not all are addressed. By breaking up the standards into LEs at smaller grain sizes, I am afforded greater precision in coding the instances of bivariate data that appear in the textbook series. Table 6 provides a list of 21 LEs taken from eight CCSSM standards related to bivariate data at the high school level.

Table 6

Learning expectation related to bivariate data at the high school level of the CCSSM

Code	LEs
S-ID.2	Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets
S-ID.5	Summarize categorical data for two categories in two-way frequency tables.
	Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies).
	Recognize possible associations and trends in the data.
S-ID.6	Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.
	Fit a function to the data (emphasize linear models).
	Fit a function to the data (emphasize quadratic models).
	Fit a function to the data (emphasize exponential models).
	Use functions fitted to data to solve problems in the context of the data.
	Use given functions or choose a function suggested by the context.
	Informally assess the fit of a function by plotting.
	Informally assess the fit of a function by analyzing residuals.
S-ID.7	Interpret the slope (rate of change) of a linear model in the context of the data.
	Interpret the intercept (constant term) of a linear model in the context of the data.
S-ID.8	Compute (using technology) the correlation coefficient of a linear fit.
	Interpret the correlation coefficient of a linear fit.
S-ID.9	Distinguish between correlation and causation.
S-IC.3	Recognize the purposes of and differences among sample surveys, experiments, and observational studies.
	Explain how randomization relates to each [type of the studies].
S-IC.5	Use data from a randomized experiment to compare two treatments.
	Use simulations to decide if differences between parameters are significant.

For each instance related to bivariate data identified, I coded one or several LEs (from the list of 21 LEs in Table 6) that match the intent of the instance. If no LE was

found to match with the instance, I left the cell for the instance blank and noted the content that the instance addressed. After this coding, I identified LEs in the CCSSM covered in textbooks and how many times the LEs appeared as counted by number of instances. I also specified the list of LEs that appear in textbook series but not in the list of 21 CCSSM LEs.

For the question regarding the GAISE Framework, the progressions of topics related to bivariate data described in GAISE Framework are reported. This information is taken from suggestion of sequencing in three developmental levels of bivariate data topics. In particular, I summarized the progressions for each combination of two variables as appeared in the GAISE Framework. I then used the progressions and examined which developmental levels were addressed and in what order they appeared in each textbook series. I also report whether specific components recommended in the framework appeared in the textbook series and which components appear in the textbook series but not in the GAISE Framework.

Question 2. Nature of Tasks and Opportunities to Access the CCSSM standards for Mathematical Practices

In order to address the nature of tasks research question, I used information from the task features portion of the coding framework. In particular, for each of the three series, I report the distribution of the level of mathematical complexity, using mean and SD. In order to illustrate the mathematical complexity framework, I also provide examples of the three levels of mathematical complexity.

For the results regarding the GAISE Framework coding, for each textbook series, I report the percentages of instances related to bivariate data that addressed each of the

four components of doing statistics. I then summarize the developmental levels using means and SDs of each component for each series. For illustration purposes, I also include examples of multiple components at the three developmental levels.

In order to measure the association between task and GAISE frameworks, I use Goodman and Kruskal's Gamma test, which is applicable for measuring association among ordinal categorical variables. That is, I examine whether the higher levels of cognitive demands according to the mathematical complexity framework were associated with higher levels of development in the GAISE Framework.

With respect to *purpose and utility*, for each the three series I report the percentages of instances that addressed each feature and then compare the percentages across the three series. I also provide illustrative examples of instances that addressed two features, one feature, and no feature of the *purpose and utility* framework.

Finally, regarding to the CCSSM mathematical practices coding, I provide the percentages of instances that addressed each of the eight practices for the three series. I then report on the prevalence of appearing, by practice and by series. In order to illustrate the features of instances, I provide specific examples related to each SMP.

The previous sections describe the coding and analysis of data. Table 7 summarizes the relationship between research questions and sources of data collected.

Table 7

The Mapping of Data Sources to Research Questions

Research Question	Sources of Data
1a	Type of data Contextual-based/abstract setting and type of association TTT Framework Learning Goals
1b	TTT Framework Learning Goals
1c	TTT framework Learning goals Analysis of the CCSSM and GAISE Framework
2a	Mathematical Complexity
2b	GAISE Framework
2c	Purpose and Utility Framework
2d	Mathematical Practices

Validity and Reliability

A validation of the coding framework was carried out before actual coding. I consulted with a mathematics educator who has expertise in statistics education, a statistician, and a Ph.D. candidate in statistics in order to decide the boundary for the inclusion of instances to be coded. In addition, the lead author of GAISE Framework also helped to clarify the GAISE Framework descriptors and to distinguish between mathematical and statistical tasks.

The reliability of coding was established through two procedures. First, two additional coders were trained and in order to establish inter-coder reliability. A concise description of the coding framework and a manual for the coding process were provided to coders for reference as they worked. The instructional instances from *CPMP's* Investigation 1 of *Course 1*, Unit 1, Lesson 1 (Hirsch et al., 2008) were used for training. Three meetings of two hours each were scheduled to develop an adequate understanding of the coding framework and application of the framework. After the first meeting, the

two coders were asked to code two other investigations in the same lesson. Their results were compared for consistency and verification and useful background information for planning of the second meeting. Discrepancies of coding were discussed in the second meeting and reconciled in the set of instances. After the two meetings, each coder independently coded 21 instances related to bivariate data from each textbook series. The third meeting was scheduled to reconcile and finalize the coding system.

Once the practice coding was completed satisfactorily, each coder was assigned another set of instances purposefully selected because they include bivariate data content. For Coder 1, CPMP *Course 2*, Unit 4 – Regression and Correlation (Hirsch et al., 2008) was used, containing a total of 69 instances related to bivariate data. For Coder 2, Chapter 2 of the UCSMP FST book – Functions and Models was used, consisting of a 66 instances related to bivariate data. As a check of test-retest reliability, I recoded 66 instances of UCSMP *FST*, Chapter 2 two months after the initial coding to ensure that I was coding the instances in a consistent manner across time.

Table 8 displays the inter-rater and test-retest reliability coefficients for each code. Some dimensions were more reliably coded than others. For example, Combination of variables, Type of association, and Contextual/Decontextualized were fairly obvious within each instance based on the high rate of agreement for these codes (more than 90%). In contrast, the code for Mathematical Complexity, The GAISE Framework components, and Standards for Mathematical Practice (SMP) required some level of interpretation to determine. Notwithstanding, the agreement was reached on all but few (4) codes related to the GAISE Framework, with reliability coefficients exceeding 70%.

There are some disagreements in coding for the GAISE Framework. For Coder 1, there were rather low rates of agreement in the assignment of codes related to Formulate Questions and Analyze Data (63% and 57%, respectively). For Coder 2, the lowest agreement was observed on the Analyze Data code (55%) followed by Interpret Results (62%). Final discussions with the coders revealed the reason for the inconsistencies. In particular, in regard to the GAISE Framework - Formulate Questions, Coder 1 did not distinguish between statistical questions and questions to analyze data and, as a result, she coded all tasks that required students to analyze data. With respect to GAISE Framework - Analyze Data, the coding framework was aimed to attend to tools (such as graphical display and numerical statistics) required when solving the problems. However, both Coder 1 and Coder 2 took the level of cognitive demands into account when coding tasks and, as a result, both lower down the developmental level of the instance. For example, instance asking students to use graphing calculator to find the linear regression is supposed to be at Level C, but the coders categorized it as Level A because not much cognitive demands are required to solve this task; students must simply input the data and use calculator features to find the equation for linear regression. With respect to the GAISE Framework - Interpret Results, Coder 2 conceptualized it slightly differently when taking into account the contextual information from the teacher's edition. For example, when an instance asked students to use the model for prediction, it should be coded as Level B of the GAISE Framework for Interpret Results. Despite this, Coder 2 did not code such an instance as Level B because the measurement unit of prediction was not reported in the teacher's edition. However, all coding differences were negotiated and

eventually reconciled. More importantly, differences in the assignment of codes were fairly uncommon overall.

Table 8

Reliability measures

Codes	Agreement with Coder 1 (N=49)	Agreement with Coder 2 (N=65)	Test-Retest Agreement (N=65)
Learning Trajectories			
Combination of variables	96%	98%	100%
Type of association	94%	92%	92%
Contextual/Decontextualized	94%	100%	100%
Mathematical Complexity	78%	86%	88%
GAISE Framework			
Formulate Questions	63%	89%	92%
Collect Data	98%	98%	98%
Analyze Data	57%	55%	92%
Interpret Results	88%	62%	89%
Purpose-Utility Framework			
Utility feature	78%	97%	94%
Purpose feature	71%	83%	91%
Standards for Mathematical Practice			
MP1	N/A	N/A	95%
MP2	78%	86%	87%
MP3	82%	92%	92%
MP4	71%	86%	91%
MP5	78%	97%	94%
MP6	76%	94%	98%
MP7	71%	78%	91%
MP8	71%	98%	100%

Summary

In this study, data were collected and recorded through an examination of the teacher's editions of three commonly accepted textbook series. Every page in the selected series that addressed bivariate data topics was analyzed using specific criteria. A coding framework was developed to identify and collect data in order to address two problems of interest: learning trajectories for bivariate data and the features of student

learning tasks associated with the content of the textbooks. I sought to address two key aspects of textbook content: *what* content is covered and *how* the content is presented. Supports for the study of these aspects of textbook analysis were suggested in a NRC report (2004) and by Stein and colleagues in the *Second Handbook of Research on Mathematics Teaching and Learning* (2007). The study approach was more descriptive than evaluative even though it included elements of several established frameworks for curriculum analyses. Selected discipline perspectives in content analysis as suggested by NRC (2004) were embedded in the process: clarity of objectives, comprehensiveness, accuracy, depth of mathematical inquiry and mathematical reasoning, organization, and balance. Finally, to support the validity of my conclusions, I made a conscientious effort to address the reliability concerns related to the application of the coding framework.

CHAPTER 4: ANALYSIS OF DATA AND RESULTS

For this study, I analyzed bivariate data tasks of three U.S. high school textbook series: *Holt McDougal Larson* series (HML), *The University of Chicago School Mathematics Project* series (UCSMP), and *Core-Plus Mathematics Project* series (CPMP). I examined the learning trajectories and the nature of tasks for bivariate data as presented in the teacher's editions. In particular, I documented how learning trajectories for bivariate relationships were introduced and developed within each series. For each series, I compared the presence of the topics with the learning expectations (LEs) emphasized at the high school level in the *Common Core State Standards for Mathematics* (CCSSM) for bivariate data. Furthermore, I compared the learning trajectories for bivariate data in the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) Framework with those in each textbook series. In addition, I analyzed each instance related to bivariate data for the following features: level of mathematical complexity, the components of doing statistics and developmental levels of the GAISE Framework, the Purpose and Utility Framework, and the eight CCSSM's Standards for Mathematical Practice (SMP).

Results are organized into two sections around the research questions. In the first section, I report findings related to the research questions about learning trajectories for bivariate data. In the next section, I share results of my analysis of the features of tasks related to bivariate data in the textbook series.

Learning Trajectories

In this section, I organize my results into four sub-sections: (1) Distribution of Instances of Bivariate Data across Three Textbook Series, (2) Description of Learning

Trajectories within Each Series, (3) Alignment of Textbook Series with the Common Core State Standards for Mathematics Learning Expectations, and (4) Comparison of Textbook Series and the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) Framework Learning Trajectories.

Distribution of Instances of Bivariate Data across Three Textbook Series

Across the three series, there were 582 instances of to-be-solved (TBS) problems (in which solutions were not present) including 699 TBS tasks¹ identified and coded for bivariate data. Table 9 summarizes the number of TBS instances and tasks across the three series. Because the structure of tasks varies across the three series, careful interpretations must be made when comparing these numbers. However, given that the structure of tasks is largely consistent within a given textbook series, the percentages of tasks related to bivariate data could provide a valid comparison across the three series. As observed in Table 9, the percentages of tasks were quite similar between HML and UCSMP and the percentage of TBS tasks within the CPMP series was about four-to-five times larger than the HML and UCSMP series percentages.

Table 9

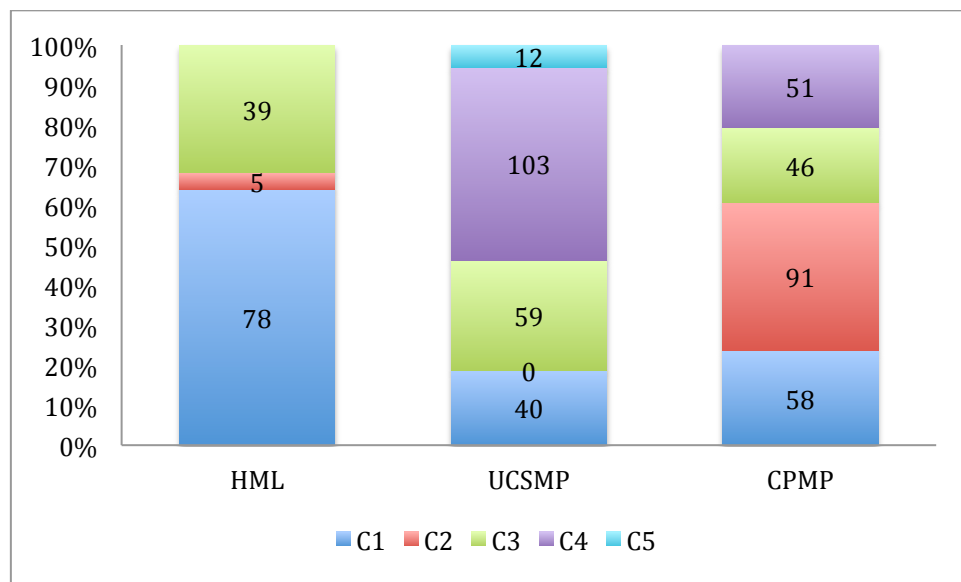
Instances related to bivariate data

Textbook Series	Number of TBS <i>instances</i> related to bivariate data	Number of TBS <i>tasks</i> related to bivariate data*	Number of TBS tasks across all topics in series	Percentage of TBS tasks related to bivariate data
HML	122	165	17404	0.9
UCSMP	214	269	20693	1.3
CPMP	246	265	4993	5.3

* More than one TBS task may be included in a single TBS instance

¹ An instance might include one or more task[s]. For an example, see chapter 3.

The number of instances related to bivariate data also varied across textbooks within a series. Figure 13 illustrates the number of bivariate data instances in each textbook for the three series. In the HML series, the fewest number of instances (5) appeared in the *Geometry* book, whereas the greatest number of instances (78) appeared in the *Algebra 1* book. Similarly, for UCSMP at the low end, no instances related to bivariate data appeared in the *Geometry* book, and at the high end, 103 instances appeared in the *Functions, Statistics, and Trigonometry* book (*FST*). In contrast, CPMP bivariate data instances were more uniformly distributed over four books (*Course 1* to *Course 4*) ranging from 46 in *Course 3* to 91 instances in *Course 2*.



Note. The letters and numerals = Course and Order followed by book title

	C1	C2	C3	C4	C5
HML	<i>Algebra 1</i>	<i>Geometry</i>	<i>Algebra 2</i>		
UCSMP	<i>Algebra</i>	<i>Geometry</i>	<i>AA</i>	<i>FST</i>	<i>PD</i>
CPMP	<i>Course 1</i>	<i>Course 2</i>	<i>Course 3</i>	<i>Course 4</i>	

Figure 13. Number of instances related to bivariate data across the textbooks in each series.

Figures 14, 15, and 16 depict the distributions of the instances related to bivariate data by chapters for each of the three series. For the HML series, about 60% of the total number of instances related to bivariate data appeared in the *Algebra 1* textbook, and 70% of those instances were in Chapter 4 of the textbook. About 35% of the bivariate data instances appeared in the HML *Algebra 2* textbook; those instances were distributed throughout Chapters 4, 5, 6, and 10. In the UCSMP series, about half of the bivariate data instances appeared in *FST*, followed by the *Advanced Algebra* (28%) and *Algebra* (22%) books. The majority of those instances appeared in Chapter 2 of *FST*, Chapter 6 of the *Algebra* book, and Chapter 2 of the *Advanced Algebra* book. The instances related to bivariate data were dispersed in many chapters more than those in the HML series. For the CPMP series, all four books contained instances related to bivariate data. The plurality of those instances appeared in Unit 4 of the *Course 2* book, Unit 5 of the *Course 4* book, and Unit 1 of the *Course 3* book. The remaining instances related to bivariate data were distributed fairly evenly across five units² of the *Course 1* book.

In regard to the combinations of variables, all three combinations—two categorical variables (CC), one categorical and one numerical variable (CN), and two numerical variables (NN)—were found in all three textbook series. However, the distributions of the instances were quite different for each combination of variable types. Generally, the vast majority of the instances used NN variables (Figure 17). This was addressed in every chapter containing instances related to bivariate data. In particular, about 77% of the total bivariate data instances are related to NN variables in CPMP to 90% of those instances appeared in UCSMP. The percentage of instances

² CPMP uses *unit* as an equivalence to *chapter* in HML and UCSMP.

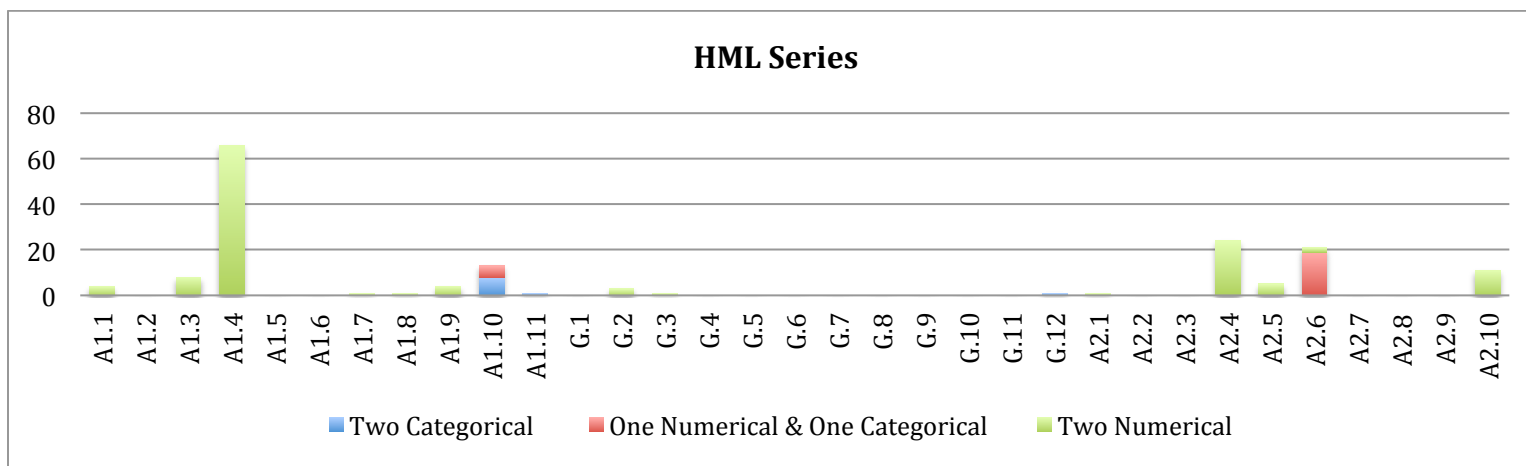


Figure 14. Distribution of the three combinations of bivariate data instances by chapters in the HML series (e.g., A1.1=Algebra 1.Chapter 1).

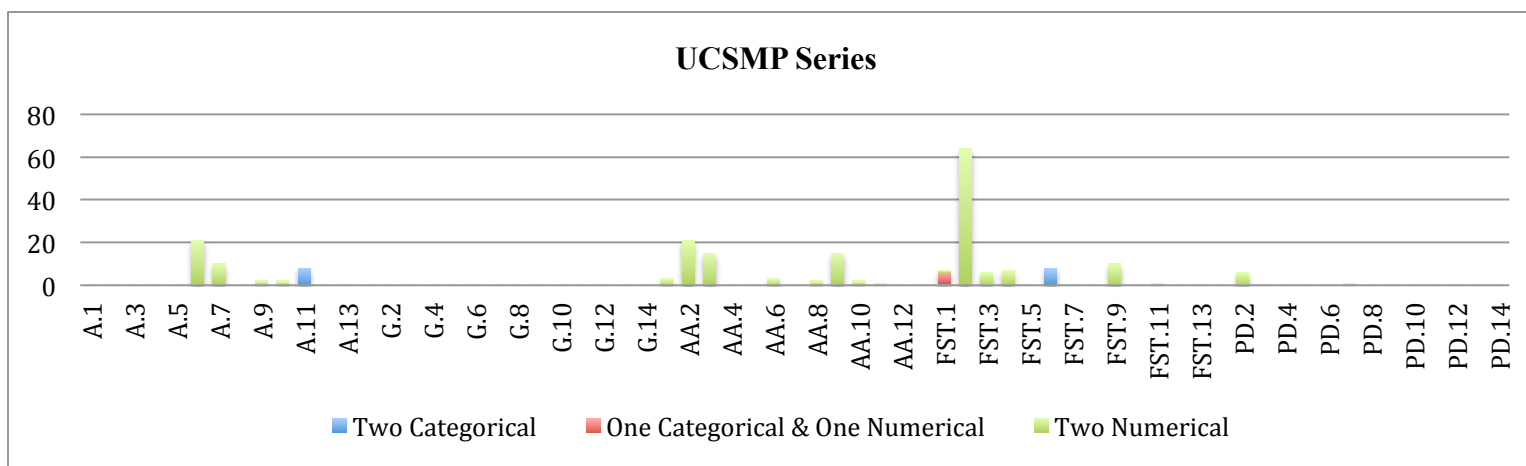


Figure 15. Distribution of the three combinations of bivariate data instances by chapter in the UCSMP series (e.g., A.1=Algebra.Chapter 1).

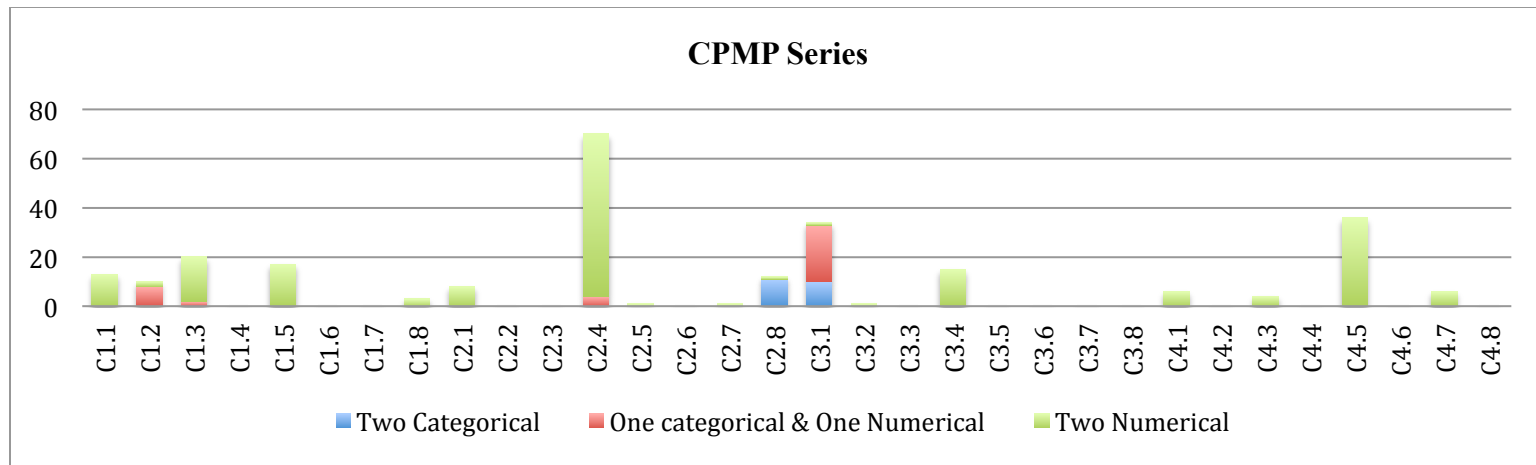


Figure 16. Distribution of the three combinations of bivariate data instances by chapter in the CPMP series (e.g., C1.1=Course 1.Unit 1).

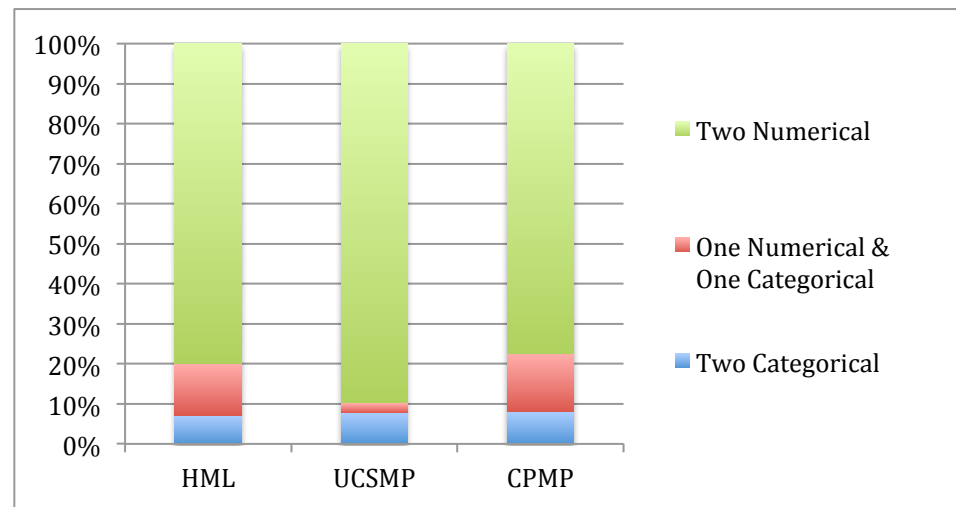


Figure 17. Percentages of instances in three combinations of bivariate data across three series.

related to CN variables was smallest for the UCSMP series (about 2%) and largest for the CPMP series (about 14%). In the CC combination, the percentage was smallest for the HML series (about 7%) and largest for the CPMP and UCSMP series (about 8%).

Of the instances related to NN combination, only 1 of 131 instances in the HML series, 13 of 244 in the UCSMP series, 29 of 214 in the CPMP series contained a set of 20 or more data points. For the CC combination, table dimensions other than 2x2 were found in the two of the three series. In particular, the UCSMP series included 2x4, 2x5, 2x6, 2x7, 2x8, 2x12, 3x4, and 3x5 tables and HML series included 2x3, 3x3, and 3x4 tables. In addition, when data were provided, the vast majority of the instances were presented in an organized table as opposed to raw data form. In particular, only 2 instances in the HML series, 2 instances in the UCSMP series, and 10 instances in the CPMP series provided data in a raw format and required students to sort, categorize, and place the data in a well-formatted table.

In the HML series, the instances involving NN variables were distributed across the textbooks, whereas the other two combinations of variable types appeared in two locations in the series. In particular, the CC combination instances appeared in Chapters 10 and 11 of the *Algebra 1* book and Chapter 12 of the *Geometry* book, whereas, the CN combination appeared in Chapter 10 of *Algebra 1* and Chapter 6 in *Algebra 2* (Figure 14). A similar pattern was found for the UCSMP series; the CN combination appeared in one location and the CC combination appeared in two locations in the series. All but two chapters (Chapter 11 of the *Algebra* book and Chapter 6 of *FST*) contained instances related to NN variables (Figure 15). For the CPMP series, the CC combination instances appeared in two units. Compared with the other two textbook series, more units (4) in

CPMP contained instances related to CN combination. Similar to the other two series, instances related to NN combination appeared regularly in nearly every unit that addressed bivariate data (Figure 16).

Figure 18 depicts the proportion of functional models, by instances. For instances involving NN variables, multiple functional models for fitting data were addressed in each of the three series. In all three series, linear modeling appeared most often, at least twice as common as the second most often models, exponential modeling. The UCSMP series placed more emphasis on quadratic modeling than the other two series; in contrast, this series placed the least emphasis on power modeling compared with the HML and CPMP series.

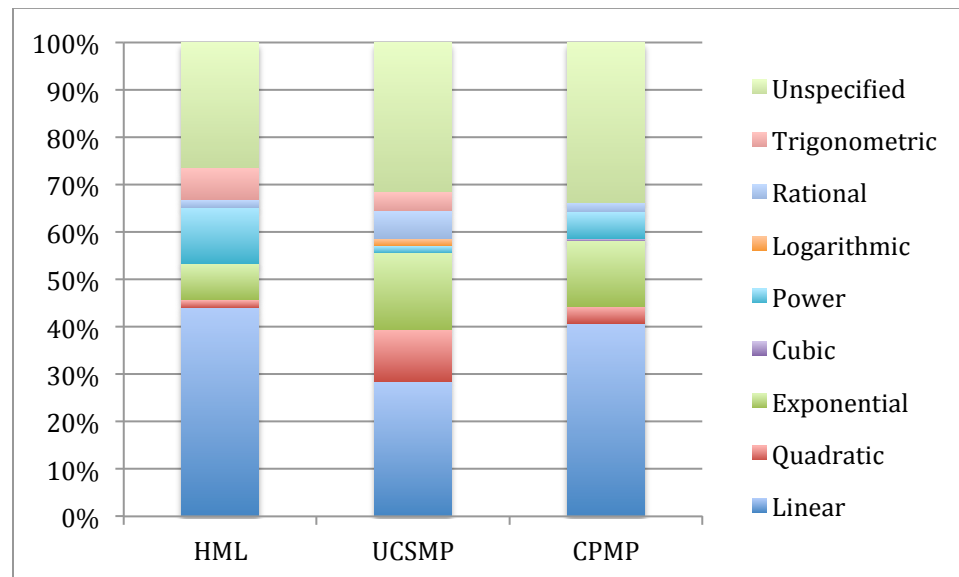


Figure 18. Types of functional models in the three series.

Description of Learning Trajectories within Each Series

The following section describes the progression of content for each of the three combinations of bivariate data. The description of learning trajectories was derived mainly from the development portion of each lesson. Instances that appear in the

homework assignment portion sometimes included learning expectations (LEs) that are somewhat different from those in the development portion. The LEs appeared in the homework assignment portion were also used to describe the trajectories. The descriptions of trajectories are organized into the three combinations of bivariate data for each series.

Holt McDougal Larson (HML) series.

HML series learning trajectory for two categorical variables. The relationship between CC variables is first introduced in Chapter 10, *Data Analysis* of the first book, *Algebra 1*, containing a total of 13 chapters. Specifically, in Lesson 10.3, *Analyze Data*, the authors of the HML series introduce two-way frequency tables as a means to determine the relationship between CC variables. They provide: (a) an example of the 2x2 table to illustrate joint and marginal frequencies, (b) an example of a two-way table addressing how to interpret the frequencies in the table, and (c) an example of developing a two-way table based on a description of data and using that table to compare the categories. Figure 19 illustrates different types of frequencies in the table while Figure 20 provides an example of making a two-way table and understanding frequencies to compare categories. The authors do not explicitly ask students to describe or quantify the relationship between the two variables, but merely to compare frequencies in the instances after the worked examples, which were referred as guided practices.

A two-way frequency table shows the number of items in various categories.
 Every element in the sample must fit into one of the categories and there must be no overlap between categories.

KEY CONCEPT

For Your Notebook

Two-way frequency table

A two-way frequency table divides the data into categories across the top and down the side.

	Apples	Oranges	Total
Boys	15	18	33
Girls	21	16	37
Total	36	34	70

The body of the table gives the **joint frequencies**.

The row and column totals give the **marginal frequencies**.

Figure 19. Illustrating frequencies in a two-way table. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.673. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

EXAMPLE 2

Make a two-way frequency table

Make a two-way frequency table for the following data.

There are 175 freshmen taking a foreign language. Of these, 88 take Spanish, 46 take French, and the rest take German. No one takes more than one language. There are 42 boys taking Spanish, 31 girls taking French, and a total of 89 girls taking a language.

Solution

The categories are Spanish, French, German, boys, and girls. Fill in the given information. Then look for ways to calculate the missing values.

For example, the number of girls taking Spanish is $88 - 42 = 46$. The number of boys taking a foreign language is $175 - 89 = 86$. The total number of students taking German is $175 - (88 + 46)$.

	Spanish	French	German	Total
Boys	42	15	29	86
Girls	46	31	12	89
Total	88	46	41	175

AVOID ERRORS

Be sure to enter the given information in the correct cells of the table.

✓

GUIDED PRACTICE

for Examples 1 and 2

1. Using the table in Example 1, tell whether more students in Mr. Cole’s homeroom prefer science or English. English
2. There are 152 students who play golf, basketball, or soccer. No one plays more than one of these sports. There are 22 who play golf, 50 who play basketball, and the rest play soccer. There are 10 boys who play golf, 26 girls who play basketball, and a total of 80 boys who play one of these sports. Make a two-way frequency table for the data. See margin.

Figure 20. An instance requiring students to read/construct two-way tables. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.674. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

In Chapter 11, *Probability*, of the *Algebra 1* textbook, a more complex use of two-way tables is introduced as a means to find probabilities, conditional probabilities, and in turn to judge the dependence of two events. It should be noted that this is a *theoretical view* of dependence and not a *statistical association* to judge the strength or statistical significance of the dependence of two events. The authors define independence as, “Two events are independent events if the occurrence of one event does not affect the occurrence of the other” (Larson et al., 2012a, p. 735). The property about the probability that both events occur equals to the product of probabilities for each event occurs is used to determine whether two events are independent (Figure 21). A nearly identical sequencing of definitions, examples, and TBS problems about the independence of two events appear in Chapter 12, *Probability*, in the *Geometry* book.

CONDITIONAL PROBABILITIES

The conditional probability of B given A can be greater than, less than, or equal to the probability of B .

The probability that event B occurs given that event A has occurred is called the **conditional probability** of B given A and is written $P(B|A)$. Note that A and B are independent events if and only if $P(B) = P(B|A)$ since the probability of event B does not depend on the occurrence of event A .

KEY CONCEPT

For Your Notebook

Probabilities of Independent and Dependent Events

Independent Events

For two independent events A and B , the probability that both events occur is the product of the probabilities of the events.

$$P(A \text{ and } B) = P(A) \cdot P(B) \quad \text{Events } A \text{ and } B \text{ are independent.}$$

Dependent Events

For two dependent events A and B , the probability that both occur is the product of the probability of the first event and the conditional probability of the second event given the first event.

$$P(A \text{ and } B) = P(A) \cdot P(B|A) \quad \text{Events } A \text{ and } B \text{ are dependent.}$$

The formulas for finding the probabilities of independent and dependent events can be extended to three or more events.

Figure 21. Probabilities of independent and dependent events. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.735. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

Summary of the LT for CC combination and its relation to other conceptual

categories. The general progression of concepts of association between CC variables is as follows: (a) know how to make a two-way table, interpret frequencies in a two-way table, and (b) use two-way tables to calculate conditional probability and examine the independence of two events. Figure 22 summarizes the learning trajectory for the CC combination.

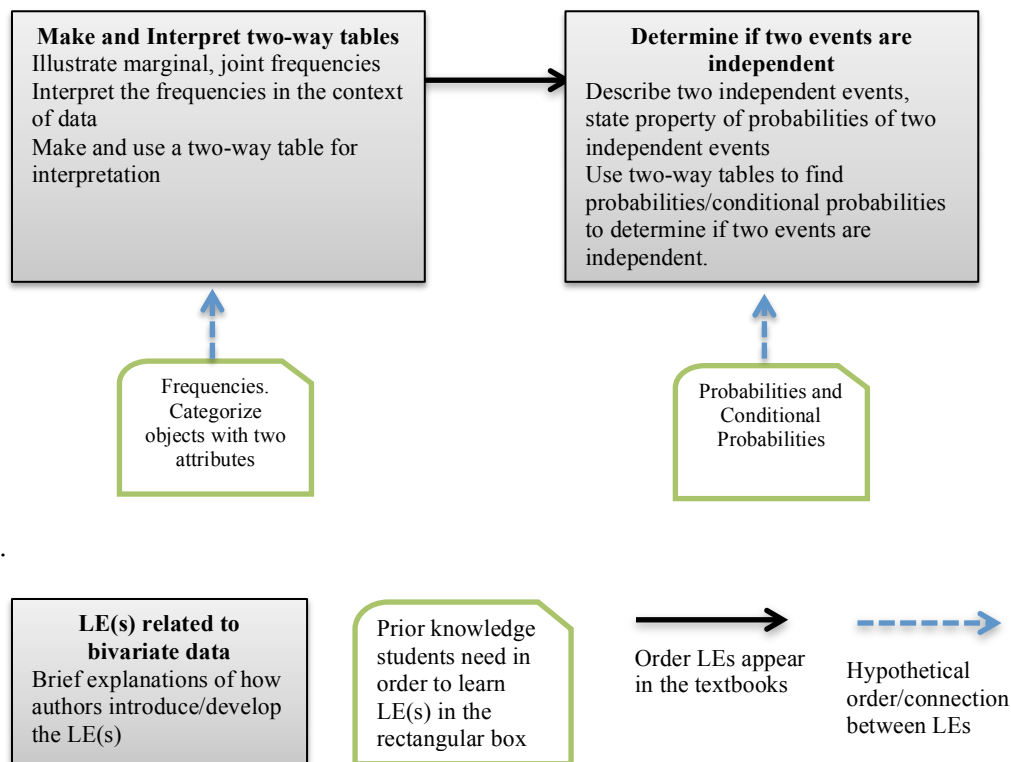


Figure 22: Learning trajectory for the CC association in the HML series.

For CC variables, the authors use conditional probability to judge the dependence of the events and used two-way tables as a means to calculate the probabilities. The connection between statistical association and conditional probabilities and frequencies is not explicitly addressed. Moreover, with the exception of the term *frequency*, no connection is made to univariate data topics for the CC combination.

HML series learning trajectory for one categorical and one numerical variable.

The association between CN variables is introduced and developed in Chapter 10, *Data Analysis* of the *Algebra 1* book. The concept of association is informally and implicitly introduced in Lesson 10.4, *Interpret Stem-and-Leaf Plots and Histograms* and Lesson 10.5, *Box-and-Whisker Plots*. In particular, the authors provide one guided practice using a back-to-back stem plot comparing the time males and females spent watching television (Figure 23).

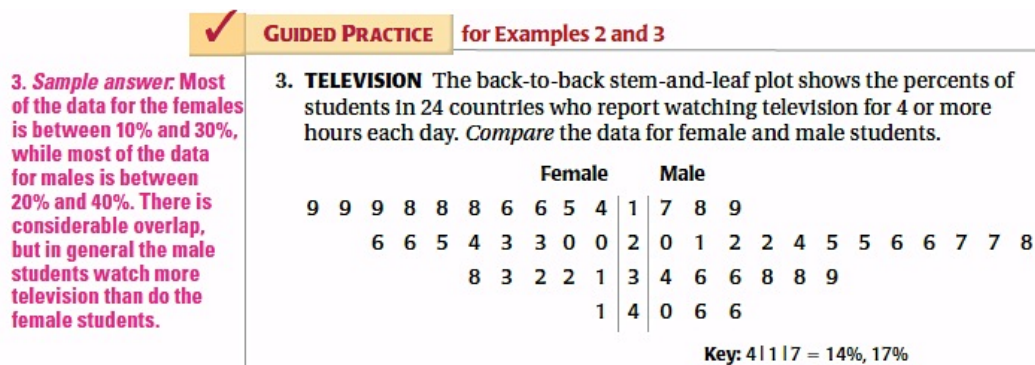


Figure 23. Comparing two distributions using back-to-back stem-and-leaf plot. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.681. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

Similarly, the authors introduce box-and-whisker plots (boxplots) and how to make and interpret the plots. They then provide an example related to comparing two distributions using two side-by-side boxplots. In one instance in the assignment portion, the authors provide side-by-side boxplots and ask students to draw a conclusion about the relationship between two variables (Figure 24). Students are expected to use the five-number summary of a distribution: median, two quartiles, maximum value and minimum value to compare the distributions. Three of four instances require students to compare two distributions of males and females in an attribute but do not explicitly expect students to use the comparison as a means to address the association between two variables.

20. **CHALLENGE** The box-and-whisker plots show the heights (in inches) of singers in a chorus, according to their voice parts. A soprano part has the highest pitch, followed by alto, tenor, and bass, respectively. Draw a conclusion about voice parts and heights. *Justify* your conclusion. **See margin.**

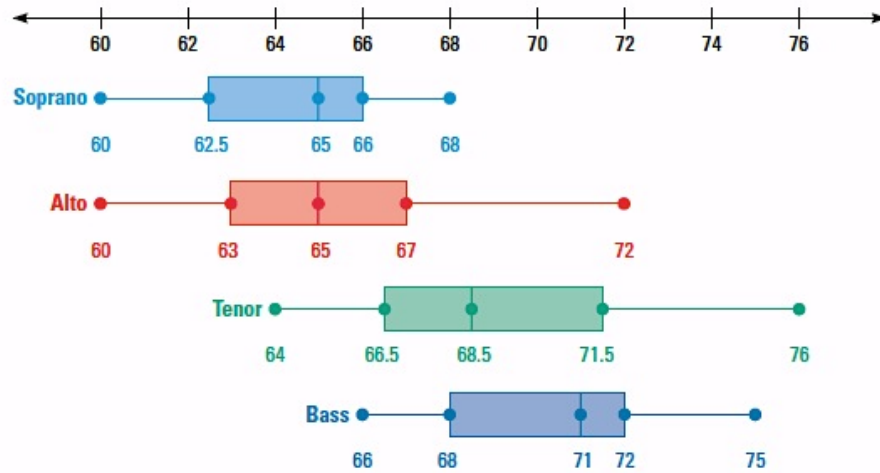


Figure 24. Comparing distributions using box-and-whisker plots. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.690. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

In Chapter 6, *Data Analysis and Statistics of Algebra 2*, in the context of addressing the methods of scientific research (e.g., surveys, experiments, and observational studies), the authors introduce an activity that involved testing the hypothesis of an experiment. In particular, the problem is to determine if the treatment, a soil supplement, resulted in a significant difference in the yield (in kilograms) of cherry tomato plants.

The authors prompt students to calculate the means of the two groups and the difference between the means, state the null hypothesis about soil supplement, conduct a simulation (collecting data in class), and display the differences between means in a histogram. They provide a histogram for 50 samples and ask students to locate the difference between the means in the histogram in order to determine the probability of the

event happening and decide to reject or accept the null hypothesis. The authors do not describe how to make the histogram using graphing calculators or computer software.

Types of studies. One issue that underlies the CC and CN combinations is distinguishing between *observational* studies and *experimental* studies. Often used to confirm causation, experimental studies involve assessing if a treatment causes a significant difference in the incidence of some attribute. In order to assess the causal relationship, one has to randomly form groups and assign the treatment randomly into groups, and then compare the groups' attribute after the treatment. Consequently, experimental studies often involve at least one categorical variable, namely groups. Furthermore, association and causation are closely related; a high correlation of two events, which is established through observational studies, helps motivate to examine the causation between them. Hence, addressing the association between CC or CN variables relates to distinguishing between observational studies and experimental studies.

The distinction between types of studies is first addressed in one instance in Chapter 10 of the *Algebra 1* book in which the authors ask students to determine which study of the link between exercise and incidence of heart attack was more reliable, a *randomized experiment* or an *observational study*. The definitions of the terms, randomized experiment and observational study, are not introduced prior to the presentation of the instances to students. Consequently, students would have to use their intuitive understanding about these terms to finish this instance.

The topic of differences among research methodologies receives greater emphasis in Chapter 6, *Data Analysis and Statistics* of *Algebra 2* book. In Lesson 6.5, *Compare Surveys, Experiments, and Observational Studies*, the authors describe experiments and

observational studies in the context of “determin[ing] whether a quantity that varies can be associated with measured differences in two different groups of individuals” (Larson et al., 2012c, p. 415). The authors define the types of studies as follows:

An experiment imposes a treatment on individuals in order to collect data on their response to the treatment....

In some cases, it may be difficult to control or isolate the variable being studied, or it may be unethical to subject people to a certain treatment or to withhold it from them. An observational study observes individuals and measures variables without controlling the individuals or their environment. (p. 415)

Using these descriptions, the authors provide examples to illustrate each type of study as well as practice for the learning expectation (Figure 25). The authors then introduce *controlled experiment and randomized comparative experiment* with illustrative examples and then link the characteristics of an experiment to the concepts of *correlation* and *causality*. The lesson ends with an example of how to choose and design a study to address a statistical question specified in the book. In addition, the authors highlight *randomization* as a characteristic of well-designed studies (including sample survey, observational study, and experiment).

EXAMPLE 4 [A]
for Exs. 13–15

17b. Use a randomized comparative experiment in which individuals are randomly assigned to a treatment group that takes the fiber supplements and a control group that does not take the supplements. [B]

Explain whether the research topic is best addressed through an *experiment* or an *observational study*. Then explain how you would design the experiment or the observational study. 13–16. See margin.

13. You want to know if homes that are close to parks or schools have higher property values than other homes.
14. You want to know if flowers sprayed twice a day with a mist of water stay fresh longer than flowers that are not sprayed.

15. Determine whether the study described in the report is a randomized comparative experiment. If it is, describe the treatment, the treatment group, and the control group. If it is not, explain why not and discuss whether the conclusions drawn from the study are valid.
16. Describe how you would set up a randomized, controlled experiment to investigate the hypothesis below. Include any precautions you would take to ensure that your conclusions are valid.

Dog food with added Omega-3 fatty acids will give dogs a shiny coat.

Early Birds Make Better Drivers

A recent study shows that adults who rise before 6:30 A.M. are better drivers than other adults. The study monitored the driving records of 140 volunteers who always wake up before 6:30 and 140 volunteers who never wake up before 6:30. The early risers had 12% fewer accidents.

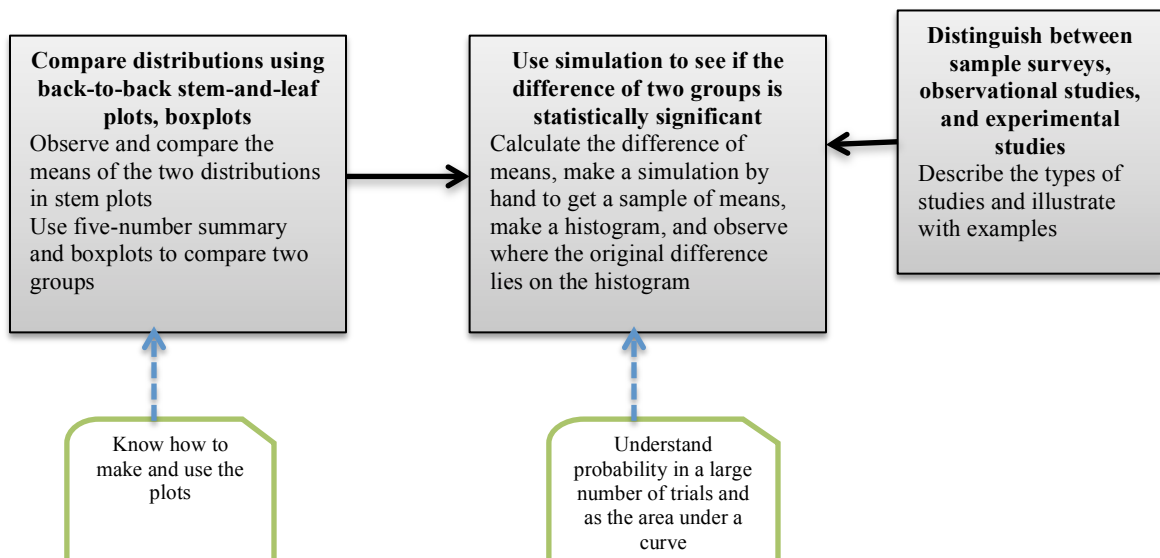


Figure 25. Distinguishing among three types of studies. From *Holt McDougal Larson Algebra 2 – Teacher’s Edition*, by R. Larson, L. Boswell, T. D. Kanold, & L. Stiff, 2012c, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.419. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

Summary of the LT for CN combination and its relation to other conceptual

categories. The progression of the relationship between CN variables is as follows: (a) compare two or more groups using graphical displays or numerical statistics of distributions, and (b) assess the significance of the mean difference between two groups.

Figure 26 summarizes the trajectory for CN combination in the HML series.



Note.

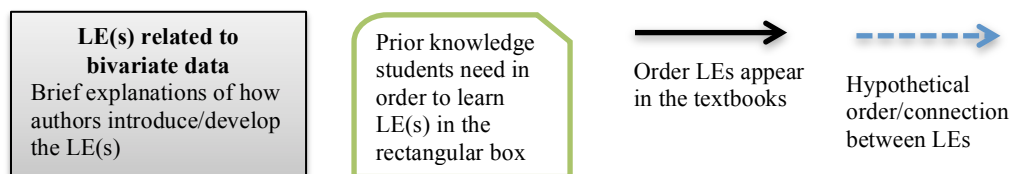


Figure 26. Learning trajectories for the CN association in the HML series.

For CN combination, graphical displays such as stem-and-leaf plots and boxplots, and numerical statistics (e.g., mean, median, interquartile range) for univariate data are used as means to address the association between CN variables. Simultaneously, comparing the differences of groups sets the stage for students to understand the utility of the statistics and the graphical displays of univariate data. In the textbooks, histograms are used to specify the probability of the difference between two groups happening in an experiment. To understand the significance of the difference, students need to understand that empirical probability, in a large number of trials, reaches the theoretical probability and graphical display of probabilities (area under a curve). It should be noted that examining the distribution of the difference of means is applicable only for two samples with the same size.

HML series learning trajectory for two numerical variables.

Linear fitting. In Chapter 1, *Expressions, Equations, and Functions* of the *Algebra 1* textbook, authors of HML first introduce the bivariate relationship by asking students to represent data of NN variables on a scatterplot, observe and informally use the pattern of increasing/decreasing in the plot for prediction. In Chapter 3, *Graphing Linear Equations and Functions* of *Algebra 1*, the authors then introduce a formula to determine rate of change and use it as a means to assess the magnitude of a specific linear model.

These same expectations are revisited in Chapter 2, *Reasoning and Proof* and in Chapter 3, *Parallel and Perpendicular Lines of Geometry*.

Fitting a line to data first appears in Lesson 4.6, *Fit a Line to Data* of Chapter 4, *Writing Linear Equations in Algebra 1*. In a decontextualized setting, the authors show scatterplots of positive correlation, negative correlation, and relatively no relationship. They demonstrate these correlations using examples of a positive correlation between hours of studying and test scores, and a negative correlation between hours of television watching and test scores. Students are expected to apply their prior knowledge (making a scatterplot) to this current lesson. They are asked to make a scatterplot using given data, observe the scatterplot and describe the correlation. Furthermore, students were introduced to using a line of fit as follows:

Step 1 Make a scatter plot of the data.

Step 2 Decide whether the data can be modeled by a line.

Step 3 Draw a line that appears to fit the data closely. There should be approximately as many points above the line as below it.

Step 4 Write an equation using two points on the line. The points do not have to represent actual data pairs, but they must lie on the line of fit. (Larson et al.,

2012a, p. 265)

The authors use an example of the number of active “red-cockaded woodpecker clusters in a part of the De Soto National Forests in Mississippi” (p. 266) over time to demonstrate the procedure.

Following the instances related to lines of fit, the authors introduce the term – *linear regression* and its relationship to *best-fitting line*, the *correlation coefficient* r ,

using the line for prediction: *interpolation* and *extrapolation*. They explain linear regression as, “the line that most closely follows a trend in data is the *best-fitting line*” (Larson et al., 2012a, p. 271). Using the total sales from women’s stores from 1997-2002 as an example, students are guided through the process of using a graphing calculator to determine the best-fitting line to model the set of data. The authors show examples of finding the best-fitting line and applying the model for prediction in other situations. In addition, the *correlation coefficient* r is introduced as “how well the best-fitting line fits the data” (Larson et al., 2012a, p. 272) and obtained when doing linear regression with graphing calculators. Although the authors do not provide a formula for r , they describe the connection between an r -value close to 1, -1 , and 0 and strength and direction of the correlation. Figure 27 provides an example of instances related to linear fitting.

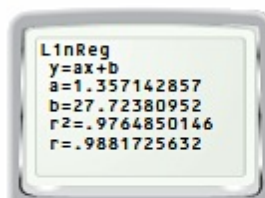
After introducing correlation coefficient and to show the difference of correlation and causation, in Chapter 4, *Extension, Correlation and Causation*, of *Algebra I*, the authors provide an example of “the number (in millions) of music album downloads and the number (in millions) of individual federal income tax returns filed electronically each year from 2004 to 2008” (Larson et al., 2012a, p.273). In this lesson, students use graphing calculators to find r and confirm that the correlation between the two variables is high; however, they are prompted to use informal knowledge to conclude that an increase in album download does not *cause* an increase in the filing of electronic tax returns.

EXAMPLE 2 Find the best-fitting line

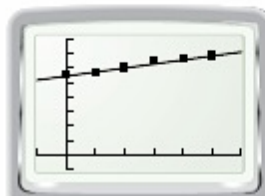
Find an equation of the best-fitting line for the scatter plot from Example 1. Determine the correlation coefficient of the data. Graph the best-fitting line.

STEP 1 Perform regression

Press **STAT**. From the CALC menu, choose LinReg(ax+b). The a - and b -values given are for an equation of the form $y = ax + b$. Rounding these values gives the equation $y = 1.36x + 27.7$. Because r is close to 1, the data have a strong positive correlation.

**STEP 2** Draw the best-fitting line

Press **Y=** and enter $1.36x + 27.7$ for y_1 .
Press **GRAPH**.

**PRACTICE**

In Exercises 1–5, refer to the table, which shows the total sales from men's clothing stores in the United States from 1997 to 2002.

Year	1997	1998	1999	2000	2001	2002
Sales (billions of dollars)	10.1	10.6	10.5	10.8	10.3	9.9

1. Make a scatter plot of the data. Describe the correlation. See art in Exercise 3; negative correlation.
2. Find the equation of the best-fitting line for the data. $S = -0.046y + 10.5$
3. Draw the best-fitting line for the data. See margin.

DRAW CONCLUSIONS

4. What does the value of r for the equation in Exercise 2 tell you about the correlation of the data? The data have relatively no correlation.
5. **PREDICT** How could you use the best-fitting line to predict future sales of men's clothing? Explain your answer. Sample answer: You cannot use the best-fitting line to predict future sales because the data do not show a strong correlation.

Figure 27. Using linear regression to make a prediction. From *Holt McDougal Larson Algebra 1 – Teacher's Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.272. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

After two approaches to finding lines of fit were addressed, the authors focus students' attention on how to assess the fit of a model. They introduce residual plots, show how to make them, and describe characteristics of a good model of fit for a set of data. They explain:

If a line is a good fit for a set of data, the absolute values of the residuals are relatively small and more or less evenly distributed above and below the x -axis in a residual plot. Residuals that are mostly positive or mostly negative imply that the line is in the wrong place. Residuals that are steadily increasing suggest the data is not linear, while wildly scattered residuals suggest that the data might have relatively no correlation. (Larson et al., 2012a, p. 283)

They use one example to illustrate how a line is not a good fit for a set of data. The TBS instances often involve providing students with data and a linear model and asking them to determine whether or not the model is a good fit for the data.

Curve fitting. Continuing to fit a model to a set of data, in Lesson 7.4, *Write and Graph Exponential Growth Functions* of Chapter 7, *Exponents and Exponential Functions* of *Algebra 1*, one instance asks students to use graphing calculators to find the exponential regression to model the relationship between the frequency of piano notes and the position of the key that creates the note. The authors do not provide a formal introduction to this kind of regression before students proceeded to accomplish the instance. This learning expectation (exponential regression) is revisited and developed in a stand-alone activity appearing in Lesson 9.8, *Compare Linear, Exponential, and Quadratic Models* of Chapter 9, *Quadratic Equations and Functions*. Several instances involve students' use of calculators to find exponential regression, quadratic regression and check the scatterplot to determine how the models fit the data. No narrative is provided in the text to define the regression; the models are found using graphing calculators.

The process of fitting various types of functions is revisited and developed in *Algebra 2*, Chapter 4, *Exponential and Logarithmic Functions*. In a stand-alone activity, *Model Data with an Exponential Function*, the authors instruct students to collect data and use the exponential regression feature of graphing calculators to model the data and assess the fit of the model by superimposing the model on a scatterplot. The authors then introduce the linearization of data by performing log and log-log transformations and asking students to observe if the transformed data points lie in a linear pattern. By solving a system of linear equations for the transformed data, a model is determined and then transferred back to find a model for the original data. Furthermore, students are prompted to use exponential and power regressions as a reference for the aforementioned methods. The chapter also includes multiple instances that asked students to find, choose or make an appropriate model based on the set of available models (linear, exponential, and power models).

The idea of fitting rational, trigonometric models to data is addressed in this series. In Chapter 5 of *Algebra 2*, *Rational Functions*, rational model-fitting appears in Lesson 5.1, *Model Inverse and Joint Variation*. By trying multiple products of $x*y$ students are prompted to observe if the products are approximately equal to each other in order to specify the model ($y = \frac{a}{x}$) for a set of data. Students use graphing calculators to find trigonometric regression in Chapter 10, *Trigonometric Graphs, Identities, and Equations*, Lesson 10.5, *Write Trigonometric Functions and Models*. The authors use data collected over time about the number of kilowatt-hours used per month by the Cape Canaveral Air Station in Florida to illustrate the process.

Summary of the LT for NN combination and its relation to other conceptual categories. The relationship between NN variables is introduced implicitly through content related to functions and explicitly in chapters related to curve fitting, regression and correlation. Other than linear fitting, other types of regressions are introduced using graphing calculators without definitions provided. For linear fitting, linear regression and correlation are obtained from calculators without the provision of formulas and their properties in related to sum of square residuals. Exponential and power fitting are also transferred to linear fitting via linearizing data by log and log-log transformations. Table 10 and Figure 28 summarize the learning trajectory for NN variables in the HML series. It should be noted that the dashed arrows are the hypothetical trajectory/connections between the LEs, which are drawn from the strategies involving when solving tasks related to LEs. For example, students can move directly from learning *patterns in scatterplots* to *eyeballed line and linear regression* and not pass through *rate of change* without any difficulties. Because of the way the authors introduce and develop content related to *eyeballed line and linear regression*, no prior knowledge related to *rate of change* is needed to learn linear fittings. This argument also holds true for other textbook series.

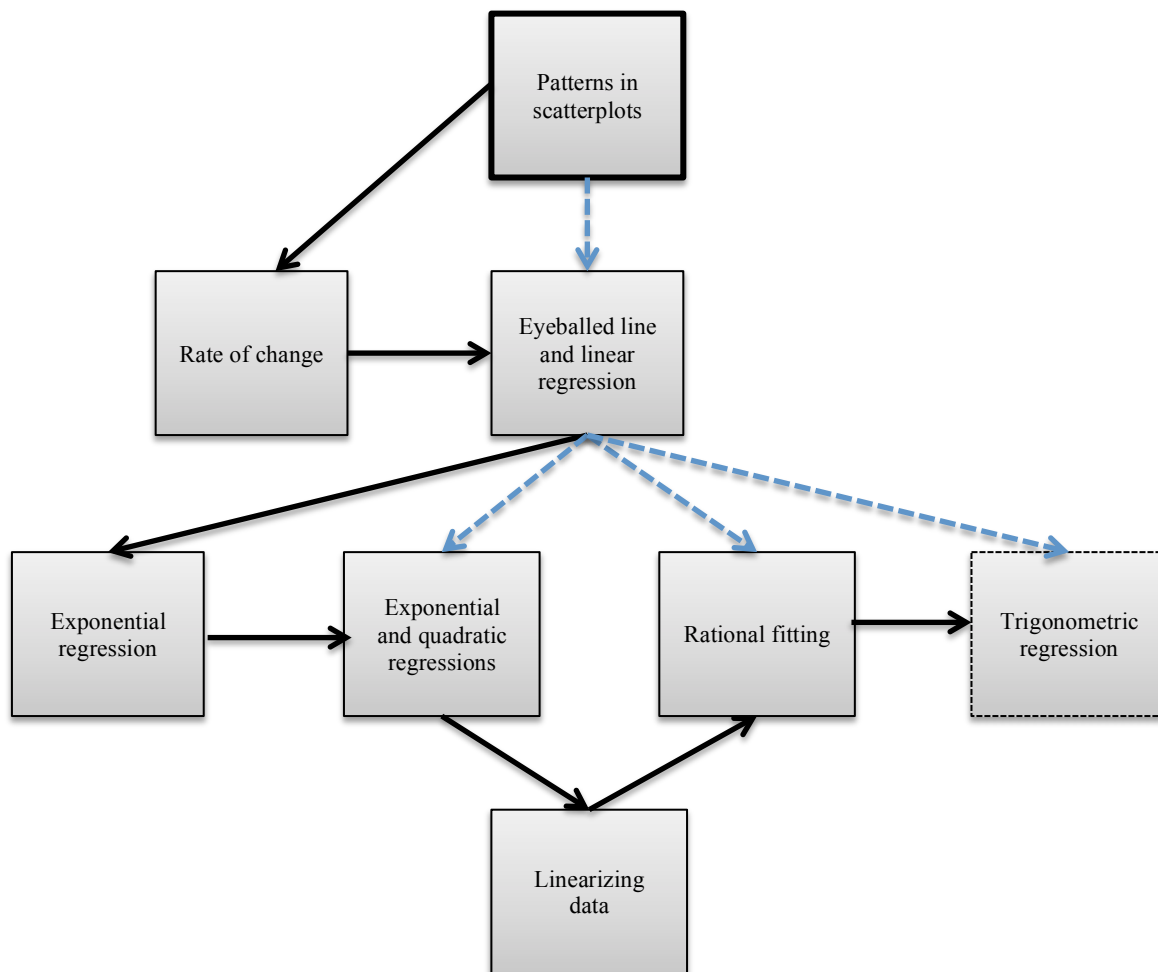
Table 10

Summary of learning trajectories for two numerical variables in the HML series

LEs related to association of bivariate data	Brief explanations of how authors introduce/develop the LEs column	Prior knowledge/ knowledge from other conceptual categories related to the LEs
Patterns in scatterplots	Prompt students to observe patterns of increasing/decreasing on a scatterplot	Make a scatterplot
Rate of change	Prompt students to use formulas to calculate rate of change	Slope of a line

Eyeballed line Linear regression	<p>Illustrate positive, negative, and no correlations in scatterplots</p> <p>Prompt students to write the equation for an eyeballed line</p> <p>Describe linear regression as best-fitting line without further explanation and asked students to obtain the regression using calculators</p> <p>r was a value obtained when doing the regression without giving a formula and property, linked to how well the best-fitting line fits the data</p> <p>Introduce residual plots and state the property of the plot for a good model, and provided residual plots/or asked students to make residual plots to assess the goodness of fit</p>	Write an equation of a line passing 2 points
Exponential regression	Prompt students to use graphing calculators to find the regression without providing any descriptions	Know exponential functions
Exponential regression Quadratic regression	<p>Prompt students to use calculators to find the regression without providing any descriptions</p> <p>Prompt students to informally assess the goodness of fit of a model using scatterplots</p>	Know behaviors of exponential and quadratic functions
Linearizing data	<p>Prompt students to linearize data using logarithmic transformations and fit a linear function for the transformed data</p> <p>Prompt students to use calculators to find the regressions</p>	<p>Know the logarithmic operation/function</p> <p>Solve a system of linear equations</p>
Rational functions fitting	Prompt students to use provided formulas and substitute a pair of data to find the function	Know behaviors of rational functions
Trigonometric regression	Prompt students to use graphing calculators to find the regression	Know behaviors of trigonometric functions

In the learning trajectory of fitting models to a set of data, functions are closely connected to the topic of model fittings. In particular, linear functions, their graphs, behavior, and slope as rate of change are used as means to study the relationship of NN variables. The behaviors of exponential, logarithmic and power functions are used to specify the kind of mathematical model that fits the data. However, students are not asked to decide which model to use until the logarithm operations are introduced to linearize data (in Chapter 4 of *Algebra 2*). Hence, the knowledge about taking the logarithm of a number and transferring it back to the original data is prerequisite to the study of curve fitting. By observing the transformed data, students are able to decide which model to choose, an exponential or power function.



Note. LTs in the textbook series begin at the bold box and end at the dashed box

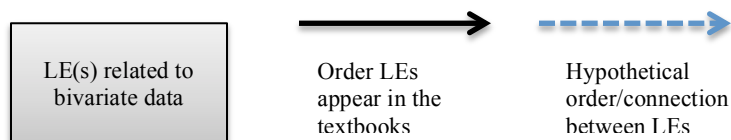


Figure 28. Learning trajectories for association between NN variables in the HML series.

The University of Chicago School Mathematics Project (UCSMP) series.

UCSMP series learning trajectory for two categorical variables. The authors first introduce the Chi-square test in Lesson 11-8, *The Chi-Square Statistics* in the *Algebra* textbook as a means to examine the association between CC variables. The test for association was situated within a larger learning goal “to determine whether a set of

discrete data differs significantly from values that would be expected when certain probabilities are given” (Brown et al., 2008, p. 697). Using an example to determine if birthdays are randomly distributed across seven days of a week, the authors demonstrate the procedure for determining the association with an explanation of how to calculate the Chi-Square value for the set of data. In particular, they show how to calculate the expected values, the deviation between the expected and the observed values, the sum of all the squares of deviations, as well as how to read the Critical Chi-Square Values table by specifying the degrees of freedom in order to decide if the test-statistic is *significant* (i.e., the probability is less than .05). Most instances in this chapter explicitly ask students to use the Chi-Square statistic to determine if the event occurs randomly.

The Chi-Square test is revisited and expanded in Lesson 6-9, *The Chi-Square Test of FST*. The authors define the test using a formula and use an example to show how to calculate the test. Next, they provide an activity simulating the likelihood of a chi-square test using a technology application to make a histogram for the chi-square values and specify where the original chi-square value lies on the histogram (Figure 29). Next, they introduce how to read the Critical Chi-Square Values table to determine if the event happens randomly without doing the simulation and to observe the probability of the event in a large number of trials. To illustrate the procedure for application of the Chi-Square test, the Titanic problem is used to determine if the unequal survivors by class happened by chance. The procedure includes: stating the hypothesis, calculating expected values, computing chi-square statistics and using calculators to specify the probability of the event to determine if it is significant. Students are also instructed to use the technology to do simulations as another approach to testing the significance.

Instances explicitly relating to the association of bivariate categorical data using frequencies appear in Chapter 1 of *FST*. By using percentage in multiple categories of the table, the association between two variables is revealed. For example, the authors show a two-way table including boys/girls vs. types of sports. They ask students to specify categorical variables in the table, compute the percentage for each category of gender in each type of sports, and “write a few sentences about the differences between boys’ and girls’ participation in these sports” (McConnell et al., 2010, p. 13).

Simulating the Likelihood of a Chi-square Value

We still have not determined how far the observed frequencies can be from expected frequencies to still be considered variation due to chance. One way to determine this is to simulate the coin-tossing experiment a large number of times, calculate the chi-square statistic for each experiment, and see how often a specific value occurs.

chi-square statistic be negative? Why or why not?

Activity 1

MATERIALS chi-square application or technology file provided by your teacher

The following simulation shows how a particular χ^2 value ($\chi^2 = 2$ in Example 2) compares to 1000 generated χ^2 values. One experiment consists of a random generation of 50 heads or tails. The difference between these simulated counts of heads and tails with the expected numbers of heads and tails are used to calculate the chi-square statistic seen in Examples 1 and 2. This statistic is computed and recorded. One thousand experiments are run in one simulation.

Step 1 Consider Example 2. Record 0.5 and 0.5 in the first two cells for probabilities. These are hypothesized probabilities because that is what χ^2 is testing: how well the observed frequencies under these probabilities match the expected count.

Step 2 Put the observed counts of heads and tails in the cells for observed frequencies. For Example 2, these are 30 and 20, respectively.

Step 3 Press the “Get Chi-Square” button and record the chi-square value.

Step 4 Run the simulation. When the simulation is finished, a distribution of the chi-square values will appear on your screen graphed as a histogram. Record the percent of experiments out of the 1000 that have χ^2 larger than the original chi-square value of 2. Compare your percent to that of others in your class.

Step 5 Summarize the results of the simulation in a sentence: If the hypothesized probabilities are accurate, a chi-square value larger than the original chi-square value would occur about ____?____ of the time by chance.

Step 6 Repeat Steps 1-5 to simulate Example 1.

Step 7 Repeat Steps 1-5 to simulate scenario (3) on page 417.

The screenshot shows a digital interface for a chi-square simulation. It features several input sections: 'Hypothesized Probabilities' with five columns (A, B, C, D, E) and one row of input boxes; 'Observed Frequencies' with five columns and one row; and 'Expected Frequencies' with five columns and one row. Below these are three buttons: 'Get Chi-Square', 'Run Simulation', and 'Clear'. Under the 'Run Simulation' button, there are labels for 'Chi:' and 'Percent of Distribution Greater than Chi-Square Value:', followed by empty space for results.

Figure 29. Simulation of the likelihood for a Chi-square value with technology. From *The University of Chicago School Mathematics Project: Functions, Statistics, and*

Trigonometry – Teacher’s Edition by J. W. McConnell et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p. 419. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

The association between CC variables is developed fully in Lesson 6-5, *Contingency Tables*. Using the aforementioned data from the Titanic, the authors provide a table about class vs. survival and show how to calculate the percentage of survival in each class. The comparison of percentages is delayed until the assignment portion of the lesson when they are used to address the CC association. In another example about the association between getting a tattoo and becoming infected with Hepatitis C (HCV), the authors prompt students to use percent of each cell to reveal the association. In addition, they note the caution about drawing conclusion of causation by identifying lurking variables, factors that are that account for the relationship between two events but not mentioned in the study.

The authors finally introduce *Averaging and Simpson’s Paradox* to show the discrepancy between the overall average and the average within categories of bivariate data. Using an example about lawsuit brought for gender discrimination in graduate school admissions the authors illustrate the paradox – the difference between the acceptance rates within each program and the overall acceptance rates (Figure 30).

Averaging and Simpson's Paradox

When data are combined or separated, the totals may not look like the parts. Here is a famous real-world example. In the 1970s, the University of California at Berkeley was sued for gender discrimination in graduate school admissions. As a result, it examined its admissions closely. In 1973, a typical year, 35% of 4321 female applicants were admitted, while 44% of 8442 male applicants were admitted. The women's acceptance rate was noticeably smaller than the men's.

The situation was actually more complicated than it looked. Applicants for graduate school apply to specific programs, not to the university as a whole, and acceptance rates vary widely from program to program. The contingency table below shows the men's and women's acceptance rates within each program.

Program	Men		Women		Overall	
	Applicants	% Admitted	Applicants	% Admitted	Applicants	% Admitted
A	825	62%	108	82%	933	64%
B	560	63%	25	68%	585	63%
C	325	37%	593	34%	918	35%
D	417	33%	375	35%	792	34%
E	191	28%	393	24%	584	25%
F	373	6%	341	7%	714	6%
Total	2691	44.5%	1835	30.4%	4526	38.8%

Within each program, men were accepted at a comparable or equal rate to women. The discrepancy arises because the programs to which the most men applied (programs A and B) had higher overall acceptance rates than the two most popular women's programs (C and E). In this case, the choice of program is a *confounding variable*, that is, an extra variable associated with one of the variables in the study (gender), that causes part of the observed effect.

The discrepancy between the overall acceptance rates and the acceptance rates within each program is an example of **Simpson's Paradox**, named after Edward H. Simpson, who wrote about the paradox in 1951. Simpson's Paradox can arise whenever different categories of data are combined (or *aggregated*) to compute an average or percent. Although Simpson's Paradox arises in a variety of situations, the best way to avoid being confused by it is to separate (or *disaggregate*) the data into the original categories.

Know that 500 boys and 400 girls were surveyed.



Jane K. Sather Tower,
University of California,
Berkeley

Figure 30. Simpson's Paradox. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher's Edition* by J. W. McConnell et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p. 391. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

Summary of the LT for CC combination and its relation to other conceptual

categories. The learning progression of this content is as follows: (a) use Chi-Square test to examine association for bivariate data, and (b) revisit the test with the aid of

simulation, and use frequencies as a means to look for association. Figure 31 summarizes the UMCSP series' learning trajectory for the CC combination of bivariate data.

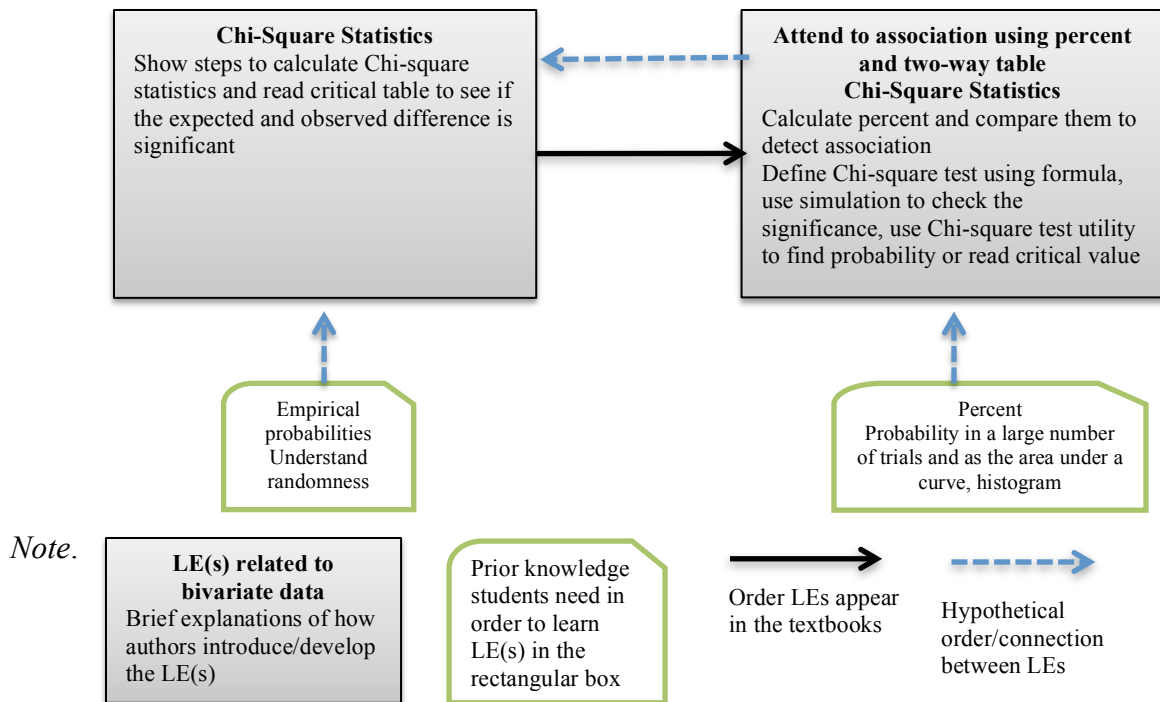


Figure 31. Learning trajectories for the CC association in the UCSMP series.

For the CC combination, the authors compare percentages to detect the association between the events and used two-way tables as a means to calculate the percentages. The connection between statistical association and percentages is explicitly addressed; that is, no connection is made to univariate data topics for the CC combination. In addition, the authors use Chi-Square test to judge the association of the events. To understand this content, students need to understand that in the long run, empirical probabilities are expected to approximate theoretical probabilities and students are able to connect probabilities to their graphical model (the area under a curve).

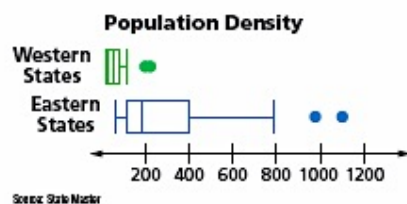
UCSMP series learning trajectory for one categorical and one numerical variable. Content related to the association between CN variables appears in Chapter 1, *Exploring Data of FST*. In Lesson 1-7, *Comparing Numerical Distributions*, graphical

displays including histograms and side-by-side boxplots and numerical statistics are used to compare two distributions. The authors use an activity comparing the height of players in the National Basketball Association and the National Football League during the 2007-2008 seasons to show the process of comparing distributions. The process includes: describing each distribution in terms of its shape, center, and spread and deciding which measures of center and spread to use to compare the two distributions. Figure 32 is an example of instances that asks students to compare groups. In the example, populations in two regions of the U.S. are compared. The comparison could lead to a conclusion about the relationship between location and population, yet this is not explicitly stated in the text.

No content related to distinguishing the types of studies is found in this textbook series.

1. **Population density** is a measure of how crowded an area is. The information below compares Western and Eastern states in the United States in this regard. (The Mississippi River is the boundary between West and East.) Write a paragraph comparing the population densities of the Western and Eastern states.

have a greater mean and median than the Western states, 301.4 and 172.3 compared to 52.4 and 38.6, respectively.



Western Density		Eastern Density	
mean	52.4	mean	301.4
median	38.6	median	172.3
standard deviation	54.4	standard deviation	298.3
interquartile range	55.4	interquartile range	299.4
Q_1	15.3	Q_1	101.7
Q_3	70.7	Q_3	401.1

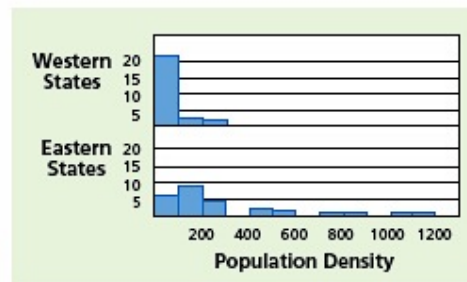


Figure 32. Comparing two distributions. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher's Edition* by J. W. McConnell et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p.419. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

Summary of the LT for CN combination and its relation to other conceptual categories. The association between CN variables is introduced once in the UCSMP series: comparing two or more distributions using graphical displays and/or numerical statistics. Figure 33 summarizes the trajectory for the CN combination.

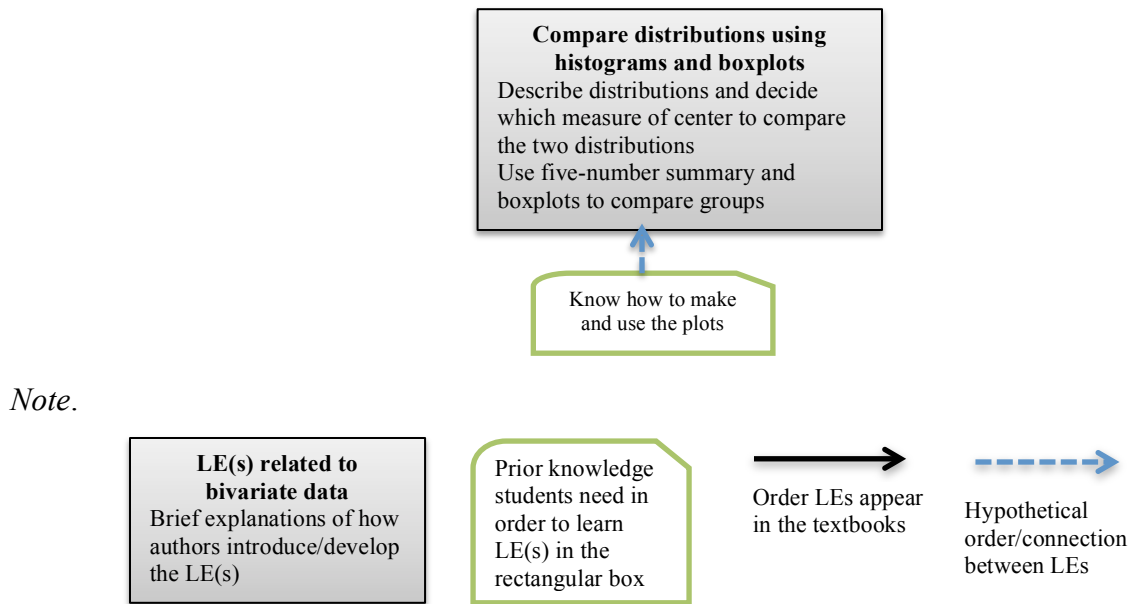


Figure 33. Learning trajectories for the CN association in the UCSMP series.

For the CN combination, graphical displays such as histograms and boxplots and numerical statistics (e.g., mean, median, interquartile range) for univariate data are used as means to examine the association. Simultaneously, comparing the differences of groups sets the stage for students to understand the utility of the statistics and the graphical displays of univariate data. Understanding about using histograms and boxplots to summarize a set of data is prerequisite for comparing distributions; however, the authors do not explicitly discuss the association between the two variables.

UCSMP series learning trajectory for two numerical variables.

Linear fitting. The relationship between NN variables first appears in Chapter 6, *Slopes and Lines* of the *Algebra* textbook. In Lesson 6-1, *Rate of Change*, rate of change

is introduced as a means to understand how a quantity varies in relation to time. Students are instructed to use the formula for rate of change to calculate the rate. They calculate the rate of change either by hand or using a spreadsheet; graphical display of the steepness of the line is used to depict the rate of change, and then to interpret in the context of data in situations related to linear modeling.

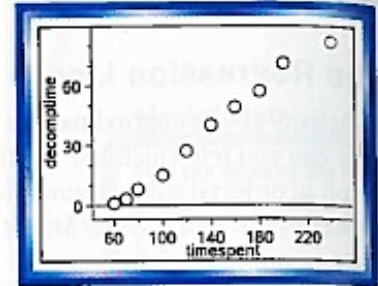
In Lesson 6-7, the authors introduce the idea of using a line to describe trends in data. Using an example about life expectancy over time in the U.S., they show the process of “eyeballing a line of fit” (Brown et al., 2008, p. 368) by writing the equation of a line passing two points. Then “the line of best fit, ... the most common way of finding a line to fit data” (p. 368), is introduced and used for prediction; students are prompted to use graphing calculators to find the line of best fit. *Deviations* are informally used to assess the fit of the two models: eyeballed (estimated) line and the line of best fit. The authors state the property of the line of best fit: “The sum of the squares of the deviations of its values from the actual values is the least of all lines” (p. 370) and illustrate the property by providing one specific example.

A consistent approach for linear fitting was found in Chapter 3, *Linear Functions and Sequence* of the *Advanced Algebra* book, but within a different context – Navy divers. Yet, a slightly new idea is incorporated: using a moveable line on the graphing calculator to assess the sum of squares of deviations (SSD) of fitting lines and empirically showing that SSD is smallest for the regression line. In the second appearance of linear fitting, one instance related to making sense of the *correlation coefficient* r and how to interpret its meaning by examining r in multiple data sets appears (Figure 34).

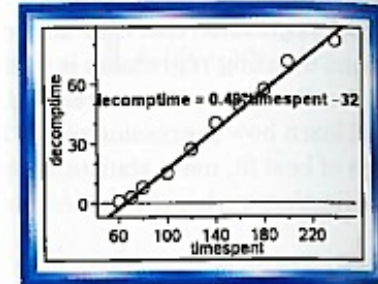
Step 1 Enter the data into columns in your statistics or spreadsheet application. Name the first column *timespent* and the second column *decomptime* to indicate which is the independent and which is the dependent variable.

A	timespent	B	decomptime	C
1	60		1	
2	70		3	
3	80		8	
4	100		15	
5	120		27	
6	140		40	
17	60			

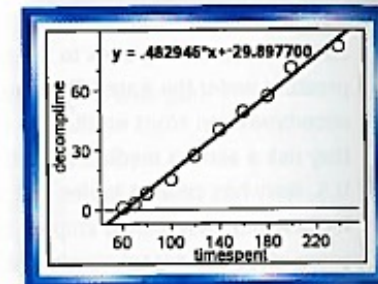
Step 2 Create a scatterplot of the data. You may be prompted to choose one column to use for the independent variable data and choose one column to use for the dependent variable data.



Step 3 If possible, add a movable line to your scatterplot. Adjust the position of the line to eyeball a line of best fit. Record the equation of your movable line. Also record the predicted values when $x = 80$ and when $x = 220$. **Answers vary. Sample: $y = 0.4829x - 29.90$**



Step 4 Remove your movable line. Choose to show a linear regression line. The application will graph the regression line on the same axes as the data. Record the equation of the regression line. How does your movable line compare to the regression line? **Answers vary. Sample: It is very close to the regression line.**



Step 5 Compare your movable line's predictions to those of the regression line. When $x = 220$, by how much do the predicted values differ from each other? When $x = 80$, by how much does the predicted value differ from the value in the table? **See margin.**

You can see from the graphs that the regression line is a reasonable model for the data.

Figure 34. Examining of linear fitting with a moveable line. From The University of Chicago School Mathematics Project: Advanced Algebra – Teacher’s Edition by J. Flanders, M. Lassak, J. Sech, M. Eggerding, P. J. Karfiol, L. McMullin, N. Weisman, & Z. Usiskin R, 2010, Chicago: IL: Wright Group/McGraw-Hill, p.177. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

Curve fitting. Linear regression is used in Chapter 7, *Using Algebra to Describe Patterns of Change of Algebra*, as a contrast to introduce exponential regression for prediction. Specifically, in Lesson 7-4, *Modeling Exponential Growth and Decay*, the

authors ask students to collect data on ball bouncing and provide two functions for students to choose to model the set of data: linear and exponential. Graphing calculators are used to find the regressions for prediction. By informally observing scatterplots, the authors suggest which model to use for that situation. The authors do not formally describe exponential regression except that is the equation of the model obtained from the graphing calculator.

A similar approach to regression is used in Lesson 9-4, *Quadratics and Projectiles* of Chapter 9, *Quadratic Equations and Functions* of *Algebra*. Through an activity asking students to collect data, the authors guide students to use graphing calculators to find the quadratic regression and superimpose it on the scatterplot to assess the fit of the model.

Exponential fitting is revisited and expanded in Lesson 9-4, *Fitting Exponential Models to Data* in Chapter 9, *Exponential and Logarithmic Functions* of *Advanced Algebra*. The authors introduce two ways to find the model: (1) by solving a system of two equations, and (2) by using graphing calculators (Figure 35). With the availability of three types of models (linear, exponential, and quadratic), the authors prompt students to try the various models for a set of data and assess the fit of the models using a scatterplot.

Revisiting multiple models fitting. Observing a pattern of increase/decrease of a quantity in a scatterplot or a table and using informal knowledge to explain the reason for the behavior are introduced in *Advanced Algebra*, Chapter 1, *Functions*. In Chapter 2, *Variation and Graph*, the authors spend two lessons addressing fitting a model to a set of

Example 1

Find an exponential model to fit the drug absorption data using the initial condition and one other point in the table.

Solution You need to find a and b in the equation $y = ab^x$. The initial condition occurs when $x = 0$, so $1000 = ab^0$.

Since $b^0 = 1$, $a = 1000$. So the equation is of the form $y = 1000b^x$.

Choose another point from the table and substitute. We chose $(20, 85)$.

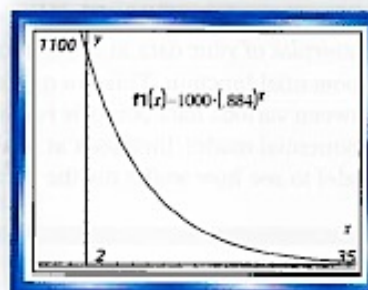
$$85 = 1000b^{20} \quad \text{Substitute.}$$

$$0.085 = b^{20} \quad \text{Divide each side by 1000.}$$

$$0.884 \approx b \quad \text{Take the 20th root of each side (or raise each side to the } \frac{1}{20} \text{ power).}$$

So, one model for the data is $y = 1000 \cdot (0.884)^x$.

Check Graph this model on a graphing utility. It looks like an exponential decay function with y -intercept 1000, as you would expect from the data.



You can also use your graphing utility to find an exponential regression model.

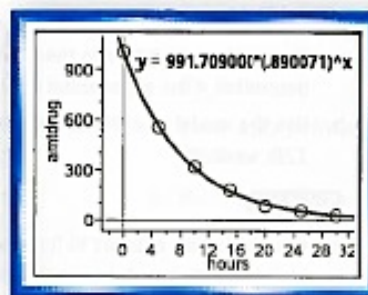
Activity

Step 1 Enter the drug absorption data into your calculator. Apply the exponential regression option to find a and b and then write an equation to model these data.

Step 2 Your calculator probably gives a and b to many digits. The accuracy of the experimental data suggests that a more sensible model rounds a and b to three digits each. Rewrite your equation after rounding.

Step 3 Use your regression model from Step 2 and the two-point model $y = 1000 \cdot (0.884)^x$ from Example 1 to estimate the amount of drug left in a body after 5 hours. Which model better fits the actual 550 mg value from the experiment?

Step 4 Compare the accuracy of the two models by making a table of values for each equation and comparing these values to the actual data. Which model appears to be more accurate overall? See margin.



Step 1: $y = 991.709 \cdot (0.890071)^x$

Step 2: $y = 992 \cdot (0.890)^x$

Step 3: The model from Step 2 yields 553.939; the one from Example 1 yields 539.835. Therefore, the Step 2 model fits better.

Figure 35. Two approaches of exponential fitting. From *The University of Chicago School Mathematics Project: Advanced Algebra – Teacher’s Edition* by J. Flanders et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p.603. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

data. In particular, in Lesson 2-7, *Fitting a Model to Data I*, direct and inverse variation models are introduced in two examples to illustrate the process of fitting an inverse

model. By specifying a general model ($y = a * x$, $y = a * x^2$, $y = \frac{a}{x}$, $y = \frac{a}{x^2}$) and using one pair of data to find the constant a in the model, the authors instruct students to find the model and use it to solve the problem in the context of data. Using informal knowledge as well as examining the scatterplot, students could specify which type of model to use and to find the model for the set of data. This skill is used and expanded in Lesson 2-8, *Fitting a Model to Data II*, to investigate situations involving more than two variables.

Linear, exponential, quadratic, and inverse fittings are revisited and expanded in Chapter 2, *Functions and Models of Functions, Statistics, and Trigonometry*. In particular, in Lesson 2-2, *Linear Models*, the authors build on the content of previous chapters and add formal investigations about the goodness of fit. In this lesson, they define *residual*, *sum of squared residuals* (SSR), which was previously referred to as SSD, and show its graphical display. They state:

The sum of squared residuals is a statistic that measures *lack of fit*. If you compare two lines, the one with the larger sum of squared residuals is not as good a model as the one with the smaller sum of squared residuals...The line that gives the smallest value provides the best fit to the data. (McConnell et al., 2010, p. 91)

Using an example about the relationship between the number of televisions per 100 and the number of unemployed per 100, the authors guide students to examine when SSR is smallest with the aid of a moveable line.

Subsequently in Lesson 2-3, *Linear Regression and Correlation*, the authors define the line of best fit, the least squares line or regression line, and state three properties of the line:

1. It is the line that minimizes the sum of squared residuals, and it is unique.
There is only one line of best fit for a set of data.
2. It contains the center of mass of the data, that is, the point (\bar{x}, \bar{y}) whose coordinates are the mean of the x -values and the mean of the y -values.
3. Its slope and intercept can be computed directly from the coordinates of the given data points. (McConnell et al., 2010, p. 94)

The authors use an example of finding the line of best fit for a set of data about the weight and the price of diamond rings to illustrate the properties (Figure 36).

The formula for r is first introduced as a statistic to describe the fit of the regression line. The range of r and illustrations of direction and strength of the value are introduced with multiples scatterplots. The authors recommend that the value could be calculated using graphing calculators. Tasks related to examining r when transforming data are addressed in the next chapter. Furthermore, a warning to be cautious about confusing correlation with causation is made using multiple examples previously investigated. In addition, the authors mention influential points impacting the model, especially on the slope, intercept, and correlation.

Contents related to exponential, quadratic, and inverse model fittings in Chapter 2 of *FST* are largely identical to the content previous introduced in the *Algebra* and

Example

- Use a statistics utility to find a line of best fit for the data about the weight in carats and the price of diamond rings from Lesson 2-2.
- According to the regression line, how much will a 0.5-carat diamond ring cost?
- Verify that the center of mass (0.212, \$793.65) is on the line.
- Find the sum of squared residuals for the linear regression.

Solution

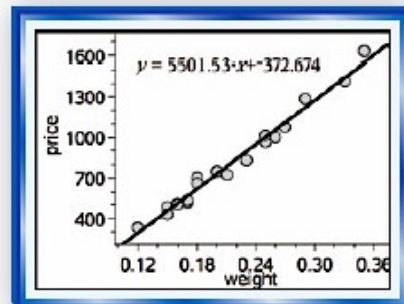
- Enter the entire data set from Lesson 2-2 into a statistics utility. Label the weight as x and the price as y . Use the linear regression feature to find an equation for y as a function of x . The linear regression model is $y = 5501.53x - 372.67$.
- Substitute 0.5 for x and calculate y .

$$y = 5501.53(0.5) - 372.67 \approx 2378.1$$

According to the linear regression model, the price would be \$2378.10.
- Substitute the coordinates of the center of mass for x and y in the regression equation.

$$y = 5501.53x - 372.67$$

Does $793.65 = 5501.53(0.212) - 372.67$?
 Yes. It checks.
- On many statistics utilities, once the regression equation is given, the residuals are calculated and stored until the next regression is found. In the screenshot at the right, the residuals are stored under the name *Resid*. So the sum of the squared residuals is 41842.9.



LinRegMx weight,price,1: CopyVar stat	
"Title"	"Linear Regression (mx+b"
"RegEqn"	"m*x+b"
"m"	5501.53
"b"	-372.674
"r ² "	0.982567
"r"	0.991245
"Resid"	"{"...}"

$\text{sum}(\text{stat.Resid}^2)$	41842.9
-----------------------------------	---------

Figure 36. Properties of the regression line. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher's Edition* by J. W. McConnell et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p.95. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

Advanced Algebra books. Nevertheless, residual plots are introduced as a means for selecting a good model. The authors use examples of making residual plots for linear, exponential, and quadratic models using calculators and attending to the pattern on the plots. Finally, they summarize steps to build a model, as follows: (a) build a model from theory (if applicable), (b) use a scatterplot to see if the model follows a pattern in the data, (c) use residual plots to check the fit of model (d) use r for linear model, and (e) be aware of interpolation and extrapolation.

Trigonometric fitting is first addressed in one instance in Chapter 10 of the *Algebra* book by observing a scatterplot and determining why the model is appropriate, and then this idea is revisited in Chapter 4 of *FST*. In particular, students are asked to sketch a curve and choose among the functions to fit the data. They are also required to write, without the aid of technology, a trigonometric function that fit the data and use calculators to find the regression. Students then compare the calculator-generated regression with the model written without the aid of technology and use the models for prediction.

The authors connect curve fitting to linear fitting, in Lesson 9-7, *Linearizing Data to Find Models*, by introducing modeling with logarithmic functions. They use an example of the effect of study time on the ability to recall the meanings of unfamiliar words. Showing the data in a scatterplot and recommending a general form of logarithmic function, the authors use the data and calculate the line of best fit for the transformed data (by taking the logarithm of the independent variable). The authors show examples taking logarithm on both variables to find a model. After making the log-log transformation, the line of best fit for the transformed data is used to specify the model for prediction. Tasks involving finding the logarithm of a dependent variable, which leads to an exponential function when transferring back to original variables, to find the linear fitting for the transformed data appear in the assignment portion of this lesson. Hence, there are three forms of transformations in this chapter $(\log(x), y)$ $(x, \log(y))$ and $(\log(x), \log(y))$. In preparation for selecting models and transformations, scatterplots of power, root, and logarithm functions families are shown as a reference. Moreover, the authors urge

students to use caution when using nonlinear models to fit data, namely to base decisions on both a prior theory and use residual plots to specify the model.

Summary of the LT for NN combination and its relation to other conceptual categories. The progression for relationships between NN variables includes several visits. It starts with examining the rate of change and linear fitting. Then exponential, quadratic regressions are introduced with graphing calculators. Observing patterns of increasing/decreasing on scatterplots is introduced as a preparation for direct and inverse fitting. Next, linear, exponential, and quadratic fittings are formally revisited with many investigations. Trigonometric fitting is then introduced and the series concludes with linearizing data to developed logarithm, exponential, and power fittings and connect with linear fitting. Table 11 and Figure 37 summarize the learning trajectories for the relationship between NN variables in the UCSMP series.

Table 11

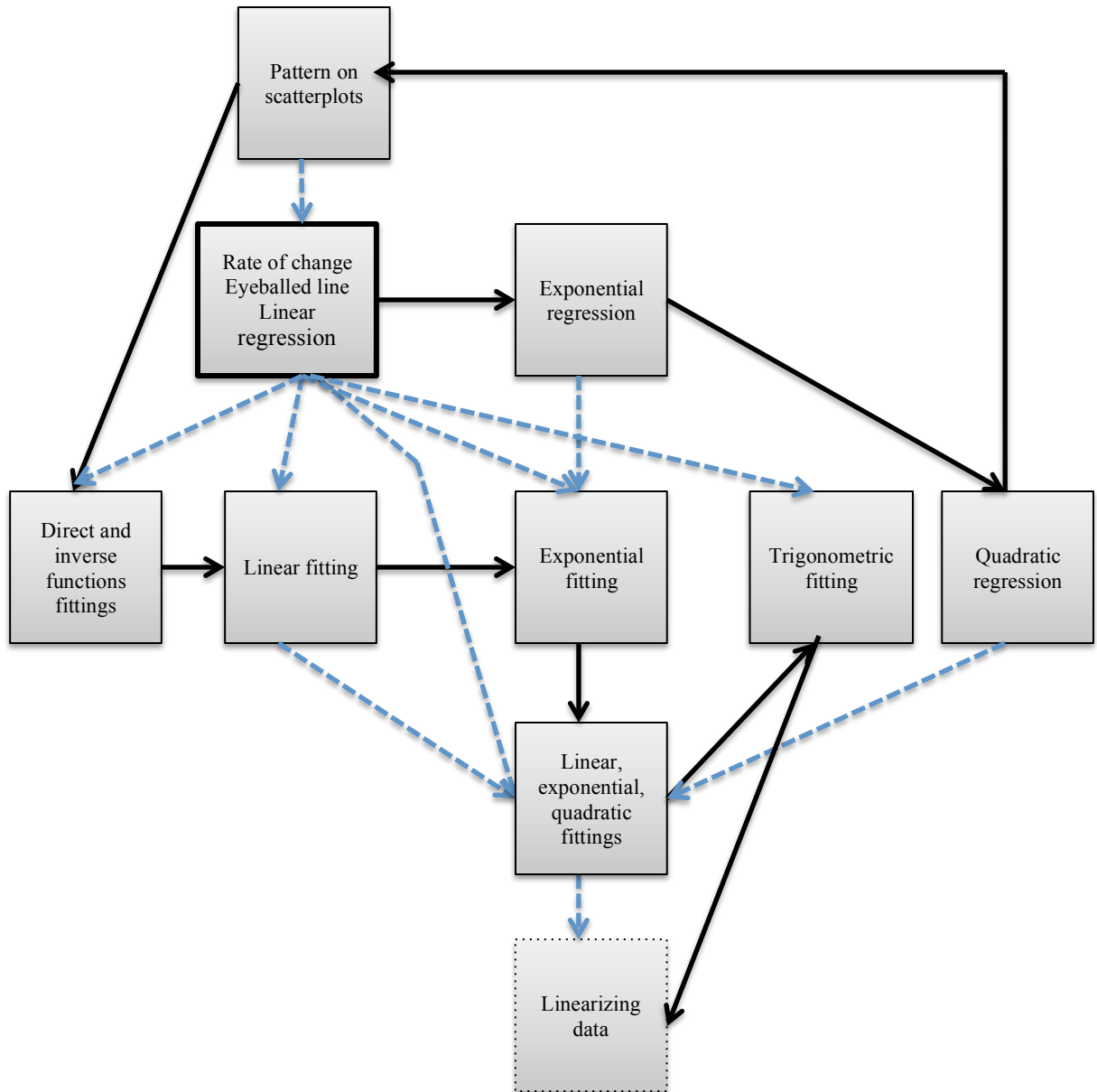
Summary of learning trajectories for two numerical variables in the UCSMP series

Main LEs related to association of bivariate data	Brief explanations of how authors introduce/develop the main learning expectations column	Prior knowledge and knowledge from other conceptual categories related to the main LEs
Rate of change Eyeballed (estimated) line Linear regression	Prompt students to use formulas to calculate rate of change (by hand or with a spreadsheet) Illustrated rated of change of a line as its steepness and meaning as how fast a quantity changes over time. Instruct students to write the equation for an eyeballed line Describe linear regression as the line of best fit, the most common way of finding a line to fit data and ask students to obtain the regression using calculators Introduce <i>deviations</i> State the property of the linear regression	Slope of a line Write an equation of a line passing 2 points
Exponential regression	Require students to use calculators to find the regression and informally use scatterplots to assess the goodness of fit	Know behaviors of exponential functions
Quadratic regression	Prompt students to use calculators to find the regression and informally use scatterplot of assess the goodness of fit	Know behaviors of quadratic functions
Pattern on	Prompt students to observe patterns of	Make a scatterplot

scatterplots	increasing/decreasing on a scatterplot and use informal knowledge to account for the behavior of the relationship	
Direct and inverse functions fittings	Guide students to use a general formula and a pair of data to find the model Prompt students to use both prior theories/common knowledge	
Linear fitting	Prompt students to write the equation for an eyeballed line Describe linear regression as the line of best fit and stated the property of the linear regression illustrated by a moveable line and ask students to obtain the regression using calculators Introduced <i>sum of square deviations</i> (SSD) Require students to make sense of r by examining r in multiple scatterplots	Write an equation of a line passing 2 points
Exponential fitting	Prompt students to use a general exponential formula and substitute two points of data to find the model Instruct students to use calculators to find the regression Require students to choose among linear, quadratic, and exponential functions to fit a set of data	Solve a system of two equations Know behaviors of the functions
Linear, exponential, and quadratic fittings	Define <i>residual</i> , <i>sum of square residuals</i> (SSR) with formulas, and stated three properties of the linear regression and examined SSR with a moveable line Define r with a formula and link to its meaning of statistic for the goodness of fit Caution about correlation and causation Instruct students to use two approaches to find exponential and quadratic models and assess the goodness of fit of the models using residual plots Recommend students to use both informal knowledge and data at hand to find a model	Write an equation of a line passing 2 points Common knowledge about relationship between events
Trigonometric fitting	Prompt students to sketch the curve for relationship and find the model by hand or use calculators	Know behaviors of trigonometric functions, the formula of trigonometric functions and its period and altitude
Logarithmic, exponential, power fittings	Instruct students to linearize data using logarithmic transformations and fit a linear function for the transformed data	Know the logarithmic operation/function Solve a system of linear equations

Through the trajectory for the NN combination, function families such as linear, exponential, power, logarithmic, trigonometric, direct and inverse models are used as means to fit the data, and the behavior of the function are used to assess the model fit. Similarly, knowing how to write an equation for a line passing through two points is a prerequisite for learning linear fitting (as it appears in the strategies the authors use to

introduce linear fitting). Furthermore, knowing how to solve systems of linear equation is required to find exponential fitting for a set of data. Finally, logarithmic transformations are used as a prerequisite to transfer exponential, logarithm, and power fitting to linear fitting.



Note. LTs in the textbook series begin at the bold box and end at the dashed box

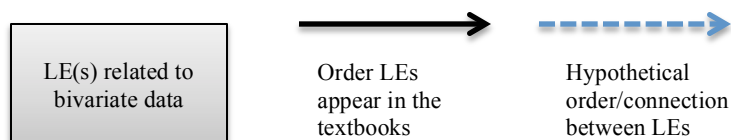


Figure 37. Learning trajectories for the NN association in the UCSMP series.

Core-Plus Mathematics Project (CPMP) Series.

CPMP series learning trajectory for two categorical variables. The relationship between CC variables first appears in Unit 8, *Probability Distributions* of the *Course 2* book. Especially in Investigation 2, *Conditional Probability* of Lesson 1, *Probability Models*, the authors pose a problem to determine if boys were more likely than girls to wear sneakers to school. They ask students to collect data related to the two variables: boys/girls and wearing sneakers/not wearing sneakers in classroom, and make a two-way table to summarize the data. The authors then require students to use the table to calculate probabilities and to examine if the Multiplication Rule holds for these two events. This table is used to find conditional probabilities and in turn to determine whether or not the two events, wearing sneakers and being a girl, are *independent*. Independence is determined by using the definition that events A and B are independent if knowing whether one of the events occurs does not change the probability that other event occurs. From this instance to determine the association between boy/girls and wearing sneakers to school, the authors lead students to draw the conclusion about conditional probability, namely that two events are independent if $P(A)=P(A|B)$ (Figure 38). After this property is established, the authors ask students use two-way tables to determine if two events are independent.

- 3 Recall that events A and B are independent if knowing whether one of the events occurs does not change the probability that the other event occurs.
- Using the data from Problem 1, suppose you pick a student at random. Find $P(\text{wearing sneakers} \mid \text{is a girl})$. How does this compare to $P(\text{wearing sneakers})$?
 - Are the events *wearing sneakers* and *is a girl* independent? Why or why not?
 - Consider this table from a different class.

	Wearing Sneakers	Not Wearing Sneakers
Boy	5	9
Girl	10	18

Suppose you pick a student at random from this class.

- Find $P(\text{wearing sneakers})$.
 - Find $P(\text{wearing sneakers} \mid \text{is a girl})$.
 - Are the events *wearing sneakers* and *is a girl* independent? Why or why not?
- d. If events A and B are independent, how are $P(A)$ and $P(A \mid B)$ related?

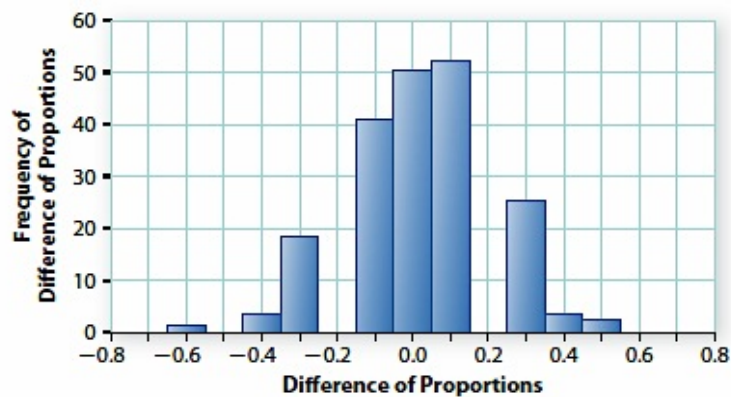
Figure 38. Property of conditional probability of independent events. From *Core-Plus Mathematics: Course 2 – Teacher’s Guide* by C. R. Hirsch, J. T. Fey, E. W. Hart, H. L. Schoen, A. E. Watkins, B. E. Ritsema, R. K. Walker, S. Keller, R. Marcus, A. F. Coxford, G. Burrill, 2008b, Columbus: OH: Glencoe/McGraw-Hill, p.530. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

The association between CC variables is revisited in Course 3, Unit 1. The comparison of proportions (joint relative frequencies) is used to detect association and to determine if the difference in proportions is significant. Randomization Distribution is used to test the significance of the difference. Figure 39 shows an example of such an instance in which authors provide CC variables and ask students to calculate the difference in the proportions and determine whether the test statistic is significant. A histogram for the differences of proportion of 195 trials is made, and students are asked

to run the experiment five more times to create histogram of 200 samples. Using this histogram, students could observe where the difference of proportions in the study lies on the histogram to judge if the treatment makes a statistical significance.

- 16** In a psychology experiment, a group of 17 female college students were told that they would be subjected to some painful electric shocks. A group of 13 female college students were told they would be subjected to some painless electric shocks. The subjects were given the choice of waiting with others or alone. (In fact, no one received any shocks.) Of the 17 students who were told they would get painful shocks, 12 chose to wait with others. Of the 13 students told they would get painless shocks, 4 chose to wait with others. (Source: Stanley Schachter. *The Psychology of Affiliation*, Stanford, CA: Stanford University Press, 1959, pp. 44–45.)

- Describe how to use a randomization test to see if the difference in the proportions of students who choose to wait together is statistically significant. You can let 0 represent a response of waiting alone and 1 represent waiting with others.
- Conduct 5 runs for your test and add them to a copy of the randomization distribution below, which shows 195 runs.



- Is the difference in the proportions who choose to wait together statistically significant?
- Why are there no differences between 0.15 and 0.25?

Figure 39. Using a randomization test to examine the difference in proportions. From *Core-Plus Mathematics: Course 3 – Teacher’s Guide* by J. T. Fey, C. R. Hirsch, E. W. Hart, H. L. Schoen, A. E. Watkins, B. E. Ritsema, R. K. Walker, S. Keller, R. Marcus, A. F. Coxford, G. Burrill, 2009, Columbus, OH: Glencoe/McGraw-Hill, p.99. Copyright 2009 by McGraw-Hill Companies, Inc. Reprinted with permission.

CPMP series learning trajectory for one categorical and one numerical variable. The relationship between CN variables is implicitly addressed in Investigation 1, *Shapes of Distributions* of Unit 2, *Patterns in Data* of Course 1. The authors provide two dot plots of the length of male and female bears and ask students to compare the overall pattern of each distribution—the shapes, centers, and spreads of the two distributions. However, the relationship between two variables is not explicitly addressed until another instance asks students to determine if silver nitrate is effective in causing more rain. Multiple graphical displays including dot plots, back-to-back stem plots, and boxplots are used to compare two distributions. Furthermore, by using the five-number summary related to a boxplot, students can compare two distributions then draw conclusion about the association between two variables. Figure 40 shows an example requiring students to compare distributions of exposure/non exposure to lead dust. In this case, students use only the boxplot for the difference of means to determine if it is significantly different.

In Investigation 2, *By Chance or from Cause?* of Unit 1, *Reasoning and Proof* in Course 3, the relationship between CN variables is revisited and expanded. The authors provide an investigation to examine if the difference between two treatments is significant. The investigation begins with a simple, but well-designed, experiment to determine “whether a calculator helps students perform better, on average, on a test about factoring numbers into primes” (Fey et al., 2009, p. 81). The authors prompt students through several steps: dividing class randomly into two equal groups, using treatment assigned to each group (use calculator vs. by hand), answering the question, and grading the tests. The test is designed such that there was predictably not any difference between

Check Your Understanding

People whose work exposes them to lead might inadvertently bring lead dust home on their clothes and skin. If their child breathes the dust, it can increase the level of lead in the child's blood. Lead poisoning in a child can lead to learning disabilities, decreased growth, hyperactivity, and impaired hearing. A study compared the level of lead in the blood of two groups of children—those who were exposed to lead dust from a parent's workplace and those who were not exposed in this way.

The 33 children of workers at a battery factory were the “exposed” group. For each “exposed” child, a “matching” child was found of the same age and living in the same area, but whose parents did not work around lead. These 33 children were the “control” group. Each child had his or her blood lead level measured (in micrograms per deciliter).

Blood Lead Level (in micrograms per deciliter)

Exposed	Control	Exposed	Control
10	13	34	25
13	16	35	12
14	13	35	19
15	24	36	11
16	16	37	19
17	10	38	16
18	24	39	14
20	16	39	22
21	19	41	18
22	21	43	11
23	10	44	19
23	18	45	9
24	18	48	18
25	11	49	7
27	13	62	15
31	16	73	13
34	18		

Source: “Lead Absorption in children of employees in a lead-related industry,”
American Journal of Epidemiology 135, 1962.

- On the same scale, produce box plots of the lead levels for each group of children. Describe the shape of each distribution.
- Find and interpret the median and the interquartile range for each distribution.
- What conclusion can you draw from this study?

Figure 40. Comparing distributions. From Core-Plus Mathematics: Course 1 – Teacher’s Guide by C. R. Hirsch et al., 2008a, Columbus: OH: Glencoe/McGraw-Hill, p.112. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

the two groups (students are told to guess when answering questions); however, the mean difference of the two groups might still be different from 0. Then students are required to

do the similar experiment 50 times and record the difference in a histogram. They are guided to observe whether the original difference was rare; stated differently, if the probability of the event was less than 5%. Consequently, students could conclude that the mean difference is relatively small and there is no significant difference of students' performance in the two groups, using calculators vs. by hand. A similar activity, if please aromas enhance task completion, is introduced related to chance and cause learning expectation. This time, the authors prompt students to use the Randomization Distribution feature of CPMP-Tools (designed by the groups of authors) to make a histogram in a large numbers of trials (100 times), to determine if the difference is significantly different. In two other situations: penny stack with left-handed and right-handed groups and stem and flower length of plants in Florida, the differences are assessed using the Randomization Distribution tool to gauge whether or not the mean differences in each situation are statistically significant. In addition, graphical displays such as two dot plots and boxplots are used to illustrate the differences.

Types of studies. Content related to types of studies first appear in *Course 2*, Unit 4. In the investigation about association and causation, the authors provide instances that attend to *lurking variables* and show multiple ways that two variables are related to a third variable. Lurking variable is specified using informal knowledge about the relationship between variables. Then, the authors describe how to draw conclusion about causation, “the only way to find out for sure [whether an association means that one variable causes the other or whether there is a lurking variable that causes both] is to conduct an experiment” (Hirsch, 2008b, p. 303). The authors then introduce *random assigned* and *treatments* in related to an experiment illustrated in the example about the

influence of a cup of tea on headache pain reduction. Students are asked to design an experiment to examine the causation in other contexts such as effect of mind games on the brain.

Types of studies are formally developed in Investigation 1, *Design of Experiments* of Lesson 4, *Statistical Reasoning* of Unit 1 of *Course 3*. In a specific example about mung beans, the authors ask students to specify treatment and response variables and provide multiple situations to determine if a conclusion about causation could be made from the design of the study. After showing these examples, the authors ask students to design an experiment to determine whether zapping mung bean seeds in a microwave makes them sprout faster (Figure 41). They then provide a formal description of three characteristics of a well-designed experiment: *random assignment*, *sufficient number of subjects*, and *comparison group or control group*. The authors ask students to use these definitions and link back to the aforementioned situations to specify if the design meets one or more of the characteristics. Furthermore, in the context of testing a polio vaccine by Jonas Salk, the authors introduce terminologies, *placebo effect*, *subject blind*, *evaluator blind*, and *double blind*. They formally introduce *lurking variable* to explain the difference between association and causation. In order to illustrate the difference, the authors use specific situations such as the study the effect of bacteria resistance to antibiotics. Informal knowledge is involved to identify lurking variables. Simulations are used to help determine if the difference in the treatments happened by chance or cause in a series of activities.

- 1 A common science experiment attempts to determine if mung bean seeds that are given a gentle zap in a microwave oven are more likely to sprout than mung bean seeds that are not given a zap.



Mung beans that were zapped in a microwave oven

- a. In such an experiment, what are the **treatments** (conditions you want to compare)? What is the **response variable** (outcome you are measuring)?
- b. For his experiment, Carlos zapped 10 mung bean seeds and 8 sprouted. Explain why Carlos should not conclude that mung bean seeds zapped in a microwave are more likely to sprout than if they had not been zapped.
- c. For her experiment, Mia took 20 mung bean seeds, picked out 10 that looked healthy and zapped them. Of the 10 that were zapped, 8 sprouted. Of the 10 that were not zapped, 3 sprouted. Explain why Mia should not conclude that mung bean seeds zapped in a microwave are more likely to sprout than if they had not been zapped.
- d. For her experiment, Julia took 4 mung bean seeds, selected 2 at random to be zapped, and zapped those 2. Both seeds that were zapped sprouted. The 2 seeds that were not zapped did not sprout. Explain why Julia should not conclude that mung bean seeds zapped in a microwave are more likely to sprout than if they had not been zapped.
- e. Design an experiment to determine if mung bean seeds are more likely to sprout if they are zapped in a microwave.

Figure 41. Experimental studies. From Core-Plus Mathematics: Course 3 – Teacher's Guide by J. T. Fey et al., 2009, Columbus, OH: Glencoe/McGraw-Hill, p.76. Copyright 2009 by McGraw-Hill Companies, Inc. Reprinted with permission.

After addressing experimental study, three types of studies are investigated in

Investigation 3, *Statistical Studies*. The authors describe three studies as follows:

sample survey or poll: you observe a random sample in order to estimate a

characteristic of the larger population from which the sample were taken. Getting

a random sample of size n is equivalent to writing the name of every member of the population on a card, mixing the cards well, and drawing n cards.

experiment: You randomly assign two (or more) treatments to the available subjects in order to see which treatment is the most effective

observational study: The conditions you want to compare come already built into the subjects that are observed. Typically, no randomization is involved. (Fey et al., 2009, p. 89)

Then the authors set up one problem: the effects of exercise on the blood pressure and of students in your school. They provide three different designs and asked students to categorize the type of the study to each one. The authors then asked students to draw conclusions from the study designs focusing on association and generalization. Several other studies are provided and students are required to classify the types of studies. A connection between types of studies and conclusions is made to clarify the characteristics of each type study.

Summary of the LT for CC and CN combinations and their relation to other conceptual categories. The progression for association between CC variables is as follows: (a) assess the independence of two events using descriptions about the relationship and link the independence to conditional probabilities, (b) use two-way tables to calculate probabilities and assess the independence, and (c) examine the difference of two proportions using simulation.

The progression for the CN includes: (a) compare distributions using graphical displays and/or numerical statistics and (b) decide if the difference between the means of two groups is statistically significant using simulation with technology. Figure 42 summarizes the trajectories for the CC and CN combinations in the CPMP series. Because content related to distinguishing types of studies involved both types of combinations—CC and CN—two learning trajectories are synthesized in one figure.

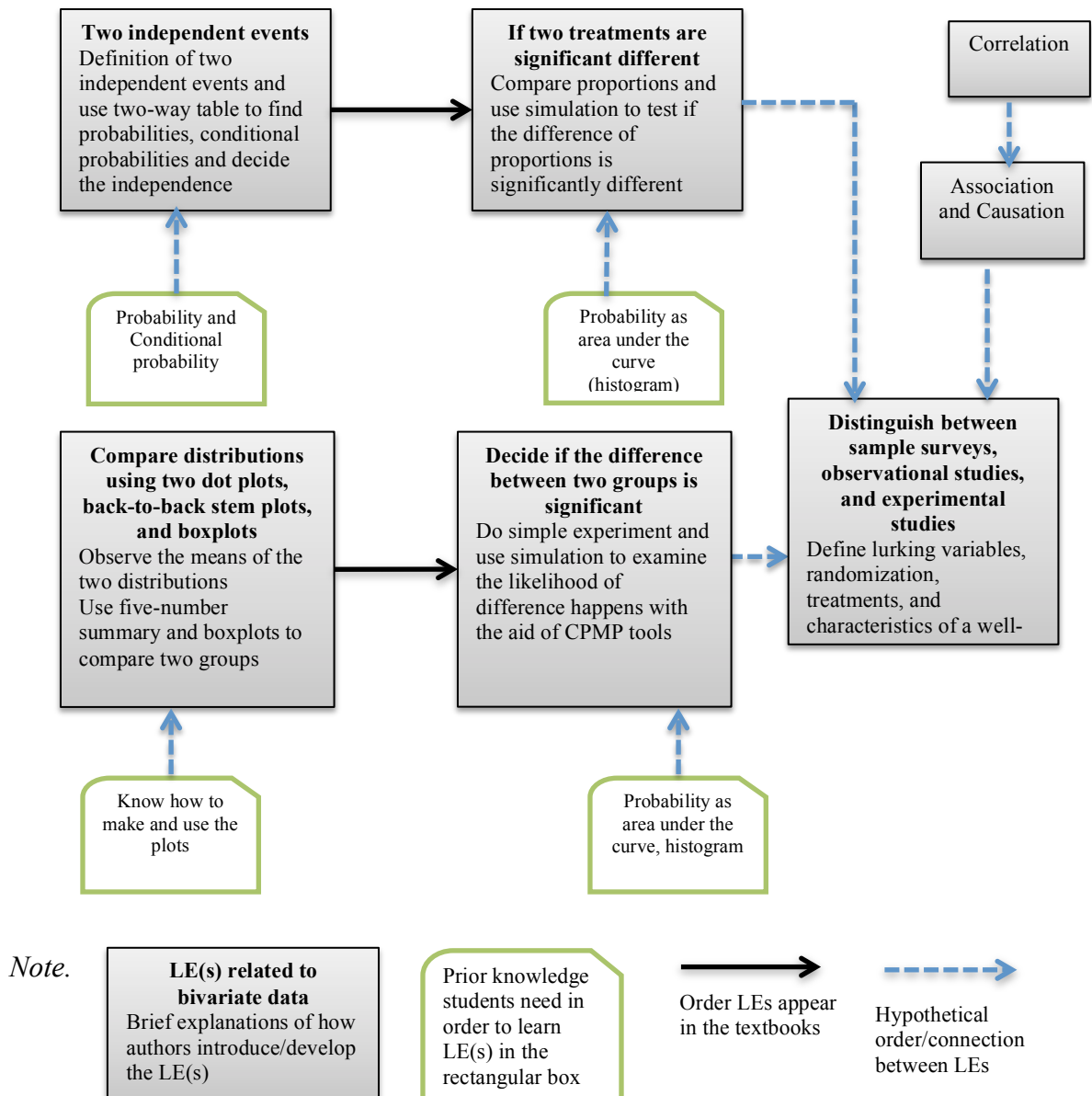


Figure 42. Learning trajectories for the CC and CN associations in the CPMP series.

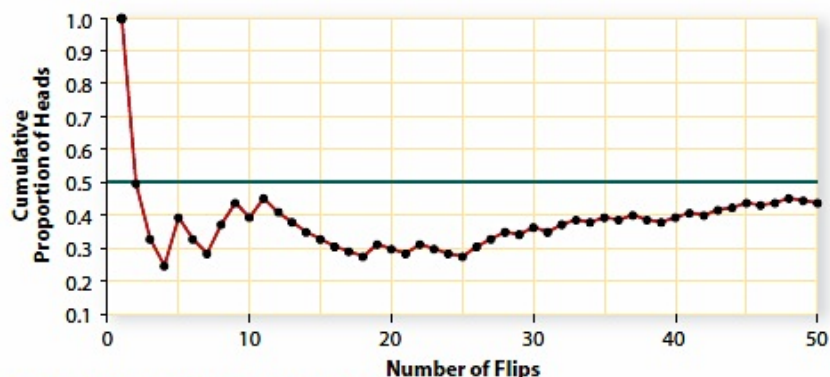
In the CN combination, understanding that empirical probability, in a large number of trials, approaches theoretical probability helps students make sense of the significance of the test. Furthermore, using histograms to express and understand the probability of an event helps students determine if the test statistic is significant. Graphical displays such as histograms, dot plots, stem plots, boxplots, and numerical statistics are used as tools to compare groups. By comparing how one attribute differs in groups, the association between two variables: the aforementioned attribute and the group attribute (e.g., gender) reveals. Vice versa, solving a problem related to the association provides a context for the numerical and graphical displays of univariate data to appear.

CPMP series learning trajectory for two numerical variables.

Observing pattern in plots. The relationship between NN variables begins in Lesson 1 Unit 1 of *Course 1*. Students are asked to collect data related to the stretch of a bungee cord in relation to attached weights. They make a scatterplot, observe a pattern, and made a prediction. Students are also required to collect data for “a hands-on experiment that leads to a random walk pattern, foreshadowing probability ideas like the Law of Large Numbers and expected value” (Hirsch et al., 2008a, p. T8). Using a scatterplot, students can informally observe the pattern of change – a pattern in a large number of trials. This content is revisited in Unit 8 – *Pattern in Chance* when students are required to do a simulation – flipping a coin and observe cumulative proportion of heads in a large number of trials (Figure 43).

Observing patterns over time is revisited in *Course 3*, Lesson 3 of Unit 4, *Statistical Process Control*. The authors use a plot over time, which is referred as a *run*

- 6 To illustrate the Law of Large Numbers, some students made the graph below. They flipped a coin 50 times and after each flip recorded the *cumulative* proportion of flips that were heads.



Source: Tim Erickson, *Fifty Fathoms*, Key Curriculum, 2002, pages 48–49.

- Was the first flip heads or tails? How can you tell from the graph whether the next flip was a head or a tail? What were the results of each of the first 10 coin flips?
- To three decimal places, what was the proportion of heads at each step in the first 10 coin flips?
- Could the graph eventually go back above 0.5? Explain.

Figure 43. Pattern of chance. From *Core-Plus Mathematics: Course 1 – Teacher’s Guide* by C. R. Hirsch et al., 2008a, Columbus: OH: Glencoe/McGraw-Hill, p.555. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

chart, to highlight how one attribute varies along a period. By observing the plot-over-time, students can detect when the mean and standard deviation (SD) change. In

Investigation 1, *Out of Control Signals*, the authors show run charts depicting a series of containers filling with milk and ask students to detect which run chart is far away from the line of mean and SD. Then the authors introduce the *three-standard-deviation rule*, “a process out of control when a single value is more than three standard deviations from the mean” (Fey et al., 2009, p. 287), to detect the out-of-control points. In addition, they illustrate several other characteristics of an out-of-control process such as too many points in a row above or below the mean. They also provide eight tests of the out-of control process illustrated in graphs and use them to detect other situations either providing a run chart or asking students to make a run chart from data provided.

Observing scatterplots attending to the zone around the mean lines provided a means to assess the statistical process control.

Linear fitting. In order to quantify a pattern, in Investigation 3, *Fitting Lines* of Lesson 1, *Modeling Linear Relationships* in Unit 3, *Linear Functions*, *Course 1* the authors require students to write a function to model a set of data. Using a situation about the relationship between flag height and shadow location, the authors ask students to eyeball (estimate) a line and write a rule for a linear function and use it to make predictions. A similar process is required for other situations such as airplane flight distance and time, altitude and temperature. When fitting a line to a set of data, students are asked to interpret rate of change, the slope and coefficient of the model in terms of the context.

After learning how to estimate a line of fit, students are required to perform a linear regression using calculators or computers and use the model for prediction. The authors also ask students to superimpose the model on the set of data and informally assess the fit of the model by observing if the model “closely matches the pattern in the data” (Hirsch et al., 2008a, p. 165). When performing a linear regression, it is informally shown that the linear regression procedure works best when data closely resembles a linear pattern. The authors do not include a formal description about linear regression to this point with the exception of the statement: “Linear regression is a branch of statistics that helps in studying relationship between variables” (Hirsch et al., 2008a, p. 166). The two approaches of fitting, line eyeballing and linear regression, are used interchangeably throughout the lesson.

Curve fitting. Following linear modeling, in Investigation 4, *Modeling Data Patterns*, Lesson 1, *Exponential Growth* in Unit 5, *Exponential Functions*, students deal with exponential regression. In particular, they are required to make a scatterplot for Midwest wolves over time since 1990, use calculators to experiment with linear and exponential regression, decide which model is a better fit for the data, and then use the model for prediction. The authors describe exponential regression for the first time in this lesson. Similarly, in Lesson 2, *Exponential Decay*, the authors ask students to collect data, make a scatterplot, and observe and write a functional rule to fit the set of data. In particular, students are required to collect data about dropping 100 coins on the table and removing all coins that landed heads up. Students are instructed to, then, roll a six-sided die and remove the number of coins equal to the number on the top face of the die.

As another type of modeling, in Lesson 1, *Direct and Inverse Variation* of Unit 1 – *Functions, Equations, and Systems* of Course 2, students are asked to do an experiment related to ramp length, platform height, and time for a ball to roll down a ramp. By making a scatterplot, students could detect the pattern of decrease/increase of the quantity. Then, students are asked to find a functional rule to model the set of data either with direct or inverse functions. For other functions including the inverse of root functions and power functions, calculators are used to find the model to fit a set of data.

In Lesson 3, *Rational Function Models and Operations*, Unit 3, *Algebraic Functions and Equations* of Course 4, rational function is visited and developed. Students study the relationship between the number of drops it takes to break a whelk and drop height by observing a scatterplot. They then validate the reason(s) quadratic and exponential models are not appropriate for the data while the inverse variation was

appropriate. Up to this lesson, three types of regressions: linear, exponential, and power are used to examine which model fits a set of data.

Expansion of linear fitting. Formal and explicit bivariate relationships are revisited and expanded in Unit 4 – *Regression and Correlation of Course 2*. In particular, in Lesson 1, *Bivariate Relationships*, the authors start with Investigation 1, *Rank Correlation*. Students are asked to rank eight different types of music (ties were not allowed) and to display the ranks in pairs in a scatterplot. They then are required to observe the pattern of the scatterplot to highlight any strong positive, strong negative associations, weak and no associations, and decide which scatterplot could be used to predict the ranks of one person using the other ranks. Students are then asked to brainstorm ways to assign a number to each scatterplot and recommended to use extreme cases of “perfect positive/negative association” to test their method. Furthermore, they are prompted to determine whether their methods are applicable to describing the direction and strength of the association with an entirely new set of data. The authors introduce the formula of Spearman’s rank correlation (r_s) and require students to calculate it by hand or by using a spreadsheet. The link between r_s and scatterplots is also addressed in this lesson. In one instance of the assignment portion, the authors introduce Kendall’s rank correlation (r_k) providing the formula and illustrating it with a set of four items. They then ask students to find the rank correlation for multiple previous sets of data and compare r_k and r_s to determine if they are equivalent (Figure 44).

In the next Investigation, *Shapes of Clouds of Points*, the authors described a good linear fit for a set of data “if they [data] form an *oval* or *elliptical cloud*” (Hirsch et al., 2008b, p. 264) illustrated with scatterplots. Students are then asked to describe the

- 17 Maurice Kendall, a British statistician (1907–1983), developed an alternative method to measure the strength of association between two rankings. **Kendall's rank correlation r_k** is given by the formula

$$r_k = 1 - \frac{2c}{\frac{n}{2}(n-1)}.$$

Here, n is the number of items being ranked. To find c , write the ranks for each item side-by-side and connect the ranks as shown below for $n = 4$.

Item	First Ranking	Second Ranking
A	1	2
B	2	4
C	3	3
D	4	1

The number of crossings of the lines is c . Here, $c = 4$.

- Does a large number of crossings indicate general agreement or general disagreement in the ranks?
- Find Kendall's rank correlation for the roller coaster data in Applications Task 1 (page 269).
- Compute Kendall's correlation when there is perfect agreement between the ranks 1 to 5. Compute Kendall's correlation when there is completely opposite ranking of five items.
- Are Spearman's and Kendall's rank correlations equivalent? That is, do they always give the same value? Explain your answer.
- Investigate whether r_k always lies between -1 and 1 .

Figure 44. Spearman's and Kendall's rank correlations. From *Core-Plus Mathematics: Course 2 – Teacher's Guide* by C. R. Hirsch et al., 2008b, Columbus: OH: Glencoe/McGraw-Hill, p.277. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

direction and *strength* of the relationships in various scatterplots (linear, curved, varying in strength) (Figure 45). They informally describe outliers using a scatterplot.

In Lesson 2, *Least Squares Regression and Correlation*, linear regression is revisited and fully developed. In an instance examining the relationship between car weight and highway mileage, the regression line gotten by calculators is used to find the model for the data. Error in prediction and residuals are defined and students are asked to calculate the error for the above model and assess the position of a data point corresponding with the value of the residual. After that, sum of squared residuals (SSR) is introduced in a set

- 1 Bivariate data are **linear** if they form an *oval* or *elliptical cloud*. Two examples of linear relationships are shown below.

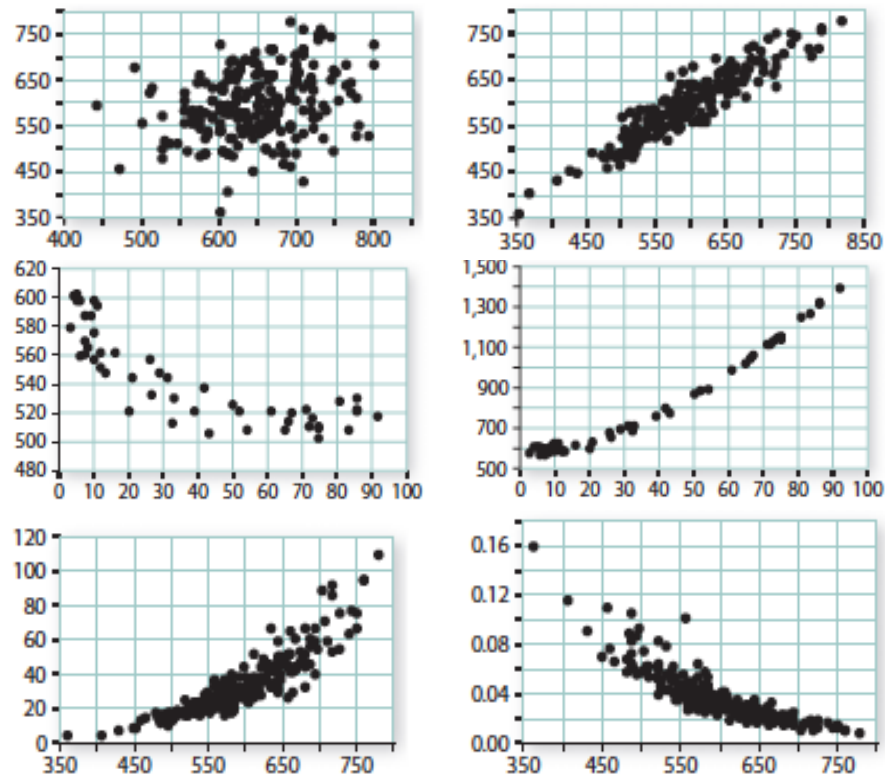


Figure 45. Different types of bivariate relationships. From *Core-Plus Mathematics: Course 2 – Teacher’s Guide* by C. R. Hirsch et al., 2008b, Columbus: OH: Glencoe/McGraw-Hill, pp. 264-265. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

of three data points for different models and compared to determine which model yielded the least SSR to lead to the conclusion about linear regression: “The regression line or least squares regression line is the line that has a smaller sum of squared errors (residuals), or SSE, than any other line” (Hirsch et al., 2008, p. 284).

In Investigation 2, *Behavior of the Regression Line*, the authors define *centroid* and prompt students to prove that the regression line passes through the centroid with a set of three data points. Influential points, which are initially considered as outlier data points, are introduced. With the aid of software, the authors prompt students to remove influential points and observe the change of slope, y -intercept of regression line. Multiple

situations such as relationships between horse height and hip angle while running, and season and World Series batting averages are the contexts in which students investigate the influential points.

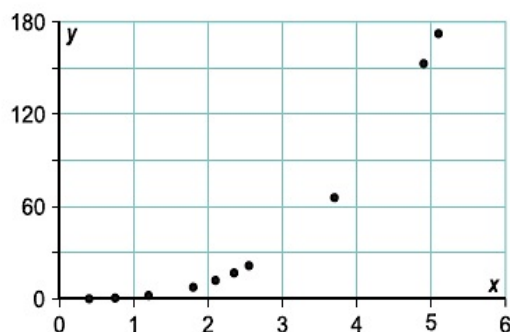
In Investigation 3, *How Strong Is the Association?*, the authors introduce Pearson's r in a formula and guided students to calculate it with a set of three data points. A geometric display dividing the scatterplot into four quadrants with two lines parallel to the coordinate axes and passing through the centroid is used to help students understand the direction of the scatterplot and sign (+ or -) of r -value. Estimating the value of r from a scatterplot is also addressed. Influential points are also visited in relation to r . The authors use multiple situations to introduce students to the idea that the value of r can be high even though the relationship is not linear. They note that caution should be exercised when interpreting r because correlation is only good for interpreting linear models. In addition students should consider "scatterplot before deciding what summary statistics are appropriate and how to interpret them" (Hirsch et al., 2008b, p. T297). Correlation is then revisited in the context of distinguishing between association and causation with the introduction of *lurking variable*, to account for the relationship between two variables. In addition, tasks asking students to examine the relationship between slope and SSE are found in the assignment portion of the lesson.

Expansion of curve fittings. Model fittings are revisited in the *Course 4* book. In Unit 1, *Families of Functions*, data transformations such as translations and reflections are made before calculating exponential regression. Tasks include asking students to add a constant to one variable and use curve fitting for the transformed data of cooling time as a function of temperature and to do a reflection and vertical translation before finding the

exponential regression for the warming time as a function of temperature. In this lesson, the goodness of fit of the models is assessed using a scatterplot.

In Unit 5, *Exponential Functions, Logarithms, and Data Modeling*, Lesson 2, *Linearization and Data Modeling*, the authors introduce residual plots. They use a problem predicting population over time in Investigation 1, *Assessing the Fit of a Linear Model*, in which they ask students to make a residual plot, a plot of independent variable on the x -axis and the difference between the actual and predicted values in the y -axis and observe the plot to determine if a linear function is an appropriate model for the data. A similar process is applied for other situations such as cost of land fertilizing and radius of the land. The authors use a logarithmic transformation in Investigation 2, *Log Transformations* to linearize data and carry out the concurrent process for the transformed data. In Investigation 3, *Log-Log Transformations*, the authors introduce taking the logarithm of two variables and find model fitting did for the transformed data. After all three models are introduced, and several instances ask students to find an appropriate model for the data. Figure 46 shows an example of experimenting with different models to fit a set of data.

- 12 When looking at the following graph and table, it is difficult to tell if the numbers come from an equation of the form $y = Ax^2$, $y = Ae^{rx}$, $y = Ax^b$, or possibly some other form.



x	y
0.4	0.1
0.75	0.55
1.2	2.2
1.8	7.6
2.1	12.0
2.35	16.9
2.55	21.6
3.7	65.8
4.9	152.9
5.1	172.4

- Experiment with transformations of the tabled x and y values until you get a linear pattern.
- Find the linear regression equation that expresses the linearized pattern and then write the resulting equation in equivalent form to express y as a function of x .

Figure 46. Linearization of data. From *Core-Plus Mathematics: Course 4 – Teacher's Guide* by C. R. Hirsch, J. T. Fey, E. W. Hart, H. L. Schoen, A. E. Watkins, B. E. Ritsema, R. K. Walker, S. Keller, R. Marcus, A. F. Coxford, G. Burrill, 2008b, Columbus: OH: Glencoe/McGraw-Hill, p.404. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

Quantification of rate of change. Instantaneous rate of change is introduced with specific data as a preparation to learn the derivative in Unit 7, *Concepts of Calculus*.

Students are asked to observe a pattern in a table, and use the table to estimate a jumper's velocity at one time point; the authors define that as *instantaneous velocity*. This concept is used to assess the relationship between two variables and for learning the concept of derivative.

Summary of the LT for NN combination and its relation to other conceptual categories. The learning progression for the NN association includes several times of content appearing. It starts with observing a pattern of increasing/decreasing in a

scatterplot, then moves to linear fitting, exponential regression, direct and inverse modeling and power regression. Exponential regression is revisited and expanded with the adding on of data transformations. Further, linearizing data is introduced to link multiple model fittings with linear fitting and residual plots are used to assess the goodness of fit. Patterns of change of a quantity over time are introduced in the pattern in chance and statistical process control. The LT ends with instantaneous rate of change as a value to measure how a quantity varies over time. Table 12 and Figure 47 summarize the CPMP series' learning trajectory for NN combination.

Table 12

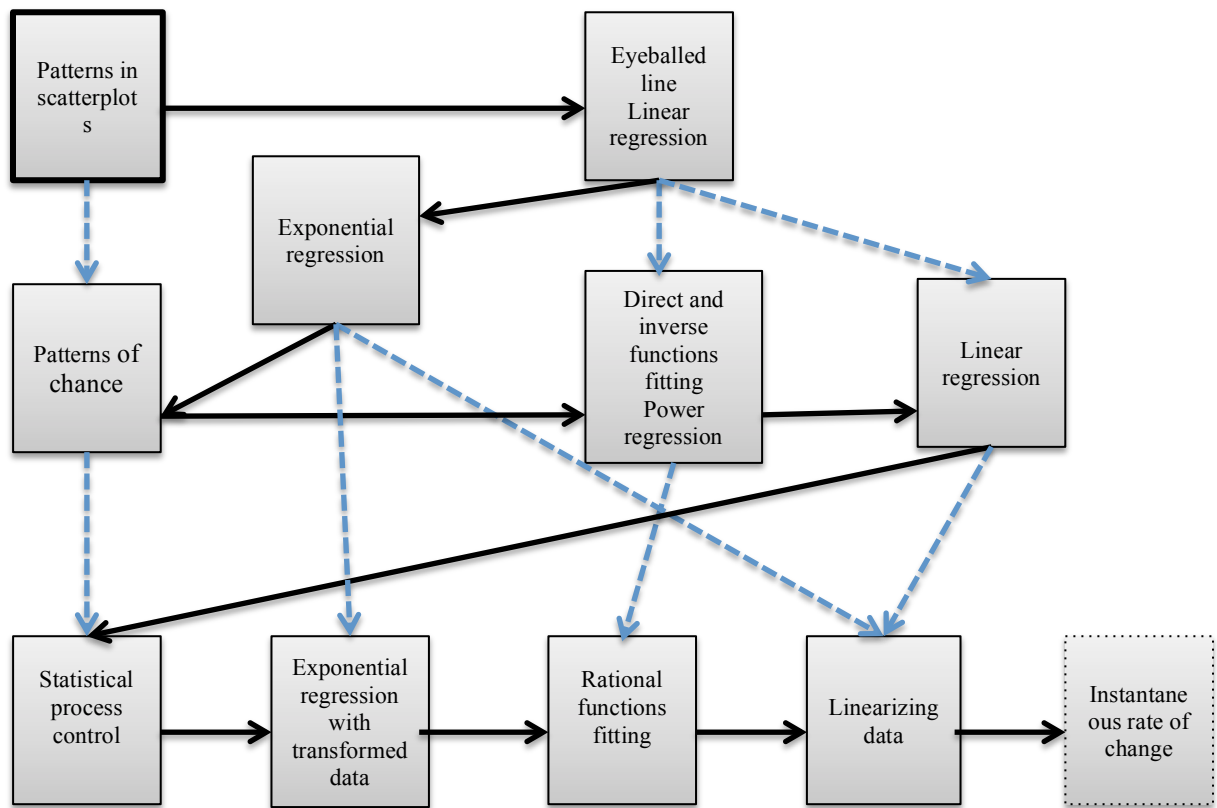
Summary of learning trajectories for two numerical variables in the CPMP series

LEs related to association of bivariate data	Brief explanations of how authors introduce/develop the LEs column	Prior knowledge and knowledge from other conceptual categories related to the LEs
Pattern on scatterplots	Prompt students to observe patterns of increasing/decreasing on a scatterplot	Make a scatterplot
Eyeballed line Linear regression	Require students to write the equation for an eyeballed line Require students to obtain the regression using calculators without formal descriptions about the regression Informally assess the goodness of fit using scatterplots Introduced <i>deviations</i> State the property of the linear regression	Slope of a line Write an equation of a line passing 2 points
Exponential regression	Prompt students to use calculators to find the regression and informally use scatterplot of assess the goodness of fit	Know behaviors of exponential functions
Pattern of chance	Prompt students to observe a pattern of data in a large number of trials	Make a scatterplot
Direct and inverse functions fittings Power regression	Guide students to use a general formula and a pair of data to find the model and superimpose the model on data Asked students to use calculators to find regressions	Know behaviors of direct and inverse functions
Linear regression	Develop and define rank correlations (Spearman's and Kendall's) as a measure to quantify association Illustrate the ideal scatterplots for a linear fitting and showed various types of association with scatterplots Define <i>SSR</i> and property of the linear regression in terms of minimizing <i>SSR</i> , passing through the centroid Introduce and investigate influential points with technology Define Pearson' correlation and ask students to	Common knowledge

	calculate, estimate r from a scatterplot, and interpret its meaning Distinguish association and causation	
Statistical process control	Ask students to make/use run charts to detect deviation from the mean and SD Show eight types of out-of-control process and asked students to categorize specific situations into the types	Make a run chart, read mean and SD from the chart
Exponential Regression with transformed data	Translate and/or reflect data and fit an exponential function to the transformed data	Know translation and reflection
Rational functions fitting	Require students to use a general formula and a pair of data to find the model and superimpose the model on data	Know behaviors of the functions
Exponential, power regressions	Introduce residual plots and how to make them to assess the goodness of fit of a model Prompt students to linearize data using logarithmic transformations and fit a linear function for the transformed data Guide students to try multiple models and use scatterplots to help specify the model	Make scatterplots
Instantaneous rate of change	Define the instantaneous rate of change and its meaning	Concept of derivatives and rate of change
Trigonometric fitting	Ask students to sketch the curve for relationship and find the model by hand or use calculators	Know behavior of trigonometric functions, know the formula of trigonometric functions and its period and altitude
Linearizing data	Instruct students to linearize data using logarithmic transformations and fit a linear function for the transformed data	Know the logarithmic operation/function Solve a system of linear equations

In NN combination, functional relationships are used as means to model data. Specifically, family of functions is used as a resource to find an appropriate model for a set of data. In addition, transformations such as translations and reflections and logarithmic transformations are used to help find an appropriate model for a set of data by finding linear model for the transformed data. For linear modeling, several concepts of linear models such as slope, rate of change, and intercept are used in the context of data in addressing the relationship between two numerical variables. With regard to examining patterns over time, the empirical approach of probability is supported. That is, students get to understand that in a large number of trials, empirical probability reaches theoretical

probability by observing plots over time. Finally, the context of data sets the stage to prepare for learning derivative.



Note. LTs in the textbook series begin at the bold box and end at the dashed box

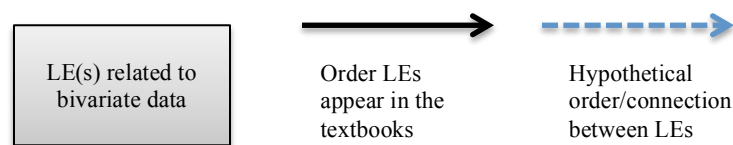


Figure 47. Learning trajectories for the NN association in the CPMP series.

Comparison of Textbook Series Learning Trajectories with Common Core State Standards for Mathematics Learning Expectations

In this section, I describe the extent to which the learning expectations in each of the analyzed textbook series align with the CCSSM. Following the narrative discussion, Table 13 provides a summary of the CCSSM Learning Expectations (LEs) that are

Table 13

Frequency of instances addressing CCSSM LEs related to bivariate data across the three series

CCSSM Code	CCSSM Learning Expectations	HML: Number of instances addressing the LEs	UCSMP: Number of instances addressing the LEs	CPMP: Number of instances addressing the LEs
S-ID.2	Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets	4	5	10
S-ID.5	Summarize categorical data for two categories in two-way frequency tables.	5	0	3
	Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies).	2	5	3
	Recognize possible associations and trends in the data.	2	6	10
S-ID.6	Represent data on two quantitative variables on a scatter plot. Describe how the variables are related.	27	33	48
	Fit a function to the data (emphasize linear models).	36	60	53
	Fit a function to the data (emphasize quadratic models).	2	28	5
	Fit a function to the data (emphasize exponential models).	11	39	31
	Use functions fitted to data to solve problems in the context of the data.	30	39	35
	Use given functions or choose a function suggested by the context.	16	40	22
	Informally assess the fit of a function by plotting.	5	37	54
	Informally assess the fit of a function by analyzing residuals.	4	20	21
S-ID.7	Interpret the slope (rate of change) of a linear model in the context of the data.	1	5	19
	Interpret the intercept (constant term) of a linear model in the context of the data.	1	0	5
S-ID.8	Compute (using technology) the correlation coefficient of a linear fit.	2	6	9
	Interpret the correlation coefficient of a linear fit.	2	4	7

CCSSM Code	CCSSM Learning Expectations	HML: Number of instances addressing the LEs	UCSMP: Number of instances addressing the LEs	CPMP: Number of instances addressing the LEs
S-IC.3	Recognize the purposes of and differences among sample surveys, experiments, and observational studies.	12	0	25
	Explain how randomization relates to each [type of the studies].	6	0	18
S-IC.5	Use data from a randomized experiment to compare two treatments.	1	0	23
	Use simulations to decide if differences between parameters are significant.	1*	0	15

* Bold is added to emphasize the LEs appeared in fewer than 3 instances

addressed in each of the three textbook series. Note that the CCSSM LEs appear in *italics* throughout the discussion of LEs and textbook series' foci.

Holt McDougal Larson series (HML). All the CCSSM LEs were found in this series (Table 13). However, the emphases given to each LE varied considerably in terms of the number of to-be-solved (TBS) instances. Only a relative few instances (ranging from 2 to 5) addressed the associations between CC and CN combinations. In particular, the LEs, *Use data from a randomized experiment to compare two treatments* (S-IC.5) and *Use simulations to decide if differences between parameters are significant* (S-IC.5), appeared only once in terms of TBS instances addressing them. More emphasis was put in fitting a function to a set of data. The focus on types of functions varied with more focus was placed on linear (36 instances) and exponential functions (11 instances) and less on quadratic functions (2 instances). Furthermore, a large number of instances (16) asked students to fit other functions such as inverse and trigonometry functions. However, relatively few instances asked students to determine if the model was a good fit for the data by using scatterplots (5 instances) or residual plots (4 instances).

At the high end in the HML series, there were 43 instances addressing the LE, *Fit a function to the data (emphasizing linear models)* (S-ID.6); however, other LEs related to linear fitting such as correlation coefficient and goodness of fit did not get much emphasis (ranging from 1 to 2 instances). These included: *Interpret the slope (rate of change) of a linear model in the context of the data* (S-ID.7); *Interpret the intercept (constant term) of a linear model in the context of the data* (S-ID.7); *Compute (using technology) the correlation coefficient of a linear fit* (S-ID.8); *Interpret the correlation*

coefficient of a linear fit (S-ID.8); and Distinguish between correlation and causation (S-ID.9).

The University of Chicago School Mathematics Project series (UCSMP).

Similar to the HML series, a lesser focus was put on CC and CN combinations. In contrast, UCSMP put a focus on the relationship between NN variables, which was related to model fittings. Across all three combinations, there were six LEs for which no instance was found in the UCSMP series: *Summarize categorical data for two categories in two-way frequency tables (S-ID.5); Interpret the intercept (constant term) of a linear model in the context of the data (S-ID.7); Recognize the purposes of and differences among sample surveys, experiments, and observational studies (S-IC.3); Explain how randomization relates to each [type of study] (S-IC.3); Use data from a randomized experiment to compare two treatments (S-IC.5); and Use simulations to decide if differences between parameters are significant (S-IC.5).*

The UCSMP series put a focus on fitting a function to a set of data. As in the HML series, linear modeling received the most emphasis (60 instances). The UCSMP authors also placed a more balanced emphasis on other types of functions including quadratic (28 instances), exponential (39 instances), and inverse functions (40). Assessing the fit of a function to data either by using scatterplots or residual plots received more emphasis in UCSMP compared with 37 and 20 instances of using scatterplots and residual plots, respectively, compared to 5 and 4, respectively in the HML series.

For linear modeling, a few instances addressed the slope of a linear model, correlation and distinguishing between correlation and causation. In particular, 5

instances asked students to *Interpret the slope (rate of change) of a linear model in the context of the data* (S-ID.7) but no instances asked students to interpret the intercept of the model. Computing and interpreting correlation received more emphasis in the UCSMP series with 6 and 4 instances, respectively, compared to 2 and 2, respectively in the HML series.

Core-plus Mathematics Project series (CPMP): The CPMP series included all LEs from the CCSSM and each of them received a more balanced emphasis. The LEs, *Summarize categorical data for two categories in two-way frequency tables* and *Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies)* (S-ID.5) received the least emphasis (3 instances for each). All other LEs appeared in at least 5 instances; specifically, LEs that addressed linear fitting and assessing the fit of a model using plots were most emphasized with more than 50 instances.

Based on the examination of the learning trajectories for bivariate data in the three series, several learning expectations related to bivariate data appeared in the series but were not addressed in the CCSSM standards. Table 14 provides the list of LEs that appeared in the three textbook series, but not in the CCSSM. On this list, three LEs were common in the three series and those are the first three LEs that were included in the HML series. The HML and UCSMP series share the same LE, *Fit a trigonometric model for a set of data*, which was not found in the CCSSM. The UCSMP and CPMP series shared four more LEs not in the CCSSM including: (1) *Properties of regression line*, (2) *Sum of square residuals*; (3) *Use common knowledge or prior theory to judge/build a model*; and (4) *Estimate correlation using scatterplots*. One LE that was found only in

Table 14

List of Learning Expectations found across the three series but not included in the CCSSM

HML Series	UCSMP Series	CPMP Series
<ul style="list-style-type: none"> • Calculate and use rate of change in the context of data • Fit an inverse model for a set of data • Linearize a set of data (using log, log-log transformations) and fit a linear model for the transformed data • Fit a trigonometric model for a set of data 	<ul style="list-style-type: none"> • Calculate and use rate of change in the context of data • Fit an inverse model for a set of data • Linearize a set of data (using log, log-log transformations) and fit a linear model for the transformed data • Fit a trigonometric model for a set of data • Properties of regression line • Sum of square residuals • Use common knowledge or prior theory to judge/build a model • Estimate correlation using scatterplots • Chi-square tests 	<ul style="list-style-type: none"> • Calculate and use rate of change in the context of data, • Fit a rational (including inverse) function to a set of data • Linearize a set of data (using log, log-log transformations) and fit a linear model for the transformed data • Properties of regression line • Sum of square residuals • Use common knowledge or prior theory to judge/build model • Estimate correlation using scatterplots • Rank correlations: Spearman's and Kendall's correlation • Outlier and Influential points • Instantaneous rate of change • Transform data and fit a model

the UCSMP series is: *Chi-square test*. Four LEs appeared in CPMP but not in the CCSSM and UCSMP are: *Transform data and fit a model*; *Rank correlations: Spearman's and Kendall's correlation*; *Outlier and Influential points*, and *Instantaneous rate of change*. Across three textbook series, there were a total of 13 LEs not found in the CCSSM for bivariate data.

Comparison of Learning Trajectories with the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Framework*

The GAISE Framework, sanctioned by the American Statistical Association in 2007, provides a guide for the development of statistical literacy, beginning with young children. The framework is comprised of three developmental levels of statistical maturity: Level A, Level B and Level C. While it is considered a developmental framework, it is neither age nor grade-level bound; rather, it articulates a developmental sequence for the learning of increasingly complex statistical concepts and processes.

Applying the GAISE Framework to the Holt McDougal Larson (HML) series.

Two categorical variables. In the HML series, Level B of understanding the CC association was addressed but only slightly in three chapters of the HML series including Chapter 10 and 11 of *Algebra*, and Chapter 12 of *Geometry*. In addition, there were no instances regarding CC variables that moved to Level C of the GAISE developmental progression. Although tasks involving two-way tables were found, learning expectation about the association between two variables was not explicitly stated in the HML textbook series. Similarly, relative frequencies (conditional probabilities) were considered, but they served the purpose of calculating the probabilities not for detecting

association. No measures for quantifying the strength of an association (e.g., ADR, QCR) were considered in this series. Figure 48 summarizes the developmental levels of the CC combination in the HML series. Each block represents the developmental levels involved in one chapter and this display holds for Figures 48 to 56.

<i>Algebra 1</i> <i>Chapter 10</i>	<i>Algebra 1</i> <i>Chapter 11</i>	<i>Geometry</i> <i>Chapter 12</i>
Level B	Level B	Level B

Figure 48. Learning trajectory of the CC association in terms of the GAISE Framework in the HML series.

One categorical and one numerical variable. In the HML series, all Levels A, B, and C were evident. Specifically, in Chapter 10 of *Algebra 1*, several instances asked students to compare the distributions of groups, but not to detect association between two variables. At Level A of the GAISE Framework, stem plots as tools to compare groups and students might estimate mean for detecting the difference of groups were used. In the same chapter, at Level B of the GAISE Framework, boxplots and five measures to compare distributions was addressed. However, the nature of the examples focused on the distribution of univariate data, and did not seek to address the association between *two* variables. Level C was evident in one instance in Chapter 6 of *Algebra 2*, when an activity of simulation appeared showing the statistical significance of the difference. However, no instances related to this expectation were found after this initial simulation activity. Furthermore, neither a test to identify outliers nor a formal test of statistical significance was found in any textbooks comprising the HML series. Figure 49 summarizes the developmental levels of association of CN variables in the HML series.

<i>Algebra 1</i> <i>Chapter 10</i>	<i>Algebra 2</i> <i>Chapter 6</i>
Levels A, B	Level C

Figure 49. Learning trajectory of the CN association in terms of the GAISE Framework in the HML series.

Different types of studies were addressed in this series including an introduction to characteristics of well-designed studies. Students were expected to learn the fundamentals of distinguishing between observational and experimental studies, correlation and causation (Level C) in Chapter 6 of *Algebra 2*. However, in HML there is no preparation for doing experiments at Levels A and B of the GAISE Framework and so, in this sense, there is no evidence of a progression across developmental levels.

Two numerical variables. The HML series addressed all three levels of the GAISE Framework in regard to the NN combination of bivariate data. Nevertheless, not all components suggested in the framework were fully considered. Level A of the GAISE Framework was addressed in two chapters (1 and 3) of *Algebra 1* when the authors introduced contents related to observing pattern in data displayed in a scatterplot and rate of change. These contents appeared again in Chapters 2 and 3 of *Geometry*. Then, in Chapter 4 of the same book, both levels B and C were addressed in relation to linear fitting including eyeballed line and linear regression. Linear modeling appeared again in Chapter 9 of *Algebra 1* to serve as a contrast to introduce quadratic and exponential regression. At level B, the sign of correlation (positive and negative) was introduced; nevertheless, the QCR did not appear. At Level C, there was no opportunity for students to quantify the goodness of fit of a model even though the authors introduced r as a “black-box number” gotten from calculator. No formal formula for r was introduced even though the authors mentioned the strength and direction of r in relation to scatterplots. In

addition, no formal definitions of the regression line were described; learning expectation related to the Sum of Square Residuals (SSR) were not found in the series. Out of the suggestions in the GAISE Framework, linearization of data and other model fittings including exponential, power, inverse, rational, and trigonometric models appeared in the series. Figure 50 summarizes the developmental levels of NN variables in the HML series.

<i>Algebra 1 Chapter 1</i>	<i>Algebra 1 Chapter 3</i>	<i>Algebra 1 Chapter 4</i>	<i>Algebra 1 Chapter 9</i>	<i>Geometry Chapter 2</i>	<i>Geometry Chapter 5</i>
Level A	Level A	Levels B, C	Levels B, C	Level A	Level A

Figure 50. Learning trajectory of the NN association in terms of the GAISE Framework in the HML series.

Applying the GAISE Framework to the University of Chicago School

Mathematics Project (UCSMP) series.

Two categorical variables. The UCSMP series begins with instances related to the CC combination at the highest level of the GAISE Framework, Level C, using a Chi-square test to assess the association between two variables in Chapter 11 of *Algebra*. The authors introduced the processes for using the test and for reading a table of critical values to determine whether the expected and observed values are significantly different. This Level was revisited in Chapter 6 of the *FST* book with the addition of simulation as another tool to check for significance. In this chapter, there were a few instances that asked students to use two-way tables for calculating percent and to informally address the association (GAISE Framework Level B). The authors have included tasks addressing Simpson’s Paradox (Level C). Other measures quantifying the strength of association of two variables at Level B such as ADR and QCR were not addressed in this series. Figure 51 summarizes the developmental levels of the CC combination in the UCSMP series.

<i>Algebra</i> <i>Chapter 11</i>	<i>FST</i> <i>Chapter 6</i>
Level C	Levels B, C

Figure 51. Learning trajectory of the CC association in terms of the GAISE Framework in the UCSMP series.

One categorical variable and one numerical variable. The UCSMP series started and stopped at Level B for this combination of bivariate data in Chapter 1 of *FST*. Specifically, box-plots, a five-number summary and two histograms were used to compare two distributions. The nature of the examples focused on univariate data to address the distribution of data, and did not seek to address the association between two variables (bivariate data). Levels A and C of the GAISE Framework were not addressed in this series. Figure 52 summarizes the developmental levels of the CN combination in the UCSMP series.

<i>FST</i> <i>Chapter 1</i>
Level B

Figure 52. Learning trajectory of the CN association in terms of the GAISE Framework in the UCSMP series.

The study of different types of studies was not investigated in the UCSMP series. However, the difference between correlation and causation (Level C) was addressed in the textbooks series in Chapter 2 of *FST*.

Two numerical variables. All three developmental levels of the GAISE Framework were addressed in the UCSMP series. In this series, tasks at Level B and C of the Framework started in the Chapter 6 of *Algebra*. At Level B, students were required to find a line of fit with eyeballed line. However, at Level C, the linear regression was first derived using graphing calculators and described as the line of best fit. Then these two

levels were addressed in Chapter 7 of the same book in which the authors use linear models to contrast with exponential regression. Then, tasks returned to Level A in Chapter 2 of *Advanced Algebra* when students were instructed to observe the pattern of association in a scatterplot. In Chapter 3 of *Advanced Algebra*, Levels B and C of the GAISE Framework were re-addressed. In this chapter, the authors added a moveable line to illustrate and define Sum of Square Residuals (SSR) and informally investigate r (Level C). Subsequently, Levels B and C were revisited in Chapter 2 of *FST*. In this chapter, SSR was formally introduced and the properties of linear regression was stated and empirically proved. Furthermore, residual plots were used as tools to assess the goodness of fit. Quadrant Count Ratio measure for the strength of association (Level B) was not found. The element of error when interpreting data at Level C was not found in the UCSMP series. At Level C tasks, the authors addressed the difference between correlation and causation with specific examples. Not explicitly stated in the GAISE Framework, linearization of data was also found and other types of regressions: exponential and power, rational, trigonometric modeling were addressed in the UCSMP series. Figure 53 summarizes the developmental levels of the NN combination in the UCSMP series.

<i>Algebra Chapter 6</i>	<i>Algebra Chapter 7</i>	<i>Advanced Algebra Chapter 2</i>	<i>Advanced Algebra Chapter 3</i>	<i>FST Chapter 2</i>
Levels B, C	Levels B, C	Level A	Level B, C	Levels B, C

Figure 53. Learning trajectory of the NN association in terms of the GAISE Framework in the UCSMP series.

Applying the GAISE Framework the Core-plus Mathematics Project (CPMP) series.

Two categorical variables. Two levels, B and C, of the GAISE Framework were addressed in the CPMP series regarding CC variables. In particular, the authors started with Level B tasks using conditional probabilities to address the CC association in Unit 8 of *Course 2*. At Level B, Two-way tables were used as a tool to assess the association of two variables by comparing relative frequencies of the cells. However, ADR and QCR measures were not found in the series. Moving to Level C, the authors introduced simulation as a way to determine whether the difference of proportions is significant in Unit 1 of *Course 3*. Phi measure (Level C) was not found in the texts. In addition, no task addressing Simpson’s Paradox was found in the series. Figure 54 summarizes the developmental levels of the CC combination in the CPMP series.

<i>Course 2</i> <i>Unit 8</i>	<i>Course 3</i> <i>Unit 1</i>
Level B	Level C

Figure 54. Learning trajectory of the CC association in terms of the GAISE Framework in the CPMP series.

One categorical variable and one numerical variable. Two developmental levels, A and B of the GAISE Framework, were addressed in Unit 2 of *Course 1* in the CPMP series. In particular, at Level A, dot plots and back-to-back stem plots were used to compare two groups and at Level B, side-by-side boxplots and a five-number summary were introduced to compare distributions. The authors did explicitly link the association between CN variables to comparing groups. Later, in Unit 1 of *Course 3*, Level C of the GAISE Framework was addressed when the CPMP-tools were used to aid in doing a simulation to determine if the mean difference between two groups was significant.

Figure 55 summarizes the developmental levels of the CN combination in the CPMP series.

<i>Course 1</i> <i>Unit 2</i>	<i>Course 3</i> <i>Unit 1</i>
Levels A, B	Level C

Figure 55. Learning trajectory of the CN association in terms of the GAISE Framework in the CPMP series.

Different types of studies were also addressed in this series in Unit 4 of *Course 2*. In particular, characteristics of well-designed studies (Level C) were introduced and addressed in the textbooks. Students began with the basic task of learning to distinguish between observational studies and experimental studies and moved to the more complex tasks of correlation and causation (Level C). No preparation for doing experiments at Levels A and B of the GAISE Framework was found in the learning trajectories.

Two numerical variables. The CPMP series addressed the order of the three developmental levels of the GAISE Framework, beginning by observing patterns at Level A, using an eyeballed line to fit data at Level B, and finding linear regression at the highest level of statistical maturity, Level C. In Unit 1 of *Course 1*, students were required to make a scatterplot and observe the pattern in data displayed in the plot for predictions (Level A). Then, Levels B and C were addressed in Unit 3 of the same book. Eyeballed line was first used (Level B) and then linear regression (Level C) was introduced to find a model for a set of data and use the model for predictions. However, the graphing calculator was used to find linear regression and no formal descriptions about the regression were provided. At Level A of the GAISE Framework, observing pattern in a scatterplot was addressed again in Unit 8 of *Course 1* when the authors introduced pattern in chance. Linear modeling (Levels B and C) was formally revisited in

Unit 4 of *Course 2*. In this chapter, at Level C, correlation coefficient, r and SSR were formally introduced and used to assess the goodness of fit for a linear model. The relation between correlation and causation was addressed after learning about correlation and was expanded to association and causation in conjunction with types of studies. The properties of linear regression, specifically in relation to minimize SSR, were shown and examined in specific examples. In Unit 4 of *Course 3*, observing pattern in scatterplot (Level A) appeared again with the introduction of run chart for statistical process control. Residual plots (Level C), as a tool to examine the goodness of fit, appeared in Unit 5 of *Course 4*.

With respect to measures quantifying the strength of association, the authors did attend to QCR (Level B) but not explicitly. Instead, they included two other correlations, Spearman's and Kendall's rank correlation appearing in Unit 4 of *Course 2*, before formal investigation of the correlation, r . Furthermore, geometric transformations such as reflections and translations were introduced in Unit 1 of *Course 1* as a preparation step to finding the model for data. Other types of regression such as exponential, power and rational were found in the CPMP series. As the two previous series, the linearization of data was addressed providing another approach for fitting (Unit 5 of *Course 4*). In addition, the topics of statistical process control (Unit 4 of *Course 3*) and instantaneous rate of change (Unit 7 of *Course 4*) were also addressed in the CPMP series; these topics were not suggested in the GAISE Framework. Figure 56 summarizes the developmental levels of NN combination in the CPMP series.

<i>Course 1</i> <i>Unit 1</i>	<i>Course 1</i> <i>Unit 3</i>	<i>Course 1</i> <i>Unit 8</i>	<i>Course 2</i> <i>Unit 4</i>	<i>Course 3</i> <i>Unit 4</i>	<i>Course 4</i> <i>Unit 5</i>
Level A	Level B, C	Level A	Levels B, C	Level A	Level C

Figure 56. Learning trajectory of the NN association in terms of the GAISE Framework in the CPMP series.

Summary of Learning Trajectories for Bivariate Data

Generally, all three series address three combinations of variables: CC, CN, and NN variables. However, the sequence of each combination differed across the series. The treatments of CN variables in the three series were quite similar, namely using graphical displays and numerical statistics to compare groups. However, the relationships between CN variables were quite different for the three series in terms of tools used to determine the relationship between the variables. Particularly, the Chi-square test was used in the UCSMP series to determine the significance of the difference between expected and observed values. For NN combination the trajectories were much different for the three series. The HML series addressed linear fitting once using calculators, without going into in-depth investigation when the UCSMP series included the topic of linear fitting several times and gradually developed depth when revisiting the same topic a second and third time. The CPMP series placed great emphasis on introducing and developing linear fitting through several times of content appearing.

The HML and CPMP series addressed all LEs related to bivariate data in the CCSSM but the level of emphasis in terms of number of instances differed markedly across the two series. Six LEs were not found in the UCSMP series. Moreover, there were three LEs common to the three series that included in the CCSSM.

In terms of the alignment to the GAISE Framework, the sequences in terms of developmental levels varied regarding the combinations of variables across the three series. In particular, the CPMP series seems to follow the developmental order suggested by the GAISE Framework for the three combinations. In contrast, the UCSMP series

frequently deviated from the GAISE prescribed order of development and this was true for all three variable combinations.

Task Features

This section is presented into five sub-sections: (1) Level of Mathematical Complexity, (2) The Guidelines for Assessment and Instruction in Statistics Education (GAISE) Framework Coding, (3) Association Between Levels of Mathematical Complexity and the GAISE Frameworks, (4) Purpose and Utility Framework Coding, and (5) The Common Core State Standards-Mathematics (CCSSM): Mathematical Practices Coding

Level of Mathematical Complexity

In the following paragraphs, I describe the distributions of the levels of mathematical complexity in each of the textbook series reviewed and provide examples of problems involving the individual levels of complexity.

Distribution of levels of mathematical complexity across the three series.

The distributions of mathematical complexity levels for the HML and UCSMP series are quite similar: about 30% of the instances were low level, 65% were moderate level and about 5% were high level. The CPMP series' distribution was different in that only about 5% of the instances were at the low-level of mathematical complexity. The highest percentage of high-level instances appeared in the CPMP series (about 15%), which is nearly three times higher than that of each of the two other series (Figure 57). Across the three series, the majority of instances required a moderate level of mathematical complexity and ranged from 65% to 80% of instances.

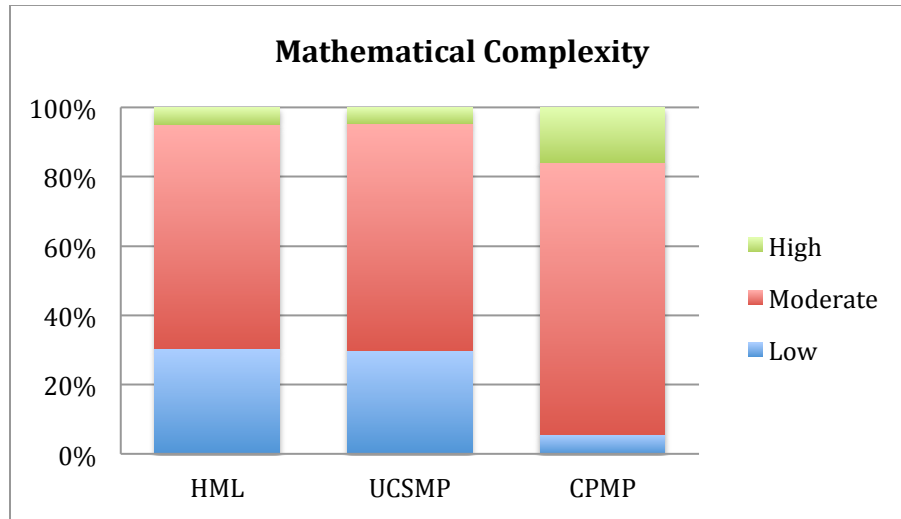


Figure 57. Proportions of levels of mathematical complexity across the three series.

By denoting a numeric code for each level (low = 1, moderate = 2, high = 3) the average of the mathematical complexity for each of the three series can be computed.

Table 15 summarizes the means and standard deviations of mathematical complexity for each of the three series. On average, CPMP instances require highest mathematical complexity among the three series.

Table 15

Means (SDs) of mathematical complexity across the three series

Series	Number of instances related to bivariate data	Mean level of complexity (SD)
HML	122	1.75 (.539)
UCSMP	214	1.75 (.532)
CPMP	246	2.1 (.455)

Examples of instances for the three levels of mathematical complexity.

About 30% of instances in the HML and UCSMP series were at a low-level mathematical complexity. Generally, these instances asked students to recall a fact or to

carry out steps without having to make decisions based on data. Figures 58, 59, and 60 are examples of such instances. In Figure 58, an instance taken from the HML series, students were required to use a calculator to find a function for the data by inputting the data into the calculator and using the trigonometric regression utility of the calculator. In Figure 59, an instance taken from the UCSMP series, students were required to determine if the data points scatter around a line and specify the direction of the lines to match with provided terms. In Figure 60, an instance taken from the CPMP series, students were required to carry out the steps such as make a scatterplot, draw the line, and make squares without much decision-making.

- 26. BICYCLISTS** The table below shows the number of adult residents R (in millions) in the United States who rode a bicycle during the months of October 2001 through September 2002. The time t is measured in months, with $t = 1$ representing October 2001. Use a graphing calculator to write a sinusoidal model that gives R as a function of t . $R = 8.94 \sin(0.520t + 2.54) + 33.3$

t	1	2	3	4	5	6	7	8	9	10	11	12
R	35	30	24	24	26	29	35	34	39	43	44	37

Figure 58. Low-level mathematical complexity instance. From *Holt McDougal Larson Algebra 2 – Teacher’s Edition*, by R. Larson et al., 2012c, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.650. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

1. Give two other names for the regression line.

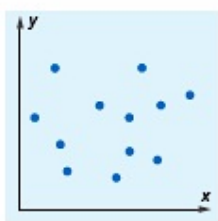
1. least squares line,
line of best fit

In 2-5, match the scatterplot with the best description.

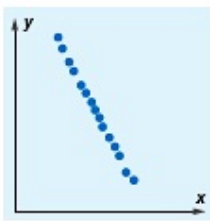
A strong negative correlation
C strong positive correlation

B weak negative correlation
D almost no correlation

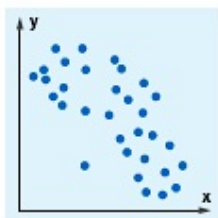
2.



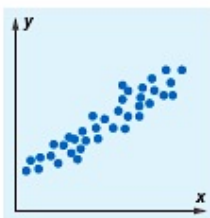
3.



4.



5.



2. D

3. A

4. B

5. C

Figure 59. Low-level mathematical complexity instance. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher's Edition* by J. W. McConnell et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p.98. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

10 Make a scatterplot of the points (1, 1), (2, 2), and (3, 5). Plot the regression line $y = 2x - \frac{4}{3}$. Draw line segments on your graph to show the residuals for each point. Illustrate the geometry of the term *squared residuals* by drawing on the graph an appropriate square for each residual.

11 Consider the set of points (1, 3), (2, 2), (3, 5), and (6, 5).

- Using the equation $y = x + 0.75$, find the sum of the residuals and the sum of the squared residuals.
- Using the equation $y = 0.5x + 2.25$, find the sum of the residuals and the sum of the squared residuals.
- One of the two equations is the regression equation. Tell which one it is and how you know.

Figure 60. Low-level mathematical complexity instances. From *Core-Plus Mathematics: Course 2 – Teacher's Guide* by C. R. Hirsch et al., 2008b, Columbus: OH: Glencoe/McGraw-Hill, p.313. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

The instances at the moderate level of mathematical complexity generally asked students to carry out multiple steps with some level of data-based decision-making.

Figures 61, 62, and 63, taken from the HML, UCSMP, and CPMP respectively, show

examples of such instances. All three instances provided a data set in a table and asked students to write a linear model for the set of data and use it for prediction. To accomplish these tasks, students had to make a scatterplot, specify a line that closely fit the data, choose two points and write an equation for the line passing through two points, and then use the equation to substitute for prediction.

EXAMPLES 3 and 4
for Exs. 17–18

17b. Sample
answer: $y = -2.2x + 111$

17. EARTH SCIENCE The mesosphere is a layer of atmosphere that lies from about 50 kilometers above Earth's surface to about 90 kilometers above Earth's surface. The diagram shows the temperature at certain altitudes in the mesosphere.

- Make a scatter plot of the data. *See margin.*
- Write an equation that models the temperature (in degrees Celsius) as a function of the altitude (in kilometers) above 50 kilometers.
- At about what rate does the temperature change with increasing altitude in the mesosphere? *Sample answer: -2.2 degrees per kilometer*

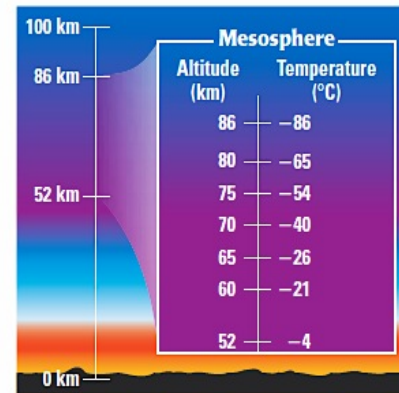


Figure 61. Moderate-level mathematical complexity instance. From *Holt McDougal Larson Algebra 1 – Teacher's Edition*, by R. Larson et al., 2012, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.269. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

In 2 and 3, use the table at the right of life expectancies for people in the United States.

- Construct a scatterplot of the ordered pairs (year, female life expectancy).
 - Eyeball a line to fit the data and find its equation.
 - What female life expectancies does your equation predict for the years 1930–2000?
 - Calculate the sum of the squares of the deviations of the predicted values in Part c from the actual values.
 - Use your equation to predict the female life expectancy in the U.S. in the year 2020.

Year	All	Females	Males
1930	59.7	61.6	58.1
1940	62.9	65.2	60.8
1950	68.2	71.1	65.6
1960	69.7	73.1	66.6
1970	70.8	74.7	67.1
1980	73.7	77.4	70.0
1990	75.4	78.8	71.8
2000	77.0	79.7	74.3

Source: U.S. Census Bureau

- Follow the directions for Question 2, but use linear regression.

Figure 62. Moderate-level mathematical complexity instance. From *The University of Chicago School Mathematics Project: Algebra – Teacher's Edition* by S. S. Brown et al., 2008, Chicago: IL: Wright Group/McGraw-Hill, p.371. Copyright 2008 by Wright Group/McGraw-Hill. Reprinted with permission.

- 2 To help in estimating the number of customers for an amusement park bungee jump, the operators hired a market research group to visit several similar parks that had bungee jumps. They recorded the number of customers on a weekend day. Since the parks charged different prices for their jumps, the collected data looked like this:

Price per Jump (in dollars)	15	20	25	28	30
Number of Customers	25	22	18	15	14

- In this situation, which variable makes sense as the independent variable and which as the dependent variable?
- Plot these data on a coordinate graph.
- Does it make sense to connect the points on your data plot? Explain your reasoning.
- Use the pattern in the table or the graph to estimate the *number of customers* if the *price per jump* is:
 - \$18
 - \$23
 - \$35
- Describe the overall pattern of change relating *price per jump* to *number of customers*.
- The market research staff suggested that the pattern could be summarized with a rule $n = 35 - 0.7p$. Does that rule produce estimates of number of customers n at various prices p like those in the survey data?

Figure 63. Moderate-level mathematical complexity instance. From *Core-Plus Mathematics: Course 1 – Teacher’s Guide* by C. R. Hirsch et al., 2008a, Columbus: OH: Glencoe/McGraw-Hill, pp. 14-15. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

Instances at the high level of mathematical complexity level require heavy cognitive demands on students, including more creative thought, analysis, and planning. For example, in Figure 64, taken from the HML series, students consider real-life constraints such as ethical issues when planning the study and also need to be aware of the properties of well-designed studies to accomplish the tasks. In Project 5 from the UCSMP series (Figure 65), students are required to complete the entire statistical process: formulate a question, collect data, analyze data, and interpret results. They had

to do careful planning and work through multiple decision points when doing the task. In the instance from the CPMP series (Figure 66), students were required to come up with a new approach to solve the problem: find a numerical statistics to represent the correlation and then check their method(s) using data at hand.

Explain whether the research topic is best addressed through an *experiment* or an *observational study*. Then explain how you would design the experiment or the observational study. 13–16. See margin.

13. You want to know if homes that are close to parks or schools have higher property values than other homes.
14. You want to know if flowers sprayed twice a day with a mist of water stay fresh longer than flowers that are not sprayed.

Early Birds Make Better Drivers

A recent study shows that adults who rise before 6:30 A.M. are better drivers than other adults. The study monitored the driving records of 140 volunteers who always wake up before 6:30 and 140 volunteers who never wake up be-

Figure 64. High-level mathematical complexity instance. From *Holt McDougal Larson Algebra 2 – Teacher’s Edition*, by R. Larson et al., 2012b, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.419. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

4 Graphing and Interpreting Statistical Data

Obtain a copy of a recent edition of either the *Statistical Abstract of the United States* or an almanac. Pick data that you find interesting or surprising (or both) that are presented in a table. Design a poster, at least 22" x 28", that interprets the data in the table, including displays to support your interpretation. Make sure that you choose a suitable headline for your poster that will attract people's attention.

5 Class Survey

Compile a database of information about the members of your mathematics class.

- a. Ask each person to complete an information sheet for the following data. Where appropriate, measure in centimeters.
 - gender
 - height
 - hand span
 - foot length
 - time spent getting ready for school each day
 - travel time to school each morning
 - time at work each day
 - time spent on homework each day
 - number of people living at home
 - language most often spoken
 - time spent on the Internet each day
 - usual mode of travel to school (school bus, walk, skateboard, bike, car, and so on)
 - time spent watching TV each day
 - time spent exercising each day
- b. Construct a computer or calculator file with the class database. The database will be used again in later chapters.
- c. Choose at least three variables. Use a statistics utility to display and summarize the data. For each variable decide which type of display is most appropriate (box plot, stemplot, bar graph, and so on). Whenever appropriate, calculate statistics such as the mean, median, standard deviation, percentiles, and range.
- d. Write a short paragraph describing a "typical" student in your class in terms of the variables you analyzed. For numerical variables, this will involve interpreting both the center and spread of the distributions.
- e. Find a variable whose value differs quite a bit by gender. Find a variable whose value doesn't differ much by gender. Justify your conclusions with statistical values or displays.

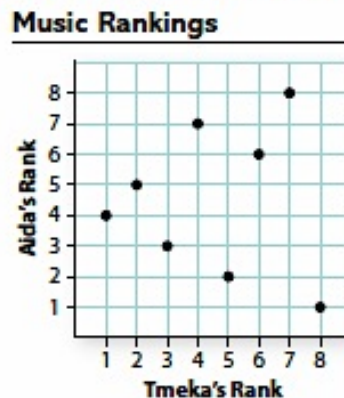
6 The Disputed Federalist Papers

Stylistic words other than those discussed in Lesson 1-8 may be used to confirm or deny the results found by Mosteller and Wallace. In a manner comparable to that carried out in Lesson 1-8, choose a stylistic word such as "of," "in," "because," "that," "and," "but," and so on, and count its occurrences in the Federalist Papers written by Hamilton and Madison, as well as the disputed papers. Confirm or deny Mosteller and Wallace's results using these data. The documents may all be found online,

Figure 65. High-level mathematical complexity instance. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher's Edition* by J. W. McConnell et al, 2010, Chicago: IL: Wright Group/McGraw-Hill, p.68. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

- 2 By looking at the scatterplots, it is fairly easy to make a decision about the *direction* (positive, negative, or none) of the association of two variables. But, as is often the case, it is helpful to have a numerical measure to aid your visual perception of the *strength* of the association.
- Use the music rankings to brainstorm about ways to assign a number to each scatterplot that indicates the direction and the strength of the association.
 - Use the method you prefer to assign a number to three of the scatterplots. Do the results make sense? If not, revise your method.
 - Describe your method to the rest of the class.
 - Where everyone can see it, post a list of the methods your class has suggested.
 - One way to test a method is to use it on extreme cases. What number does each method give for a **perfect positive association**—two rankings that are identical? For a **perfect negative association**—two rankings that are opposite of each other? As a class, decide on a method that seems to make the most sense.
 - Use the following pairs of ranks prepared by Tmeka and Aida to test the method your class selected.

(1, 4), (2, 5), (3, 3), (4, 7), (5, 2), (6, 6), (7, 8), (8, 1)



- Is there an obvious positive or negative association between
- Use your class' method from Part b to assign a number to this association.
- How does your number indicate the direction and strength of this association?

Figure 66. High-level mathematical complexity instance. From *Core-Plus Mathematics: Course 2 – Teacher's Guide* by C. R. Hirsch et al., 2008b, Columbus: OH: Glencoe/McGraw-Hill, p. 260-261. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

The Guidelines for Assessment and Instruction in Statistics Education (GAISE)

Framework Coding

In this section, I describe how the components of doing statistics are addressed in the three textbook series. Several examples are shown to illustrate the framework. These examples supplement the aforementioned description of progression of the GAISE Framework.

Four components of doing statistics.

Table 16 summarizes the percentages of instances that address zero to four components of doing statistics for each of the three series. Across the three series, very few instances address zero or four components of doing statistics (less than 8% for all three series). In contrast, about 35% or more of the instances address two components.

Table 16

Percentages of instances address no to four components of doing statistics

Number of components	HML (%)	UCSMP (%)	CPMP (%)
0	4.2	1.6	1.2
1	17.8	36.9	13.8
2	50.0	38.5	41.9
3	20.1	18.9	35.4
4	7.9	4.1	7.7

Table 17 shows the percentage of instances related to bivariate data that addressed each component of doing statistics for each series. Generally across the three series, the lowest percentage (4.9% to 12.1%) of the instances addressed the *Collect Data* (CD) component. The second lowest percentage of instances addressed *Formulate Questions* (FQ) (27.1% for UCSMP to 45.5% for CPMP). On average, 91.6% of all instances

related to bivariate data addressed *Analyze Data* (AD). The majority of the instances across the three series (63.1% to 87.7%) addressed *Interpret Results* (IR).

Table 17

Percentages of instances addressing the components of doing statistics in each of the three series

	HML (%)	UCSMP (%)	CPMP (%)
Formulate questions	30.3	27.1	45.5
Collect Data	4.9	12.1	11.9
Analyze Data	88.5	95.3	91
Interpret Results	63.1	75.2	87.7

For the *Formulate Question* component, in all the instances, students were not required to generate their own statistical questions for a given problem context; in other words, the statistical question was always provided for them. As a result, all such instances were coded at developmental Level A of the GAISE Framework. Most of the instances provided questions.

In some cases, a statistical question was stated, for example: “How is the stretch of bungee cord related to the weight of the bungee jumper?” (Hirsch et al., 2008a, p.4). There is a need to collect data to answer this question. In other cases, a statistical question was not present, but the description provided a context for a statistical investigation: “In the Activity below, exponential regression models how high a ball bounces” (Brown et al., 2008, p. 419). After this statement, students engage in the entire process of doing statistics. In other cases, data were available to students, but the statement or the question provided a context for statistical investigation. For example, “Use the data in Applications Task 2 to study the relationship between price per bungee jump and income from one day’s operation at the five parks that were visited.” (Hirsch et al., 2008a, p. 15).

More often in the textbooks, questions were available, but they were questions leading to data analysis; that is, a graphical display or numerical statistics were already present. For example, Figure 67 shows an example of two questions that were asked but were not coded as *Formulate (statistical) Questions*.

19. ★ **SHORT RESPONSE** A field biologist collected and measured alligator snapping turtle eggs. The graph shows the mass m (in grams) of an egg as a function of its length ℓ (in millimeters).

a. **Describe** As the lengths of the eggs increase, what happens to the masses of the eggs? **increases**

b. **Estimate** Is 27.5 g a reasonable estimate for the mass of an egg that is 38 mm long? **Explain.**

Yes; 27.5 grams is between the mass of an egg that is just under 38 millimeters long and an egg that is just over 38 millimeters long.

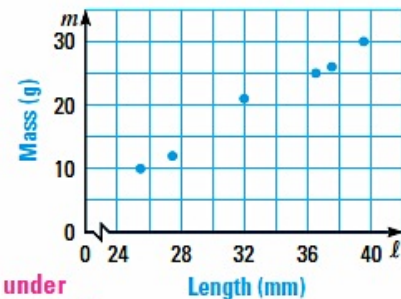


Figure 67. Questions for data analysis. From *Holt McDougal Larson Algebra I – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.54. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

Situations that provided students opportunities to *Collect Data* did not include “*design for difference*” (Franklin et al., 2007, p. 14) meaning that students need to plan to collect data representing different groups. Instead students were required to conduct a *simple experiment* or *compile a census of classroom*, which were coded as Level A. Generally, the *Collect Data* instances dealt with an in-class experiment (Figure 68), doing a census survey of their classmates, or an extended project outside the classroom (in the UCSMP series). In only a few situations (one in the UCSMP series and three in the CPMP series) were students required to *begin awareness of design for differences* and use random selection, which were coded at Level B of the GAISE Framework. The following example, taken from the project, taken from UCSMP *Algebra* book, illustrates an instance coded at Level B.

Testing Astrology

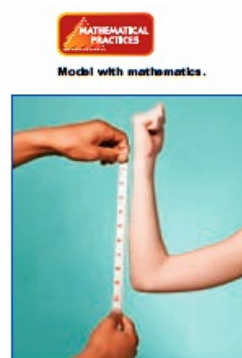
From a book (like *Who's Who*) or online source, find at least 100 famous people in a field and note their birthdays. Identify the astrological sign of each person. Then tabulate the number of people with each sign. Do these data lead you to believe that certain birth signs are more likely to produce famous people? Use a chi-square statistics assuming a random distribution of the birthdays among the 12 astrological signs is expected. Although $n - 1 = 11$ here, refer to row 10 from the chi-square table on page 699. (Brown et al., 2008, p. 704)

Scatter Plots and Functions

MATERIALS • tape measure • graph paper

QUESTION How can you tell whether a graph represents a function?

A *scatter plot* is a type of display for paired data. Each data pair is plotted as a point. In this activity, you will work in a group to make a scatter plot. You will measure the height of each student in your group and the length of his or her forearm. The length of the forearm is the distance from the elbow to the wrist.

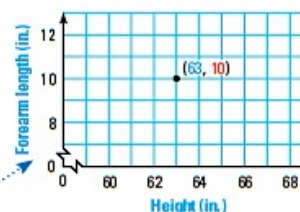



EXPLORE Collect data and make a scatter plot

STEP 1 *Collect data* Measure the height of each student in your group and the length of his or her forearm. Record the results for each student in one row of a table like the one shown.

Height (inches)	Forearm length (inches)
63	10
?	?

STEP 2 *Make a scatter plot* Use graph paper to draw axes labeled as shown. Then plot the data pairs (*height*, *forearm length*). For example, plot the point (63, 10) for a student with a height of 63 inches and a forearm length of 10 inches.



The symbol  on an axis represents a break in the axis.

DRAW CONCLUSIONS Use your observations to complete these exercises

- Examine your scatter plot. What does it suggest about the relationship between a person's height and the person's forearm length?
Sample answer: The greater the person's height, the greater the length of the forearm.
- Compare your table with those of the other groups in your class. Determine which of the tables represent functions and which do not.
Answers will vary. Tables that are functions should have only one forearm length for each height.
- Is it possible to determine whether a table represents a function by looking at the corresponding scatter plot? *Explain.*
Yes; if there is more than one y for an x it is not a function.

Figure 68. A full statistical investigation process. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson et al., 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.48. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

No instance requiring Level C of the GAISE Framework for *Collect Data* was found in any of the three series.

Developmental levels of the GAISE Framework.

For the *Analyze Data* and *Interpret Results*, Figures 69 and 70 show the percentages of the GAISE developmental levels that appeared in the instances found in the three series. Generally, most instances were at Level B of the GAISE Framework for both *Analyze Data* (ranging from 45.9 to 73.1%) and *Interpret Results* (ranging from 61 to 78.9%). On average, the least percentage of instances was coded at Level A across the three series for *Analyze Data*. In contrast, for *Interpret Results*, the lowest percentage of instances was at Level C.

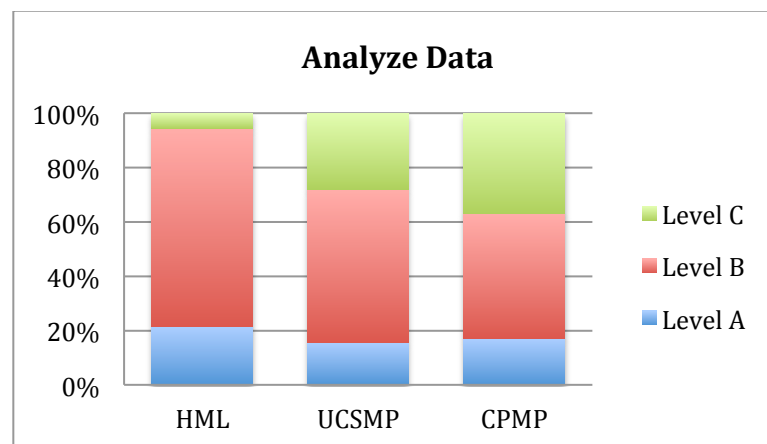


Figure 69. Percentages of *Analyze Data* instances across the three series.

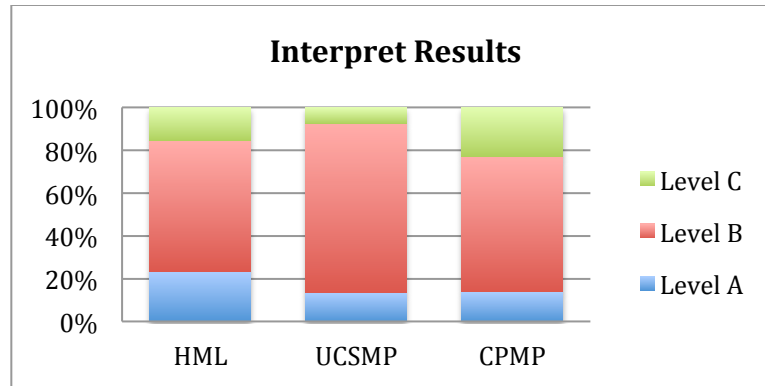


Figure 70. Percentages of *Interpret Results* across the three series.

In order to compare distributions of the statistical components, an approach similar to assigning numerical values to the Levels of Mathematical Complexity was applied to the developmental levels of the GAISE Framework (Level A = 1, Level B = 2, Level C = 3). Few instances were coded for *Collect Data* and *Formulate Question* and most were at Level A of the GAISE Framework; therefore, I only compared the distributions of *Analyze Data* and *Interpret Results* for the three series. Tables 18 and 19 show the number of instances that were coded for each statistical component, the means and standard deviations of the codes across the three series. As observed from the tables, on average, the CPMP series has the highest value, followed by the UCSMP series, and then the HML series for both components: *Analyze Data* and *Interpret Results*.

Table 18

Means (SDs) of Analyze Data across the three series

Series	Number of bivariate data instances	Mean of GAISE Developmental Levels (S.D)
HML	108	1.84 (.496)
UCSMP	204	2.12 (.651)
CPMP	222	2.2 (.710)

Table 19

Means (SDs) of Interpret Results across the three series

Series	Number of instances related to bivariate data	Mean of GAISE Developmental Levels (S.D)
HML	77	1.92 (.623)
UCSMP	161	1.94 (.457)
CPMP	214	2.09 (.602)

It should be noted that *Analyze Data* and *Interpret Data* codes were not always the same for a specific instance. For example, Figure 68 shows an example of an instance that was coded at GAISE Level A for both *Analyze Data* and *Interpret Results* (observe association between two variables). In another case, Figure 58 shows an instance that was coded at Level C for *Analyze Data* (use trigonometric regression in analyzing data) but was not coded for *Interpret Results* because there was no requirement for students to connect back to the context. Similarly, Figure 64 shows an example of an instance that was coded at GAISE Level C for *Interpret Results* (Understand the difference between observational studies and experiments), but was not coded for *Analyze Data* because no data was involved in this instance.

One question of interest naturally arises: Are the two coding components, *Analyze Data* and *Interpret Results*, associated? To explore this question, I applied the Goodman and Kruskal's Gamma test to determine whether there is a significant association among ordinal categorical variables. Even though several instances were coded as high for *Analyze Data* and a lower code for *Interpret Result* and vice versa, Table 20 shows that the two codes are statistically accordant, $\gamma = .993$, $p = .000$, $N = 421$ meaning that

increasing codes in *Analyze Data* were associated with increasing codes in *Interpret Results*.

Table 20

Association between Analyze Data and Interpret Results

		Interpret Results (IR)			Total
		1	2	3	
Analyze Data (AD)	Count	65	5	0	70
	1 % within AD	92.9%	7.1%	0.0%	100.0%
	% within IR	92.9%	1.7%	0.0%	16.6%
	Count	5	236	2	243
	2 % within AD	2.1%	97.1%	0.8%	100.0%
	% within IR	7.1%	80.8%	3.4%	57.7%
	Count	0	51	57	108
	3 % within AD	0.0%	47.2%	52.8%	100.0%
	% within IR	0.0%	17.5%	96.6%	25.7%
	Count	70	292	59	421
	% within AD	16.6%	69.4%	14.0%	100.0%
	Total % within IR	100.0%	100.0%	100.0%	100.0%

Association Between Levels of Mathematical Complexity and the GAISE

Frameworks

Two frameworks, the NAEP levels of mathematical complexity and the GAISE Framework, were used to code the same set of instances. Because both frameworks essentially address levels of sophistication of tasks, I explored whether there is a concordance of the sets of two codes. That is, does an instance of getting a high code in one framework tend to get a high code in the other framework? Because only a few instances address *Formulate Questions* and *Collect Data*, and the lowest developmental levels of the components were addressed, I sought to examine the association between the coding of two other components of the GAISE Framework, *Analyze Data* and *Interpret Results*, and the mathematical complexity levels.

In Figures 71 and 72, a large portion of instances were coded at the middle level for both frameworks, no instance was coded at the highest level for mathematical complexity and at the lowest level for *Analyze Data*; few instances received a low level code for *Analyze Data* and a high level code for mathematical complexity. Similarly, few instances received a high level code for Interpret Results and a low level code for mathematical complexity and vice versa.

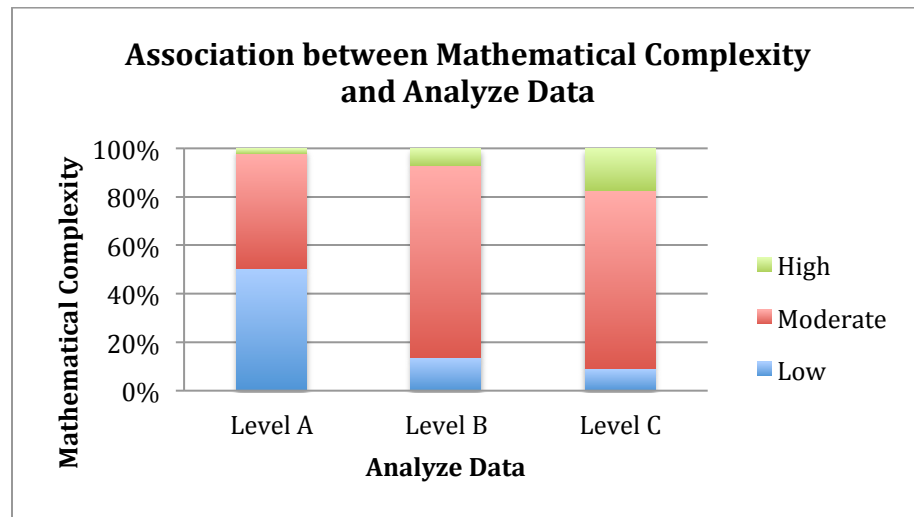


Figure 71. Association between mathematical complexity and *Analyze Data*.

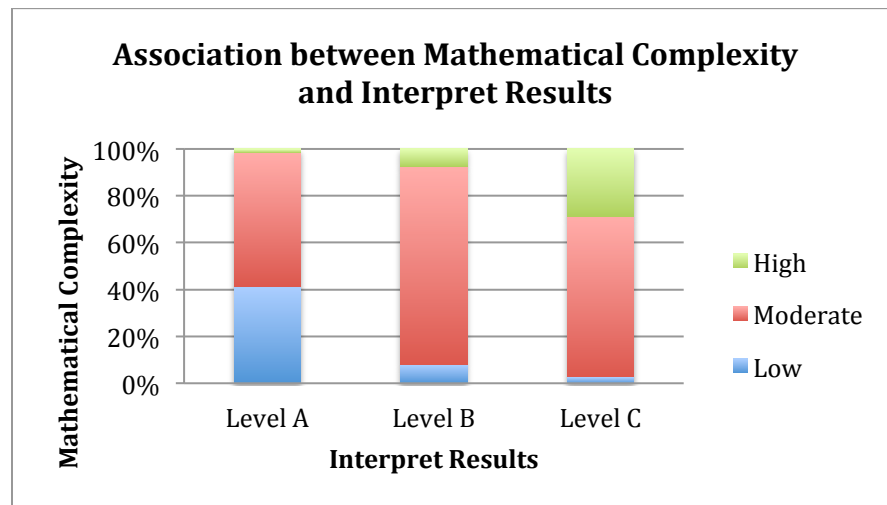


Figure 72. Association between mathematical complexity and *Interpret Results*.

As a formal approach, I used the Goodman and Kruskal 's Gamma to examine the association, and results showed that there was positive accordance between levels of mathematical complexity and both components of *Analyze Data and Interpret Results* of the GAISE Framework, $\gamma = .566, p = .000, N = 534$ and $\gamma = .721, p = .000, N = 452$

Table 21

Association between mathematical complexity and Analyze Data

			Analyze Data (AD)			Total
			1	2	3	
Mathematical Complexity (MC)	Low	Count	47	40	13	100
		% within MC	47.0%	40.0%	13.0%	100.0%
		% within AD	50.5%	13.5%	9.0%	18.7%
	Moderate	Count	44	235	107	386
		% within MC	11.4%	60.9%	27.7%	100.0%
		% within AD	47.3%	79.4%	73.8%	72.3%
	High	Count	2	21	25	48
		% within MC	4.2%	43.8%	52.1%	100.0%
		% within AD	2.2%	7.1%	17.2%	9.0%
Total		Count	93	296	145	534
		% within MC	17.4%	55.4%	27.2%	100.0%
		% within AD	100.0%	100.0%	100.0%	100.0%

Table 22

Association between mathematical complexity and Interpret Results

			Interpret Results (IR)			Total
			1	2	3	
Mathematical Complexity (MC)	Low	Count	29	25	2	56
		% within MC	51.8%	44.6%	3.6%	100.0%
		% within IR	41.4%	8.1%	2.7%	12.4%
	Moderate	Count	40	261	50	351
		% within MC	11.4%	74.4%	14.2%	100.0%
		% within IR	57.1%	84.5%	68.5%	77.7%
	High	Count	1	23	21	45
		% within MC	2.2%	51.1%	46.7%	100.0%
		% within IR	1.4%	7.4%	28.8%	10.0%
Total		Count	70	309	73	452
		% within MC	15.5%	68.4%	16.2%	100.0%
		% within IR	100.0%	100.0%	100.0%	100.0%

Purpose and Utility Framework Coding

In this section, I highlight the task features in relation to the Purpose and Utility Framework (Ainley et al. 2006). In addition to highlighting the *purpose* and *utility* of

learning tasks, I provide examples of instances addressing the features. Figures 73 and 74 show the percentages of instances that address each feature of the Purpose and Utility Framework. As illustrated in the figures, across the three series, half of the instances in the HML series and more than 60% of the instances in the UCSMP and CPMP series addressed the *purpose* feature of the framework. That is, when solving the problems presented in these instances, students see a virtual or actual product after accomplishing the tasks. More than 95% of instances related to bivariate data address the *utility* feature (for all series). This means that most of the instances provided a context for students to see when, how, and why the mathematical ideas were used.

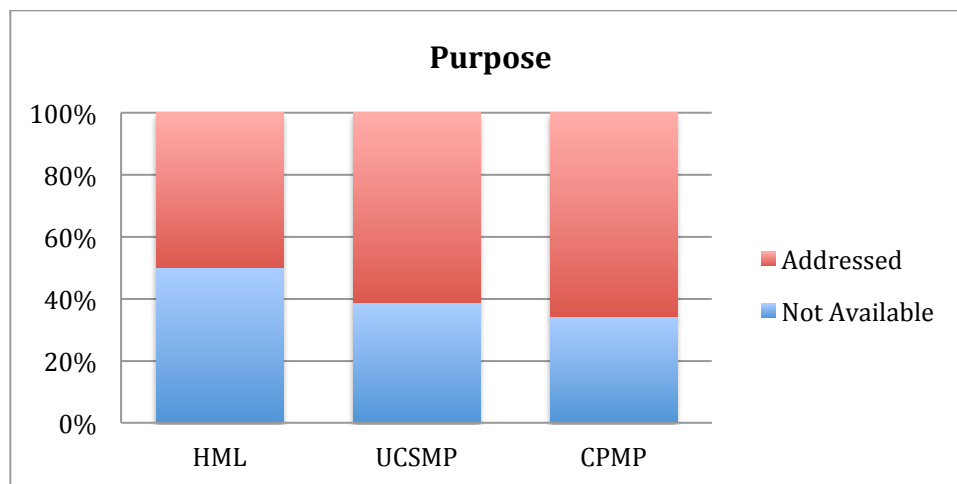


Figure 73. Percentages of instances that address the *purpose* feature across the three series.

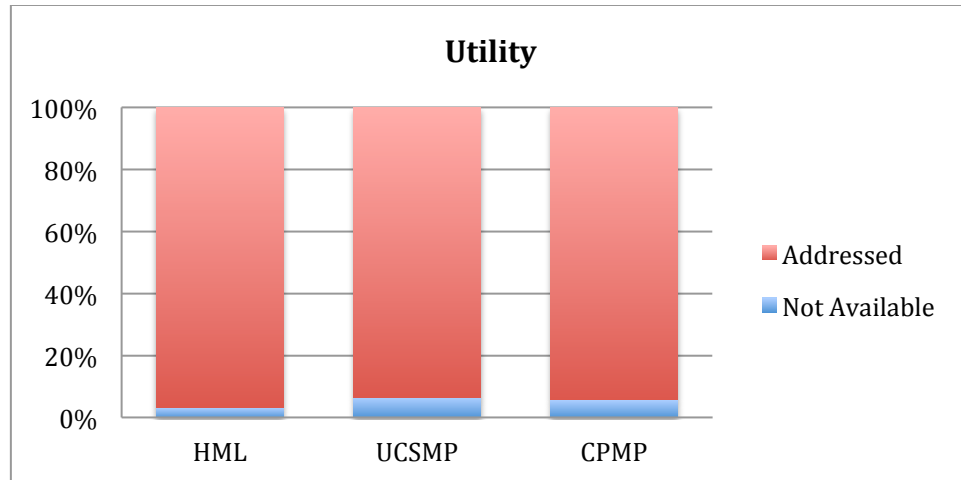


Figure 74. Percentages of instances that address the *utility* feature across the three series.

Tables 23, 24, 25 show the number of instances that address zero, one, and two features. Across the three series, most instances addressed both features (ranging from 50% to 65.3%). In contrast, very rare instances (only one across three series) addressed *purpose* but not *utility* feature. A smaller amount of instances did not address any feature (ranging from 3.3% to 6.5%). Finally, about 29% (for CPMP) to 46.7% (for HML) addressed *utility* but not *purpose* feature.

Table 23

Number of instances address across purpose and utility features in HML (N=122)

Purpose	Utility		
		No	Yes
	No	4	57
	Yes	0	61

Table 24

Number of instances address across purpose and utility features in UCSMP (N=216)

Purpose	Utility		
		No	Yes
	No	14	69
	Yes	0	131

Table 25

Number of instances address across purpose and utility features in CPMP (N=246)

Purpose	Utility		
		No	Yes
	No	14	71
	Yes	1	160

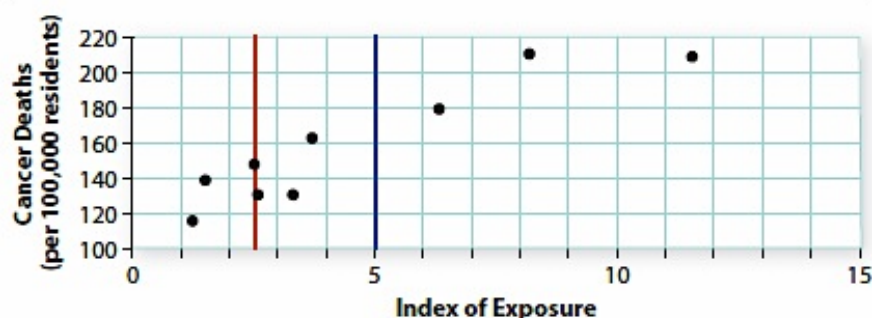
Following are examples of instances that (a) address *purpose* and *utility* features, (b) address *purpose* but not *utility* feature, (c) address *utility* but not *purpose* feature, and (d) do not address either feature.

Figure 50 shows an example of an instance that addresses both *purpose* and *utility* features. That is, when solving this task, students could see when, how, and why scatterplots, linear fitting are used in daily life (*utility* of mathematics processes) and when doing the prediction, they see the *purpose* of solving this problem: make a model for prediction.

Only one instance across the three series was categorized as addressing *purpose* but *not utility* features. Generally, if *purpose* was addressed, *utility* was addressed also. Figure 75 shows an example of such instance. In this instance, after solving this task, students know “why the regression line is sometimes called the ‘line of averages’” (Hirsch et al., 2008b, p. 318), but they do not see why, when and how mathematical ideas such as linear regression are used.

- 24 In this task, you will explore why the regression line is sometimes called the “line of averages.” Shown below is a plot of the data from the Check Your Understanding on page 285 of the exposure to radioactive waste from a nuclear reactor in Hanford, Washington, and the rate of deaths due to cancer in these communities. The vertical lines divide the set of data points into thirds.

Radioactive Waste Exposure



- Compute the centroid of the points in each of the three vertical strips using the data given on page 285 and plot them on a copy of the plot.
- Draw in the regression line.
- How close are the three centroids to the regression line? Could you have done a good job of approximating the regression line just by drawing in a line that goes almost through the three centroids?

Figure 75. An instance that addresses the *purpose* but not *utility* feature. From *Core-Plus Mathematics: Course 2 – Teacher’s Guide* by C. R. Hirsch et al., 2008b, Columbus, OH: Glencoe/McGraw-Hill, p.318. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

The example “Collect data on the number of cases reported for some disease in recent years for your city and state. Which type of function (linear, exponential, power, or logarithm) best models the data over time” (McConnell, 2010d, p. 606) shows an instance that addresses the *utility* but not the *purpose* feature. When doing this task, students see how scatterplots and functional models are used. However, they did not get to see the purpose of solving the problem because accomplishing the task did not yield such a product.

Figure 76 shows an example of an instance in which neither the *purpose* nor *utility* feature was addressed. Students were asked to make a scatterplot, draw a line of fit, and write an equation of the line. They were not given an opportunity to understand *why* and *when* to use the mathematical concepts and no product resulted from accomplishing the task.

Make a scatter plot of the data. Draw a line of fit. Write an equation of the line.

15.

x	0	1	2	3	4
y	15	35	53	74	94

16.

x	0	2	4	8	10
y	-2	6	15	38	50

15, 16. See margin.

Figure 76. An instance that addressed neither the *purpose* nor the *utility* feature. From *Holt McDougal Larson Algebra 1 – Teacher’s Edition*, by R. Larson, L. Boswell, T. D. Kanold, & L. Stiff, 2012a, Orlando, FL: Houghton Mifflin Harcourt Publishing Company, p.291. Copyright 2012 by Houghton Mifflin Harcourt Publishing Company. Reprinted with permission.

The Common Core State Standards-Mathematics: Mathematical Practices Coding

Developing students’ awareness and utilization of the eight mathematics practices provides a set of skills that are important to develop mathematics proficiency. The Mathematical Practice Standards are:

MP1: Make sense of problems and persevere in solving them.

MP2: Reason abstractly and quantitatively.

MP3: Construct viable arguments and critique the reasoning of others.

MP4: Model with mathematics.

MP5: Use appropriate tools strategically.

MP6: Attend to precision.

MP7: Look for and make use of structure.

MP8: Look for and express regularity in repeated reasoning.

Table 26 summarizes the percentages of instances that address zero to eight mathematical practices for each series. Generally, across three series, no instances addressed all eight practices; rarely did instances address seven practices (ranging from 0% to 6.5%) and none of the practices (ranging from 0.4% to 4.0%). In contrast, about 11.4% (for CPMP) to 19.6% (for HML) of the instances addressed three practices, about 20% (for CPMP) to 37.9% (for UCSMP) addressed four practices, and about 22.4% (for UCSMP) to 26.1% (for CPMP) addressed five practices.

Table 26

Percentages of instances address zero to eight MPs for each series

	Zero MPs	One MP	Two MPs	Three MPs	Four MPs	Five MPs	Six MPs	Seven MPs	Eight MPs
HML	4.0	13.9	13.1	19.6	24.6	23.0	1.6	0.0	0.0
UCSMP	3.7	6.5	4.7	14.0	37.9	22.4	10.3	0.5	0.0
CPMP	0.4	2.0	8.2	11.4	20.0	26.1	25.3	6.5	0.0*

*The percentages do not sum to 100 due to rounding

Figure 77 summarizes the percentages of instances that address the CCSSM mathematical practices across the three series. Generally, most instances related to bivariate data addressed MP4 *Model with mathematics* (ranging from 76.3% to 89%) and MP6 *Attend to precision* (81.1% to 87.8%). These instances often involved data and building a mathematical model to understand the situations or to solve real-life problems (addressing MP4). Furthermore, they often involved measuring, calculating, and making scaled graphing, which were coded as MP6. More than half of the instances offer potential to apply MP2 *Reason abstractly and quantitatively* (ranging from 57.4% to 82.4%) because most of those instances include data and students needed an understanding of the quantity and to be able to find the relationship between two quantities. Only a few instances addressed MP1 *Make sense of problems and persevere in*

solving them (3.3% to 10.6%) and MP8 *Look for and express regularity in repeated reasoning* (0% to 7.8%); in fact, no instance was found to address MP8 in the HML series.

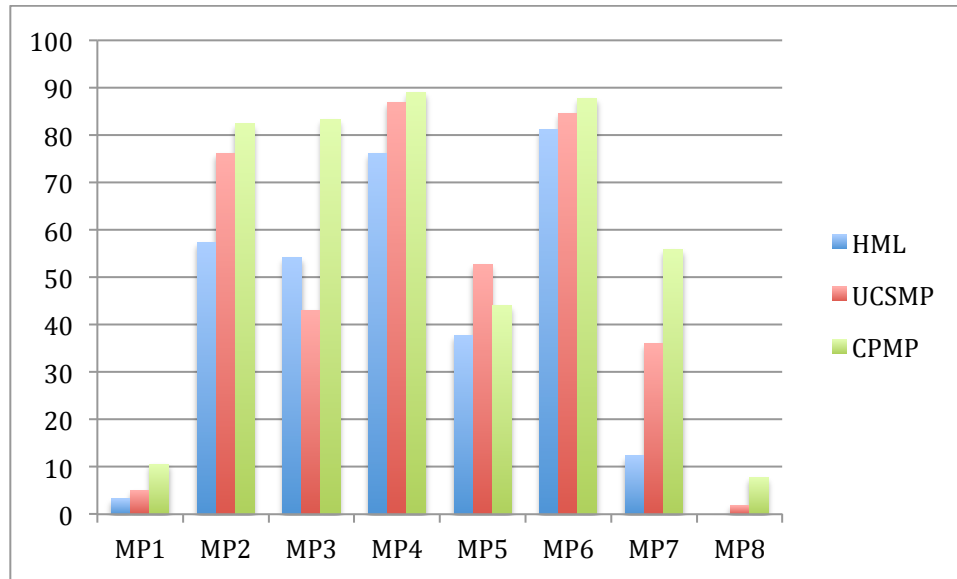


Figure 77. Percentages of instances that address the SMP in each of the three series.

By series, generally, the percentages were highest for the CPMP series for all practices except MP5 *Use appropriate tools strategically*. In contrast, the percentages of opportunities for students to apply the mathematical practices were lowest for the HML series with the exception of MP3 *Construct viable arguments and critique the reasoning of others*. The HML series offered very few opportunities for students to access MP7 *Look for and make use of structure* (12.3%) and a fair portion of instances were found to address MP5 (37.7%). Less than half of the instances in the UCSMP series addressed MP3 (43%) or MP7 (36%). The CPMP consisted of a large portion of instances that addressed all mathematical practices except MP1 (10.6%) and MP8 (7.8%).

Examples of the practices are shown below. Figure 78 shows an example of instance that addressed MP1 *Make sense of problems and persevere in solving them*.

Students use different approaches when solving the problem: predict and compare with data at hand as a way of triangulation. Further, this instance addressed other mathematical practices, including MP2 *Reason abstractly and quantitatively* – students need to examine the relationship between quantities; MP3 *Construct viable arguments and critique the reasoning of others* – students need to “be prepared to explain [their] choice”; MP4 *Model with mathematics* – students are building a mathematical model for the situation; MP6 *Attend to precision* – students need to be precise when collecting data and making a scatterplot, and MP7 *Look for and make use of structure* – students need to look for a pattern for the set of data.

- 1 Suppose that you were asked to conduct this experiment:
 - Get a collection of 100 coins, shake them well, and drop them on a tabletop.
 - Remove all coins that are lying heads up and record the number of coins left.
 - Repeat the shake-drop-remove-record process until 5 or fewer coins remain.
 - a. If you were to record the results of this experiment in a table of (*drop number*, *coins left*) values, what pattern would you expect in the data? What function rule would probably be the best model relating drop number n to number of coins left c ?
 - b. Conduct the experiment, record the data, and then use your calculator or curve-fitting software to find a function model that seems to fit the data pattern well.
 - c. Compare the model suggested by logical analysis of the experiment to that found by fitting a function to actual data. Decide which you think is the better model of the experiment and be prepared to explain your choice.

Figure 78. An instance that addressed MP1, MP2, MP3, MP4, MP6, and MP7. From *Core-Plus Mathematics: Course 1 – Teacher’s Guide* by C. R. Hirsch et al., 2008a, Columbus: OH: Glencoe/McGraw-Hill, p.330. Copyright 2008 by McGraw-Hill Companies, Inc. Reprinted with permission.

As mentioned in chapter 3 of this study, students can approach most tasks by using paper and pencil or technology. For the reason MP5 *Use appropriate tools strategically* was coded using the guidelines from the teacher’s edition of the three series. Figure 79 shows examples of tasks that address MP5 because the textbook authors noted about the use of technology as a signal for teachers to attend to when implementing the task in classroom.

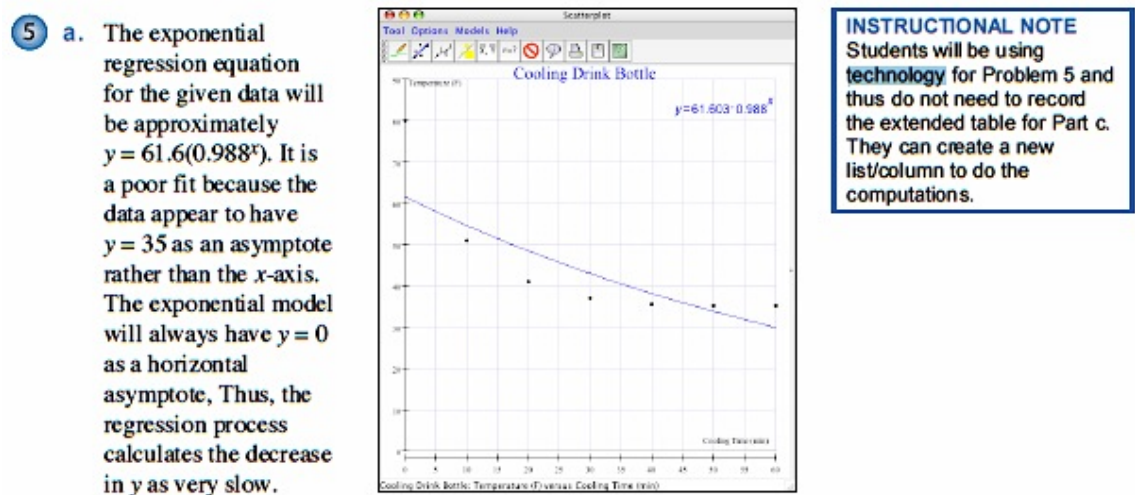


Figure 79. An instance that addressed MP5. From *Core-Plus Mathematics: Course 4 – Teacher’s Guide* by C. R. Hirsch et al., 2010, Columbus: OH: Glencoe/McGraw-Hill, p.T30. Copyright 2010 by McGraw-Hill Companies, Inc. Reprinted with permission.

Most instances (more than 80%) were coded to allow students to apply MP6 *Attend to precision* because generally, they required students to collect data, measure, make a scaled graph, and/or calculate. Still another case, students need to use specialized terms for communication when solving tasks, such tasks were also coded to for MP6 (Figure 66).

Many instances asked students to find a model for a set of data; however, these were coded for MP7 if students were to decide what pattern to look for and use it. Figure 80 shows an example of an instance that asks students to find a line of best fit for

predicting, which was coded to meet MP8 but not MP7 because students were not looking for a pattern, instead they were told to use the linear model for prediction. This instance offered the potential for students to look for a shortcut to realize the invariance of quantities when changing scale.

13. Consider the following data, which give the height h in cm and weight w in kg of twelve students.

h (cm)	144	168	140	157	153	162	160	160	166	166	173	165
w (kg)	68	86	69	77	85	74	78	84	84	83	82	82

- Enter these data into a statistics utility. Create a scatterplot with h on the horizontal axis. **See margin.**
- Find the line of best fit for predicting weight from height.
- Compute the correlation coefficient. **$r = 0.7499$**
- Use a statistics utility to convert the height to inches (1 in. = 2.54 cm) and weight to pounds (1 lb = 0.454 kg). Draw a new scatterplot. How is the scatterplot different from that in Part a?
- Compute the correlation coefficient and the regression equation for predicting weight from height in Part d.
- Which of the following statistics remain invariant under scale changes of the data?
 - correlation coefficient **Invariant**
 - slope of the regression line **not Invariant**
 - y-intercept of the regression line **not Invariant**

Part c: $d(M - m)$
 $= |d|r$

13b. $w = 3.628 + 0.4746h$

13d. The scatterplot appears the same but the scale of the axes is different.

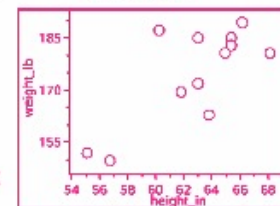


Figure 80. An instance that addressed MP8. From *The University of Chicago School Mathematics Project: Functions, Statistics, and Trigonometry – Teacher’s Edition* by J. W. McConnell et al., 2010, Chicago: IL: Wright Group/McGraw-Hill, p.191. Copyright 2010 by Wright Group/McGraw-Hill. Reprinted with permission.

Summary of Task Features for Bivariate Data

Generally, the distributions of mathematical complexity were similar for the HML and UCSMP series, but they were different from that of the CPMP series. On average, the level of mathematical complexity was highest for the CPMP series. All three series include a majority of instances coded at the moderate level and only few of instances at the high level of complexity.

For the GAISE Framework, few instances address *Formulate Questions* and *Collect Data* components. When the two components were addressed, the instances were

not at the most sophisticated developmental level, Level C. Most instances include *Analyze Data* at all three GAISE developmental levels. However, the distributions of the GAISE developmental levels were different across the three series. Similar for *Interpret Results*, all developmental levels were addressed in all three series. However, the distributions of the developmental levels were different: the average of GAISE developmental levels of the CPMP series was highest, followed by those of UCSMP and then HML series.

Tests of association for ordinal categorical data showed that two components of the GAISE Framework, *Analyze Data* and *Interpret Results* are concordant. Similarly, NAEP mathematical complexity levels and *Analyze Data* are concordant. Finally, mathematical complexity codes and *Interpret Results* are also concordant, meaning one code tends to increase as the other one increases.

The majority of the instances of bivariate data addressed the *utility* feature and a smaller proportion of instances address the *purpose* feature. Although all four combinations of *utility* and *purpose* features (*purpose* and *utility*, *purpose* not *utility*, *utility* not *purpose*, not *purpose* and not *utility*) were found across the three series, only one instance was found to address the *purpose* but not *utility* feature.

Finally, the potential for accessing and applying mathematical practices was evident in all three series. Across three series, most instances address three to five MPs often including MP4 and MP6. Generally, the percentages of instances that addressed the practices were highest for the CPMP series and lowest for the HML series. On average, most instances addressed MP2, MP4, and MP6. In contrast, few instances address MP1 and MP8, while a moderate amount of instances addressed MP3, MP5, and MP7.

CHAPTER 5: DISCUSSION, SUMMARY, AND RECOMMENDATIONS

In this study, I analyzed the learning trajectories (LTs) and task features for bivariate data within three high school textbook series currently used in the U.S. In this chapter, I summarize the study and discuss the findings as well as provide implications for mathematics education researchers and practitioners. This chapter is organized into five sections: (a) Summary of the Study, Findings and Discussion, (b) Limitations, (c) Implications and Recommendations, and (d) Summary.

Summary of the Study, and Findings and Discussions

For more than two decades, statistics has been advocated for and emphasized in the school mathematics curriculum (Franklin et al., 2007; CCSSI, 2010; Jones et al., 2004; NCTM, 1989, 2000). Collectively, the goals of these efforts are to equip students to be wise citizen-consumers of the myriad of data omnipresent in their daily lives (Ben-Zvi & Garfield, 2004; Shaughnessy, 2007). Bivariate relationships play a critical role in statistics and, in fact, many mathematical problems are addressed through applications of bivariate data analyses (Shaughnessy, 2007). In addition, covariational reasoning (i.e., reasoning about bivariate relationship) plays a crucial role in everyday life because it is closely related to the concept of causal relationships, which allows people “to explain the past, control the present and predict the future” (Crocker, 1981, p. 272). However, research has consistently shown that people’s judgments of relationships between events are generally poor (e.g., Crocker, 1981).

Over generations of school reform, there have been numerous efforts to improve student learning but ultimately “a school mathematics curriculum is a strong determinant of what students have the opportunities to learn and what they do learn (NCTM, 2000, p.

14). Among multiple aspects of curriculum development, textbooks are clearly one of the most common tools for the teaching and learning of mathematics (Grouws & Smith, 2000; Robitaille, & Travers, 1992; Tarr et al., 2006). Robitaille and Travers (1992) cite textbooks as a “significant factor in determining students’ opportunity to learn and their achievement” (p. 706).

The National Research Council (2004) called for rigorous approaches when analyzing curricular materials. Among these, learning trajectories are at the heart of every phase of mathematics education: standards, curriculum development, instruction, and assessment (Batista, 2004; Confrey et al., 2009; Daro et al., 2011). Hence, documenting the LTs for bivariate data has the potential to identify strengths and weaknesses of curricular materials that students use to learn this content. In addition, examining the features of tasks provides an essential source of information for curriculum developers to improve textbook materials, and enhance the teaching and learning of mathematics.

Purpose of the Study

Given that textbooks play central roles in teaching and learning mathematics, this study analyzed contemporary high school mathematics textbooks in order to address the following research questions:

1. What learning trajectories for topics in bivariate data are represented in high school mathematics textbooks in the U.S.? How are those learning trajectories similar to and different from those embedded within the *Common Core State Standards for Mathematics* (CCSSM)? To what degree are the learning trajectories found in textbooks aligned with the developmental levels described in

the Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report?

- a. What topics in bivariate data are addressed in the textbooks and what learning trajectories for bivariate data are evident in the textbooks?
 - b. How are connections made between the topics in bivariate data and topics in univariate data in the textbooks? How are connections made between bivariate data and other conceptual categories in mathematics at the high school level?
 - c. To what extent are the CCSSM standards for bivariate data at the high school level evident in the textbook materials, and to what extent does the approach to, and sequence of, content in the materials reflect the developmental progressions of the topics described by the GAISE Framework?
2. With respect to bivariate data, what is the nature of the instructional tasks presented in textbooks? Do the tasks provide opportunities for students to access the CCSSM's Standards for Mathematical Practice (SMP)?
- a. What levels of mathematical complexity are required by the tasks related to bivariate data?
 - b. To what degree are the GAISE developmental levels reflected in the tasks related to bivariate data?
 - c. What is the quality level of the tasks in terms of purpose and utility?
 - d. How do the tasks provide opportunities for students to access the CCSSM's Standards for Mathematical Practice?

Methodology

Three current high school mathematics textbook series—the *Holt McDougal Larson* series (HML), the *University of Chicago School Mathematics Project* (UCSMP), and the *Core-plus Mathematics Project* (CPMP)—comprised the sample for this study. In order to examine the LTs and task features for bivariate data, I analyzed teacher's editions.

I examined the three series from multiple perspectives that are grounded in research literature: presence of combinations of variables (two categorical variables [CC], two numerical variables [NN], and one categorical and one numerical variable [CN]); learning goals; techniques and theories; NAEP mathematical complexity; components and developmental levels of the GAISE Framework; purpose and utility features, and the Standards for Mathematical Practice of the CCSSM. The LTs for bivariate data were described using two aspects: (1) the distributions of bivariate tasks, and (2) qualitative detailed descriptions of how bivariate data content was introduced and developed in each series. I then compared these LTs with the CCSSM learning expectations (LEs) and learning progressions for bivariate data that are advocated in the GAISE Framework. For the task features, I report descriptive statistics and distributions of data. In order to examine the association between components of the GAISE Framework, and the GAISE and task frameworks, I applied Goodman and Kruskal's Gamma tests. Additionally, I provided examples to illustrate each of the key task features in my analysis.

Results of the Study and Discussions

Learning trajectories.

Distributions of bivariate data instances. Results showed that LTs for the three combinations of variables appeared in all three series. By far, the most emphasis was placed on NN combination, comprising more than 80% of bivariate data instances across three series. Compared to all tasks in each series, the proportion of bivariate data tasks was highest for the CPMP series (5.3%), followed by the UCSMP (1.3%), and lowest for the HML series (0.9%). Content related to bivariate data was generally distributed evenly across the four textbooks in the CPMP series, whereas the bivariate content was more clustered in the UCSMP *Algebra*, *Advanced Algebra*, and *Functions, Statistics, and Trigonometry* books and the HML *Algebra 1* and *Algebra 2* books. Given that the attention given to bivariate data in the HML and UCSMP series is rather small, students utilizing these curricula likely will have limited opportunities to learn the content, and this might explain students' generally poor judgments between two variables (e.g., Batanero et al., 1997; Crocker, 1982; Jennings et al., 1982; Smedslund, 1963; Snyder, 1981). On the other hand, research has found that students tend to make better judgments when using graphs rather than tables (Lane et al., 1985); coincidentally, nearly all the instances for NN variables across the three series involved graphs.

Of the instances related to NN combination, only 1 of 131 instances in the HML series, 13 of 244 in the UCSMP series, 29 of 214 in the CPMP series contained a set of 20 or more data points. In addition, when data were provided, the vast majority of the instances were presented in an organized table as opposed to raw data form. In particular, only 2 instances in the HML series, 2 instances in the UCSMP series, and 10 instances in the CPMP series provided data in a raw format and required students to sort, categorize, and place the data in a well-organized table. Consequently, students using these textbooks

do not have to confront the messiness of data analysis, as is recommended by statistics educators (e.g., Ben-Zvi & Garfield, 2004).

Two categorical variables combination of bivariate data. All three textbook series addressed the CC combination in relation to two-way tables and conditional probabilities; however, the manner in which each series developed the content was qualitatively different. Specifically, the HML series began by introducing how to make and read two-way tables but did not use comparing absolute frequencies to examine the association between two variables. The authors then asked students to judge the independence of two events using conditional probabilities. In contrast, the UCSMP authors started with use of the Chi-Square test to judge the dependency of two events and subsequently revisited the test using data simulations. The UCSMP series LT ended by comparing the relative frequencies and using two-way tables to judge associations. In further contrast, the CPMP series began with the independence of two events and introduced conditional probability. As a formal tool to examine the association, CPMP then examined the significance of the difference between proportions (probabilities).

Even though the CC combination was addressed, it appeared very little across the three series (less than 5% of the total bivariate instances for each of the three series). Coincidentally, research has found that students perform worse in judging covariation of categorical variables compared with numerical variables, and covariational reasoning cannot be transferred from one type of combination (e.g., NN combination) to the other combination (e.g., CC combination) (Beach & Scopp, 1966; Erlick & Mills, 1967, Jennings et al., 1982). By placing so little emphasis on the CC combination, textbook

authors might unknowingly inhibit students' learning, and misconceptions they often encounter when solving tasks related to CC variables might not be addressed.

The sequence of content related to judging the association of CC variables in UCSMP was reversed in relation to the GAISE Framework. In particular, the formal tool used at the highest developmental level in the GAISE Framework, was introduced quite early (i.e., in *Algebra*) and treated as a procedure. The early introduction of sophisticated tools without first acquiring more foundational understanding (in lower levels) runs counter to researchers' recommendations that students move through increasingly higher levels of sophisticated statistical reasoning (Cobb, 2000; Franklin et al., 2007).

All three series devoted little attention to quantifying the strength of association (e.g., Agreement and Disagreement Ratio [ADR], Quadrant Count Ratio [QCR]), as advocated in the GAISE Framework (Franklin et al., 2007). In all three series, most instances related to CC variables included 2x2 tables. Measures such as ADR and QCR are suitable for helping students move from informal and intuitive quantification of the strength of association to the formal statistical tests. However, despite the benefits of ADR and QCR, these measures simply were not present. Moreover, multiple strategies and misconceptions found in research about the association of CC variables (e.g., Batanero et al., 1997, 1998) were not explicitly addressed in any of the three series.

Table dimensions other than 2x2 were found in two of the three series. In particular, the UCSMP series included 2x4, 2x5, 2x6, 2x7, 2x8, 2x12, 3x4, and 3x5 tables. This might be because the Chi-Square test was used as a tool to judge the association and is versatile for use in a variety of table dimensions. The HML series included 2x3, 3x3, and 3x4 tables and often asked students to utilize these to compare

frequencies or to calculate conditional probabilities, but not for judgments about association. Research has shown that, when judging association, students' performance varies with respect to the dimensions of contingency tables (Batanero et al., 1996). Hence, it is helpful to attend to different techniques and strategies that are applicable for solving problems with multiple dimensions of tables to expose students to the plethora of table dimensions. However, the authors of the textbook series in this study did not include different approaches, techniques, and strategies for multiple types of tables.

One categorical and one numerical variable combination of bivariate data. All three series addressed the relationship between CN variables by comparing distributions. However, the HML and UCSMP series did not ask students to judge the association but instead students were to merely compare distributions. In addition, the three textbook series differed in their use of graphical displays and numerical statistics to compare groups. Among the three series, the UCSMP series placed the least emphasis on CN combination (about 2% of total bivariate data instances in the series) compared with HML (about 13%) and CPMP (14%). The HML and CPMP series were similar in their introduction and development of the content; however, the examination of statistical difference between two groups received little emphasis (1 instance) in HML. Research provides evidence that sampling is a complicated concept (Chance, delMas, & Garfield, 2004) and requires multiple prior concepts of statistics such as distribution, sample, population, variability, and sampling to understand. Thus, merely one instance probably does not provide an adequate opportunity for students to develop an understanding of this content. A careful introduction to and development of data simulations, as it is done in the CPMP series, is essential to helping students develop an understanding the

sophisticated concepts of testing for significance and to prepare them for formal reasoning later (Cobb, 2000; Franklin et al., 2007).

Research suggests that the comparison of two or more groups has the advantage of helping students understand characteristics of a single distribution (Cobb, 1999; Cobb, McClain, & Gravemeijer, 2003; Konold & Higgins, 2003; Watson & Moritz, 1999). Simultaneously, all series include tasks related to the comparison of two or more groups in the introduction of distribution. However, HML and UCSMP did not explicitly connect the comparison of groups to the judgment of association between one categorical and one numerical. It is possible that teachers of these curricula might not be aware of the connection between comparing distributions and association judgment and, in turn, this might compromise students' understanding.

Types of studies. The HML and CPMP series included appropriate content related to types of studies and distinguishing between correlation and causation. However, the developments of the content were different. The HML series included examples for each type of study and asked students to design studies to investigate particular issues; whereas, the CPMP series provided one situation with contrasting designs of studies and linked the results to the important distinction between *correlation* and *causation*. Efforts to make connections between multiple topics within mathematics can enhance students' overall learning of mathematics (NRC, 2001).

Surprisingly absent from the UCSMP series was content related to the types of studies, because these are recommended in both the CCSSM and GAISE Framework. It is possible that this exclusion of content is due to a perception that types of studies are not *mathematical* content per se. Although authors of UCSMP did not mention about

types of studies in the entire series, they did address the distinction between correlation and causation.

Even though all three series addressed correlation and causation, tasks asking students to examine the validity of study conclusions related to correlation and causation were absent. If offered, such tasks would force students to confront the common misconception that a *correlation* is synonymous with *causation* (e.g., Crocker, 1981; Ross & Cousins, 1993).

Two numerical variables combination of bivariate data. Even though all three series contained similar topics related to bivariate data, the manner in which authors' of the three series introduced and developed the content was quite different. For the NN combination, the HML series began by requiring students to make a scatterplot and observe patterns of increase/decrease then moved to linear modeling, exponential regression, and quadratic regression. The HML series then included exponential fitting with the aid of linearizing data with log and log-log transformations and added on rational function fitting and trigonometric regression. All model regressions but linear and rational were introduced using calculators with no formal descriptions about the properties of the regressions such as minimizing SSR. When obtained from calculators without any definitions or formal investigations, regression equations and coefficients such as r are viewed as a black box (McCallum, 2003), and such an approach might not support student learning.

The UCSMP series began with linear fitting of NN variables and then moved to exponential, quadratic regressions, and direct and inverse fittings. The authors then revisited linear, exponential and quadratic fittings and added trigonometric fitting.

Logarithmic, exponential, and power fittings were introduced last with the aid of log and log-log transformations. The authors followed a *spiral development* of learning, with content repeated several times with increasingly sophisticated understanding of the content.

In further contrast, the CPMP series began with an informal introduction of the content and then developed the content through more formalism. The CPMP authors began by asking students to observe patterns of data increasing/decreasing in a scatterplot, continued to linear fitting, exponential regression and then to direct, inverse fittings and power regressions. During students' first experience with regression, they used calculators to find models for the regression. The authors then introduced formal investigations to explore the properties of the regressions in the second presentation of the regression. They included data transformations such as translations, reflections, log and log-log transformations to connect multiple model fittings to linear models. Although trigonometric fitting was not found in the CPMP series, the authors instead included tasks related to *patterns of chance* and *statistical process control* to study changes in one quantity as a function of time. In addition, CPMP is the only series that introduced other measures of correlation to assess the goodness of fit, such as Spearman's rank correlation. While these various measures of association help prepare students to move through different levels of understanding of quantification of the goodness of fit, they are not suggested in either the GAISE Framework or the CCSSM.

Each of the three series emphasized NN association, which is closely related to the concept of functions and is an important content in school mathematics (e.g., CCSSI, 2010; NCTM, 2000; Tall & Bakar, 1992). The emphasis on the NN combination over

other combinations coincides with the finding that subjects perform better in covariational reasoning with numerical data than categorical data (e.g., Beach & Scopp, 1966; Erlick & Mills, 1967, Jennings et al., 1982). In the NN combination, linear functions received much greater emphasis than the other models (ranging about 30% to 40% of all the models). Linear functions include important connections to key mathematical concepts such as slope and rate of change, are emphasized in the CCSSM (CCSSI, 2010). Exponential fitting was emphasized second-most. The emphasis on linear and exponential fittings is consistent with the recommendations of professional organizations that identify these as the two most important types of mathematical function (CCSSI, 2010; NCTM, 2000). Quadratic fitting was addressed in only one situation, projectile balls, but this result might be due to the relative scarcity of everyday life phenomena that can be modeled with quadratic fitting.

Several models such as trigonometric, logarithmic, and cubic fittings were rarely found (or were not found) and this might be because the series required students to use technology to find the models as opposed to calculating by hand. An exception was the instances using logarithmic and trigonometric models in the UCSMP series.

Related to model fitting, informally assessing the goodness of fit using scatterplots was addressed in all three series but with different degrees of emphasis. In particular, estimating and interpreting a correlation coefficient from a scatterplot were emphasized in UCSMP and CPMP. The authors of these UCSMP and CPMP series advised students to use both informal knowledge about relationships and data at hand to specify models. In contrast, tasks in the HML series typically asked students to find a

specific model for a set of data using calculators, without first asking them to estimate the fit of the model using scatterplots.

Alignment with the Common Core State Standards-Mathematics learning expectations. I generated a list of 21 CCSSM learning expectations (LEs) from eight CCSSM standards for bivariate data and used the LEs with the content of the three series. All 21 of the LEs were addressed in HML and CPMP, but the emphasis given to each of the LEs varied by series. Some LEs get much emphasis such as S-ID.6. *Represent data on two quantitative variables on a scatter plot. Describe how the variables are related* (with 27 instances in HML and 48 instances in CPMP) whereas some was addressed in only very few instances such as S-IC.5. *Use simulations to decide if differences between parameters are significant* (1 instance in HML and 15 instances in CPMP). Publishers can claim their books are aligned with the CCSSM but it might be that this is based on a single instance. For example, each of 4 CCSSM LEs was addressed with only one instance and 6 CCSSM LEs was addressed by two instances in the HML series. This finding is consistent with Seeley's (2003) argument that publishers try to stuff textbooks with content in order to meet many state standards and therefore maximize adoption. By way of contrast, 6 LEs were not addressed in UCSMP.

At the same time, there were 3 bivariate data LEs common to all series that were not found in the CCSSM. In addition, 6 LEs in the UCSMP series and 8 LEs related to bivariate data in the CPMP series were not found in the CCSSM. Some of these LEs such as *Estimate correlation using scatterplots* were also suggested in the GAISE Framework. Researchers emphasize the role of LTs in supporting coherence and rigor in standards and curriculum development (Confrey et al., 2010, Daro et al., 2011) and the CCSSM is

purportedly based on research about LTs and the logic of mathematics (Heck et al., 2011). However, it is not clear which learning expectations were developed using research related to LTs, particularly for the content related to bivariate data.

Comparison of GAISE Framework and textbook series learning trajectories.

The results of learning trajectories for three combinations showed that the HML and CPMP series generally followed the developmental-level sequence prescribed in the GAISE Framework for bivariate data. Notwithstanding the developmentally appropriate sequencing of content, statistical content such as defining and investigating SSR was altogether missing in the HML series. The UCSMP series was least in accordance with the GAISE Framework's guidelines for the sequencing of content. Specifically, for the CC combination, the authors introduced Chi-Square test (Level C) to examine the association very early in *Algebra* then move down to detect the association by comparing percentages (Level A) in *FST*.

Task features.

Mathematical complexity. The vast majority of instances (ranging from 65% to 80%) related to bivariate data in the three series were at a moderate level of mathematical complexity. Across the textbook series, the average level of mathematical complexity was highest for CPMP (2.10) and lowest for HML (1.75) and UCSMP (1.75). This finding was somewhat different from the literature indicating that the majority of tasks included in U.S. textbooks require a low level of cognitive demands (e.g., Jones, 2004; Schmidt et al., 1997; Valverde et al., 2002). The scales used to determine the level of cognitive demands or the difference of analyzed content might explain this discrepancy in findings. The tasks coded in this study often involved data and required students to do

multiple steps with some levels of decision making to model the data, and these inherently involve higher levels of mathematical complexity. Hence, it might be informative to examine how the cognitive demands of tasks vary by content strands within a set of curriculum materials.

The GAISE Framework coding. An examination of the components of the GAISE Framework across the three series showed that across three series, a sizable proportion (ranging from 27.1% to 45.5%) of bivariate data instances addressed *Formulate Questions* (FQ). The FQ component affords the opportunity for students to gradually become aware of the difference between mathematical and statistical problems (Franklin et al., 2007). Introducing or reviewing the process of formulating questions at the beginning of a task, provides students with an opportunity to develop an understanding of the purpose of statistics and make connections to the other statistical components on which they may be working. However, despite the prevalence of FQ instances, students nevertheless were not afforded opportunities to generate statistical questions for themselves but instead were provided them by authors.

Few bivariate data instances (ranging 4.9% to 12.1%) provided students opportunities to *Collect Data* (CD), one of the crucial components of doing statistics as suggested in the GAISE Framework. In addition, when students were required to *collect data* they were not required to engage in higher levels of mathematical maturity to “*design for difference*” (Franklin et al., 2007, p. 14) and to use random selection. Notwithstanding, there were several instances related to bivariate data in the CPMP and UCSMP series that included the CD component in the form of a project at the end of the chapters or as an investigation that could be carried out in class. By doing projects and/or

statistical investigations, students can work with authentic data and become aware of the measurement error and the inherent messiness of data analysis (Ben-Zvi & Garfield, 2004).

Almost all bivariate data instances (88.5% to 95.3%) addressed the *Analyze Data* (AD) component, and the majority of the instances (63.1% to 87.7%) included the component *Interpret Results* (IR). Similarly, almost all the instances addressing CD were at the lowest developmental level (Level A) with the exception of a few in CPMP and UCSMP that addressed Level B of the GAISE framework. On average, CPMP included the greatest number of opportunities (2.20) for students to *Analyze Data*, this was followed by the UCSMP (2.12) and HML (1.84) series. For IR, on average, the CPMP's developmental level (2.09) was higher than those of the UCSMP (1.94) and HML series (1.92).

An attempt was made to examine the relationship between the four components of the GAISE Framework as well as explore associations between the developmental levels of GAISE and levels of mathematical complexity. When aggregating the data from the three series, Goodman and Kruskal's Gamma tests show that AD and IR codes were *accordant*; meaning that, tasks ranked with a high code in the AD component were likely to be assigned a high code in the IR component, and vice versa. The Goodman and Kruskal's Gamma test also showed that two frameworks, the GAISE Framework and the NAEP mathematical complexity framework, were comparable. That is, an increase in the level of coding in one framework generally led to an increased coding level in the other framework.

Despite the accordant ratings of two frameworks, there were several bivariate data instances that got a high code in one framework but a low code in the other. Generally in these situations, complicated knowledge was simplified by providing a procedure to follow or by having students use calculators without providing explanation. Therefore, the two frameworks are largely concordant and yet serve different purposes when addressing two interrelated aspects of tasks. In particular, the NAEP mathematical complexity framework is used to identify the mathematical complexity for *all* mathematical tasks (including statistical tasks) while the GAISE Framework is used for the purpose of identifying the developmental level of statistical thinking and reasoning and is applicable to *statistical* tasks only. Furthermore, the GAISE Framework suggests that students should move through increasingly sophisticated reasoning without skipping over any level whereas the mathematical complexity framework is not built on the premise of the continuous development of three levels.

Purpose and utility task features. In terms of *purpose* and *utility* features across the three series, almost all (more than 90%) instances addressed the *utility* feature, meaning that students explained how, when, and why the mathematical ideas are useful. On average, more than half (ranging from 50% to 65%) of the instances attended to the *purpose* feature, meaning that a virtual or actual product was available for students as a result of the accomplished task. This might be explained by the nature of tasks in the sample; they involved data, required interpretation, and were connected to real life. Results also showed that if a task addressed the *purpose* feature, it almost always addressed the *utility* feature.

Common Core State Standards Mathematics Mathematical Practices coding.

Results related to SMP showed that the three series offered potential for students to access the eight practices. However, the extent of access varied by each SMP and by series. On average for the percentages of instances addressing the SMPs, the CPMP series offered the most (57.6%) instances addressing the practices followed by the UCSMP (48.3%) and then HML (40.3%) series. Specifically, on average the three series offered students the greatest number of opportunities to learn MP2 *Reason abstractly and quantitatively* (72%), MP4 *Model with mathematics* (84%), and MP6 *Attend to precision* (84.5%). The textbooks offered a moderate number of opportunities to learn MP3 *Construct viable arguments and critique the reasoning of others* (60.1%), MP5 *Use appropriate tools strategically* (44.9%), and MP7 *Look for and make use of structure* (34.7%). In contrast, the least number offer opportunities to learn MP1 *Make sense of problems and persevere in solving them* (6.3%) and MP8 *Look for and express regularity in repeated reasoning* (3.2%). The lowest percent for MP1 might be because *persevere in solving [problems]* can rarely be evident in textbook and instead should be observed in practice (Stein & Lane, 1996). Notwithstanding, I looked for the feature in the tasks asking students to solve one problem with using multiple approaches to code for this SMP; however, they were rarely present in the three series (less than 10.6%). The low percent of bivariate instances addresses MP8 might relate to the design features of the curriculum or nature of the content analyzed, statistics. Generalization is essential when learning mathematics (Sasman, Olivier, & Linchevski, 1999). However, the three series analyzed did not afford the opportunity for students to generalize strategies for a set of similar problems or use “analogous problems, and try special cases and simpler forms of

the original problem” (CCSSI, 2010, p. 6) to help solve problems.

Limitations of the Study

Personal bias is inherent in any qualitative study when selecting topics, coding, analyzing data and reporting the results. Nevertheless, I have worked diligently to maintain the integrity of this research study. To be trained in advanced mathematics, working toward a Ph.D. with a minor in statistics, and taking extensive courses in mathematics education provides a strong background to conduct my analyses with integrity. I argue that I am qualified to conduct the curriculum analyses, as suggested by the NRC (2004), namely “Content analyses ... should be conducted by a variety of scholars ... who should identify their qualifications, values concerning mathematical priorities, and potential sources of bias regarding their execution of content analyses” (p. 197).

One potential limitation is inherent in framing coding scheme and the inclusion and exclusion of tasks for coding. However, I made an effort to minimize personal bias by integrating consultations with researchers whose expertise was in statistics, statistics education, or mathematics education to maintain the trustworthiness of this research study. I made strong efforts to establish the reliability of coding and develop a coding manual that was grounded in the research literature.

The sample of tasks might bring about another limitation for this study related to findings for tasks features. To be included in the sample for coding, tasks must often include data and require statistical reasoning. Therefore, the findings related to mathematical complexity, purpose and utility, and mathematical practices might not be representative of population of *all* tasks for the whole series.

The final limitation relates to the small sample size, namely three textbook series. In order to address this limitation, I chose textbooks with different design principles, classroom expectations, and approaches to the sequencing of content and, in doing so, was able to document and represent the diversity of learning trajectories for bivariate data, in spite of the small sample size. I did not aim to provide extensive results for the trajectories in U.S. curriculum with this research, but instead intended to let it serve as an illustration of an uncharted research approach with implications for future use.

Implications and Recommendations for Future Research

The findings of this study provide several insights in mathematics education, specifically related to curriculum design and curriculum research. In this section, I discuss implications that are drawn from the results and offer recommendations for future research.

Curriculum Development

A criticism of U.S. curriculum is that it covers too many topics without going in-depth in any of them (Valverde et al., 2002) and is highly repetitive (Flanders, 1987, 1994), thus sacrificing the coherence and rigor of the content. The findings of this study partially support this criticism. For example, HML covered some complex topics (e.g., examining the significance of statistics test) with exactly one instance or only a few instances. Furthermore, in the HML series, some topics (e.g., independent events) in *Algebra I* were repeated exactly without any subsequent expansion in *Geometry* books. Learning trajectories could inform curriculum design to reduce the fragmentation of content and build rigor by providing a systematic and sequential map of important content (Simon & Tzur, 2004). The use of learning trajectories as a basic construct

enables curriculum developers to use concepts such as prior knowledge to form matrices of connection among strands, thereby promoting coherence of the curriculum.

The GAISE Framework provides a robust description of the sequence of statistical content in increasing levels of sophistication, and therefore it can serve as a guide for developing learning progressions for statistical content such as bivariate data. However, the findings related to LTs of this study show that the suggestions were not fully addressed in the textbook series. For example, the quantification of CC association (e.g., QCR, ADR) at Level B were absent in the three series, and Chi-Square Statistics (Level C) appeared too early in the UCSMP series in *Algebra*. Future curriculum development might include these quantifications to help students to move to the most sophisticated content and embody the developmental levels in the curriculum.

This study found several instances that simultaneously addressed more than one component of doing statistics, more than one SMP, and multiple learning expectations. Using such super-instances promote the connectedness of statistical and mathematical topics. In addition, by incorporating multiple LEs within a single task, which requires students to be able to make decisions in major points, curriculum developers have the potential to raise the level of cognitive demands required by the task.

In the years since the release of the CCSSM, there have been efforts to revise or rewrite curricula to align with the new standards (Heck et al., 2011). For the content analyzed, HML addressed all LEs in the CCSSM and yet some LEs were given little emphasis (in terms of number of instances addressed) and, moreover, some mathematical practices were not evident in the HML series. In contrast, CPMP gave appropriate attention to both the CCSSM content and mathematical practices. Hence, it would be

wise to invest in curriculum development so that more textbook series embody the CCSSM content and practice standards as well as address the GAISE Framework.

Teachers' editions are an essential component in a publisher's wide array of resources. It is critical that they go beyond merely providing answers to problems found in the students' texts. Two of the three series include features that help teachers in lesson planning such as chapter overviews, and linking contents to the end-of-chapter projects (in UCSMP) and discussion of students' misconceptions (in CPMP). Thus it is important that curriculum developers also include in the teachers' editions information related to learning trajectories for interconnections between mathematical ideas, students' common misconceptions and faulty strategies students might use.

Content Analysis

This study has demonstrated how learning trajectories are useful in the analysis of curriculum materials. In particular, (a) examining the learning goals the authors convey in the curriculum materials, (b) the activities related to the content as presented by the order of tasks, and (c) the connection between the learning goals and learning activities and among activities collectively provide a valid picture of how bivariate data content is introduced and developed in the three curriculum materials. Researchers can use the construct of LT to capture the nuance the introduction and development of content as presented in different curriculum materials.

Results of this study show that the GAISE Framework and NAEP mathematical complexity frameworks are concordant. Although the frameworks are concordant, future researchers can decide to use one or both of them to serve different purposes of content analysis and curriculum evaluation. Specifically, if researchers are concerned about the

cognitive demands required by mathematical tasks, then the NAEP mathematical complexity framework is better suited for that purpose. In contrast, if researchers wish to document the developmental levels related to statistics content, the GAISE Framework better meets this demand. Notwithstanding, using both frameworks for statistics content has the advantage to provide different perspectives for one issue and help compare each other.

Another purpose of content analysis is to assess the alignment of a curriculum to a given set of student learning goals or standards. The findings in this study show that the approach is sensible to capture the alignment of multiple curriculum materials with learning standards (i.e., the CCSSM and the GAISE Framework). Specifically, comparing the *learning trajectories* of content in the textbooks and in the standards (e.g., the GAISE Framework) can contribute to the research related to alignment that focuses on the extent of inclusion and exclusion of content (e.g., Dingman, 2007; Porter, 2002).

Teacher Education

According to the Conference Board of Mathematical Sciences (2001, 2011), teachers are least prepared to teach statistics and probability. Effective teachers must deeply understand mathematical content to teach it well (Ball, Thames, & Phelps, 2008). The learning trajectories presented in this study offer pre-service teachers opportunities to make the connections between and among bivariate data content. Hence, *mathematics courses* should heighten awareness about the sequencing of topics and their interrelatedness. Moreover, *methods courses* should provide pre-service teachers the opportunity to learn about and understand the construct of learning trajectories and,

ultimately, gain deeper understanding and be prepared for their future roles in the classroom.

As indicated in this study, the GAISE Framework offers a systematic way to analyze curricula. In-service teachers might benefit from learning the framework and applying its components to an analysis of current curricular materials. In particular, in-service teachers can trace learning trajectories in one textbook or compare textbooks in relation to these aspects. Using the analysis method of this study can prepare in-service teachers to make informed decisions regarding the selection of mathematics textbooks as well as assist them in adapting them to meet with needs of students in their classrooms.

Future Research

Given that textbooks play an important role in the educational system, it is imperative to conduct content analysis of textbooks at all levels of education. The methodology used in this study can be readily applied to the study of bivariate data content in other grade levels. Therefore, I recommend extending this research to curricular materials used at the elementary and middle school levels. Additionally, although this study focused on bivariate data content, by modifying the boundaries for the inclusion of other statistical concepts such as distributions, variation, future studies can apply this methodology to analyze other school mathematics content.

Results of this study showed that statistical content can be introduced and developed quite differently in U.S. high school mathematics textbooks. For example, with regard to linear fitting, the HML authors provide a procedure for determining the line of best fit and show examples to illustrate the procedure. Students were expected to follow the strategies introduced in the examples related to the content. In contrast, with

the same content of linear fitting, the CPMP authors build an investigation and a series of questions to guide student learning without providing specific worked examples for students to imitate. This difference can be captured using another important construct that needs to be addressed, namely the *pedagogical storyline* (Heck, Chval, Weiss, & Ziebarth, 2012). Therefore, future curriculum evaluation research needs to examine both content and pedagogical storylines in order to provide a more complete or holistic characterization of curricula.

Perhaps it is predictable that the findings related to the CCSSM mathematical practices varied for different strands/conceptual categories in school mathematics. The framework developed for this study to code for mathematical practices worked for the tasks related to bivariate data; however the need still remains for other mathematical content to be examined and it is likely that the framework coding for SMP would need to be revised in order to fit plethora of topics in the school mathematics curriculum.

This study did not include a comparison of learning trajectories in the curriculum materials and in the CCSSM. This is because learning trajectories of the CCSSM at high school level were not available at the time of this study. Hence, future research can expand LTs of K-8 CCSSM (e.g., Confrey et al., 2012) to continue the development of LTs at the high school level. In turn, researchers can examine the relationship between LTs evident in the CCSSM and those present in curriculum materials.

Findings from this study provide hypothetical learning trajectories for bivariate data found in the textbooks. Future research can synthesize HLTs, research related to bivariate data, and recommendations of the GAISE Framework to form efficient LTs and, in turn, examine how the LTs play out in actual classrooms. In addition, data gained from

students' learning might inform modifications to the learning trajectories and in turn provide a "best case" LT for future learning.

Summary

This study, categorized as a vertical analysis, was the first attempt to use the construct of learning trajectories to analyze bivariate data in textbooks. This study does not aim to evaluate curricula and provide the depth or quality of curricular treatments, but to capture the differences in the introduction and development of bivariate data content between textbook series as well as to highlight the opportunities that students might have when learning the content. In addition, multiple facets of tasks were first examined to point out important aspects for developers to consider when designing and for teachers to attend to when selecting tasks.

Among the findings was the existence of multiple learning trajectories for the content analyzed. The CPMP series best reflects knowledge about learning trajectories and aligns well with the suggestions in the GAISE Framework. In addition, even though it was developed before the release of the CCSSM, CPMP addresses the content related to bivariate data and mathematical practices of the CCSSM document. The remaining issues for future research are to form hypothetical learning trajectories taking incorporate LTs from the textbooks, findings from research in bivariate data, the suggestions from the GAISE Framework, and try them out in different classroom settings for future revision.

References

- Adi, H., Karplus, R., Lawson, A., & Pulos, S. (1978). Intellectual development beyond elementary school VI: Correlational reasoning. *School Science and Mathematics*, 78(8), 675-683. doi: 10.1111/j.1949-8594.1978.tb18270.x
- Ainley, J., Bills, L., & Wilson, K. (2003). *Designing tasks for purposeful algebra*. Paper presented at the 3rd Conference of the European Society for Research in Mathematics Education, Bellaria.
- Ainley, J., & Pratt, D. (2002). *Purpose and utility in pedagogic task design*. Paper presented at the 26th Annual Conference of the International Group for the Psychology of Mathematics Education, UK.
- Ainley, J., Pratt, D., & Hansen, A. (2006). Connecting engagement and focus in pedagogic task design. *British Educational Research Journal*, 32(1), 23-38. doi: 10.1080/01411920500401971
- Artique, M. (2000). *Instrumentation issues and the integration of computer technologies into secondary mathematics teaching*. Paper presented at the Annual Meeting of GDM, Potsdam, Germany.
- Artique, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, 7(3), 245-274. doi: 10.1023/A:1022103903080
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, November/December(59), 389-407. doi: 10.1177/0022487108324554
- Baroody, A. J., Cibulskis, M., Lai, M.-I., & Li, X. (2004). Comments on the use of learning trajectories in curriculum development and research. *Mathematical Thinking and Learning*, 6(2), 227-260. doi: 10.1207/s15327833mtl0602_8
- Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer based teaching environmen. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics: Proceedings of the 1996 IASE Round Table Conference* (pp. 191-205). Voorburg, The Netherlands: International Statistical Institute.
- Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27(2), 151-169.
- Batanero, C., Godino, J. D., & Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 221-236). Stellenbosh, South Africa: University of Stellenbosh.
- Battista, M. T. (2004). Applying cognition-based assessment to elementary school students' development of understanding of area and volume measurement. *Mathematical Thinking and Learning*, 6(2), 185-204. doi: 10.1207/s15327833mtl0602_6

- Beach, L. R., & Scopp, T. S. (1966). Inferences about correlations. *Psychonomic Science*, 6(6), 253-254.
- Bell, A., Brekke, G., & Swan, M. (1987). Diagnostic teaching: 4 graphical interpretations. *Mathematics Teaching*, 120, 56-60.
- Bell, A., Brekke, G., & Swan, M. (1987). Diagnostic teaching: 5 graphical interpretation teaching styles and their effects. *Mathematics Teaching*, 120, 50-57.
- Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics*, 2(1), 34-42.
- Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-15): Springer Netherlands.
- Beyth-Marom, R. (1982). Perception of correlation reexamined. *Memory & Cognition*, 10(6), 511-519. doi: 10.3758/bf03202433
- Bobko, P., & Karren, R. (1979). The perception of Pearson product moment correlations from bivariate scatterplots. *Personnel Psychology*, 32(2), 313-325. doi: 10.1111/j.1744-6570.1979.tb02137.x
- Brown, S. A., Breunlin, R. J., Wiltjer, M. H., Degner, K. M., Eddins, S. K., Edwards, M. T., . . . Usiskin, Z. (2008). *The University of Chicago School Mathematics Project: Algebra - Teacher's edition* (3rd ed.). Chicago, IL: McGrawHill - Wright Group.
- Carlson, M. (1998). A cross-sectional investigation of the development of the function concept. In E. Dubinsky, A. H. Schoenfeld & J. J. Kaput (Eds.), *Research in collegiate mathematics education III, Issues in mathematics education* (Vol. 7, pp. 115-162).
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 33(5), 352-378. doi: 10.2307/4149958
- Carlson, M. P. (2002). Physical enactment: A powerful representational tool for understanding the nature of covarying relationships. In F. Hitt (Ed.), *Representations and mathematics visualization* (pp. 63-77). Mexico: CINVESTAV.
- Carlson, M. P., Larsen, S., & Jacobs, S. (2001). An investigation of covariational reasoning and its role in learning the concepts of limit and accumulation. In R. Speiser, C. Maher & C. Walter (Eds.), *Proceedings of the 23rd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 145-153). Snowbird, UT: PME-NA.
- Catena, A., Maldonado, A., & Cándido, A. (1998). The effect of frequency of judgment and the type of trials on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance*, 24(2), 481-495. doi: 10.1037/0096-1523.24.2.481
- Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In Ben-Zvi & J. Garfield (Eds.), *The challenges of developing statistical literacy, reasoning, and thinking* (pp. 295-324). Netherlands: Kluwer Academic.
- Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning*

- and Verbal Behavior*, 6(1), 151-155. doi: 10.1016/S0022-5371(67)80066-5
- Chapman, L. J., & Chapman, J. B. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74(3), 271-280.
- Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 72(3), 193-204. doi: 10.1037/h0024670
- Charalambous, C. Y., Delaney, S., Hsu, H.-Y., & Mesa, V. (2010). A comparative analysis of the addition and subtraction of fractions in textbooks from three countries. *Mathematical Thinking and Learning*, 12(2), 117-151. doi: 10.1080/10986060903460070
- Clement, J. (1989). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics*, 11(1-2), 77-87.
- Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, 6(2), 81-89. doi: 10.1207/s15327833mtl0602_1
- Clements, D. H., Wilson, D. C., & Sarama, J. (2004). Young children's composition of geometric figures: A learning trajectory. *Mathematical Thinking and Learning*, 6(2), 163-184. doi: 10.1207/s15327833mtl0602_5
- Cleveland, W. S., Diaconis, P., & McGill, R. (1982). Variables on scatterplots look more highly correlated when the scales are increased. *Science*, 216, 1138-1141.
- Cobb, P. (2000). The importance of a situated view of learning to the design of research and instruction. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning*. Westport, CT: Ablex.
- Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction*, 21(1), 1-78. doi: 10.1207/s1532690xci2101_1
- Common Core State Standards Initiative [CCSSI]. (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.
- Conference Board of Mathematical Sciences. (2012). *The mathematical education of teachers II*. Providence, Rhode Island: Mathematical Association of America.
- Confrey, J., & Krupa, E. E. (2012). The Common Core State Standards for Mathematics: How did we get here, and what needs to happen next? In C. R. Hirsch, G. T. Lappan & B. J. Reys (Eds.), *Curriculum issues in an era of Common Core State Standards for Mathematics*. Reston, VA: NCTM.
- Confrey, J., & Maloney, A. (2010). *The construction, refinement, and early validation of the equipartitioning learning trajectory*. Paper presented at the 9th International Conference of the Learning Sciences - Volume 1, Chicago, Illinois.
- Confrey, J., Maloney, A., Nguyen, K., Mojica, G., & Myers, M. (2009). *Equipartitioning/splitting as a foundation of rational number reasoning using learning trajectories*. Paper presented at the 33rd Conference of the International Group for the Psychology of Mathematics Education, Thessaloniki, Greece.
- Corcoran, T., Mosher, F. A., & Rogat, A. (2009). Learning progressions in science: An evidence-based approach to reform (Research report #RR-63). Philadelphia, PA: Consortium for

Policy Research in Education.

- Coulombe, W. N., & Berenson, S. B. (2001). Representations of patterns and functions: Tools for learning. In A. A. Cuoco & F. R. Curcio (Eds.), *The roles of representation in school mathematics (2001 Yearbook)* (pp. 166-172). Reston, VA: National Council of Teachers of Mathematics.
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90(2), 272-292. doi: 10.1037/0033-2909.90.2.272
- Crocker, J. (1982). Biased questions in judgment of covariation studies. *Personality and Social Psychological Bulletin*, 8, 214-220.
- Daro, P., Mosher, F. A., & Corcoran, T. (2011). Learning trajectories in mathematics: A foundation for standards, curriculum, assessment, and instruction. Philadelphia, PA: Consortium for Policy Research in Education.
- Dierdorp, A., Bakker, A., Eijkelhof, H., & van Maanen, J. (2011). Authentic practices as contexts for learning to draw inferences beyond correlated data. *Mathematical Thinking and Learning*, 13(1-2), 132-151. doi: 10.1080/10986065.2011.538294
- Dingman, S. W. (2007). *Mathematics textbooks and state curriculum standards: an analysis of the alignment between the written and intended curricula*. (Doctoral dissertation), University of Missouri, Columbia, MO. Retrieved from <https://mospace.umsystem.edu/xmlui/handle/10355/4690> MOspace database.
- Donnelly, J. F., & Welford, A. G. (1989). Assessing pupils' ability to generalize. *International Journal of Science Education*, 11(2), 161-171. doi: 10.1080/0950069890110205
- Erlick, D. E. (1966). Human estimates of statistical relatedness. *Psychonomic Science*, 5(10), 365-366.
- Erlick, D. E., & Mills, R. G. (1967). Perceptual quantification of conditional dependency. *Journal of Experimental Psychology*, 73(1), 9-14. doi: 10.1037/h0024138
- Estapa, A., & Batanero, C. (1996). Judgments of correlation in scatter plots: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, 4, 21-41.
- Fey, J., & Hirsch, C. (2007). The case of Core-Plus mathematics. *Perspectives on the design and development of school mathematics curricula*, 129-142.
- Fiedler, K. (1991). The tricky nature of skewed frequency tables: An information loss account of distinctiveness-based illusory correlation. *Journal of Personality and Social Psychology*, 60, 24-36.
- Flanders, J. R. (1987). How much of the content in mathematics is new? *Arithmetic Teacher*, 35(1), 18-23.
- Flanders, J. R. (1994). Student opportunities in grade 8 mathematics: Textbook coverage of the SIMS Test. In I. Westbury, C. A. Ethington, L. A. Sosniak & D. P. Baker (Eds.), *In Search of more effective mathematics education: Examining data from the IEA Second International Mathematics Study* (pp. 61-93). Norwood, NJ: Ablex.
- Flanders, J. R. (1994). Textbooks, teachers, and the SIMS Test. *Journal for Research in Mathematics Education*, 25(3), 260-278. doi: 10.2307/749338

- Flanders, J. R., Lassak, M., Sech, J., Eggerding, M., Karafiol, P. J., McMullin, L., . . . Usiskin, Z. (2010). *The University of Chicago School Mathematics Project: Advanced Algebra - Teacher's edition* (3rd ed.). Chicago, IL: McGrawHill - Wright Group.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines and assessment for instruction in statistics education (GAISE) report: A Pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association.
- Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372-396. doi: 10.1111/j.1751-5823.2007.00029.x
- Gravemeijer, K. (2004). Local instruction theories as means of support for teachers in reform mathematics education. *Mathematical Thinking and Learning*, 6(2), 105-128. doi: 10.1207/s15327833mtl0602_3
- Gray, C. W. (1968). Predicting with intuitive correlations. *Psychonomic Science*, 11(2), 41-42.
- Grouws, D. A., & Smith, M. S. (2000). Findings from NAEP on the preparation and practices of mathematics teachers. In E. A. Silver & P. A. Kenney (Eds.), *Results from the seventh mathematics assessment of the national assessment of education progress* (pp. 107-141). Reston, VA: National Council of Teachers of Mathematics.
- Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, 12(4), 392-407. doi: 10.1016/S0022-1031(76)80006-6
- Hamilton, D. L., & Rose, T. L. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology*, 39(5), 832-845. doi: 10.1037/0022-3514.39.5.832
- Heck, D. J., Chval, K. B., Weiss, I. R., & Ziebarth, S. W. (2012). Developing measures of fidelity of implementation for mathematics curriculum materials enactment. In D. J. Heck, K. B. Chval, I. R. Weiss & S. W. Ziebarth (Eds.), *Approaches to studying the enacted mathematics curriculum* (pp. 67-87). Charlotte, NC: Information Age Publishing.
- Heck, D. J., Weiss, I. R., & Pasley, J. D. (2011). A priority research agenda for understanding the influence of the Common Core State Standards for Mathematics. Chapel Hill, NC.
- Herbst, P. (1995). *The construction of the real number system in textbooks: A contribution to the analysis of discursive practices in mathematics*. (Unpublished Master' Thesis), University of Georgia, Athens.
- Hiebert, J. S., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In J. Frank K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 1). Charlotte, NC: Information Age Publishing.
- Hirsch, C. R. (2007). *Perspectives on the design and development of school mathematics curricula*. Reston, VA: National Council of Teachers of Mathematics.
- Hopfensperger, P., Jacobbe, T., Lurie, D., & Moreno, J. (2012). *Bridging the gap between Common Core State Standards and teaching statistics*. Alexandria, VA: American Statistical Association.

- Inhelder, B., & Piaget, J. (1958). Random variations and correlations (A. Parsons & S. Milgram, Trans.). In B. Inhelder & J. Piaget (Eds.), *The growth of logical thinking from childhood to adolescence* (pp. 224-242). London: Routledge & Kegan Paul.
- Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, 79(1), 1-17. doi: 10.1037/h0093874
- Jennings, D., Amabile, T., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211-230). Cambridge, England: Cambridge University Press.
- Johnson, G. J., Thompson, D. R., & Sharon, L. S. (2010). Proof-related reasoning in high school textbooks. *Mathematics Teacher*, 103(6), 410-417.
- Jones, D. L. (2004). *Probability in middle grades mathematics textbooks: An examination of historical trends, 1957-2004*. (Doctoral dissertation), University of Missouri at Columbia, MO. Retrieved from <http://laurel.lso.missouri.edu/record=b5299484~S1>
- Jones, G. A., Langrall, C. W., Mooney, E. S., & Thornton, C. A. (2004). Models of development in statistical reasoning. In D. Ben-Zvi & G. Joan (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 97-117). Netherlands: Kluwer Academic Publishers.
- Kaput, J. J. (1992). Patterns in students' formalization of quantitative patterns. In G. Harel & E. Dubinsk (Eds.), *The concept of function: Aspects of epistemology and pedagogy*, MAA Notes (Vol. 25, pp. 290-318). Washington, DC: Mathematical Association of America.
- Kelley, H. H. (1967). Attribution theory in social psychology In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.
- Konarski, R. (2005). Judgments of correlation from scatterplots with contaminated distributions. *Polish Psychological Bulletin*, 36(51-62).
- Konold, C. (2002). Alternatives to scatterplots. In B. Phillips (Ed.), *Proceedings of the sixth international conference on teaching statistics, Cape Town, South Africa*. Voorburg, The Netherlands: International Statistical Institute.
- Konold, C., & Higgins, T. L. (2003). Reasoning with data. In J. Kilpatrick, G. W. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193-215). Reston, VA: NCTM.
- Krabbendam, H. (1982). *The non-qualitative way of describing relations and the role of graphs: Some experiments*. Paper presented at the Conference on functions Enschede, The Netherlands.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96(4), 674-689. doi: 10.1037/0033-295x.96.4.674
- Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. Orlando, FL: Academic Press.
- Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., Sheldon, H. W., Klahr, D., & Sharon, M. C.

- (1995). Strategies of Knowledge Acquisition. *Monographs of the Society for Research in Child Development*, 60(4), i-157. doi: 10.2307/1166059
- Lane, D. M., Anderson, C. A., & Kellam, K. L. (1985). Judging the relatedness of variables: The psychophysics of covariation detection. *Journal of Experimental Psychology: Human Perception and Performance*, 11(5), 640-649.
- Larson, R., Boswell, L., Kanold, T. D., & Stiff, L. (2012). *Holt McDougal Larson Algebra 1 - Teacher's Edition*. Orlando, FL: Holt McDougal - Houghton Mifflin Harcourt.
- Larson, R., Boswell, L., Kanold, T. D., & Stiff, L. (2012). *Holt McDougal Larson Geometry - Teacher's Edition*. Orlando, FL: Holt McDougal - Houghton Mifflin Harcourt.
- Larson, R., Boswell, L., Kanold, T. D., & Stiff, L. (2012). *Holt McDougal Larson Algebra 2 - Teacher's Edition*. Orlando, FL: Holt McDougal - Houghton Mifflin Harcourt.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60(1), 1-64. doi: 10.3102/00346543060001001
- Lesh, R., & Yoon, C. (2004). Evolving communities of mind-in which development involves several interacting and simultaneously developing strands. *Mathematical Thinking and Learning*, 6(2), 205-226. doi: 10.1207/s15327833mtl0602_7
- Li, Y. (1999). *An analysis of algebra content, content organization and presentation, and to-be-solved problems in eighth-grade mathematics textbooks from Hong Kong, Mainland China, Singapore, and the United States*. (Doctoral dissertation), University of Pittsburgh, Pennsylvania. Retrieved from <http://search.proquest.com/docview/304519739>
- Lipe, M. G. (1990). A lens model analysis of covariation research. *Journal of Behavioral Decision Making*, 3(1), 47-59. doi: 10.1002/bdm.3960030105
- Lovell, K. (1961). A follow-up study of Inhelder and Piaget's the growth of logical thinking. *British Journal of Psychology*, 52(2), 143-153. doi: 10.1111/j.2044-8295.1961.tb00776.x
- McCallum, W. G. (2003). Thinking out of the box. In J. T. Fey, A. Cuoco, C. Kieran, L. McMullin & R. M. Zbiek (Eds.), *Computer Algebra Systems in Secondary School Mathematics Education* (pp. 73-86). Reston, VA: NCTM.
- McConnell, J. W., Brown, S. A., Karafiol, P. J., Brouwer, S., Ives, M., Lassak, M., . . . Usiskin, Z. (2010). *The University of Chicago School Mathematics Project: Functions, statistics, and trigonometry - Teacher's Edition* (3rd ed.). Chicago, IL: McGrawHill - Wright Group.
- McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., & Cooney, T. J. (1987). *The underachieving curriculum: Assessing U. S. School mathematics from an international perspective*. Champaign, Illinois: Stipes Publishing Company.
- Mesa, V. (2004). Characterizing practices associated with functions in middle school textbooks: An empirical approach. *Educational Studies in Mathematics*, 56(2-3), 255-286. doi: 10.1023/b:educ.0000040409.63571.56
- Mesa, V. M. (2000). *Conceptions of function promoted by seventh- and eighth-grade textbooks*

- from eighteen countries*. Unpublished doctoral dissertation. University of Georgia. Athen, GA.
- Mevarech, Z., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, 32(3), 229-263. doi: 10.1023/a:1002965907987
- Meyer, J., Taieb, M., & Flascher, I. (1997). Correlation estimates as perceptual judgments. *Journal of Experimental Psychology: Applied*, 3(1), 3-20. doi: 10.1037/1076-898x.3.1.3
- Monk, S. (1992). Students' understanding of a function given by a physical model. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy*, *MAA Notes* (Vol. 25, pp. 175-193). Washington, DC: Mathematical Association of America.
- Moore, D. S. (1990). Uncertainty. In L. Steen (Ed.), *On the shoulders of giants: A new approach to numeracy* (pp. 95-137): National Academy of Sciences.
- Moritz, J. B. (2000). Graphical representations of statistical associations by upper primary students. In J. B. A. Chapman (Ed.), *Mathematics education beyond 2000: The 23rd Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 440-447). Perth: Mathematics Education Research Group of Australasia.
- Moritz, J. B. (2002). Study times and test scores: What students' graphs show. *Australian Primary Mathematics Classroom*, 7(1), 24-31.
- Moritz, J. B. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227-256). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- National Assessment Governing Board. (2006). Mathematics framework for the 2007 National Assessment of Educational Progress. Washington, DC: The National Academies Press.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: The National Academy Press.
- National Research Council. (2004). *On evaluating curricular effectiveness: Judging the quality of K-12 mathematics evaluations*. Washington, DC: The National Academies Press.
- Neimark, E. D. (1975). Longitudinal development of formal operations thought. *Genetic Psychology Monographs*, 91(2), 171-225.
- Nemirovsky, R. (1996). A functional approach to algebra: Two issues that emerge. In N. Dedrarg, C. Kieran & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 295-313). Boston: Kluwer Academic Publishers.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Olson, T. (2010). *Articulated learning trajectories related to the development of algebraic*

- thinking that follow from patterning concepts in middle grades mathematics*. (Doctoral dissertation), University of Missouri Columbia. Retrieved from <https://mospace.umsystem.edu/xmlui/handle/10355/8318> MOspace database.
- Peressini, A. L., DeCraene, P. D., Rockstroh, M. A., Viktora, S. S., Canfield, W. E., Wiltjer, M. H., & Usiskin, Z. (2010). *The University of Chicago School Mathematics Project: Precalculus and Discrete Mathematics - Teacher's Edition* (3rd ed.). Chicago, IL: McGrawHill - Wright Group.
- Perez Echvarria, M. P. (1990). *Psicologia del razonamiento probabilistico*. [The psychology of probabilistic reasoning]. Madrid: Ediciones de la Universidad Autenoma de Madrid.
- Peterson, C. (1980). Recognition of noncontingency. *Journal of Personality and Social Psychology*, 38(5), 727-734. doi: 10.1037/0022-3514.38.5.727
- Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, 31(7), 3-14. doi: 10.3102/0013189x031007003
- Reys, B. J., Reys, R. E., & Chavez, O. (2004). Why mathematics textbooks matter. *Educational Leadership*, 61(4), 61-66.
- Reys, B. J., Thomas, A., Tran, D., Dingman, S., Kasmer, L., Newton, J., & Teuscher, D. (2013). State-level actions following adoption of Common Core State Standards for Mathematics. *NCSM Journal*, 14(2), 5-13.
- Robitaille, D. F., & Travers, K. J. (1992). International studies of achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* Reston, VA: National Council of Teachers of Mathematics.
- Ross, D. J. (2011). *Functions in contemporary secondary mathematics textbook series in the United States*. (Doctoral dissertation), University of Missouri, Columbia, MO. Retrieved from <https://mospace.umsystem.edu/xmlui/handle/10355/14225> MOspace database.
- Ross, J. A., & Cousins, J. B. (1993). Pattern of student growth in reasoning about correlational problems. *Journal of Educational Psychology*, 85(1), 49-65.
- Sasman, M. C., Olivier, A., & Linchevski, L. (1999). Factors influencing student's generalisation thinking processes. In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 161-168). Haifa, Israel: PME.
- Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht, The Netherlands: Kluwer.
- Schustack, M. W., & Sternberg, R. J. (1979). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, 110, 101-120.
- Seeley, C. L. (2003). Mathematics textbook adoption in the United States. In G. Stanic & J. Kilpatrick (Eds.), *A history of school mathematics* (pp. 957-988). Reston, VA: NCTM.
- Seggie, J. L., & Endersby, H. (1972). The empirical implications of Piaget's concept of correlation. *Australian Journal of Psychology*, 24(1), 3-8. doi: 10.1080/00049537208255778

- Senk, S. L., & Thompson, D. R. (2003). *Standards-based school mathematics curricula: What are they? What do students learn?* Mahwah, NJ: Erlbaum.
- Shaklee, H., & Mims, M. (1981). Development of rule use in judgments of covariation between events. *Child Development*, 52(1), 317-325. doi: 10.2307/1129245
- Shaklee, H., & Paszek, D. (1985). Covariation judgment: systematic rule use in middle childhood. *Child Development*, 56(5), 1229-1240. doi: 10.2307/1130238
- Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory & Cognition*, 8(5), 459-467. doi: 10.3758/bf03211142
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age.
- Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26(2), 114-145.
- Simon, M. A., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning*, 6(2), 91-104. doi: 10.1207/s15327833mtl0602_2
- Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4(1), 165-173. doi: 10.1111/j.1467-9450.1963.tb01324.x
- Snyder, M. (1981). Seek and ye shall find: Testing hypotheses about other people. In E. T. Higgins, D. C. Herman & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium on Personality and Social Psychology*. Hillsdale, NJ Erlbaum.
- Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202-1212.
- Sotos, A. E. C., Vanhoof, S., Noortgate, W. V. D., & Onghena, P. (2009). The transitivity misconception of Pearson's correlation coefficient. *Statistics Education Research Journal*, 8(2), 33-55.
- Stanton, T. (Producer). (2010, 08/01/2012). McGraw-Hill and University of Chicago School Mathematics Project announce first continuous pre-K-12 math curriculum. Retrieved from <http://pressrelated.com/press-release-mcgraw-hill-authors-receive-presidential-awards-for-excellence-in-mathematics-and-science-teaching.html>
- Steffe, L. P. (2004). On the construction of learning trajectories of children: The case of commensurate fractions. *Mathematical Thinking and Learning*, 6(2), 129-162. doi: 10.1207/s15327833mtl0602_4
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, 2(1), 50-80. doi: 10.1080/1380361960020103
- Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 319-369): Charlotte, North Carolina: Information Age Publishing.

- Stevenson, H. W., & Bartsch, K. (1992). An analysis of Japanese and American textbooks in mathematics. In R. Leetsma & H. Walberg (Eds.), *Japanese educational productivity* (pp. 103-133). Ann Arbor: Center for Japanese Studies, University of Michigan.
- Stockburger, D. W. (1982). Evaluation of three simulation exercises in an introductory statistics course. *Contemporary Educational Psychology*, 7(4), 365-370. doi: 10.1016/0361-476X(82)90021-2
- Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning*, 11(4), 258-288. doi: 10.1080/10986060903253954
- Swan, M. (1985). *The language of functions and graphs*. University of Nottingham: Shell Center.
- Swan, M. (1988). Learning the language of functions and graphs *Mathematics Interfaces: Proceedings of the 12th Biennial Conference of the Australian Association of Mathematics Teachers* (pp. 76-80). Newcastle, NSW: The New England Mathematical Association.
- Swatton, P. (1994). Pupils' performance within the domain of data interpretation, with particular reference to pattern recognition. *Research in Science & Technological Education*, 12(2), 129-144. doi: 10.1080/0263514940120203
- Swatton, P., & Taylor, R. M. (1994). Pupil performance in graphical tasks and its relationship to the ability to handle variables. *British Educational Research Journal*, 20(2), 227-243. doi: 10.1080/0141192940200207
- Tarr, J. E., Chavez, O., Reys, R. E., & Reys, B. J. (2006). From the written to the enacted curricula: The intermediary role of middle school mathematics teachers in shaping students' opportunity to learn. *School Science and Mathematics*, 106(4), 191-201.
- Thompson, P. W. (1994). Images of rate and operational understanding of the Fundamental Theorem of calculus. *Learning Mathematics*, 125-170. doi: 10.1007/978-94-017-2057-1_5
- Trolier, T. K., & Hamilton, D. L. (1986). Variables influencing judgments of correlational relations. *Journal of Personality and Social Psychology*, 50(5), 879-888. doi: 10.1037/0022-3514.50.5.879
- Truran, J. M. (1997). Understanding of association and regression by first year economics students from two different countries as revealed in responses to the same examination question. In J. Garfield & J. M. Truran (Eds.), *Research papers on stochastics educations from 1997* (pp. 205-212). Minneapolis, MN: University of Minnesota.
- Usiskin, Z. (2007). The case of the University of Chicago School Mathematics Project - Secondary Component. In C. R. Hirsch (Ed.), *Perspectives on the design and development of school mathematics curricula* (pp. 173-182). Reston, VA: NCTM.
- Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Boston: Kluwer Academic Publishers.
- Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 19(3),

231-241. doi: 10.1037/h0082908

- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. London: Batsford.
- Wasserman, E. A., Dorner, W. W., & Kao, S. F. (1990). Contributions of specific cell information to judgments of interevent contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 509-521. doi: 10.1037/0278-7393.16.3.509
- Watson, J., & Moritz, J. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, 37(2), 145-168. doi: 10.1023/a:1003594832397
- Watson, J. M., & Moritz, J. B. (1997). *Student analysis of variables in a media context*. Paper presented at the International Congress on Mathematics Education - 8 (ICME), Seville, Spain.
- Wavering, M. J. (1989). Logical reasoning necessary to make line graphs. *Journal of Research in Science Teaching*, 26(5), 373-379. doi: 10.1002/tea.3660260502
- Willoughby, S. (2010). Reflections on five decades of curriculum controversies. In B. J. Reys & R. E. Reys (Eds.), *Mathematics curriculum: Issues, trends, and future directions (72nd Yearbook of the National Council of Teachers of Mathematics)* (pp. 77-85). Reston, VA: National Council of Teachers of Mathematics.
- Woodward, A. (1994). Textbook. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2 ed., Vol. 2, pp. 6366-6371): BPC Wheatons, Exeter.
- Wright, J. C., & Murphy, G. L. (1984). The utility of theories in intuitive statistics: The robustness of theory-based judgments. *Journal of Experimental Psychology: General*, 113(2), 301-322. doi: 10.1037/0096-3445.113.2.301
- Zieffler, A. S. (2006). *A longitudinal investigation of the development of college students' reasoning about bivariate data during an introductory statistics course*. (Doctoral dissertation), University of Minnesota, Minneapolis, MN. Retrieved from <http://search.proquest.com/docview/305305165>

APPENDIX A: CCSSM'S MATHEMATICAL PRACTICES CODING FRAMEWORK

CCSSM Statement	Descriptors of the coding framework
<p><i>Make sense and persevere on solving problems</i> (MP1). [S]tudents start by explaining to themselves the meaning of a problem and looking for entry points to its solution. They analyze givens, constraints, relationships, and goals. They make conjectures about the form and meaning of the solution and plan a solution pathway rather than simply jumping into a solution attempt. They consider analogous problems, and try special cases and simpler forms of the original problem in order to gain insight into its solution. They monitor and evaluate their progress and change course if necessary... they check their answers to problems using a different method, and they continually ask themselves, "Does this make sense?" They can understand the approaches of others to solving complex problems and identify correspondences between different approaches. (CCSSI, 2010, p. 6)</p>	<p>Perseverance when solving a problem is only observable in practice, not in the written direction of a task itself. For this study, I considered the first component: make sense of problem. I identified if the task offers opportunities for students to do the preparatory work before jumping into solving the task. I also examined if the task asks students to approach one problem with multiple methods to check their solution. Particularly, I determined if the task requires students to formulate and clarify problems and situations such as to: (a) construct a verbal or symbolic statement or a question in which a mathematical problem goal can be specified, (b) design an appropriate statistical experiment to solve a stated problem or to specify the data and range of data needed.</p>
<p><i>Reason abstractly and quantitatively</i> (MP2). [S]tudents make sense of quantities and their relationships in problem situations. They bring two complementary abilities ...: the ability to decontextualize—to abstract a given situation and represent it symbolically and manipulate the representing symbols as if they have a life of their own, without necessarily attending to their referents—and the ability to contextualize, to pause as needed during the manipulation process in order to probe into the referents for the symbols involved. Quantitative reasoning entails habits of ... considering the units involved; attending to the meaning of quantities, not just how to compute them; and knowing and flexibly using different properties of operations and objects. (CCSSI, 2010, p. 6)</p>	<p>I examined whether a task provides opportunities for students to: (a) make sense of quantities (not merely numbers), (b) reason about the relationship between two quantities, and (c) reason abstractly with symbols and formula particularly in algebra.</p>
<p><i>Construct viable arguments and critique the reasoning of others</i> (MP3). [S]tudents understand and use stated assumptions, definitions, and previously established results in constructing arguments. They make conjectures and build a logical progression of statements to explore the truth of their conjectures. They are able to analyze</p>	<p>I examine whether a task provides opportunities for students to argue or critique. Specifically, I look for the performance expectations from the task that ask students to (a) verify the computational correctness of a solution, or justify a step in the solution, (b) identify information relevant to verify or disprove a conjecture, (c) argue the truth of a</p>

<p>situations by breaking them into cases, and can recognize and use counterexamples. They justify their conclusions, communicate them to others, and respond to the arguments of others. They reason inductively about data, making plausible arguments that take into account the context from which the data arose. [S]tudents are also able to compare the effectiveness of two plausible arguments, distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is. ...Students learn to determine domains to which an argument applies. Students ... can listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments. (CCSSI, 2010, pp. 6-7)</p>	<p>conjecture or construct a plausible argument, (d) identify a contradiction (something that is never true), (e) critique a written or spoken mathematical idea, solution, result, or method for solving a problem and the efficiency of the method or similarly critique an algorithm and its efficiency.</p>
<p><i>Model with mathematics</i> (MP4). [S]tudents can apply the mathematics they know to solve problems arising in everyday life, society, and the workplace. [S]tudents who can apply what they know are comfortable making assumptions and approximations to simplify a complicated situation, realizing that these may need revision later. They are able to identify important quantities in a practical situation and map their relationships using such tools as diagrams, two-way tables, graphs, flowcharts and formulas. They can analyze those relationships mathematically to draw conclusions. They routinely interpret their mathematical results in the context of the situation and reflect on whether the results make sense, possibly improving the model if it has not served its purpose. (CCSSI, 2010, p. 7)</p>	<p>I considered if a task provides opportunities for students to use mathematics to deal with real-life problems. I look for performance expectations from the task that attends to formulating and clarifying problems and situations such as tasks that ask students to (a) construct a verbal or symbolic statement of a real world or other situations, (b) simplify a real world or other problem situation by selecting aspects and relationships to be captured in a representation modeling the situation, (c) select or construct a mathematical representation of a problem (real-world or other problem situation plus a related question/goal), and (d) develop notations or terminologies to record actions and results of real-world or other mathematizable situations.</p>
<p><i>Use appropriate tools strategically</i> (MP5). [S]tudents consider the available tools when solving a mathematical problem. These tools might include pencil and paper, concrete models, a ruler, a protractor, a calculator, a spreadsheet, a computer algebra system, a statistical package, or dynamic geometry software. Proficient students are sufficiently familiar with tools appropriate for their grade or course to make sound decisions about when each of these tools might be helpful, recognizing both the insight to be gained and their limitations. They detect possible errors by strategically using estimation and other mathematical knowledge. When making mathematical models, they know that</p>	<p>Inherent in almost every task are different approaches in terms of using tools such as paper and pencil or technological tools to solve the task. In order to avoid every task being coded as offering students exposure to this practice, I examined if the task mentions about the selection of tools used to solve the task. In addition, a task will be coded to meet this practice if inherent in the task, students have to use a technological tool, but not merely for calculation. For example, the task asks students to carry out a simulation (e.g., find a probability of a Monte Carlo simulation). In this situation, the technological tool is used to deepen understanding.</p>

<p>technology can enable them to visualize the results of varying assumptions, explore consequences, and compare predictions with data. [S]tudents ... are able to identify relevant external mathematical resources, such as digital content located on a website, and use them to pose or solve problems. They are able to use technological tools to explore and deepen their understanding of concepts. (CCSSI, 2010, p. 7)</p>	
<p><i>Attend to precision (MP6).</i> [S]tudents try to communicate precisely to others. They try to use clear definitions in discussion with others and in their own reasoning. They state the meaning of the symbols they choose, including using the equal sign consistently and appropriately. They are careful about specifying units of measure, and labeling axes to clarify the correspondence with quantities in a problem. They calculate accurately and efficiently, express numerical answers with a degree of precision appropriate for the problem context. (CCSSI, 2010, p. 7)</p>	<p>I looked at the intent of tasks in the set-up phase in the materials. I determine if the task asks students to: calculate, to measure, to use specialized terms, symbols. In particular, I look for performance expectations that ask students to: (a) use equipment to measure, (b) compute/calculate with or without instruments, (c) graph with scale with or without technology/device, (d) collect data by surveys, samples, measurements, etc., (e) develop or select, using notations, terminologies to record actions and results in dealing with real-world or other mathematizable situations, and (f) describe the characteristics of a formal algorithm or solution procedure.</p>
<p><i>Look for and make use of structure (MP7).</i> [S]tudents look closely to discern a pattern or structure. They also can step back for an overview and shift perspective. They can see complicated things, such as some algebraic expressions, as single objects or as being composed of several objects. (CCSSI, 2010, p. 8)</p>	<p>In this study, I determined if a task provides opportunities for students to look for a structure within the task. For example, I examined task performance expectations that ask students to: (a) fit a curve of given type to a set of data (only if students are not told what kind of curve to fit), (b) classify mathematical objects by implicit criteria (e.g., geometric shapes), (c) predict a number, pattern, outcome, etc., that will result from an operation, procedure or experiment before it is actually performed.</p>
<p><i>Look for and express regularity in repeated reasoning (MP8).</i> [S]tudents notice if calculations are repeated, and look both for general methods and for shortcuts. As they work to solve a problem, mathematically proficient students maintain oversight of the process, while attending to the details. They continually evaluate the reasonableness of their intermediate results. (CCSSI, 2010, p. 8)</p>	<p>I determined if a task offers students opportunities to draw from a series of similar situations a general technique, strategy or algorithm to use in a class of problems. I analyzed task performance expectations that ask students to: (a) describe the effect of a change in a situation (e.g., the effect on its graph of changing a parameter), (b) develop a formal algorithm for computation or a formal solution procedure for problems of a specified class or type, (c) identify a class of problems for which a formal solution procedure is appropriate, (d) generalize the solution, the strategy, or the algorithm of a specific problem, and (e) abstract the common elements from multiple related situations.</p>

VITA

Dung Tran was born on June 23, 1981 in Hue city, Vietnam, where he pursued most of his education through the Master's degree. He is a son of Trần Ninh and Mai Thị Én in a big family including ten siblings. He earned a Bachelor in Mathematics and Master of Education in Mathematics Education at Hue University, Vietnam. He also finished a Ph.D. in Learning, Teaching, and Curriculum specializing in mathematics education and a Ph.D. Minor in Statistics at the University of Missouri – Columbia, U.S. His interest includes: mathematical modeling, learning trajectories, statistical reasoning and thinking, and secondary curriculum development. Dung taught at the Mathematics Department in Hue University College of Education for six years before coming to U.S. to pursue his Ph.D. program. He has accepted a position as a Research Associate at Friday Institution for Educational Innovation at North Carolina State University.