

HOW DO MATHEMATICS TEACHERS TEACH STATISTICS? A
COMPREHENSIVE EXAMINATION OF STATISTICAL PROBLEM-SOLVING
PROCESS AND INSTRUCTIONAL QUALITY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

VUSLAT ŞEKER AKIN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF ELEMENTARY EDUCATION

JULY 2023

Approval of the thesis:

HOW DO MATHEMATICS TEACHERS TEACH STATISTICS? A
COMPREHENSIVE EXAMINATION OF STATISTICAL PROBLEM-SOLVING
PROCESS AND INSTRUCTIONAL QUALITY

submitted by **VUSLAT ŞEKER AKIN** in partial fulfillment of the requirements for
the degree of **Doctor of Philosophy in Elementary Education, the Graduate
School of Social Sciences of Middle East Technical University** by,

Prof. Dr. Sadettin KİRAZCI
Dean
Graduate School of Social Sciences

Prof. Dr. Semra SUNGUR
Head of Department
Department of Elementary Education

Prof. Dr. Erdiñ ÇAKIROĞLU
Supervisor
Department of Mathematics and Science Education

Assoc. Prof. Dr. Engin ADER
Co-Supervisor
Boğaziçi University
Department of Mathematics and Science Education

Examining Committee Members:

Prof. Dr. Mine İŞİKSAL BOSTAN (Head of the Examining
Committee)
Middle East Technical University
Department of Mathematics and Science Education

Prof. Dr. Erdiñ ÇAKIROĞLU (Supervisor)
Middle East Technical University
Department of Mathematics and Science Education

Prof. Dr. Emine ERKTİN
Boğaziçi University
Department of Mathematics and Science Education

Prof. Dr. Didem AKYÜZ
Middle East Technical University
Department of Mathematics and Science Education

Assist. Prof. Dr. Elçin EMRE AKDOĞAN
TED University
Department of Mathematics and Science Education



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Vuslat ŞEKER AKIN

Signature:

ABSTRACT

HOW DO MATHEMATICS TEACHERS TEACH STATISTICS? A COMPREHENSIVE EXAMINATION OF STATISTICAL PROBLEM-SOLVING PROCESS AND INSTRUCTIONAL QUALITY

ŞEKER AKIN, Vuslat

Ph.D., The Department of Elementary Education

Supervisor: Prof. Dr. Erdinç ÇAKIROĞLU

Co-supervisor: Assoc. Prof. Dr. Engin ADER

July 2023, 516 pages

This qualitative case study examines the statistical teaching practices in middle school mathematics classrooms. It investigates discipline-specific teaching practices by analyzing the statistical content covered and the approaches to the statistical problem-solving processes. Moreover, it explores the quality of statistics instruction within the broader context of mathematics, considering four dimensions: The depth of mathematical and statistical concepts covered, the richness of mathematics and statistics, the occurrence of errors and imprecision, and student engagement in meaning-making and reasoning during statistics instruction within the context of mathematics. The study focuses on the teaching practices of two middle school mathematics teachers, where data was collected through classroom observations, interviews, and classroom artifacts. The data was analyzed using the Transcript Coding Protocol Using GAISE Framework and Mathematical Quality of Instruction Framework. The findings show that the statistical teaching practices were primarily traditional. The emphasis was on the data analysis element of the statistical problem-solving process, focusing on calculating measures of central tendency and constructing

graphs, while other elements received less attention. Moreover, the teaching practices lacked a comprehensive approach to the statistical problem-solving cycle. The study also reveals that the quality of statistics instruction in the observed classrooms was generally typical, with specific lessons demonstrating higher quality. Furthermore, the teaching practices centered around mathematical and statistical activities with mostly error-free and precision. The richness of mathematics and statistics was generally low to mid-quality. Working with students and mathematics and statistics generally indicated lower quality. However, opportunities for students to engage in meaning-making and reasoning were limited.

Keywords: Statistical Teaching Practices, Statistical Problem-Solving Processes, Instructional Quality, Middle School Mathematics Teachers, Qualitative Case Study

ÖZ

MATEMATİK ÖĞRETMENLERİ İSTATİSTİĞİ NASIL ÖĞRETİYOR? İSTATİSTİKSEL PROBLEM ÇÖZME SÜRECİ VE ÖĞRETİM KALİTESİ ÜZERİNE KAPSAMLI BİR İNCELEME

ŞEKER AKIN, Vuslat

Doktora, İlköğretim Bölümü

Tez Yöneticisi: Prof. Dr. Erdiñ ÇAKIROĞLU

Ortak Tez Yöneticisi: Doç. Dr. Engin ADER

Temmuz 2023, 516 sayfa

Bu nitel durum çalışması, ortaokul matematik sınıflarındaki istatistiksel öğretim uygulamalarını incelemektedir. Çalışma, kapsanan istatistiksel içeriği ve istatistiksel problem çözme süreçlerine yaklaşımları analiz ederek disipline özgü öğretim uygulamalarını araştırmaktadır. Ayrıca, istatistik öğretiminin kalitesini, matematiğin daha geniş bağlamı içinde, kapsanan matematiksel ve istatistiksel kavramların derinliği, matematik ve istatistiğin zenginliği, hataların ve belirsizliklerin ortaya çıkması ve matematik bağlamında istatistik öğretimi sırasında öğrencilerin anlam oluşturma ve akıl yürütmeye katılımı şeklinde dört boyut göz önünde bulundurarak araştırmaktadır. Çalışma, iki ortaokul matematik öğretmenin öğretim uygulamalarına odaklanmakta olup, veriler sınıf içi gözlemler, görüşmeler ve sınıf içi eserler aracılığıyla toplanmıştır. Veriler, GAISE Çerçevesi ile Oluşturulmuş Transkript Kodlama Protokolü ve Matematiksel Öğretim Kalitesi Çerçevesi ile analiz edilmiştir. Bulgular, merkezi eğilim ölçülerinin hesaplanması ve grafiklerin oluşturulmasına odaklanılarak istatistiksel problem çözme sürecinin veri analizi unsuruna vurgu yapılırken, diğer unsurlara daha az dikkat etmesiyle istatistik öğretim

uygulamalarının daha geleneksel olduğunu göstermektedir. Öğretim uygulamaları istatistiksel problem çözme döngüsüne yönelik kapsamlı bir yaklaşımdan yoksundu. Çalışma gözlemlenen sınıflardaki istatistik öğretiminin kalitesinin genel olarak tipik olduğunu, belirli derslerin daha yüksek kalite gösterdiğini ortaya koymaktadır. Öğretim uygulamaları çoğunlukla hatasız ve belirsiz olmayan matematiksel ve istatistiksel etkinlikler etrafında toplanmıştır. Matematik ve istatistiğin zenginliği genellikle düşük ila orta kalitedeydi. Öğrenciler ile matematik ve istatistik çalışmaları genellikle daha düşük kaliteye işaret etmektedir. Bununla birlikte, öğrencilerin anlam oluşturma ve akıl yürütme fırsatları sınırlıydı.

Anahtar Kelimeler: İstatistiksel Öğretim Uygulamaları, İstatistiksel Problem Çözme Süreçleri, Öğretim Kalitesi, Ortaokul Matematik Öğretmenleri, Nitel Durum Çalışması

To my family



ACKNOWLEDGMENTS

I would like to take this opportunity to express my deepest gratitude to the people who have contributed significantly to the completion of my dissertation. Their support, guidance, and encouragement have been invaluable throughout this challenging journey.

First and foremost, I would like to express my heartfelt appreciation to my supervisor, Prof. Dr. Erdiñ Çakırođlu. His unwavering dedication, expertise, and constant encouragement have been crucial in shaping the direction of my research. I am truly grateful for his invaluable guidance, insightful feedback, and tireless efforts to help me reach my full potential. You have always been available to answer my questions and provide me with support, even when I was feeling overwhelmed or discouraged.

I would also like to express my gratitude to my co-supervisor, Assoc. Prof. Dr. Engin Ader, whose expertise and valuable insights have contributed to the development of my dissertation. His support and guidance have been instrumental in refining my ideas and improving the quality of my work.

I would also like to express my sincere gratitude to my committee members, Prof. Dr. Emine Erktin, Prof. Dr. Mine Işıksal Bostan, Prof. Dr. Didem Akyüz, and Assist. Prof. Dr. Elçin Emre Akdođan. Their expertise, critical feedback, and constructive suggestions were invaluable in shaping the final outcome of this research. I am deeply grateful for their time, dedication, and contributions to this project.

I would like to extend my sincere appreciation to my friends, who have been there for me throughout this research journey. Your constant support, friendship, and positive energy have made a significant impact on my success. To my ODTU friends, Esra, Tuğçe, thank you for your unwavering encouragement, insightful discussions, and shared experiences. Your friendship has been a source of inspiration and motivation,

and I am grateful for the laughter and support we have shared. Nurbanu, I am so grateful for your belief in me and your unwavering support. You have been there for me through thick and thin, and I know that I can always count on you. I believe that we will be on our path together for the rest of our lives. To my Boğaziçi friends, Kübra, Merve, Beyhan, and Melek, thank you for your friendship, companionship, and steadfast belief in my abilities. Banu, my sister-like cousin, encouragement, comforting words, and belief in my capabilities have been invaluable.

I would like to express my deepest gratitude to my husband, Nejat, whose support, patience, and understanding have been the pillars of strength throughout this journey. His willingness to listen to my concerns, provide comfort during difficult times, and offer insightful perspectives has been invaluable in keeping me motivated and focused.

Lastly, I would like to express my heartfelt appreciation to my beloved family members: my father, İbrahim, and my mother, Emine, whose love, encouragement, and never-ending belief in me have been the driving force behind my accomplishments. I am eternally grateful for their continuous support, understanding, and sacrifices they have made on my behalf. I would also like to thank my brothers Murat and Fuat for their constant support and encouragement and for being a source of inspiration in my life. I am grateful for your presence and the special bond we share. I know that I can always count on you to be there for me, no matter what. You are my biggest supporters, and I am so lucky to have you in my life. I love you so much!

TABLE OF CONTENTS

PLAGIARISM	iii
ABSTRACT	iv
ÖZ	vi
DEDICATION	viii
ACKNOWLEDGMENTS	ix
TABLE OF CONTENTS	xi
LIST OF TABLES	xiv
LIST OF FIGURES	xvi
LIST OF ABBREVIATIONS	xxi
CHAPTERS	
1. INTRODUCTION	1
1.1. Statistics and Mathematics	2
1.1.1. Contrasting Mathematics and Statistics: Context, Variation, Reasoning, and Problem-Solving	3
1.1.2. The Complementary Nature of Mathematics and Statistics in School Mathematics	5
1.2. Mathematical and Statistical Teaching Practices	7
1.2.1. Big Ideas in Statistics	7
1.2.2. Statistical Problem-Solving Process	8
1.2.3. Instructional Quality in Mathematics Education	9
1.3. The Purpose of The Study	11
1.4. Research Questions	11
1.5. The Significance of the Study	12
1.6. Definition of Terms	16
1.7. Delimitations	19
1.8. Organizations of the Study	20
2. LITERATURE REVIEW	22
2.1. Teaching and Learning Statistics	22
2.1.1. Statistical Literacy, Reasoning, and Thinking	23
2.1.2. Big Ideas in Statistics	25
2.1.3. Statistical Investigations	39
2.1.4. Curriculum and Curriculum Expectations	55

2.1.5. Related Literature on Statistical Investigations.....	68
2.2. Mathematics and Statistics Classroom Practices	87
2.2.1. Classroom Discourse.....	92
2.2.2. Teaching Triad as the Conceptual Background	94
2.2.3. Instructional Quality in Mathematics Education.....	96
2.3. Summary of Literature Review	133
3. METHODOLOGY	136
3.1. Research Design.....	136
3.2. Participants and the Context of The Study.....	140
3.2.1. Selection of the Participants.....	140
3.2.2. Data Collection Tools for the Context of the Study.....	142
3.2.3. Teachers and Teaching Context	144
3.3. Data Collection.....	162
3.3.1. Data Collection Procedure	162
3.3.2. Gaining Entry to the Field and Ethical Considerations.....	164
3.3.3. The Role of the Researcher	166
3.3.4. The Data Collection Tools	167
3.4. Data Analyses.....	171
3.4.1. Schematic Diagram	172
3.4.2. Video Analysis	173
3.5. Validity and Reliability	187
4. FINDINGS	192
4.1. Within Case Analyses	193
4.1.1. Case 1: Ali's Teaching Practices.....	193
4.1.2. Case 2: Esra's Teaching Practices.....	272
4.2. Cross-Case Analysis.....	350
4.2.1. Statistical Content	350
4.2.2. Statistical Problem-Solving Process.....	352
4.2.3. Mathematical and Statistical Quality of Instruction: Segment Level Analysis.....	357
4.2.4. Mathematical and Statistical Quality of Instruction: Lesson Level Analysis.....	379
5. DISCUSSION AND CONCLUSION	380
5.1. Statistical Teaching Practices.....	381
5.1.1. Statistical Content	381
5.1.2. Statistical Problem-Solving Process.....	385

5.2. The Quality of Statistics Instruction within Broader Mathematics Instruction.....	394
5.2.1. The Richness of Mathematics and Statistics	395
5.2.2. Working with Students and Mathematics and Statistics	406
5.2.3. Errors and Imprecision	411
5.2.4. Student Participation in Meaning-Making and Reasoning	418
5.2.5. Overall discussion on MQI	426
5.3. Implications for the Practice and Recommendations	428
5.4. Limitations of the Study	431
5.5. Implications for Future Research	433
REFERENCES.....	435
APPENDICES	
A. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS COMMITTEE.....	470
B. DATA COLLECTION TOOLS	471
B1. Field Notes (Adapted from Jacobbe, 2007, p.122).....	471
B2. Interviews	471
C. TRANSCRIPT OF AN EPISODE	478
D. CURRICULUM VITAE	485
E. TURKISH SUMMARY / TÜRKÇE ÖZET	490
F. THESIS PERMISSION FORM / TEZ İZİN FORMU	516

LIST OF TABLES

Table 2.1 Three Developmental Levels in the Formulate Statistical Investigative Questions Process Component.....	45
Table 2.2 Three Developmental Levels in the Collect Data/Consider Data Process Component.....	48
Table 2.3 Exploratory Data Analysis vs. Formal Probability-Based Inference	51
Table 2.4 Three Developmental Levels in the “Analyze the Data” Process Component.....	52
Table 2.5 Three Developmental Levels in the “Interpret Results” Process Component.....	55
Table 2.6 Grade-level Expectations in Data Analysis and Probability Content Standard	56
Table 2.7 Learning Outcomes Related to Statistics in the Primary School Mathematics Curriculum.....	61
Table 2.8 Learning Outcomes Related to Statistics in The Middle School Mathematics Curriculum.....	63
Table 2.9 The Overview of The Studies on Formulating Statistical Investigative Questions.....	69
Table 2.10 The Overview of the Studies on Collecting/Considering Data	74
Table 2. 11 The Overview Of The Studies on Data Analysis	77
Table 2.12 The Overview Of The Studies on Interpreting The Results.....	80
Table 2.13 The Overview of The Studies on The Entire Statistical Problem-Solving Process	83
Table 2.14 The Overview of The Studies on The Entire Statistical Problem-Solving Process (Teacher-Focused).....	86
Table 2.15 Mathematics Teaching Practices and Corresponding Statistics Teaching Practices	89
Table 2.16 Process of Development and Theoretical Foundations of Example Frameworks.....	101
Table 2.17 The Characteristics of Example Frameworks	102
Table 2.18 Description of the Four Main Dimensions of the Mathematical Quality of Instruction	110
Table 2.19 The Overview of the Studies conducted on the Mathematical Quality of Instruction	115
Table 2.20 The Overview of the Studies Conducted on the Relationship between Teacher Knowledge and Instructional Quality	129
Table 2.21 The Overview of the Studies Conducted on the Relationship between Statistical Knowledge for Teaching and Teaching Practices.....	131
Table 3.1 Demographic Background of 7A Class.....	157
Table 3.2 The Descriptive Statistics for Classroom Context Variables.....	158
Table 3.3 Demographic Background of 7D Class.....	160
Table 3.4 The Descriptive Statistics and Correlations for Classroom Context Variables	161

Table 3.5 Timeline of Data Collection Process	163
Table 3.6 Weekly Observation Period	164
Table 3.7 Transcript Coding Protocol Using GAISE Framework	176
Table 3.8 Coding Decisions for the Episode with the Transcript Coding Protocol using the GAISE Framework.....	179
Table 3.9 Example of Scoring Decisions of Codes in MQI Corresponded to a Segment.....	185
Table 3.10 The Scoring Criteria for the Whole-Lesson MQI Code.....	187
Table 4.1 Statistical Content in Ali's Instruction	195
Table 4.2 Ali's Statistical Problem-Solving Process: The Occurrences of the Level of Elements	196
Table 4.3 Ali's Statistical Problem-Solving Process: The Co-occurrences of the Level of Elements	203
Table 4.4 "The Richness of the Mathematics and Statistics" Dimension Scores of Ali's Instructional Segments	207
Table 4.5 "Working with Students and Mathematics and Statistics" Dimension Scores in Ali's Instructional Segments.....	230
Table 4.6 The "Errors and Imprecision" Dimension Scores in Ali's Instructional Segments.....	242
Table 4.7 "Student Participation in Meaning Making and Reasoning" Dimension Scores in Ali's Instructional Segments	256
Table 4.8 Statistical Content in Esra's Instruction.....	274
Table 4.9 Esra's Statistical Problem-Solving Process: The Occurrences of the Level of Elements	276
Table 4.10 Esra's Statistical Problem-Solving Process: The Co-occurrences of the Level of Elements	282
Table 4.11 "The Richness of the Mathematics and Statistics" Dimension Scores in Esra's Instruction	287
Table 4.12 "Working with Students and Mathematics and Statistics" Dimension Scores in Esra's Instruction	310
Table 4.13 "Errors and Imprecision" Dimension Scores in Esra's Instruction	321
Table 4.14 "Student Participation in Meaning-Making and Reasoning" Dimension Scores in Esra's Instruction	335

LIST OF FIGURES

Figure 2.1 Representations of Statistical Literacy, Reasoning, and Thinking	23
Figure 2.2 Big Ideas of Statistics	27
Figure 2.3 First Dimension of the 4-Dimensional Framework for Statistical Thinking in Empirical Enquiry	41
Figure 2.4 Statistical Problem-Solving Process	42
Figure 2.5 Instructional Triangle.....	96
Figure 2.6 Prototypical Process in Developing and Using a Framework	98
Figure 2.7 A Continuum of Classifying Frameworks and Observational Instruments.....	98
Figure 2.8 Research Model	120
Figure 3.1 Multiple-Case Study Procedures	139
Figure 3.2 Procedures for Preparing Observations	168
Figure 3.3 An Example from a Teacher's Schematic Diagram	173
Figure 3.4 Coding Example of Episodes in a Lesson on Transcript Coding Protocol Using GAISE Framework	178
Figure 3.5 Coding Example of Segments in A Lesson on The MQI	184
Figure 4.1 Example Task of “Formulation of Questions” Element at Level 1 (Varışlı, 2017, p.153)	198
Figure 4.2 “The Richness of the Mathematics and Statistics” Dimension in Ali's Instruction	208
Figure 4.3 The Distribution of Scores Assigned to Lesson Segments for the “Linking between Representation” Code in Ali's Instruction	209
Figure 4.4 Graph Task 1 (Varışlı, 2017, p.150)	210
Figure 4.5 Graph Task 2.....	211
Figure 4.6 Graph Task 3.....	212
Figure 4.7 The Distribution of Scores Assigned to Lesson Segments for “Explanation” Code in Ali's Instruction.....	214
Figure 4.8 Average Task 1	216
Figure 4.9 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Sense-Making” Code in Ali's Instruction ..	217
Figure 4.10 The Distribution of Scores Assigned to Lesson Segments for “Multiple Procedures or Solution Methods” Code in Ali's Instruction	221
Figure 4.11 The Distribution of Scores Assigned to Lesson Segments for “Patterns and Generalization” Code in Ali's Instruction.....	223
Figure 4.12 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Language” Code in Ali's Instruction.....	224
Figure 4.13 The Distribution of Scores Assigned to Lesson Segments for “Overall Richness of Mathematics and Statistics” Code in Ali's Instruction	226
Figure 4.14 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “The Richness of Mathematics and Statistics” Dimension in Ali's Statistics Instruction.....	228

Figure 4.15 “Working with Students and Mathematics and Statistics” Dimension in Ali's Instruction.....	231
Figure 4.16 The Distribution of Scores Assigned to Lesson Segments for “Remediation of Student Errors and Difficulties” Code in Ali's Instruction	232
Figure 4.17 Graph Task 4 (Varışlı, 2017, p.155)	234
Figure 4.18 The Distribution of Scores Assigned to Lesson Segments for “Teacher Uses Student Mathematical and Statistical Contributions” in Ali's Instruction	236
Figure 4.19 The Distribution of Scores Assigned to Lesson Segments for “Overall Working with Students and Mathematics and Statistics” in Ali's Instruction	239
Figure 4.20 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “The Working with Students and Mathematics and Statistics” Dimension in Ali's Statistics Instruction	240
Figure 4.21 “Errors and Imprecision” Dimension in Ali's Instruction	243
Figure 4.22 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Content Errors” Code in Ali's Instruction	244
Figure 4.23 Graph Task 5 (Varışlı, 2017, p.162)	245
Figure 4.24 Teacher Construction of a Graph of a Data Set “4, 4, 4, 4, 4, 4”	247
Figure 4.25 The Distribution of Scores Assigned to Lesson Segments for “Imprecision in Language or Notation” Code in Ali's Instruction	248
Figure 4.26 The Distribution of Scores Assigned to Lesson Segments for “Lack of Clarity in the Presentation of Mathematical and Statistical Content” Code in Ali's Instruction	250
Figure 4.27 Average Example: The Number of Shoes (Varışlı, 2017, p.162)	250
Figure 4.28 A Circle Graph Example of Ömer's Games [Ali 07.05.2018-1 00:00-15:00]	252
Figure 4.29 The Distribution of Scores Assigned to Lesson Segments for “Overall Errors and Imprecision” Code in Ali's Instruction	253
Figure 4.30 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in The” Errors and Imprecision” Dimension in Ali's Instruction	254
Figure 4.31 “Student Participation in Meaning Making and Reasoning” Dimension in Ali's Instruction	257
Figure 4.32 The Distribution of Scores Assigned to Lesson Segments for “Students Provide Explanations” Code in Ali's Instruction	258
Figure 4.33 The Distribution of Scores Assigned to Lesson Segments for “Student Mathematical and Statistical Questioning and Reasoning” Code in Ali's Instruction	261
Figure 4.34 The Distribution of Scores Assigned to Lesson Segments for the “Students Communicate about the Mathematics and Statistics of the Segment” Code in Ali's Instruction	263
Figure 4.35 The Distribution of Scores Assigned to Lesson Segments for “Task Cognitive Demand” Code in Ali's Instruction	265
Figure 4.36 The Distribution of Scores Assigned to Lesson Segments for ‘Students Work with Contextualized Problems’ Code in Ali's Instruction	267

Figure 4.37 The Distribution of Scores Assigned to Lesson Segments for “Overall Student Participation in Meaning-Making and Participation” Code in Ali’s Instruction	269
Figure 4.38 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Student Participation in Meaning-Making and Reasoning” Dimension in Ali’s Instruction	270
Figure 4.39 “The Richness of the Mathematics and Statistics” Dimension in Esra’s Instruction	288
Figure 4. 40 The Distribution of Scores Assigned to Lesson Segments for “Linking Between Representation” Code in Esra’s Instruction.....	289
Figure 4.41 The Distribution of Scores Assigned to Lesson Segments for “Explanations” Code in Esra’s Instruction	292
Figure 4.42 Example Task 1: Course Types (Varışlı, 2017, p.162).....	295
Figure 4.43 Example Task 2: Hospitalization (MoNe, 2017, p.234)	295
Figure 4.44 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Sense-making” Code in Esra’s Instruction	297
Figure 4.45 The Distribution of Scores Assigned to Lesson Segments for “Multiple Procedures or Solution Methods” Code in Esra’s Instruction...	301
Figure 4.46 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Language” Code in Esra’s Instruction	304
Figure 4.47 The Distribution of Scores Assigned to Lesson Segments for the “Overall Richness of Mathematics and Statistics” Code in Esra’s Instruction	306
Figure 4.48 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “The Richness of Mathematics and Statistics” Dimension in Esra’s Instruction	308
Figure 4.49 “Working with Students and Mathematics and Statistics” Dimension in Esra’s Instruction	311
Figure 4.50 The Distribution of Scores Assigned to Lesson Segments for “Remediation of Student Errors and Difficulties” Code in Esra’s Instruction	312
Figure 4.51 The Distribution of Scores Assigned to Lesson Segments for the “Teacher Uses Student Mathematical and Statistical Contributions” Code in Esra’s Instruction.....	315
Figure 4.52 The Distribution of Scores Assigned to Lesson Segments for the “Overall Working with Students and Mathematics and Statistics” Code in Esra’s Instruction	317
Figure 4.53 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Working with Students and Mathematics. and Statistics” Dimension in Esra’s Instruction.....	319
Figure 4.54 The “Errors and Imprecision” Dimension in Esra’s Instruction.....	322
Figure 4.55 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Content Errors” Code in Esra’s Instruction	323
Figure 4.56 Esra’s First Drawings for Describing the Mode	325
Figure 4.57 Esra’s Second Drawing for Describing the Mode	326
Figure 4.58 Esra’s Second and Third Drawings for Describing the Mode	327
Figure 4.59 The Distribution of Scores Assigned to Lesson Segments for “Imprecision in Language or Notation” Code in Esra’s Instruction	328

Figure 4.60 The Distribution of Scores Assigned to Lesson Segments for “Lack of Clarity in the Presentation of Mathematical and Statistical Content” Code in Esra’s Instruction.....	330
Figure 4.61 The Distribution of Scores Assigned to Lesson Segments for the “Overall Errors and Imprecision” Code in Esra’s Instruction	332
Figure 4.62 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Errors and Imprecision” Dimension in Esra’s Instruction	333
Figure 4.63 “Student Participation in Meaning Making and Reasoning” Dimension in Esra’s Instruction.....	336
Figure 4.64 The Distribution of Scores Assigned to Lesson Segments for “Students Provide Explanations” Code in Esra’s Instruction.....	337
Figure 4.65 The Distribution of Scores Assigned to Lesson Segments for “Student Mathematical and Statistical Questioning and Reasoning” Code in Esra’s Instruction.....	339
Figure 4.66 The Distribution of Scores Assigned to Lesson Segments for the “Students Communicate about the Mathematics and Statistics of the Segment” Code in Esra’s Instruction.....	340
Figure 4.67 The Distribution of Scores Assigned to Lesson Segments for “Task Cognitive Demand” Code in Esra’s Instruction.....	343
Figure 4.68 The Distribution of Scores Assigned to Lesson Segments for “Students Work with Contextualized Problems” Code in Esra’s Instruction	345
Figure 4.69 The Distribution of Scores Assigned to Lesson Segments for “Overall Student Participation in Meaning Making and Reasoning” Code in Esra’s Instruction.....	347
Figure 4.70 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Student Participation in Meaning-Making and Reasoning” Dimension in Esra’s Instruction	348
Figure 4.71 Graphs as the Statistical Content in Both Teachers’ Statistics Instruction	351
Figure 4.72 Measures of Central Tendency as the Statistical Content in Both Teachers’ Statistics Instruction	352
Figure 4.73 “Formulation of Questions” Element in Both Teachers’ Statistics Instruction	353
Figure 4.74 “Data Collection” Element in Both Teachers’ Statistics Instruction.....	354
Figure 4.75 “Data Analysis” Element in Both Teachers’ Statistics Instruction	354
Figure 4.76 “Interpretation” Element in Both Teachers’ Statistics Instruction	355
Figure 4.77 “The Classroom Work is Connected to Mathematics and Statistics” Code in Both Teachers’ Instruction	358
Figure 4.78 “Linking Between Representations” Code in Both Teachers’ Statistics Instruction.....	359
Figure 4.79 “Explanations” Code in Both Teachers’ Statistics Instruction.....	360
Figure 4.80 “Mathematical and Statistical Sense-Making” Code in Both Teachers’ Statistics Instruction	361
Figure 4.81 “Multiple Procedures or Solution Methods” Code in Both Teachers’ Statistics Instruction.....	362
Figure 4.82 “Patterns and Generalization” Code in Both Teachers’ Statistics Instruction	363

Figure 4.83 “Mathematical and Statistical Language” Code in Both Teachers’ Statistics Instruction	364
Figure 4.84 “Overall Richness of the Mathematics and Statistics” Code in Both Teachers’ Statistics Instruction	365
Figure 4.85 “Remediation of Student Errors and Difficulties” Code in Both Teachers’ Statistics Instruction	366
Figure 4.86 “Teacher Uses Student Mathematical and Statistical Contributions” Code in Both Teachers’ Statistics Instruction	367
Figure 4.87 “Overall Working with Students and Mathematics and Statistics” Code in Both Teachers’ Statistics Instruction	369
Figure 4.88 “Mathematical and Statistical Content Errors” Code in Both Teachers’ Statistics Instruction	370
Figure 4.89 “Imprecision in Language or Notation” Code in Both Teachers’ Statistics Instruction	371
Figure 4.90 “Lack of Clarity in Presentation of Mathematical and Statistical Content” Code in Both Teachers’ Statistics Instruction	372
Figure 4.91 “Overall Errors and Imprecision” Code in Both Teachers’ Statistics Instruction	373
Figure 4.92 “Students Provide Explanation” Code in Both Teachers’ Statistics Instruction	374
Figure 4.93 “Student Mathematical and Statistical Questioning and Reasoning” Code in Both Teachers’ Statistics Instruction	375
Figure 4.94 “Students Communicate about the Mathematics and Statistics of the Segment” Code in Both Teachers’ Statistics Instruction	376
Figure 4.95 “Task Cognitive Demand” Code in Both Teachers’ Statistics Instruction	377
Figure 4.96 “Students Work with Contextualized Problems” Code in Both Teachers’ Statistics Instruction	378
Figure 4.97 “Overall Student Participation in Meaning-Making and Reasoning” Code in Both Teachers’ Statistics Instruction	378

LIST OF ABBREVIATIONS

CAQDAS	Computer-Assisted Qualitative Data Analysis
CCSSI	Common Core State Standards Initiative
CCSSM	Common Core State Standards for Mathematics
CLASS	Classroom Assessment Scoring System
DA	Data Analysis
DC	Data Collection
DMEE	Dynamic Model of Educational Effectiveness
EDA	Exploratory Data Analysis
Exp	Explanations
FQ	Formulation of Questions
GAISE	Guidelines for Assessment and Instruction in Statistics Education
I	Interpretation
ILN	Imprecision in Language or Notion
IQA	Instructional Quality Assessment
ISTOF	International System of Teacher Observation and Feedback
LBR	Linking between Representations
LCPM	Lack of Clarity in Presentation of Mathematical and Statistical Content
LMT	Learning Mathematics for Teaching Project
M-Scan	Mathematics Scan
MCE	Mathematical and Statistical Content Errors
ML	Mathematical and Statistical Language
MoNE	Ministry of National Education
MPSM	Multiple Procedures and Solution Methods
MQI	Mathematical Quality of Instruction
MSM	Mathematical and Statistical Sense Making
NCTE	National Center for Teacher Effectiveness

NCTM	National Council of Teachers of Mathematics
OCASP	Overall Student Participation in Meaning Making and Reasoning
OE&I	Overall Errors and Imprecision
ORM	Overall Richness of Mathematics and Statistics
OWM	Overall Working with Students and Mathematics and Statistics
P&G	Patterns and Generalization
QIFL	The Quality of Instruction for Fraction Lessons
RSE&D	Remediation of Student Errors and Difficulties
RTOP	Reformed Teaching Observation Protocol
S	Student
SCM	Students Communicate about the Mathematics and Statistics of the Segment
SMQ&R	Student Mathematical and Statistical Questioning and Reasoning
SPE	Students Provide Explanations
SWCP	Students Work with Contextualized Problems
T	Teacher
TBD	Three Basic Dimensions
TCD	Task Cognitive Demand
TEDS-Instruct	Teacher Education Study in Mathematics-Instruct
TRU	Teaching for Robust Understanding
TUSMC	Teacher Uses Student Mathematical and Statistical Contributions
UTOP	UTeach Observation Protocol
PPDAC	Problem, Plan, Data, Analysis, Conclusion

CHAPTER 1

INTRODUCTION

Teaching and learning statistics have gained increased attention in mathematics education research in recent decades (Kazak et al., 2023). Statistics help students understand the world by processing quantitative information (Ben-Zvi & Garfield, 2008). Statistics education aims to make students “read, analyze, critique, and make inferences from distributions of data” (Shaughnessy, 2007, p.968). Therefore, statistics curricula worldwide focus on key ideas and concepts and statistical investigations, which are crucial for developing students’ statistical thinking, literacy, and reasoning (Ben-Zvi & Garfield, 2008; MacGillivray & Pereira-Mendoza, 2011). Recent research shows that employing statistical investigations and key ideas in teaching and learning statistics contributes to making sense of the world through statistics education rather than theoretical statistics, including p-values and confidence intervals (Watson & Smith, 2022). Hence, a statistics curriculum should allow students to engage in statistical practices with the key concepts from the early years (Bargagliotti et al., 2020).

Despite the increasing emphasis on statistical investigations, many statistics curricula and practices still prioritize procedures and mathematical aspects, which poses challenges for both teachers and learners (Leavy, 2015; Makar & Fielding-Wells, 2011). Although mastering procedural skills and comprehending mathematical relationships and probability are essential components of teaching statistics, the curricular goals should first focus on understanding statistical investigations in the context of their purpose and cyclic nature (Newton et al., 2011). Students can develop statistical thinking and reasoning skills by engaging in statistical investigations, ultimately achieving statistical literacy (Bargagliotti et al., 2020; Garfield & Ben-Zvi, 2008). Therefore, it is crucial to determine effective teaching practices in statistics education that correspond to curricular goals.

A written curriculum is implemented in classrooms through teachers' intentions and enactments, leading to student learning (Stein et al., 2007). Classroom practices originate from the intended curriculum, influenced by teachers' beliefs and knowledge (Eichler, 2011). Teachers' beliefs and knowledge as a part of teacher competence are critical factors in promoting high-quality instruction, which includes effective teaching practices (Blömeke et al., 2020). Several studies have explored the relationship between teachers' beliefs and knowledge of mathematics and the enacted curriculum referring to the classroom practices (e.g., Charalambous et al., 2012; Charalambous & Hill, 2012; Hill et al., 2008; Hill & Charalambous, 2012b). In recent years, there has been a growing interest in examining the knowledge and beliefs of teachers in statistics education, similar to mathematics education research (e.g., Burgess, 2011; Groth, 2007, 2013; Leavy, 2015). Although the related studies provide insights into statistics education research, they do not necessarily investigate how such beliefs and knowledge translate into teaching practices in statistics classrooms, which need to be explored to promote better statistical learning experiences.

Aside from the factors shaping classroom practices, several studies have focused on exploring and improving mathematical teaching practices (e.g., Boaler & Staples, 2008; Charalambous & Litke, 2018; Schoenfeld, 2013). However, there is a lack of literature on statistics teaching practices that promote better student learning (Eichler, 2011). Gerstenschlager (2019) suggests statistics teaching practices that complement the mathematics teaching practices proposed by NCTM (2014) to promote high-quality instruction. However, before comparing teaching practices in mathematics and statistics, it is essential to examine the similarities and differences between the two disciplines to enhance classroom practices in statistics education.

1.1. Statistics and Mathematics

The place of statistics in school mathematics has been discussed in literature due to the nature of statistics (Wild et al., 2018). Statistics is a discipline rather than a part of mathematics (Ben-Zvi & Garfield, 2004). The nature of statistics differs from mathematics in terms of context, variation, types of reasoning, and problem-solving (Burrill & Biehler, 2011; Groth & Meletiou-Mavrotheris, 2018). However, statistics is

one strand in school mathematics. Therefore, it is crucial to understand the advantages of teaching and learning statistics in mathematics classrooms. In the following section, I will discuss the contrasting features of statistics and mathematics and the complementary nature of mathematics and statistics in school mathematics.

1.1.1. Contrasting Mathematics and Statistics: Context, Variation, Reasoning, and Problem-Solving

This section presents a comparison between mathematics and statistics, focusing on their contrasting features in terms of context, variation, reasoning, and problem-solving.

1.1.1.1. Context

The context has been used in both fields. However, the purpose of using context is different for each field (Wild et al., 2018). Finding patterns, generalization, and abstraction are the primary aims of mathematics as a field (Cobb & Moore, 1997; Gattuso & Ottaviani, 2011). The ultimate purpose is to reach abstraction to understand the structure (Cobb & Moore, 1997). Mathematics uses context as a motivation for learners or a problem source for applications to fulfill the purpose of mathematics (Burrill & Biehler, 2011; Cobb & Moore, 1997). At one point, context could be minimized or is not used to understand mathematics' structure (Cobb & Moore, 1997). However, context is inevitable in statistics (Wild et al., 2018).

Like mathematics, statistics also deals with numbers, looks for patterns, and needs generalizations (Cobb & Moore, 1997). However, statistics uses data as numbers in a context and structured or unstructured complex data in sounds or pictures (Bargagliotti et al., 2020). The context is necessary to provide meaning to the data in statistics (Cobb & Moore, 1997; Gattuso & Ottaviani, 2011). Gattuso and Ottaviani (2011) emphasized the need for context to explore the nature of data for analyzing and reaching meaningful results. Watson et al. (2018) stated that researchers underline the necessity of context in statistics. However, most resources for teaching statistics do not consider context, and they consist of numbers without context to carry out the algorithms related to statistical concepts (e.g., arithmetic mean). Therefore, it is vital to understand the

nature of context for both disciplines. The made-up context could work in the mathematics classroom because of the purpose of mathematics but does not work in a statistics classroom due to the nature of statistics (Cobb & Moore, 1997). The nature of statistics depends on context, but mathematics gets rid of the context in the end for abstraction and generalization (Shaughnessy, 2007).

1.1.1.2. Variation

Variation is another concept creating differences between mathematics and statistics. The nature of mathematics is deterministic, while statistics is stochastic; indeterminist (Burrill & Biehler, 2011; Gal & Garfield, 1997). However, both disciplines consist of variation or variability concepts, but their nature is different (Gattuso & Ottaviani, 2011). In mathematics, variables depend on each other; the relationship among variables is subject to study within a part of mathematics. The variability in mathematics expects the exact fit of data on, for instance, a “function graph” (Burrill & Biehler, 2011, p.64). In statistics, variability is the core concept that the field deals with (Reading & Shaughnessy, 2004). The variation can be in different forms like measurement variation, variation in groups or individuals, or simply data (Cobb & Moore, 1997; Shaughnessy, 2007). Wild and Pfannkuch (1999) described considering variation as one fundamental issue in statistical thinking. They pointed out that noticing, acknowledging, measuring, modeling, explaining, and dealing with variation are necessary components of statistical thinking. The precise nature of mathematics does not deal with the noise in statistical problems (Burrill & Biehler, 2011).

1.1.1.3. Types of Reasoning

The types of reasoning are also different in mathematics and statistics. The nature of mathematics deals with patterns and abstraction (Delmas, 2004). The reasoning in mathematics is context-free; the context is used for reaching abstraction (Gattuso & Ottaviani, 2011). Mathematical reasoning is based on deduction. While working on mathematics, one needs to use “axioms and definitions” for solving mathematical problems through the processes of “modeling, optimizing, and proving” (p.126). In addition, certainty is essential to mathematical reasoning (Begg, 1997). However, statistical reasoning is based on data; it includes inductive reasoning (Delmas, 2004).

Uncertainty is expected in statistics. Different solutions can be induced from the same data (Gattuso & Ottaviani, 2011; Watson & Smith, 2022). Begg (1997) stated the processes used in statistical reasoning as classifying, describing, arguing, inferring, and proving. Delmas (2007) noted that statistical reasoning is complex because it also requires deductive reasoning due to the mathematical concepts used in statistics.

1.1.1.4. Problem-solving

Problem-solving is also different for mathematics and statistics (Shaughnessy, 2007). Polya defined four steps in the mathematical problem-solving process: Understand, plan, execute, and review. Mathematical problems are generally well-defined; the problem-solving process is mostly linear (Delmas, 2004). In addition, Delmas (2004) stated that real-world problems are present in mathematics instruction and are used as vehicles for making sense of abstract mathematical concepts due to the nature of mathematics. In statistical problem-solving, the problems are poorly defined because they deal with real problems or are a part of real problems (Shaughnessy, 2007; Wild & Pfannkuch, 1999). Therefore, the nature of the statistical problem is not linear. In addition, the statistical problem-solving process consists of the data collection part after the problem and planning sections. It is based on the data with variation and context, unlike problems in mathematical problem-solving (Shaughnessy, 2007). The statistical problem-solving process includes anticipating variability, while the mathematical problem-solving process requires deterministic answers (Bargagliotti et al., 2020).

1.1.2. The Complementary Nature of Mathematics and Statistics in School

Mathematics

Statistics and mathematics are separate disciplines, even if statistics are covered in school mathematics in K-12. Recognizing both disciplines' necessities and characteristics is necessary to learn and teach statistics effectively (Gattuso & Ottaviani, 2011). As mentioned, both fields differ in terms of the role of context, the problem-solving process, the presence of variability, and the types of reasoning (Burrill & Biehler, 2011; Groth & Meletiou-Mavrotheris, 2018). Recognizing all

differences and considering how both disciplines contribute to each other is crucial for better learning outcomes in statistics and mathematics.

Gattuso and Ottaviani (2011) discuss the importance of recognizing the advantages of covering statistics in school mathematics. First, statistics are contextualized and provide real problems to solve for students, increasing students' motivation in mathematics classrooms. Therefore, teachers can benefit from students' positive attitudes toward statistics to develop such attitudes toward mathematics. Also, communicating results is an essential part of the statistical problem-solving process under the interpretation of findings to answer the statistical investigative questions using collected data analyzed within the process. Communication as a part of curriculum standards in mathematics can benefit from the nature of statistics involving communicating findings. Furthermore, specific skills and understandings of mathematics are required in the statistical problem-solving process, including "arithmetic procedures, measurement skills, and proportional reasoning related to fractions, decimals, ratios, and percentages" (Gattuso, 2006; Gattuso & Ottaviani, 2011; Watson, 2016, p.133).

Gattuso and Ottaviani (2011) also focus on the contributions of statistical processes in improving students' mathematical skills (e.g., classifying). In both disciplines, the use of representations is crucial. Especially graphical representations are necessary to represent the data for the analysis in statistics (e.g., understanding the distributions) and to model the problems for the analysis of the problems (e.g., understanding the problems and leading to solutions). In addition, Wild et al. (2018) also discussed that mathematical thinking is necessary for dealing with statistics, especially when analyzing data due to the presence of mathematical concepts and abstraction in statistics. Statistics make use of mathematical practices like patterns and generalizations in data or data represented in graphs (Delmas, 2004). All in all, teachers should recognize the advantages of covering statistics in their mathematics lessons, but they should also understand the unique characteristics of each subject (Gattuso & Ottaviani, 2011).

It is essential first to examine the underlying similarities and differences between these disciplines to make sense of mathematics and statistics teaching practices. By exploring the distinct characteristics of mathematics and statistics as fields, it is possible to lay a solid foundation for understanding the subsequent discussion of their respective teaching practices. Therefore, in the following section, the mathematical and statistical teaching practices will be discussed in terms of big statistical ideas, statistical problem-solving processes, and the quality of instruction.

1.2. Mathematical and Statistical Teaching Practices

Research has identified effective teaching practices for mathematics, and it is crucial to understand these practices to improve student learning outcomes (Charalambous & Kyriakides, 2017; Copur-Gencturk, 2012). However, there is a lack of research on statistics teaching practices, likely due to the unique nature of statistics compared to mathematics (Eichler, 2011). Therefore, a better understanding of statistics teaching practices in mathematics classrooms is needed. Statistical teaching practices differ from mathematical education practices, including task goals, problem-solving approaches, representations, discourse issues, questioning methods, building fluency, productive struggle, and using evidence (Gerstenschlager, 2019). Understanding the similarities and differences between teaching practices in mathematics and statistics, including defining big ideas in statistics and the statistical problem-solving process, is essential to providing high-quality instruction.

1.2.1. Big Ideas in Statistics

The study of statistics in schools is focused on a set of big ideas, identified and discussed in the literature, that provide a conceptual framework for understanding data and drawing conclusions from it (Garfield & Ben-Zvi, 2004, 2008; Shaughnessy, 2019a; Watson et al., 2013). Garfield and Ben-Zvi (2004, 2008) have provided a list of eight big ideas in statistics that students typically encounter during their K-12 education, including data, distribution, trend, variability, models, association, samples and sampling, and inference, with randomness added in their later work. Watson et al. (2013) have also identified five big, interrelated statistical ideas: expectation, variation, distribution, informal inference, and randomness.

Understanding these big ideas is crucial for students to provide better learning and teaching experiences in statistics (Watson et al., 2013). It has been argued that big ideas, instead of mere procedures, are essential for students to learn to conduct statistical investigations and draw meaningful conclusions (Callingham et al., 2021). Shaughnessy (2019a) suggests that creating a list of big ideas in statistics can help students organize their understanding of statistical concepts and processes. Callingham et al. (2021) also indicate a curriculum based on a learning progression in statistics by considering key concepts or ideas using mathematical proficiencies like fluency, reasoning, problem-solving, and understanding to improve teaching and learning statistics. In conclusion, understanding the big ideas in statistics is essential for students to become statistically literate. Therefore, teaching practices need to be shaped by big ideas in statistics.

1.2.2. Statistical Problem-Solving Process

As discussed earlier, there has been a growing emphasis on the importance of an investigative approach to statistics education. This approach involves engaging students in complete statistical investigations or specific steps within statistical investigations to develop big ideas in statistics and gain a deeper understanding of the world around them (Watson & Callingham, 2015; Wild & Pfannkuch, 1999; Pfannkuch, 2018). By encouraging students to ask questions, gather data, analyze and interpret results, and draw conclusions, an investigative approach helps students develop statistical thinking, reasoning, and literacy.

One framework for statistical thinking in empirical inquiry is the Investigative Cycle developed by Wild and Pfannkuch (1999). The Investigative Cycle includes the PPDAC cycle - problem, plan, data, analysis, and conclusions - and provides a systematic approach to the statistical problem-solving process. The other framework GAISE (Guidelines for Assessment and Instruction in Statistics Education), developed by the American Statistical Association (ASA), incorporates a two-dimensional curriculum framework. The first dimension includes the stages of the statistical problem-solving process (Bargagliotti et al., 2020, p.13):

- (1) Formulating statistically investigative questions*
- (2) Collecting/considering data*
- (3) Analyzing the data*
- (4) Interpreting the results*

The second dimension consists of three developmental levels (A, B, and C) that focus on students' development of foundational statistics skills throughout the statistical problem-solving process, emphasizing the importance of statistical literacy. The GAISE framework is a curriculum document designed to guide the teaching and assessment of statistics education in K-12 settings. Watson et al. (2018) described the framework as the practice of statistics. Litke and Hill (2020) developed a protocol using the GAISE framework to explore discipline-related practices in teaching statistics. The protocol includes four components of the investigative process with three categories with increasing presence of the practices recommended by the GAISE framework. The protocol particularly suits elementary school statistics content and focuses on developmental level A.

Research has shown two possible ways to approach statistical investigations in classroom practices (Burgess, 2011). The first involves giving students data to explore and focusing on the data analysis and interpretation steps (e.g., Akkoç & Selçuk, 2017; Chick & Watson, 2002; Makar, 2014; Makar & Confrey, 2005; Watson & Moritz, 2000). The second involves starting with statistical investigative questions and following the remaining phases of the Investigative Cycle (e.g., Lavigne & Lajoie, 2007; Öz & Işıksal Bostan, 2021; Silva et al., 2022; Watson & English, 2015; 2017b). Both approaches engage students in a statistical problem-solving process to develop their statistical thinking, reasoning, and literacy (Burgess, 2011; Litke & Hill, 2020). An investigative approach helps students develop a deeper understanding of statistical concepts and their real-world applications.

1.2.3. Instructional Quality in Mathematics Education

Instructional quality in mathematics education is crucial to students' mathematics learning (Griffin et al., 2009) and is defined by the observable behavior of the teacher in the classroom (Blömeke et al., 2022). Good mathematics instruction is composed of various components which support the development of mathematical proficiencies like

reasoning, understanding, fluency, and problem-solving (Callingham et al., 2021). However, there is a continuous debate in the literature regarding what constitutes good teaching practices and how to achieve them (Fraenke et al., 2007; Schoenfeld, 2014). The main goal of examining instructional quality is to improve teaching to create opportunities for better student learning. Therefore, measuring and understanding teaching practices is necessary to achieve this goal (Bell et al., 2019; Blömeke et al., 2022; Hill & Grossman, 2013; Learning Mathematics for Teaching, 2011; Schlesinger & Jentsch, 2016; Schoenfeld, 2014; Stein & Lane, 1996).

The Mathematical Quality of Instruction (MQI) Framework was developed to improve mathematics instruction by understanding the quality of content-specific teaching practices in mathematics. It considers instruction as the interaction between the teacher, students, and content, described by the instructional triangle by Cohen et al. (2003). The MQI framework provides a comprehensive view of mathematics instruction through 21 segment-level codes that identify specific teaching practices in four dimensions: the richness of mathematics, working with students and mathematics, errors and imprecision, and student participation in meaning-making and reasoning. The MQI instrument helps mathematics educators assess classroom practices in school mathematics, allowing for a comprehensive approach to understanding and improving mathematics teaching (Charalambous & Litke, 2018; Charalambous & Praetorius, 2018).

The literature shows a limited study examining the actual classroom practices in statistics education (Eichler, 2011; Litke & Hill, 2020). Teaching statistics demands both discipline-related practices (Litke & Hill, 2020) and subject-specific practices within broader mathematics due to the place of statistics in school mathematics and the cooperation of two fields (Delmas, 2004; Gattuso, 2006; Gattuso & Ottaviani, 2011; Moore & Cobb, 2000; Wild et al., 2018). Based on reviews of current issues in statistics education, including the similarities and differences between statistics and mathematics and their respective teaching practices, as well as instructional quality in mathematics education, the following section will outline the purpose and significance of the current research.

1.3. The Purpose of The Study

This qualitative case study aims to understand mathematics teachers' statistics teaching practices in mathematics classrooms. To achieve this aim, two distinct frameworks are employed. The first framework explores whether teachers are engaged in the four-step model of statistical problem-solving processes by the Transcript Coding Protocol using the GAISE Framework. The second framework uses the Mathematical Quality of Instruction (MQI) instrument to broadly reveal the quality of statistics instructional practices within broader mathematics.

The study uses both frameworks to provide a comprehensive understanding of mathematics teachers' statistical practices and the quality of their instruction in mathematics classrooms. Specifically, the study will examine the statistical content covered by teachers and how they approach the four-step model of statistical problem-solving processes. In addition, the quality of instruction will be explored through four dimensions: the richness of mathematics and statistics, how teachers work with students to support their understanding of mathematics and statistics, and what mathematical and statistical errors and imprecision are demonstrated by teachers. Also, the study will explore how students participate in meaning-making and reasoning during statistics instruction within the broader mathematics context.

1.4. Research Questions

In light of the purpose of the study, the research questions guiding this study are designed to provide a comprehensive understanding of the statistical teaching practices employed by mathematics teachers in the classroom. Therefore, the qualitative case study seeks to answer the following research questions:

1. What are middle school mathematics teachers' statistical practices in the classroom?
 - 1.1 What statistical content do mathematics teachers cover in their instruction?
 - 1.2 How do mathematics teachers approach the four-step model of statistical problem-solving processes in their instruction?

2. What is the quality of mathematics teachers' statistics instruction within their broader mathematics instruction?

2.1 To what extent is classroom instruction connected to both mathematics and statistics?

2.2 How rich is the mathematics and statistics content covered in the lessons?

2.3 How do mathematics teachers work with students to support their understanding of both mathematics and statistics?

2.4 What are the mathematical and statistical errors and imprecision demonstrated by mathematics teachers?

2.5 How do students participate in meaning-making and reasoning during statistics instruction within the broader mathematics context?

These research questions help to explain the statistical practices of mathematics teachers in the classroom and how the quality of their statistics instruction is within the broader mathematics instruction. By examining the statistical content covered and how the statistical content is aligned with the curricular expectations and the approach to the four-step model of statistical problem-solving processes, a better understanding of mathematics teachers' statistical practices could be obtained. Examining the quality of statistics instruction in broader mathematics instruction will provide insights into the strengths and weaknesses of mathematics instruction as a whole.

1.5. The Significance of the Study

Statistics education has gained increasing attention in mathematics education research. Curricular documents around the world, including The Guidelines for Assessment and Instruction in Statistics Education (GAISE), National Council of Teachers of Mathematics (NCTM), Common Core State Standards for Mathematics (CCSSM), and Australian and New Zealand Mathematics and Statistics Curricula, have moved towards emphasizing the statistical problem-solving process. Turkish mathematics curriculum learning objectives are consistent with other global curricula, with a growing emphasis on statistical problem-solving (Ader, 2016, 2018; Frischemeier et al., 2022; MoNE, 2018, Özmen & Baki, 2019). However, curricular documents have been updated rapidly through collective work to suggest teaching practices to teachers

without considering how much they can implement them. There has been little research on how classroom practices are in line with the curricular documents (Litke & Hill, 2020). Therefore, it is necessary to bridge the gap between recommended statistical teaching practices and their implementation in the classrooms by gathering information from the field to improve statistics instruction.

First, this study aims to explore middle school mathematics teachers' statistics teaching practices, specifically addressing their alignment with curricular objectives in the Turkish Middle School Mathematics Curriculum and their approaches to the four-step model of statistical problem-solving processes. It supports with Eichler's (2011) assertion that exploring the transition from the written to the enacted curriculum is crucial for improving statistics teaching practices and ultimately enhancing student outcomes. Although some studies have explored students' and teachers' knowledge and beliefs of statistics as shaping factors of enacted curriculum, there remains a significant gap in the literature regarding the actual practices of statistics in mathematics classrooms (e.g., Burgess, 2011; Eichler, 2011; Eichler & Zapata-Cardona, 2016; Groth, 2013).

The curriculum and curricular resources serve as guiding frameworks for teachers, and investigating the alignment between teachers' practices and the prescribed objectives is essential to ensure effective statistical teaching practices. In the context of Turkish school mathematics, where teachers are required to cover all the objectives outlined in the curriculum, it becomes necessary to examine teachers' coverage of statistical objectives. Batur et al. (2021) revealed a limited depth and coverage of statistical objectives in the Turkish curriculum compared to international standards. By addressing the research question, this study aims to shed light on the extent to which teachers cover the specific statistical content outlined in the curriculum. By examining the enacted curriculum in terms of statistical content covered, the study intends to gain insights into the areas where the written curriculum may need further refinement or adjustments to better correspond to the realities of classroom implementation. Furthermore, linking the big ideas with the specific statistical content teachers cover offers a comprehensive perspective on statistics instruction. It ensures that students learn the procedural aspects and develop a deep conceptual understanding of the

statistics they learn. This study provides in-depth insights into the alignment between teachers' instructional practices and the coverage of statistical content, focusing on the big ideas in statistics. These insights can inform curriculum development, instructional materials, and professional development opportunities for teachers.

By employing the four-step model of statistical problem-solving processes, this study provides a discipline-related approach to examining mathematics teachers' statistical teaching practices (Watson et al., 2018). Examining whether teachers use all four elements of the statistical problem-solving process enables the identification of potential gaps in their instruction, providing a more nuanced understanding of mathematics teachers' practices beyond assessing whether they teach statistics as a discrete discipline (Litke & Hill, 2020; Visnovska & Cobb, 2019). Understanding how mathematics teachers approach statistics instruction is essential as their practices might significantly impact students' statistical thinking, reasoning, and literacy (Bargagliotti et al., 2020; Ben-Zvi & Garfield, 2008; MacGillivray & Pereira-Mendoza, 2011).

Overall, by exploring middle school mathematics teachers' statistical teaching practices through examining the enacted curriculum, this study contributes to advancing the understanding of the relationship between the written and enacted curricula in the context of statistics instruction in mathematics classrooms. Moreover, this research, which observes teaching practices in natural settings over an extended duration, provides a broader understanding of mathematics teachers' statistical practices. The insights gained from this study's findings have the potential to inform policymakers about areas that require improvement in the curriculum and curricular resources, ultimately supporting the enhancement of statistics education.

Furthermore, the significance of this study is the contribution to understanding the quality of statistics instruction in mathematics classrooms. As discussed earlier, statistics is often a part of the mathematics curriculum in many countries. Exploring the similarities and differences between both disciplines helps to understand how they can contribute to each other (Gattuso & Ottaviani, 2011). Therefore, looking at the teaching practices related to statistics within the broader context of mathematical

practices is necessary. The Mathematical Quality of Instruction (MQI) framework is used in the current study in order to explore the quality of instruction. Using the MQI instrument in conjunction with the protocol related to the statistical problem-solving process, the quality of statistics instruction can be examined more broadly to gain insights into how teachers' statistical practices relate to their overall instructional practices (Hannigan et al., 2013). This study is significant due to beyond the discussion on the nature of statistics as a discipline, the study considers statistics as a discipline with the fact that its place in school mathematics. Therefore, it explores the statistical teaching practices in natural settings and mathematics classrooms. It looks at both discipline-related practices in statistical investigations and subject-specific teaching practices in mathematics, as mathematics is considered a broader umbrella subject. This study considers the demands on mathematics teachers to integrate statistical practices into their teaching and the overall demands of the mathematics curriculum. Developing a more comprehensive understanding of the statistics teaching practices within the broader context of mathematics inform the literature on how effectively integrate statistics in school mathematics within current conditions as the other theoretical contribution of the current study. Ultimately, this study has the potential to contribute to the literature on effective statistics instruction in mathematics classrooms.

Studies on teaching practices in statistics education are limited (Litke & Hill, 2020; Tarr et al., 2006). This study uses an observational study approach that involves real-time practice within the classroom context, including interaction with teachers and students over content and video analysis, allowing for repeated engagement with the practice. Therefore, this study can provide valuable insights into what teachers do in terms of discipline-related practices and the quality of their practices in their mathematics classrooms, helping to understand where and how to improve statistics teaching in mathematics classrooms. Furthermore, the MQI consists of comprehensive subject-specific teaching practices in segment-level codes examining each practice's quality level. Determining the mathematical and statistical teaching practices through quality indicators inform teacher educators about the strengths and weaknesses of such practices as well as to what extent they are present during teaching statistics within mathematics classrooms. By identifying areas of strength and weakness in current

teaching practices, teacher educators can develop targeted interventions to improve the quality of mathematics and statistics instruction in schools, ultimately leading to improved student learning outcomes.

Overall, it is crucial to explore classroom practices, including the activities teachers use, which key ideas they cover, and which steps of the statistical problem-solving process are focused on, to better understand the quality of statistical instruction within broader mathematics instruction. This will allow us to identify areas where improvements can be made in statistics education to better equip students with the statistical thinking, literacy, and reasoning skills necessary to make sense of the world.

1.6. Definition of Terms

The following key terms will be defined and clarified to ensure a common understanding between the researcher and the reader.

Middle school mathematics teachers refer to individuals who teach mathematics to students in a classroom setting from fifth grade to eighth grade. This study uses middle school mathematics teachers for seventh-grade mathematics teachers.

Statistics is considered “a methodological discipline. It exists not for itself but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data (Cobb & Moore, 1997, p.801).” Wild et al. (2018) defined the statistics provided by the American Statistical Association (ASA) as “the science of learning from data and of measuring, controlling, and communicating uncertainty” (p.6).

Statistics education refers to “an interdisciplinary field that is focused on the teaching and learning of statistics.” (Zieffler et al., 2018, p.37)

Statistical content refers to the specific topics and concepts related to the statistics covered in mathematics classrooms.

Big Ideas in Statistics are interrelated concepts in statistics, including expectations, variation, distribution, informal inference, and randomness (Van de Walle et al., 2010; Watson et al., 2018). Watson et al. (2018) pointed out the recognition of big ideas to contribute conceptually to the practice of statistics rather than calculation.

Statistical teaching practices refer to the methods and techniques mathematics teachers use to teach statistics.

Statistics literacy is described as having skills necessary for “understanding statistical information or research results” (Ben-Zvi & Garfield, 2004, p.7).

Statistical reasoning is “the way people reason with statistical ideas and make sense of statistical information. It involves making interpretations based on sets of data, representations of data, or statistical summaries of data” (Ben-Zvi & Garfield, 2004, p.7).

Statistical thinking is conceptualized as “an understanding of why and how statistical investigations are conducted and the “big ideas” that underlie statistical investigations” (Ben-Zvi & Garfield, 2004, p.7).

The statistical problem-solving process is “defined as a four-step process of (1) formulating a statistical investigative question, (2) collecting or considering data, (3) analyzing data, and (4) interpreting results” (Bargagliotti et al., 2020, p.7). This process also refers to the investigative Cycle in A Framework for Statistical Thinking in Empirical Enquiry by Wild and Pfannkuch (1999) and the practice of statistics by Watson et al. (2018). The current study also uses the term “statistical investigation.”

Statistical tasks are “the instructional sequences that the teachers were learning to use and adapt in their classrooms focused on the analysis of univariate and bivariate data sets and involved classroom activities in which students analyzed data to address the problem at hand.” (Visnovska & Cobb, 2019, p.288)

Episodes are “the class’s interactions with each unique set of data encountered by the class” (Litke & Hill, 2020, p.487)

Teaching is “a collection of practices, including pedagogy, learning, instructional design, and managing organization. There are more practitioners than teachers, more practices than pedagogy, and the environments of teaching and learning are implicated in the interactions” (Cohen et al., 2003, p.124). It could also be defined as *instruction* containing “interactions among teachers and students around content, in environments.” (Cohen et al., 2003, p.122).

Interaction “refers to no particular form of discourse but to teachers’ and students’ connected work, extending through days, weeks, and months.” (Cohen et al., 2003, p.122)

Mathematics teaching practices refer to the use of the methods and techniques of mathematics teachers to teach mathematical concepts and skills.

Mathematical refers to the disciplinary nature of the MQI derived from the literature on mathematics knowledge for teaching, including subject-specific practices in mathematics. Still, it encompasses more general mathematics than specific topics (LMT, 2011).

Quality refers to “the distinctive character of something, considering which features are worth considering and examining constellations of features that constitute the overall character of particular instruction” (LMT, 2011, p.30).

The mathematical quality of instruction means “the nature of the mathematical content available to students during instruction. This is distinct from a focus on instructional strategies or pedagogical style” (LMT, 2011, p. 30).

The richness of mathematics and statistics “captures the depth of the mathematics and statistics offered to students” by focusing on the meaning of facts and procedures or key mathematical practices (Charalambous & Litke, 2018, p.447)

Working with students and mathematics and statistics is about “teachers’ ability to appropriately interpret and respond to students’ mathematical and “statistical” ideas and errors, as well as their capability to capitalize on these ideas and errors during instruction” (Charalambous & Litke, 2018, p.447)

Errors and imprecision refer to “the degree to which instruction is mathematically and “statistically” incorrect (Charalambous & Litke, 2018, p.447)

Student Participation in Meaning Making and Reasoning refers to “the focus on student involvement with the mathematical and ‘statistical’ content through cognitively demanding activities” (Charalambous & Litke, 2018, p.447).

1.7. Delimitations

Delimitation refers to defining the scope of a study, which is described as a “researcher-controlled map of the research process” (Coker, 2022, p.143). It is a way to set boundaries that make the research transparent (Coker, 2022). Delimitations point out the specific decisions of the researchers concerning the research process (Simon & Goes, 2013, as cited in Coker, 2022). This section presents particular delimitations of the study.

First, the study is delimited by the number of participants. Two practicing teachers are participating in the study to explore middle school mathematics teachers’ statistics teaching practices. Moreover, it is delimited by time and place in a way that the study was conducted within the statistics unit taking approximately 15 lessons in 7th grade of public schools. Although the timeframe covers the whole statistics unit in 7th grade, the study might be limited to the specific grade level, which may not capture the overall picture of the middle school mathematics teachers’ statistical practices.

The study focuses on mathematics teachers’ statistical practices during classroom observations without considering factors affecting classroom processes, including teacher knowledge and beliefs. Those factors may influence their teaching practices through different data collection methods, such as interviews or surveys. However, specific shreds of evidence related to teacher beliefs and knowledge are provided in the

context of the study for understanding the teaching practices through an in-depth description of participants. Furthermore, the observational data is supported by video and audio recordings as well as descriptive and evaluative field notes.

1.8. Organizations of the Study

This section presents the organization of this dissertation consisting of five chapters. Chapter 1 introduces the study, including a background of the research problem, the purpose statement, the research questions, the significance of the study, key terms in the study, and the delimitation of the study.

Chapter 2 presents a review of the relevant literature, focusing on teaching and learning statistics as a field and statistical and mathematical teaching practices with the ultimate purposes of statistical literacy, reasoning, and thinking. The first section consists of in-depth descriptions of key statistical ideas, statistical problem-solving processes, and related studies regarding statistical investigations. The second section provides literature on statistical and mathematical teaching practices, including the descriptions of different instructional quality frameworks and the details regarding the Mathematical Quality of Instruction (MQI) framework and possible factors affecting instructional quality with a comprehensive review of related studies.

Chapter 3 describes the research methodology, including the research design, participants with the context of the study, data collection methods, and data analysis process with the descriptions of two frameworks called “A Transcript Coding Instrument using the GAISE Framework” and “the Mathematical Quality of Instruction” with their coding decisions as well as the validity and reliability issues.

Chapter 4 presents the results of the study in line with answering the research questions. First, the within-case analyses are provided by presenting the results of each teacher’s teaching practices in terms of the alignment of the statistical content covered in the classroom context with the objectives in the curriculum, the teaching practices in terms of the statistical problem-solving process, and the quality of mathematics and statistics instruction closing with the overall mathematical quality of instruction. Then the results

regarding the cross-case analyses are provided following the same structure as the within-case analyses.

The last chapter, Chapter 5, discusses the study's key findings with related literature, implications of the results, limitations of the study, and suggestions for future research.



CHAPTER 2

LITERATURE REVIEW

This study aims to investigate mathematics teachers' statistical practices in the classroom and explore the quality of their statistics instruction relating to their broader mathematics instruction. To accomplish these goals, this chapter comprehensively reviews the related literature in two main sections: (a) teaching and learning statistics and (b) mathematics teachers' classroom practices in mathematics and statistics education. This chapter offers insights into mathematics teachers' statistical practices by examining the key ideas in statistics and the approach to the four-step model of statistical problem-solving processes. In addition, the chapter explores the quality of statistics instruction within broader mathematics instruction to identify the strengths and weaknesses of mathematics instruction as a whole. By synthesizing and analyzing the findings from previous research, this chapter aims to provide a comprehensive understanding of mathematics teachers' statistical practices and their implications for mathematics instruction as a whole. The insights gained from this literature review will provide a foundation for the study method.

2.1. Teaching and Learning Statistics

This section aims to provide a comprehensive literature review on the learning and teaching of statistics in K-12 education. The literature review will include the importance of statistical reasoning, literacy, and thinking, the big ideas in statistics, and the role of statistical investigations in developing statistical thinking, reasoning, and literacy. In addition, the review will examine the curriculum and curricular expectations for statistics education in K-12 settings and the research conducted in statistical investigations to provide evidence-based best practices for teaching and learning statistics.

2.1.1. Statistical Literacy, Reasoning, and Thinking

Statistical literacy, reasoning, and thinking are widely recognized as critical aspects of statistics education. Therefore, statistics education research strongly emphasizes these components in the design and implementation of teaching strategies and curricula (Ben-Zvi & Garfield, 2004; Burgess, 2009; Gomez-Blancarte & Chavez-Aguilar, 2022; Shaughnessy, 2007). The researchers discussed the distinction between the three important concepts albeit their overlapping nature (Ben-Zvi & Garfield, 2004; Shaughnessy, 2007). However, Gomez-Blancarte and Chavez-Aguilar (2022) summarize different models regarding the interaction of three concepts as can be seen in Figure 2.1. The first two models are discussed in the literature as recognizing the presence of distinct features of all three concepts. Gomez-Blancarte et al. (2021) added the last representation as a “unidimensional representation” that considers all three together in an instructional environment without their distinct features (as cited in Gomez-Blancarte & Chavez-Aguilar, 2022, p.1). This section briefly discusses statistical literacy, reasoning, and thinking, respectively.

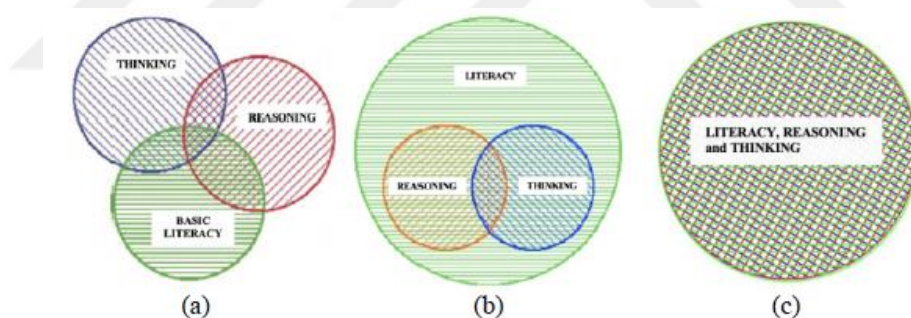


Figure 2.1 Representations of Statistical Literacy, Reasoning, and Thinking

Note: From “The Intersection of Statistical Literacy, Reasoning, And Thinking” by A.L. Gomez-Blancarte and R.D. Chavez-Aguilar, 2022, *Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11 2022)*, p.1 (DOI: 10.52041/iase.icots11.T8G2)

Statistics literacy is described as having skills necessary for “understanding statistical information or research results” (Ben-Zvi & Garfield, 2004, p.7). It is important to acknowledge that most students who learn statistics in school will not produce formal statistics in their professional lives. However, they will be consumers of statistics in their daily lives, for example, while reading articles or deciding what to buy or social

media data. Statistical literacy helps people make informed decisions by interpreting, evaluating, and communicating data or research findings (Bargagliotti et al.,2020; Garfield & Ben-Zvi, 2008; Shaughnessy, 2007; Wild et al., 2018). Also, statistical literacy is necessary for people to be good at their work; to progress in their professional life by following the research done in their area of study to improve their knowledge and practices (Bargagliotti et al.,2020). Ben-Zvi and Garfield (2004) summarized the skills needed for statistical literacy: “organize data, construct and display tables, and work with different representations of data” and emphasized the importance of “understanding of concepts, vocabulary, and symbols, and includes an understanding of probability as a measure of uncertainty” (p.7). Those skills contribute to people’s understanding of the world, consisting of vast amounts of data (Bargagliotti et al.,2020).

Statistical reasoning is another important concept in teaching and learning statistics. Ben-Zvi and Garfield (2004) defined statistical reasoning as “the way people reason with statistical ideas and make sense of statistical information. It involves making interpretations based on sets of data, representations of data, or statistical summaries of data” (p.7). In addition, statistical reasoning requires people to relate different statistical concepts and explain the statistical processes they engage in. Garfield (2002) explained different reasoning skills in statistics: reasoning about data, representations of data, statistical measures, uncertainty, samples, and associations. Improving statistical reasoning is required for making sound decisions based on evidence and is necessary for all study and work areas (Bargagliotti et al.,2020).

Statistical thinking consists of an understanding of the problem-solving processes using data (Wild et al., 2018). More specifically, Ben-Zvi and Garfield (2004) defined statistical thinking as “an understanding of why and how statistical investigations are conducted and the “big ideas” that underlie statistical investigations” (p.7). They also emphasized the importance of recognizing variation, using descriptive and inferential analysis methods, being aware of the context for statistical investigations, and evaluating the results obtained through the statistical investigations (Ben-Zvi & Garfield, 2004). Wild and Pfannkuch (1999) explained five essential types of statistical thinking as (a) recognition of the need for data instead of “personal experiences,” (b)

transnumeration -being present in all parts of data analysis because it is described as “a dynamic process” where learners need to use different representations of data for the understanding real data, (c) consideration of variation- acknowledging the presence of uncertainty and recognizing the presence of variation in many aspects in the statistical investigation process, (d) reasoning with statistical models- using statistical models to make sense of investigative process, and (e) integrating the statistical and contextual – making links between statistics and the context to reach a meaning through the analyses of the results (p.227). Altogether, statistical thinking helps people to “make sense of information in which variation and/or uncertainty is present” (MacGillivray & Pereira-Mendoza, 2011, p.110). In addition, it consists of all the necessary knowledge for learning and teaching statistics (Shaughnessy, 2007).

Statistical literacy, reasoning, and thinking are ultimately the main goals of learning and teaching statistics. While these concepts are distinct, they often overlap in practice (Gomez-Blancarte & Chavez-Aguilar, 2022). Specifically, statistical literacy, reasoning, and thinking contribute to teaching and learning big ideas or statistical concepts and processes. For example, the variability concept is fostered with statistical literacy by recognizing unpredictable events, with statistical reasoning while understanding the concept of variability through sampling methods, and with statistical thinking by explaining the variability that occurs during an investigation with (Gomez-Blancarte & Chavez-Aguilar, 2022). Conversely, the teaching and learning of these elements contribute to the enhancement of statistical literacy, reasoning, and thinking (Bargagliotti et al., 2020; Gomez-Blancarte & Chavez-Aguilar, 2022; MacGillivray & Pereira-Mendoza, 2011).

2.1.2. Big Ideas in Statistics

School mathematics aims to provide students with big ideas about mathematics (Yeo & Toh, 2019). Big ideas are the basis of teaching and learning mathematics and contribute to one’s mathematical content knowledge (Charles, 2005). Charles (2005) defines a big idea as: “A Big Idea is a statement of an idea that is central to the learning of mathematics, one that links numerous mathematical understandings into a coherent whole.” (p.10). Van de Walle et al. (2010) also define big ideas as “large networks of

interrelated concepts” (p.26). These definitions are in line with NCTM (2000)’s emphasis on the importance of seeing “mathematics as a coherent and connected enterprise” (p.17) through the big ideas in which it is necessary to form and understand “terminology, definitions, notation, concepts, and skills emerging in the process” (p.15). Considering mathematics as a whole and connecting concepts and skills through topics take place with the help of finding out big ideas through research and teaching environment (Charles, 2005; Van de Walle et al., 2010). Also, Van de Walle et al. (2010) underlined how crucial it is to arrange the instruction around big ideas because it helps learners “see that mathematics is an integrated whole, not a collection of isolated bits and pieces” (p.2).

Statistics curriculum as a part of school mathematics focuses on specific big ideas in statistics. Garfield and Ben-Zvi (2004) provided eight big ideas in statistics students experience during their K-12 education. These big ideas are data, distribution, trend, variability, models, association, samples and sampling, and inference. Garfield and Ben-Zvi (2008) also explained the central ideas of statistics in their book, but they added one more big idea to their list: randomness. In addition, they explicitly remark on “center” instead of explaining the center under a broader big idea, “trend.”

Furthermore, they renamed the “association” big idea into covariation (Garfield & Ben-Zvi, 2004; Garfield & Ben-Zvi, 2008). In addition to Garfield and Ben-Zvi (2004; 2008), Watson et al. (2013) provided five big ideas at the request of the Australian Association of Mathematics Teachers (AAMT) (as cited in Watson et al., 2018). The five big ideas in statistics are variation, expectation, distribution, informal statistics, and randomness (Watson, 2015; Watson et al., 2013; 2018). Shaughnessy (2019) spotlights the two most significant ideas in statistics as distribution and inference due to the purpose of statistics as working through uncertainty while making decisions based on data. He also described expectation and variation as elements of distribution and inference.

Figure 2.2 illustrates the relationships between the five big ideas in statistics described by Watson et al. (2018) within the context of classroom statistics. The diagram visually

represents how the concepts are interconnected and interact in the statistics curriculum (Callingham et al., 2021).

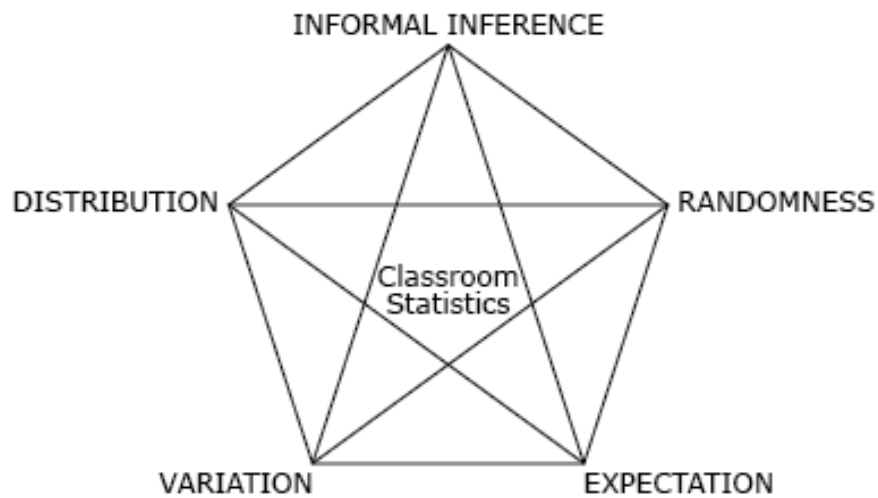


Figure 2.2 Big Ideas of Statistics

Note: From “The Practice of Statistics” by J. Watson, N. Fitzallen, J. Fielding-Wells, and S. Madden, 2018, In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education*, p. 126 (https://doi.org/10.1007/978-3-319-66195-7_4). Copyright 2018 by Springer International Publishing AG

Watson et al. (2013) provide brief definitions of the interconnectedness of big ideas in classroom activities as follows:

Variation is the term that describes the differences we observe around us in every aspect of life.

Expectation arises when we wish to harness variation and summarise data, e.g., by finding a typical value in a data set or the chance of a random event.

Distribution is the lens through which we look at variation, used to identify and describe variation and look for and confirm expectations. Randomness describes a phenomenon in which the outcome of a single repetition is uncertain, but there is nonetheless a regular distribution of relative frequencies in a large number of repetitions.

Informal Inference is the evidence-based process balancing the variation and expectation found in sample data when answering a meaningful population-based question, acknowledging uncertainty (as cited in Callingham et al., 2021, p.332-333)

In the upcoming sections, I will provide a comprehensive overview of the five key concepts in statistics education: Expectation, Variation, Distribution, Informal

Inference, and Randomness. The sections will begin by delving into the concept of Expectation, followed by an examination of Variation, Distribution, Informal Inference, and finally, Randomness. Each section will provide a detailed analysis of the concept and its importance in understanding statistical thinking. Additionally, the interconnectedness of the five big ideas will be highlighted throughout the discussion.

2.1.2.1. Expectation

The term *expectation* is one of the core concepts in statistics. The concept consists of research conducted on measures of central tendencies, also called average or mean, and probability (Shaughnessy, 2007; Shaughnessy, 2019; Watson et al., 2007; Watson, 2009; Watson et al., 2018). In literature, the expectation is defined as “the probabilities of outcomes or means of data sets” (Watson et al., 2007, p.84). Watson et al. (2007) provided a mathematical definition of expectation as “expected value a continues functions” (p.84). They also stated that the students are engaged in expectation concepts with the concepts of “probabilities, averages, “caused” differences and random distributions of outcomes” in schools rather than the technical definition (p.84). Watson (2009) points out that the term expectation is a correspondence of “middle or average for frequency distributions” (p.33). Center is the corresponding big idea in Garfield and Ben-Zvi’s lists of big ideas in statistics. Shaughnessy (2019) points out that Watson et al. (2013) mention expectation instead of center in Garfield and Ben Zvi’s lists of big ideas in statistics. Even though both concepts seem different, and “center” big ideas evoke the idea of the measures of central tendency, both groups of researchers aim to explain similar ideas with different names. Garfield and Ben-Zvi’s (2008) “center” big idea is more comprehensive than just calculating measures of central tendencies and describes the same idea of expectation. They explain the concept of center with the concepts related to distribution by using the notion of “a signal in a noisy process” by Konold and Pollatsek (2002).

Watson et al. (2007) state that the mean, as expected value, has been in school curricula for a long time. Even before focusing on statistics as a field, the average concepts like mean, median, and mode were in the mathematics curricula of many countries (Jacobbe & Carvalho, 2011). Much research has been conducted on how the average

concept needs to be defined and how students and teachers conceptualize averages due to the long history of the average concept in school curricula (e.g., Goodchild, 1988; Jacobbe, 2012; Konold & Pollatsek, 2002; Landtblom & Sumpter, 2021; Makar, 2014; Mokros & Russell, 1995; Strauss & Bichler, 1988; Watson & Moritz, 2000). Average concepts provide summary information of a data set with a single value, an estimated mean for a population when the sample is representative (Shaughnessy, 2007; Jacobbe & Carvalho, 2011). However, the single value obtained from a data set also referred to as the typical value, does not consider variation and is compatible with the deterministic view of mathematics (Garfield & Ben-Zvi, 2008; Jacobbe & Carvalho, 2011; Konold & Pollatsek, 2002; Shaughnessy, 2007).

Mokros and Russell (1995) describe average as “a tool for summarizing and describing as well as for comparing data” (p.21). Their research showed that the students consider average as (i) mode, (ii) algorithm, (iii) reasonable, (iv) midpoint, and (v) mathematical point of balance. The first two views of average do not consist of the idea of representativeness, unlike the remaining three views. Garfield and Ben-Zvi (2008) support Mokros and Russell's (1995) findings regarding children's conceptualizations of average view measures of central tendency as a typical value. Furthermore, Konold and Pollatsek (2002) provide different interpretations of average as (i) data reduction, (ii) fair share, (iii) typical value, and (iv) signal in the noise. The “signal in noise” interpretation contributes to understanding the average concept within its place in statistics as a field. They point out the relationship between average and variation, stating that “By central tendency, we refer to a *stable* value that (a) represents the signal in a variable process and (b) is better approximated as the number of observations grows” (Konold & Pollatsek, 2002, p.262). They consider the mean and median as stable features, the certainty, in a variable system that involves uncertainty (Konold & Pollatsek, 2002). Therefore, average and variation should be considered together to understand the data (Garfield & Ben-Zvi, 2008; Konold & Pollatsek, 2002). As a further consideration, the research shows that variation should be introduced earlier than the concepts of expectation, contrasting the history of introducing average concepts first in school curricula (Watson, 2005; Watson et al., 2007; Watson, 2009). Garfield and Ben-Zvi (2008) also emphasize that the summary of the data sets depends not only on average but also variability around the center.

2.1.2.2. Variation

Variation is another big idea that comes right after expectation. Two terms are used in this big idea: variability and variation (Burrill & Biehler, 2011; Reading & Shaughnessy, 2004). The term variability is about “something is apt or liable to vary or change” as “the characteristic of entity that is observable” (Reading & Shaughnessy, 2004, p. 201-202). The term variation is about “the act of varying or changing condition” as “the describing or measuring of that characteristic” (Reading & Shaughnessy, 2004, p. 201-202). Variation, as the big idea, differentiates statistics from mathematics (Reading & Shaughnessy, 2004; Watson, 2006). It is the core aspect of statistics that provides the foundation for drawing sound inferences when making decisions based on data (English & Watson, 2016). Wild and Pfannkuch (1999) state that the decision-making process in statistics occurs under the circumstance of uncertainty which drives by “omnipresent variation” (p.227). Based on the omnipresent nature, variation in data could be real or induced through data collection due to measurement, sampling, or accident (Wild & Pfannkuch, 1999).

Various models related to understanding variation are present in the literature (Shaughnessy, 2019a). For example, Wild and Pfannkuch's (1999) model of statistical thinking includes four considerations regarding variation. First, people react to variation by ignoring its existence, investigating its nature, and changing it to reach more desirable conditions (Reading & Shaughnessy, 2004; Wild & Pfannkuch, 1999). Therefore, it is necessary to notice and acknowledge variation. Second, it is crucial to measure and model the variation to predict, explain, or control (Burrill & Biehler, 2011, Wild & Pfannkuch, 1999). Third, variation needs to be explained and dealt with, and lastly use investigative strategies for understanding variation (Wild & Pfannkuch, 1999). Reading and Shaughnessy (2004) add two more considerations related to variation as describing and representing variation. All six issues related to variation and its omnipresent nature are significant contributing factors to statistical thinking (Reading & Shaughnessy, 2004). In addition, Garfield and Ben-Zvi (2005) present seven areas of knowledge of variability parallel to Wild and Pfannkuch (1999)'s considerations of variation. The areas of knowledge of variation are (a) developing intuitive ideas of variability, (b) describing and representing variability, (c) using variability to make comparisons, (d) recognizing variability in special types of

distributions, (e) identifying patterns of variability in fitting models, (f) using variability to predict random samples or outcomes, and (g) comparing variability as a part of statistical thinking. Reading and Reid (2010) adds two more components: (h) recognizing sources of variation and (i) resolving expectations with observed variations. Peters (2011) states that Garfield and Ben-Zvi's descriptions regarding variability knowledge start from "intuitive ideas of variability to statistical thinking that incorporates consideration of variability" (p.79).

Shaughnessy (2007) identified eight types of conceptions of variability in the literature: (a) variability in particular values, including extremes or outliers, (b) variability as change over time, (c) variability as the whole range of possible values, (d) variability as the likely range of a sample, (e) variability as distance or difference from a fixed point, (f) variability as the sum of residuals, (g) variation as covariation or association, and (h) variation as distribution. Shaughnessy (2019a) classifies the first four conceptions of variability as falling under the umbrella of "variability" in an exploratory data analysis approach, while the latter four are related to the concept of "variation," which is often used to measure variability (Reading & Shaughnessy, 2004; Shaughnessy, 2019a). The development of the concept of variation has also been investigated in the literature (Watson & Kelly, 2002; Watson et al., 2007; Watson et al., 2020). Watson et al. (2007) presented a model of the developmental progression of variation, which involves the interaction of expectation at six levels. In this model, the first stage involves appreciating either variation or expectation. Students can understand the concepts of expectation and variation and their relationship at this stage by using "established links between them in comparative contexts" (p.113).

In school curricula, variation is often introduced with standard deviation or measures of variability (Sanchez et al., 2011; Shaughnessy, 2007). Watson et al. (2007) also note that variation encompasses ideas such as "uncertainty, change, anticipated change, unanticipated change, and random behavior" (p.112). In addition, all steps in the statistical problem-solving process involve addressing variability by anticipating, acknowledging, accounting for, and allowing for it (Bargagliotti, 2020; Franklin et al., 2007). However, knowledge related to variability has often been assessed through the understanding of "definitions, calculation, and simple interpretations of standard

measures of dispersions (e.g., computing and interpreting the range, SD, and IQR)” in traditional settings (Garfield & Ben-Zvi, 2004b, p.96). Therefore, the literature emphasizes the importance of a conceptual understanding of variation by connecting it to various statistical contexts (e.g., Canada, 2006; Lehrer et al., 2007; Makar & Confrey, 2005; Reading, 2004; Reading & Reid, 2006; Watson et al., 2007; Watson et al., 2022).

2.1.2.3. Distribution

The concept of distribution is fundamental in K-12 statistics education (NCTM, 2009; Salinas-Herrera & Salinas-Hernández, 2022). Due to the nature of statistics, it is necessary to differentiate “data as individual values and distribution as a conceptual entity” (Bakker & Gravemeijer, 2004, p.148) and comprehend the set of data as a distribution (Garfield & Ben-Zvi, 2008). Moore and McCabe (1993) define distribution as “the pattern of variation of a variable is called its *distribution*. The distribution records the numerical values of the variable and how often each value occurs” (p. 6) (as cited in Watson, 2009, p.32).

Representations of data distributions are an integral component of data analyses (NCTM, 2009). Graphical representations, such as tables and graphs, are particularly useful in understanding the variation in data, with distribution being one of the critical forms of variation (Callingham et al., 2019). Garfield and Ben-Zvi (2008) illustrate the importance of using different representations to reveal various aspects of a distribution from the same dataset. Similarly, Moore and McCabe (1993) emphasize the significance of using graphical representations to comprehend distributions (as cited in Watson, 2009). The literature presents two types of distributions: empirical and theoretical (Salinas-Herrera & Salinas-Hernández, 2022). According to Burrill and Reading (2011), distributions form the basis for "reasoning about statistical variables from empirical distributions, random variables from theoretical distributions, and summaries in sampling distributions" (p.63). Additionally, Moore and McCabe (1993) describe the basic representations of distributions as frequency distributions (as cited in Watson, 2009). The shape of frequency distributions can be analyzed through various graphs, such as histograms, dot plots, stem-and-leaf plots, or boxplots (NCTM,

2000). Therefore, descriptive statistics include empirical distributions (Salinas-Herrera & Salinas-Hernández, 2022). Furthermore, theoretical distributions with continuous variables and binomial ones can also be discussed within the context of probability (Watson, 2009; Salinas-Herrera & Salinas-Hernández, 2022). Comparing and contrasting empirical distributions with theoretical distributions is essential to “make predictions and generalizations” (Salinas-Herrera & Salinas-Hernández, 2022, p.577)

Researchers have highlighted the complexity of the distribution concept, as it encompasses a wide range of ideas such as variation, sampling, expectation, shape, and density (Bakker & Gravemeijer, 2004; Biehler et al., 2018; Garfield & Ben-Zvi, 2008; Makar & Confrey, 2005). The distribution concept is often characterized by its center, shape, spread, density, skewness, and outliers (Bakker & Gravemeijer, 2004; Biehler et al., 2018; Makar & Confrey, 2005; Reading & Reid, 2006). Bakker and Gravemeijer (2004) stress that these big ideas are crucial when understanding data as an entity. However, focusing solely on a big idea does not contribute to understanding the distribution and only considers data as individual values (Makar & Confrey, 2005). As stated by Biehler et al. (2018), “reasoning about distributions involves integrating reasoning about these multiple statistical concepts” (p.155). For example, expectation and variation are necessary components for examining patterns in data distributions and making meaningful interpretations (Burrill & Reading, 2011; Garfield & Ben-Zvi, 2008; NCTM, 2000, 2009; Watson, 2009).

Furthermore, Garfield and Ben-Zvi (2008) emphasize the significance of comprehending the concept of distribution through its relationship to the ideas of center and spread, which are essential for interpreting results. Thus, graphical displays should be constructed and selected based on their intended purpose to make sense of the data, perceive the center and spread within the distribution, and discern the shape of the distribution, including possible outliers (NCTM, 2000). In addition, understanding the center in relation to the variation in distribution provides a conceptual foundation for the idea of the center as a signal amidst noise (Garfield & Ben-Zvi, 2008). The idea of variation plays a crucial role in distribution. Without variation in a dataset, a distribution could not be formed, and distribution allows individuals to grasp the omnipresence of variability (Bakker & Gravemeijer, 2004;

Cobb & Moore, 1997). Distribution is also considered a visual representation of variation in data (Makar & Confrey, 2005). Therefore, understanding variation is closely tied to conceptualizing distribution. As a result, by grasping the distribution concept, students can acquire the necessary skills to analyze and interpret data effectively (Biehler et al., 2018).

2.1.2.4. Informal Inference

Statistical inference is a foundational statistical concept and an integral part of everyday life (Kazak et al., 2016; Makar & Rubin, 2018; Zieffler et al., 2008). As described by NCTM (2009), “statistical inference is drawing conclusions from and making decisions based on data obtained” (p.87). Burrill and Reading (2011) refer to sampling and inference as a fundamental idea in statistics in which the relationship between samples and the population is established through drawing conclusions based on samples with a degree of certainty. The importance of uncertainty is highlighted due to the presence of data from random samples, which may not provide a complete picture of populations and their characteristics (Wild et al., 2011). Makar (2013) claims that making inferences based on the data obtained and making predictions and estimations about the population make statistics powerful.

Conventionally, statistical inference encompasses a variety of concepts such as probability, random sampling, distributions, confidence intervals, critical values, and p-values (Harradine et al., 2011; Wild et al., 2011). Harradine et al. (2011) outlined statistical inference as having three essential components: “the reasoning process, the concepts, and the associated computations” (p.236). They emphasized the interactive nature of these components and the difficulties that people may encounter when learning the first two components, particularly in understanding sampling, sampling distributions, the null and alternative hypotheses, statistical significance, p-values, and confidence intervals. Wild et al. (2011) proposed a pathway to a formal inference that helps learners relate the concepts learned during the path, contributing to the complex and challenging nature of statistical inference. This pathway includes that (Wild et al., 2011, p.252):

- *we should work from a minimal set of the biggest ideas of statistical inference.*
- *the mechanics of the inferential step should not be at all demanding.*
- *inferences should be able to be performed without students taking their eyes off their graphs.*
- *methods should have connections to the more formal methods to be used later.*

Efforts have been made to expose students to different ways of understanding and applying statistical inference, as students have difficulties in this area, by introducing them to scenarios that enable them to use informal methods of drawing conclusions (Zieffler et al., 2008). Watson and English (2015) underline that informal inference is the foundation of formal inference. The informal inference is also considered “a bridge between exploratory data analysis and formal statistical inference” (Gil & Ben-Zvi, 2011, p.88). Makar et al. (2011) add an “informal” word to statistical inference to make statistical inference more comprehensive than hypothesis testing and make it more accessible for learners without depending on “formal statistical measures and procedures to formulate their inference” (p.153). Harradine et al. (2011) describe informal inference as the reasoning involving formal inference without using probability distributions and calculations. Makar and Rubin (2009) provide essential concepts to make statistical inference more comprehensive (p.85):

- *Notion of uncertainty and variability articulated through language that broke from the mathematical convention of claims of certainty.*
- *Reliance on the concept of aggregate (as opposed to individual points) through the use of generalizations about the group.*
- *Acknowledgment of a mechanism or tendency that extended beyond the data at hand; and*
- *Evidence for reasoning based on purposeful use of data.*

The informal inference is one of the big ideas, characterized as more complex than the previously mentioned ones because it demands the expectation and a level of confidence in the expectation because of the uncertainty of the results (Watson & Fitzallen, 2022). As such, three types of reasoning are needed during informal inferential reasoning, also known as informal inference (Zieffler et al., 2008, p. 45):

- *Reasoning about possible characteristics of a population (e.g., shape, center) based on a sample of data.*

- *Reasoning about possible differences between two populations based on observed differences between two samples of data (i.e., are differences due to an effect as opposed to just due to chance?); and*
- *Reasoning about whether or not a particular sample of data (and summary statistic) is likely (or surprising) given a particular expectation or claim.*

Based on the critical concepts and reasoning needed for informal inference, Makar and Rubin (2009) provide three characteristics of informal inference in which the types of reasoning are needed (p.85):

- *generalization, including predictions, parameter estimates, and conclusions that extend beyond describing the given data.*
- *the use of data as evidence for those generalizations; and*
- *employment of probabilistic language in describing the generalization, including informal reference to levels of certainty about the conclusions drawn.*

In summary, statistical inference is the process of drawing conclusions and generalizations from data (Makar & Rubin, 2009; NCTM, 2009; Zieffler et al., 2008). While formal statistical inference employs established methods and procedures, informal statistical inference is a less structured yet more comprehensive approach that does not adhere to established statistical procedures (Zieffler et al., 2008). It is crucial to note that the term "informal" does not denote an inferior approach but emphasizes the broad spectrum of inferential reasoning and the possibility of viewing statistical inference beyond formal methods (Makar & Rubin, 2009).

2.1.2.5. Randomness

Randomness is a fundamental concept in probability and statistics education, but effective teaching can be challenging (Chick, 2018; Engel & Sedlmeier, 2005; Watson & Fitzallen, 2019). Researchers have defined randomness as a phenomenon characterized by a degree of predictability and certainty (e.g., Chick, 2018; Garfield & Ben-Zvi, 2008; Moore, 1990; Watson et al., 2018). Watson and Fitzallen (2019) provide a formal definition of randomness on the Top-Drawer Teachers: Statistics website, stating that “randomness describes a phenomenon in which the outcome of a single repetition is uncertain, but there is nonetheless a regular distribution of relative

frequencies in a large number of repetitions.” This definition is consistent with Moore's (1990) definition of randomness as a non-deterministic concept that is not synonymous with haphazardness. Engel and Sedlmeier (2005) further elaborate on the concept, stating that randomness is characterized by the possibility of multiple outcomes and “that the realized outcome is not predictable with certainty” (p.176). Watson et al. (2018) note that randomness arises in processes that display patterns over the long term but have unpredictable individual outcomes. Garfield and Ben-Zvi (2008) posit that while the outcome of a specific random event cannot be predicted, long-term patterns can be predicted.

Biehler (1994) posits two types of statistical thinking: deterministic and probabilistic. The latter, which relies on drawing a random sample from a population to make accurate inferences (as cited in Pfannkuch & Wild, 2004), encompasses the concept of randomness. Biehler also contends that Exploratory Data Analysis (EDA) is a component of deterministic thinking in statistics (Pfannkuch & Wild, 2004). Engel and Sedlmeier (2005) emphasize the importance of probabilistic thinking in developing statistical thinking without an excessive focus on computations. Batareno (2015) offers various conceptualizations of randomness, such as equiprobability, uncertainty, and as a model, and among others. Considering these conceptualizations, teachers should provide a learning environment that enables students to develop their probabilistic reasoning skills (Batareno, 2015). Children should be provided with opportunities to experience chance and randomness through hands-on experimentation with concrete objects (NCTM, 2000). As Moore (1990) asserts, providing children with guided experiences with randomness is crucial for learning formal probability in later years. Additionally, it is essential to note that statistics and probability are conceptually interconnected rather than distinct fields (Kazak & Confrey, 2007).

Individuals experiencing uncertainty during their childhood are particularly in need of comprehending the random occurrences happening around them (Batanero, 2015). As Batanero (2015) notes, “the omnipresence of randomness implies our need to understand random phenomena in order to make adequate decisions when confronted with uncertainty” (p. 34). Making decisions in the face of uncertainty created by random phenomena is crucial to drawing statistical conclusions (Batanero, 2015;

Common Core State Initiative [CCSSI], 2010; NCTM, 2000). CCSSI (2010) highlights the significance of randomization in statistical inferences, as it allows for drawing valid conclusions about a population by collecting data from a random sample. Similarly, Moore (1990) emphasizes the importance of utilizing random sampling and experimental randomization in inferential statistics and formal probability. Random sampling helps to eliminate any potential bias that may arise from “personal choice” in data collection during statistical problem-solving processes (Moore, 1990). Moreover, randomization involves assigning subjects to treatment groups to make causal inferences from statistical investigation without any form of bias (Moore, 1990; CCSSI, 2010). These concepts of random sampling and randomization are incorporated into the statistics strand of school curricula, linking probability and statistics topics together (Batareno, 2015).

Watson et al. (2018) argue that the five big ideas in statistics education, which include variation, expectation, distribution, informal inference, and randomness, are interconnected. Without understanding variation and expectation, it can be difficult to grasp the concept of randomness (Watson & Fitzallen, 2019). As previously mentioned, understanding variation involves utilizing variability to predict “random” samples and outcomes (Chick, 2018; Reading & Reid, 2010). Watson and Fitzallen (2019) further connect the definition of randomness to variation through the uncertainty of a single repetition and expectation through the regularity of distribution arising from multiple repetitions. Overall, randomness is central to statistics education and linked to other key concepts (Watson et al., 2018). Effective teaching of randomness requires understanding these interconnected concepts and providing students with hands-on experiences to develop their probabilistic reasoning skills and make decisions in uncertain situations (Batareno, 2015; Engel & Sedlmeier, 2005; Moore, 1990; NCTM, 2000).

2.1.2.6. Summary of Big Ideas in Statistics

The core concepts of teaching and learning statistics, known as big ideas, have been widely discussed in the literature (Garfield & Ben-Zvi, 2004; Garfield & Ben-Zvi, 2008; Shaughnessy, 2007; Shaughnessy, 2019a; Shaughnessy, 2019b; Watson et al.,

2013). Watson et al. (2013) identified five big, interrelated ideas in statistics: expectation, variation, distribution, informal inference, and randomness. Researchers have examined the conceptual development of these big ideas in children to enhance statistical literacy, reasoning, and thinking (Garfield & Ben-Zvi, 2004).

It is essential for students to understand the big ideas in statistics, rather than simply learning a set of procedures, to effectively conduct statistical investigations and draw meaningful conclusions (Callingham et al., 2021; Shaughnessy, 2019a; Watson et al., 2018). Watson et al. (2018) emphasize the importance of students being aware of the role of big ideas in statistical processes. Shaughnessy (2019a) suggests creating a list of big ideas in statistics that encompasses both statistical concepts and processes as “two perspectives on organizing what is most important in statistics interact with and support one another” (p. 5).

2.1.3. Statistical Investigations

Statistics educators often stress the importance of an investigative approach in developing statistical literacy, reasoning, and thinking skills (Ben-Zvi & Garfield, 2008; Gal & Garfield, 1997; MacGillivray & Pereira-Mendoza, 2011). Moreover, developing big ideas in statistics necessitates students' engagement in complete statistical investigations or specific steps within statistical investigations (Watson & Callingham, 2015). Wild and Pfannkuch (1999) highlight the crucial role of statistical investigations in learning within context, as students can use data collected through these investigations to gain a deeper understanding of the world around them. The statistical investigation is a fundamental aspect of teaching and learning statistics.

Different frameworks cover mainly similar components in the process (Arnold, 2008). Four-phase cycle frameworks include posing questions, collecting, analyzing, and interpreting the results. Lee et al. (2022, p.2-3) state that examples of the four-cycle framework are the PCAI Model of Statistical Investigations (Graham, 1987), the Process of Statistical Investigations (Friel et al., 2006), Mobilize Introduction to Data Science Curriculum (Gould et al., 2016), and Statistical Problem Solving as an Investigative Process and Developmental Levels of Statistical Literacy (Franklin et al.,

2007; Bargagliotti et al., 2020). Five-phase cycle frameworks consider “planning and exploring data” along with four components, as in the four-cycle frameworks (Lee et al., 2022, p.2). Lee et al. (2022, p.2-3) also state that examples of the five-cycle framework are A 4-Dimensional Framework for Statistical Thinking in Empirical Enquiry (Wild & Pfannkuch, 1999), IDSSP model (IDSSP, 2019) and Practices of Statistics (Watson et al., 2018).

In this section, I will briefly explain two frameworks because they are closely related to the theoretical background of the current study. First, I will introduce the 4-Dimensional Framework for Statistical Thinking, focusing on the Investigative Cycle by Wild and Pfannkuch (1999), which is an example of a five-phase cycle framework. Then I will introduce Statistical Problem Solving as an Investigative Process and Developmental Levels of Statistical Literacy in The Guidelines for Assessment and Instruction in Statistics Education (GAISE) reports by Bargagliotti et al. (2020) and Franklin et al. (2007), which is an example of a four-phase cycle framework. Then I will briefly compare both frameworks. In the following parts of the current study, I name the latter framework, the GAISE or statistical problem-solving process. After all, I will explain the two-dimensional framework concerning four components in the statistical problem-solving process from GAISE reports in detail.

First, Wild and Pfannkuch (1999) describe “A Framework for Statistical Thinking in Empirical Enquiry” with four dimensions a statistical thinker engages in. The first dimension is called as Investigative Cycle, followed respectively by Types of Thinking (Dimension 2), Interrogative Cycle (Dimension 3), and Dispositions (Dimension 4) (p.225). The first dimension is a starting point for engaging in investigations and statistical thinking and is described as “the most widely referenced dimension” (Watson & English, 2015, p.582). They call the dimension PPDAC (Problem, Plan, Data, Analysis, Conclusions) cycle, which is interested in solving statistical problems as can be seen in Figure 2.3. The cycle starts with a “problem” statement followed by “planning” various parts of the investigation (e.g., sampling, measurement). The “data” step consists of collecting and handling data, followed by the “analysis” of data. Then, the “conclusions” are made based on the data analysis.

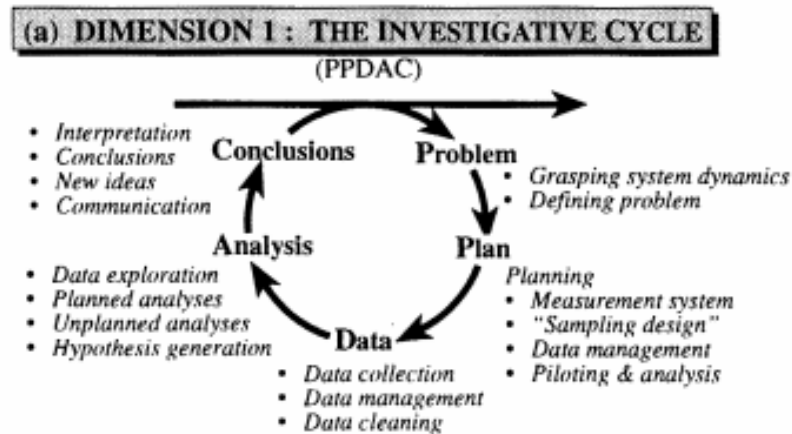


Figure 2.3 First Dimension of the 4-Dimensional Framework for Statistical Thinking in Empirical Enquiry

Note: From "Statistical Thinking in Empirical Enquiry" by C. Wild, and M. Pfannkuch, 1999, *International Statistical Review*, 67, p.226 (<https://iase-web.org/documents/intstatreview/99.Wild.Pfannkuch.pdf>)

The GAISE is a Pre-K curriculum framework that aims to provide statistical literacy for all, developed by the American Statistical Association (ASA) (Franklin et al., 2007). The framework was updated in 2020 during the outbreak of COVID-19 by re-emphasizing the significance of statistical literacy for all. Then GAISE II is called a framework for statistics and data science education (Bargagliotti et al., 2020). Both GAISE reports (Bargagliotti et al., 2020; Franklin et al., 2007) present a two-dimensional framework for statistics education.

The first dimension consists of the statistical problem-solving process components, also called the investigative process as can be seen in Figure 2.4. According to GAISE reports, the statistical investigations start with "formulating statistically investigative questions," followed by "collecting/considering data," then "analyzing the data," and lastly, "interpreting the results" (Bargagliotti et al., 2020, p.13). The second dimension includes three developmental levels (A, B, and C), which are described as "recommendations for developing students' foundational skills in statistical reasoning in three levels" (p.9). In the GAISE I report, the levels seem parallel to grade levels, but the focus moved to the sequential nature of all levels in terms of the development of statistical literacy in the GAISE II report (Bargagliotti et al., 2020; Franklin et al., 2007). Bargagliotti et al. (2020) state that the concepts in Level A are a prerequisite to

Level B. Similarly, Level A and Level B concepts are prerequisites to the concepts in Level C. Therefore, it is important to state that when students start to learn concepts in Level A in, even high school, they do not engage in Level C concepts before meeting the prerequisites. When students follow three developmental levels regardless of their grade levels, their progress into more complex concepts can be observed (Groth, 2008).



Figure 2.4 Statistical Problem-Solving Process

Note: From *Prek-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education* (p.13), by A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson., L. Perez, and D. Spangler, 2020, National Council of Teachers of Mathematics. Copyright © 2020 by American Statistical Association.

When both frameworks are compared, GAISE (Bargagliotti et al., 2020) does not address data collection as two separate actions, “plan” and “data,” like Wild and Pfannkuch (1999). In the GAISE’s statistical problem-solving process, the data collection component includes planning of measurement and sampling processes and collecting and handling the data. The remaining steps or components overlap with each other. In detail, the problem phase is the same as posing statistically investigative questions. The analysis phase overlaps with analyzing the data component. Lastly, the conclusion phase is compatible with interpreting the results.

Understanding the statistical investigation process, also called the statistical problem-solving process, is one aim of the statistics curriculum (Arnold & Franklin, 2021; Gal & Garfield, 1994). Gal and Garfield (1994) address the importance of the nature of the process because all steps in the investigation are related. Statistical investigations start with a statistical investigative question which needs to be answered by collecting data, analyzing the collected data, and interpreting the results (Bargagliotti et al., 2020; Gattuso & Ottaviani, 2011; Newton et al., 2011). Bargagliotti et al. (2020) emphasize

the questioning throughout the investigative cycle and point out one of the end products of the cycle as a new statistical investigative question. From a curricular point of view, many countries consider the investigative cycle as the core of teaching statistics (Leavy, 2015).

Burgess (2011) described two possible ways of statistical investigations in classroom practices. First, the students can be given the data to explore and pose questions based on the data or focus on the data analysis and interpretation steps of statistical investigations. One step in the statistical investigation can be the focus on the practice rather than the whole process (Litke & Hill, 2020). Even then, the learners need to engage in the step by considering other steps in the investigative cycle (MacGillivray & Pereira-Mendoza, 2011). Second, the statistical investigations start with statistical investigative questions and are followed by the remaining phases. Learners need to be familiar with and engage in a statistical problem-solving process to develop students' statistical thinking, reasoning, and literacy (Burgess, 2011; Gal & Garfield, 1997).

Furthermore, the relationship between the main concepts and the stage of a statistical investigation is crucial (Watson & Callingham, 2015). Watson and Callingham (2015) explained how big ideas are present in each step of statistical investigation. They stated that variation and expectation play a role in every stage of the investigation process. Without variation and expectation, questions would not be meaningful. Additionally, data collection is carefully planned to consider both variation and expectation. The concept of randomness is vital to ensure the quality of data collection. In data analysis, the expectation within the variation is determined; distribution becomes essential due to organizing variation for people to understand the data through expectation. In the statistical investigation, the expectation needed to be determined to reach a conclusion about the research question; reaching a conclusion involves informal and formal inferences with a level of confidence based on the school level.

All in all, the two-dimensional framework, including the statistical problem-solving process and the developmental levels described in the GAISE reports, is considered one theoretical background of this study (Bargagliotti et al., 2020). In the following section, the statistical problem-solving process with the following four components is

presented separately: “(1) formulating statistical investigative questions, (2) collecting or considering data, (3) analyzing data, and (4) interpreting results” (p.13) with three developmental levels. After describing the two-dimensional framework, I will explain the place of statistical investigations in the Turkish curriculum and then provide related literature on statistical investigations.

2.1.3.1. Formulating Statistical Investigative Questions

The first component of statistical investigation is about the problem or the statistical investigative questions. Arnold (2013) emphasizes the “starting point” of statistical investigation based on the PPDAC cycle (p. 18). Frischemier and Leavy (2020) state that there is limited research on the “problem” step of statistical investigations compared to the data, analysis, and conclusions steps. Recent literature focuses on the importance of formulating statistical investigative questions for conducting the statistical investigation (e.g., Arnold, 2013; Arnold & Franklin, 2021; Bargagliotti et al., 2020; Frischemier & Leavy, 2020)

Arnold and Franklin (2021) discuss that all steps in the statistical problem-solving process include statistical questions. They describe four types of statistical questions, which are interrogative questions, investigative questions, data collection questions, and analysis questions. Interrogative questions, data collection questions, and analysis questions are statistical questions that appear in one or more steps in the statistical problem-solving process. Investigative questions are the ones that need to be solved throughout the problem-solving process (Arnold & Franklin, 2021). Additionally, Frischemeier and Leavy (2020) also describe the statistical questions based on the number of variables: one-variable and two-variable statistical questions. One-variable statistical questions are categorized into three types: yes/no questions, questions that focus on a specific value, and questions that focus on variable distribution. On the other hand, two-variable statistical questions are categorized into three types: yes/no questions, group comparison questions, and open and complex questions involving two variables (Frischemeier & Leavy, 2020). These categorizations highlight the different ways students can be engaged in statistical problem posing considering the number of variables.

Statistical investigative questions have specific characteristics. First, the investigative questions need to anticipate variability. Both GAISE reports discuss the difference between a deterministic question and a statistical investigative question anticipating variability (Bargagliotti et al., 2020; Franklin et al., 2007). In addition, the questions need to clearly state the variables, the participants, and the intent (Bargagliotti et al., 2020, p. 14). In addition, the investigative questions are followed by data collection or considering secondary data. All in all, the investigative questions need to be worthwhile to investigate and lead to statistical investigations, which consist of other types of statistical questions (Bargagliotti et al., 2020; Arnold & Franklin, 2021).

Table 2.1 illustrates students' progression through developmental levels from Level A to Level C in the “Formulate Statistical Investigative Questions” process component identified by Bargagliotti et al. (2020, p. 16).

Table 2.1

Three Developmental Levels in the Formulate Statistical Investigative Questions Process Component

Process Component	Level A	Level B	Level C
I. Formulate Statistical Investigative Questions	Understand when a statistical investigation is appropriate Pose statistical investigative questions of interest to students where the context is such that students can collect or have access to all required data Pose summary (or descriptive) statistical investigative questions about one variable regarding small, well-defined groups (e.g., subset of a classroom, classroom, school, town) and extend these to include comparison and association statistical investigative questions between variables Experience different types of questions in statistics: those used to frame an investigation, those used to collect data, and those used to guide analysis and interpretation	Recognize that statistical investigative questions can be used to articulate research topics and that multiple statistical investigative questions can be asked about any research topic Understand that statistical investigative questions take into account context as well as variability present in data Pose summary, comparative, and association statistical investigative questions about a broader population using samples taken from the population Pose statistical investigative questions that require looking at a variable over time Understand that there are different types of questions in statistics: those used to frame an investigation, those used to collect data, and those used to guide analysis and interpretation Pose statistical investigative questions for data collected from online sources and websites, smartphones, fitness devices, sensors, and other modern devices	Formulate multivariable statistical investigative questions and determine how data can be collected and analyzed to provide an answer Pose summary, comparative, and association statistical investigative questions for surveys, observational studies, and experiments using primary or secondary data Pose inferential statistical investigative questions regarding causality and prediction

Note: From *Prek-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education* (p.16), by A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson., L. Perez, and D. Spangler, 2020, National Council of Teachers of Mathematics. Copyright © 2020 by American Statistical Association.

The table also provides detailed recommendations at each level. It is important to note that each level is a prerequisite for the next level in posing statistical investigative questions, highlighting the importance of building on prior knowledge and skills in developing statistical thinking.

The GAISE reports (Bargagliotti et al., 2020; Franklin et al., 2007) state that Level A requires students to recognize a statistical investigative question and the different types of questions in statistics. They start to pose descriptive statistical investigative questions and experience comparative and associational questions. Data collected needs to be from one variable with a specific group close to the students in general. Level B adds another layer to Level B because students are expected to pose multiple statistical investigative questions for a context or a research topic, consider variability in the questions, and consider broader populations. The students also start to engage in secondary data at this level. In Level C, the students need to pose multivariable and inferential statistical investigative questions and use different research designs with primary and secondary data in addition to Level A and B requirements.

2.1.3.2. Considering Data/ Data Collection

The second component of the statistical problem-solving process is called data collection. The statistical investigative questions need to be answered through data collected or considered (Bargagliotti et al., 2020; Franklin et al., 2007; Gattuso & Ottaviani, 2011). One main aim of statistical practices is to collect sample data to make generalizations about the populations (Watson et al., 2018).

The component is covered in different frameworks. First, Wild and Pfannkuch (1999) described data collection in the plan and data phases of PPDAC. Franklin et al. (2007) described this step as data collection in the first GAISE report. Then, in the latest GAISE report, the researchers named this step “collect/consider the data” due to the change in the current world (Bargagliotti et al., 2020). They discussed the different types of data based on adding the word “consider” to this step. The data collected firsthand is called primary data, and the data already present, not collected firsthand, is called secondary data. Both types of data are handled in this step. The data need to

be collected or considered based on a statistical investigative question followed by the data analysis and interpretation phases. In addition, the planning phase of data collection needs to consider the variability issue in which different research designs need to understand, control, reduce or recognize and explain the variability (Bargagliotti et al., 2020).

It is also discussed the necessity of students' experiencing this step to make sense of the data in general (Ben-Zvi et al., 2018; Hall, 2011). Students can gain skills and knowledge related to data by planning and carrying out the data collection. This step also helps students to understand the following stages of the statistical problem-solving process. By making sense of the data in this step, the students understand the quality of the data, the analysis needed, and the findings (Ben-Zvi et al., 2018). The students could make appropriate inferences about their initial problem with a firm understanding of the data (Hall, 2011).

Table 2.2 illustrates students' progression through developmental levels from Level A to Level C in the "Collect Data/Consider Data" process component identified by Bargagliotti et al. (2020, p. 17). The table also provides detailed recommendations for instruction at each level. It is important to note that each level is a prerequisite for the next level in the data collection component, highlighting the importance of building on prior knowledge and skills in developing statistical thinking.

The GAISE reports (Bargagliotti et al., 2020; Franklin et al., 2007) state that the students get to know the data as information and collect or consider data through surveys to understand a specific group of interest in Level A. Students at Level B can use the data collected or considered to answer a statistical investigative question to understand a population. In addition, Level B requires students to have a sense of random assignment to grasp the concept of controlling variables. At Level C, students can first plan the data collection or consideration process and consider further issues like random sampling and random assignment with different types of data collection methods. The random assignment issue at Level C requires students to understand the cause-and-effect relationship.

Table 2.2*Three Developmental Levels in the Collect Data/Consider Data Process Component*

Process Component	Level A	Level B	Level C
II. Collect Data/ Consider Data	Understand that data are information; recognize that to answer a statistical investigative question, a person may collect data themselves specifically for that purpose, or a person may use data that have been collected by other people for another purpose Understand how to collect and record information from the group of interest using surveys and measurements collected from observations and simple experiments Understand that a variable measures the same characteristic on several individuals or objects and results in data values that may fluctuate Understand that within a data set there can be different types of variables (e.g., categorical or quantitative) Interrogate the data set to understand the context of the variables as they may relate to statistical investigative questions Understand that data are not always pristine but may contain errors, have missing values, etc., and that decisions have to be made about how to account for these issues	Understand that data are information collected and recorded with a purpose and can be organized and stored in a variety of structures (e.g., spreadsheets) Understand that a sample can be used to answer statistical investigative questions about a population. Recognize the limitations and scope of the data collected by describing the group or population from which the data are collected Understand that data can be used to make comparisons between different groups at one point in time and the same group over time Recognize that data can be collected using surveys and measurements, and develop a critical attitude in analyzing data collection methods Understand that quantitative variables may be either discrete or continuous Understand how to interrogate the data to determine how the data were collected, from whom they were collected, what types of variables are in the data, how the variables were measured (including units used), and the possible outcomes for the variables Understand that data can be collected (primary data) or existing data can be obtained from other sources (secondary data) Understand how random assignment in comparative experiments is used to control for characteristics that might affect responses	Word as: Apply an appropriate data collection plan when collecting primary data or selecting secondary data for the statistical investigative question of interest. Distinguish between surveys, observational studies, and experiments Understand what constitutes good practice in designing a sample survey, an experiment, and an observational study Understand the role of random selection in sample surveys and the effect of sample size on the variability of estimates Understand the role of random assignment in experiments and its implications for cause-and-effect interpretations Understand the issues of bias and confounding variables in observational studies and their implications for interpretation Understand practices for handling data that enhance reproducibility and ensure ethical use, including descriptions of alterations, and an understanding of when data may contain sensitive information Understand how concerns about privacy and human subjects may affect the collection and distribution of data Understand that in some circumstances, the data collected or considered may not generalize to the desired population, or this data may be the entire population

Note: From *Prek-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education* (p.17), by A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson., L. Perez, and D. Spangler, 2020, National Council of Teachers of Mathematics. Copyright © 2020 by American Statistical Association.

Furthermore, while the students at Level A understand the existence of different variables that “measure the same characteristic on several individuals” (p.17), students at Level B use data for comparison purposes and are required to differentiate the types of quantitative variables. In Level C, students consider confounding variables. Lastly, students at Level A question the data in the context. At Level B, the students question the whole process of data collection. At Level C, the students question further issues like reproducibility and ethical aspects.

2.1.3.3. Analyze the Data

The data analysis is the third component of the statistical problem-solving process. Wild and Pfannkuch (1999) define the analysis step with four issues: “Data exploration, planned and unplanned analyses, and hypothesis generation” (p.226). In GAISE reports, the data analysis component consists of “selecting appropriate graphical and numerical methods” (Franklin et al., 2007, p.11). Those tools for data analysis are necessary to “explore, describe, and compare variability in distributions” (p.14). Both exploratory and confirmatory data analysis (also known as inferential statistics) are essential components of statistical investigations.

First, the literature discusses the importance of exploratory data analysis (EDA) (Ben Zvi et al., 2018; Burgess, 2011; Cobb & Moore, 1997; Makar, 2013; Wild et al., 2018). John Tukey introduced exploratory data analysis in the 1970s (Ben Zvi et al., 2018; Garfield & Ben-Zvi, 2008). EDA is “the discipline of organizing, describing, representing, and analyzing data, with a heavy reliance on informal analysis methods, visual displays, and, in many cases, technology” (Garfield & Ben-Zvi, 2008, p.126). EDA consists of descriptive statistics like the concepts of variation, distributions, aggregation, and average and graphs like stem and leaf plots, box plots, and bar graphs (Garfield & Franklin, 2011; Konold & Pollatsek, 2002; Makar, 2013). These are the tools for describing the distributions (Makar, 2013).

Tukey’s introduction of EDA has changed the statistical practices (Ben Zvi et al., 2018; Burrill & Biehler, 2011). EDA helps students understand the data through more simple data tools like graphs for visual techniques and arithmetic calculations to discover the patterns in the data (Wild et al., 2018). EDA helps statistics education be more accessible to “the school level” by not dealing with probability and inferential statistics first (Burrill & Biehler, 2011, p.61). Burgess (2011) discuss EDA’s contributions to school curricula which focus on students’ work on real data through investigations. They also stated that working with real data in learning and teaching statistics is a foundation for “more formal statistical ideas” (p.260). In addition, Braham and Ben Zvi (2019) underlined that working with real data through EDA also helps learners to form statistical investigative questions, make assumptions related to

the problem, relate those conjectures with the data in context, and lastly, make formal inferences.

Tukey (1977) defined EDA as detective work due to its nature of understanding data in different aspects before going into details of planned data analysis. EDA is also called “statistics without probability” (Burrill & Biehler, 2011, p.11). Tukey (1977) discussed that EDA is a necessary component of data analysis, like the first thing to do during data analysis. Still, he also discussed that EDA is insufficient alone because it does not give the whole picture of the problem we would like to solve.

Data analysis relies on inferential statistics to make inferences from samples to populations and includes more complex mathematical computations and probabilistic thinking (Ben Zvi et al., 2018; Cobb & Moore, 1997; Gattuso & Ottaviani, 2011). Inferential statistics is also called “statistics with probability” (Burrill & Biehler, 2011, p.11). It was present even before Tukey introduced EDA (Tukey, 1977). Tukey (1980) discussed that both data analyses should be taught, and it is better to introduce EDA first due to the techniques it contains for understanding the data. It helps to see “go beyond the data” through predictions and estimates about the data (Makar, 2013, p.35).

Currently, the practice of statistics follows Tukey’s suggestions. Students not introduced to probability engage in more descriptive statistics and graphs included in EDA (Makar, 2013; Watson et al., 2018). After introducing probability, students are expected to make formal inferences through confirmatory data analysis techniques (Burrill & Biehler, 2011).

Table 2.3 compares Exploratory Data Analysis (EDA) and standard inference (or inferential statistics) in three aspects: the purpose, the audience for the conclusions, and the formality of the conclusions. Cobb and Moore (1997) have distinguished the two types of data analysis based on these aspects.

Furthermore, children mainly deal with data in textbooks, so they engage in statistical investigations' data analysis and interpretation steps (Hall, 2011). Children’s

experiences with the data analysis remain procedural, with a minimum focus on statistical thinking and reasoning (Newton et al., 2011).

Table 2.3

Exploratory Data Analysis vs. Formal Probability-Based Inference

Exploratory Data Analysis	Statistical Inference
Purpose is unrestricted exploration of the data, searching for interesting patterns.	Purpose is to answer specific questions, posted before the data were produced.
Conclusions apply only to the individuals and circumstances for which we have data in hand	Conclusions apply to a larger group of individuals or broader class of circumstances
Conclusions are informal, based on what we see in the data.	Conclusions are formal, backed by a statement of our confidence in them.

Note: From “Mathematics, Statistics, and Teaching” by G. W. Cobb and D. S. Moore, 1997, *American Mathematical Monthly*, 104(9), p. 808 (<https://doi.org/10.2307/2975286>). Copyright 1997 by Mathematical Association of America

Statistics education has specific goals, like mastering procedural skills, understanding mathematical relationships, and understanding probability and chance (Gal & Garfield, 1994, p.4). Those aims are precursors of understanding EDA and inferential statistics. The core of statistics education is to understand statistical investigations as a whole process (Gal & Garfield, 1994) and to understand each component, like data analysis, as a part of the whole investigation (Garfield & Franklin, 2011).

Table 2.4 illustrates students' progression through developmental levels from Level A to Level C in the “Analyze the Data” process component, as identified by Bargagliotti et al. (2020, p.18). The table also provides detailed recommendations for each level. In addition, each level is a prerequisite for the next level in the data analysis component, highlighting the importance of building on prior knowledge and skills in developing statistical thinking.

GAISE reports (Bargagliotti et al., 2020; Franklin et al., 2007) state that understanding the distribution is a starting point for analyzing the data. Then students use different types of graphs; “tables, picture graphs, dot plots, and bar graphs” (p.18) at Level A and “dot plots and boxplots” at Level B to represent variability. Students are expected to understand key characteristics of distributions in quantitative variables such as center, variability, and shapes. Students use mean as the center, range as variability,

and clusters as shapes at Level A. In addition, students use mean and median as a center, interquartile, and mean absolute deviation as variability, symmetric and asymmetrical shape, and modes number shape at Level B.

Table 2.4

Three Developmental Levels in the “Analyze the Data” Process Component

Process Component	Level A	Level B	Level C
III. Analyze the Data	Understand that the distribution of a categorical variable or quantitative variable describes the number of times a particular outcome occurs Represent the variability of categorical variables or quantitative variables using appropriate displays (e.g., tables, picture graphs, dotplots, bar graphs) Describe key features of distributions for quantitative variables, such as: - center: mean as the equal share, and median as the middle-ordered value of the data - variability: range as the difference between the greatest and least value, and dispersion as how many units from the equal share value - shape: number of clusters, symmetric or not, and gaps Recognize distributions can be used to compare two groups Observe whether there appears to be an association between two variables	Represent the variability of quantitative variables using appropriate displays (e.g., dotplots, boxplots) Learn to use the key features of distributions for quantitative variables, such as: - center: mean as a balance point, and median as the middle-ordered value - variability: interquartile range and mean absolute deviation (MAD) - shape: symmetric or asymmetric and number of modes Use reasoning about distributions to compare two groups based on quantitative variables Explore patterns of association between two quantitative variables or two categorical variables: - measures of correlation: quadrant count ratio (QCR) - comparison of conditional proportions across categorical variables	Use technology to subset and filter data sets and transform variables, including smoothing for time series data Identify appropriate ways to summarize quantitative or categorical data using tables, graphical displays, and numerical summary statistics, which includes using standard deviation as a measure of variability and a modified boxplot for identifying outliers Summarize and describe relationships among multiple variables Understand how sampling distributions (developed through simulation) are used to describe the sample-to-sample variability of sample statistics Develop simulations to determine approximate sampling distributions and compute p-values from those distributions Describe associations between two categorical variables using measures such as difference in proportions and relative risk Describe the relationship between two quantitative variables by interpreting Pearson's correlation coefficient and a least-squares regression line Use simulations to investigate associations between two categorical variables and to compare groups Construct prediction intervals and confidence intervals to determine plausible values of a predicted observation or a population characteristic

Note: From *Prek-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education* (p.18), by A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson., L. Perez, and D. Spangler, 2020, National Council of Teachers of Mathematics. Copyright © 2020 by American Statistical Association.

Students at Level C choose appropriate graphs and measures of central tendency and use standard deviation as a measure of variability. Furthermore, while students at Level A understand distributions can be used for comparing two groups, students at Level B can use the distributions for comparison. Students at Level C use sampling distribution

with an understanding of sample-to-sample variability and calculate p-values based on the distributions.

Lastly, students at Level A are expected to understand if there is a relationship among variables. Students at Level B seek patterns in both categorical and quantitative variables by using quadrant count ratio and conditional proportions. Students at Level C use different association measures for categorical and quantitative variables (e.g., Pearson's correlation coefficient, a least-squares regression line, differences in proportions, and relative risk). In addition, students at Level C can use confidence intervals to understand the population characteristic.

2.1.3.4. Interpret the Results

The last component of the statistical problem-solving process is the “interpretation of the results.” The component corresponds to the “conclusions” step in the investigative cycle of Wild and Pfannkuch's (1999) framework described through interpretations, conclusions, new ideas, and communications. People need to interpret findings through “developing interpretive skills and statistical literacy” and “communicating statistically” to discuss the findings (Gal & Garfield, 1994, p.4-5). Verbal and written communication is necessary for all steps in statistical investigations. However, this step requires particular attention to teaching and developing communication skills because all steps eventually connect to this stage (MacGillivray & Pereira-Mendoza, 2011).

Furthermore, it is necessary to consider variability while interpreting the results. Especially acknowledging variability allows people to generalize and look beyond the data (Bargagliotti et al., 2020). While generalizing findings, possible biases or limitations must be considered (Gal & Garfield, 1994). Of necessity, people need to learn about commenting on or considering variations while commenting on the results because they can face incorrect comments and interpretations (MacGillivray & Pereira-Mendoza, 2011). In addition, MacGillivray and Pereira-Mendoza (2011) also emphasized the importance of recognizing this step as not searching for “correct

answers” to the questions but requiring “commenting, interpreting, and discussing” with the focus on variation (p.116)

The interpretation component requires people to consider all other steps (Bargagliotti et al., 2020; Gattuso & Ottaviani, 2011). For example, Gattuso and Ottaviani (2011) emphasized its close relation with the statistical investigative question because “statistics comes in touch with the questions” with the interpretation of the findings (p.125). The data analysis and interpretation steps have “a back and forth” relationship because interpreting the results might lead to critiques of current data analysis and the necessity of further analysis to answer the research question in depth (Wild et al., 2011, p.11). In addition, the interpreted results need to be connected to the data collected for the data collection process to make meaningful inferences and explain the findings better by assessing the process critically (Hall, 2011; Makar & Fielding-Wells, 2011).

Table 2.5 shows students’ progress through developmental levels from Level A to Level C in the “Interpret Results” process component. Detailed recommendations for each level were provided (Bargagliotti et al., 2020, p.19). The table also shows that each level is a prerequisite for the next level in the interpretation component. The GAISE reports (Bargagliotti et al., 2020; Franklin et al., 2007) describe the use of statistical analysis in answering investigative questions as a critical component across all levels of statistics education. However, effective communication of results requires guidance from teachers at Level A, less guidance but more comprehensiveness at Level B, and more formality with increased independence at Level C.

In addition, students at Level A can communicate the results within the limits of the group or population they used. At Level B, students are expected to communicate results beyond the sample with evidence and consider uncertainty, sampling variability, and representativeness issues. At Level C, students should be able to consider outliers, the estimate of a population characteristic, chance variation, the margin of error, issues related to missing values, the relationship among variables, and p values while interpreting the results

Table 2.5*Three Developmental Levels in the “Interpret Results” Process Component*

Process Component	Level A	Level B	Level C
IV. Interpret Results	Use statistical evidence from analyses to answer the statistical investigative questions and communicate results through structured answers with teacher guidance Make statements about the group or population from which the data were collected, recognizing that conclusions are limited to these groups and cannot be generalized to other groups Describe the difference between two groups with different conditions	Use statistical evidence from analyses to answer the statistical investigative questions and communicate results with comprehensive answers and some teacher guidance Acknowledge that looking beyond the data is feasible Generalize beyond the sample providing statistical evidence for the generalization and including a statement of uncertainty and plausibility when needed Recognize the uncertainty caused by sample to sample variability State the limitations of sample information (e.g., a sample may or may not be representative of the larger population, measurement variability) Compare results for different conditions in an experiment	Use statistical evidence from analyses to answer the statistical investigative questions and communicate results through more formal reports and presentations Evaluate and interpret the impact of outliers on the results Understand what it means for an outcome or an estimate of a population characteristic to be plausible or not plausible compared to chance variation Interpret the margin of error associated with an estimate of a population characteristic Acknowledge the presence of missing values and understand how missing values may add bias to an analysis Use multivariate thinking to understand how variables impact one another Communicate statistical reasoning and results to others in a variety of formats (verbal, written, visual) Understand how to interpret simulated p-values appropriately

Note: From *Prek-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education* (p.19), by A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson., L. Perez, and D. Spangler, 2020, National Council of Teachers of Mathematics. Copyright © 2020 by American Statistical Association.

2.1.4. Curriculum and Curriculum Expectations

GAISE is considered a curriculum document that provides recommendations for teaching and assessing statistics education at the K-12 level. GAISE aims to promote statistical literacy among students and help them develop the skills needed to make informed decisions based on data (Bargagliotti et al., 2020, Franklin et al., 2007). The details provided in GAISE are provided in the previous section. In addition, The National Council of Teachers of Mathematics (NCTM), a professional organization in the US focusing on improving mathematics education, developed a set of standards for teaching and learning mathematics called the “*Principles and Standards for School Mathematics*” (NCTM, 2000). The ten standards, providing directions for what students need to know and do in mathematics in pre-K-12 grade levels, are divided into two categories. Five are content standards; Data Analysis and Probability are a part of the content standards along with Number and Operations, Algebra, Geometry, and Measurement. The other five are called Process Standards: “problem-solving, reasoning and proof, communication, connections, and representations” (NCTM,

2000, p.29). In the Data Analysis and Probability Content Standard, NCTM (2000) provide some recommendations on what students need to learn in statistics throughout K-12 education (p.48):

Instructional programs from prekindergarten through grade 12 should enable all students to—

- *formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them;*
- *select and use appropriate statistical methods to analyze data;*
- *develop and evaluate inferences and predictions that are based on data;*
- *understand and apply basic concepts of probability.*

Furthermore, NCTM (2000) also provided grade-level correspondence for each recommendation stated above. For example, the correspondence expectation for the first recommendation, “formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them can be seen in Table 2.6 (Grades Pre-K-2, p.108; Grades 3-5, p. 176; Grades 6-8, p.248, Grades 9-12, p.324).

Table 2.6

Grade-level Expectations in Data Analysis and Probability Content Standard

PreK-2	Grades 3-5	Grades 6-8	Grades 9-12
pose questions and gather data about themselves and their surroundings;	design investigations to address a question and consider how data-collection methods affect the nature of the data set;	formulate questions, design studies, and collect data about a characteristic shared by two populations or different characteristics within one population;	understand the differences among various kinds of studies and which types of inferences can legitimately be drawn from each;
sort and classify objects according to their attributes and organize data about the objects;	collect data using observations, surveys, and experiments; represent data using tables and graphs such as line plots, bar graphs, and line graphs;	select, create, and use appropriate graphical representations of data, including histograms, box plots, and scatterplots.	know the characteristics of well-designed studies, including the role of randomization in surveys and experiments;
represent data using concrete objects, pictures, and graphs.	recognize the differences in representing categorical and numerical data.		understand the meaning of measurement data and categorical data, of univariate and bivariate data, and of the term variable;
			understand histograms, parallel box plots, and scatterplots and use them to display data;
			compute basic statistics and understand the distinction between a statistic and a parameter.

Note: Adapted from *Principles & Standards for School Mathematics* (p.108, 176, 248, 234). The National Council of Teachers of Mathematics, 2000. Copyright © 2000 by The National Council of Teachers of Mathematics.

NCTM also published books called Essential Understandings on each content area, including Statistics. Shaughnessy (2019b) summarized the big ideas provided in the books by grade level in Essential Understanding of Statistics- Grades 6-8 (Kader et al., 2013) and in Essential Understanding of Statistics- Grades 9-12 (Peck et al., 2013). Kader et al.(2013) stated that the big ideas in statistics for Grades 6-8 are “variability in data and distributions,” “comparing distributions,” “associations between two variables,” and “samples and populations” (as cited in Shaughnessy, 2019b, p.48). Also, Peck et al. (2013) provided five interrelated essential understandings for Grades 9-12 (as cited in Shaughnessy, 2019b, p.49):

- *Data consists of structure and variability.*
- *Distributions describe variability.*
- *Hypothesis tests answer the question, “Do I think this could have happened by chance?”*
- *The way data are collected matters.*
- *Evaluating an estimator involves considering bias, precision, and sampling method.*

In addition, in the United States (U.S.), the Common Core State Standards for Mathematics (CCSSM) are developed as a set of academic standards for mathematics education by the National Governors Association and the Council of Chief State School Officers (Common Core State Initiative [CCSSI], 2010; NCTM, 2014). These standards were developed by a state-led initiative and have been adopted by most states in the U.S. The CCSSM provides a clear and consistent framework for what students should know and be able to do in mathematics at each grade level, like NCTM’s principal and standards for school mathematics. Eight Standards for Mathematical Practice are present in the CCSSM (CCSSI, 2010, p.6-8):

- S1 Make sense of problems and persevere in solving them.*
- S2 Reason abstractly and quantitatively.*
- S3 Construct viable arguments and critique the reasoning of others.*
- S4 Model with mathematics.*
- S5 Use appropriate tools strategically.*
- S6 Attend Precision.*
- S7 Look for and make use of structure.*
- S8 Look for and express regularity in repeated reasoning.*

CCSSI (2010) stresses the importance of connecting the standards for mathematical practices to the standards for mathematical content. Therefore, each grade level from Pre-K to Grade 8 consists of a statistics section. During the Grade Pre-K-5 period, the statistics content is under the Measurement and Data section and contains “representing and interpreting data” content. During the Grade 6-8, the complexity increases, and the big ideas like variability, distribution, random sampling, informal inference, and chance are covered under Statistics and Probability content. At the High School level, CCSSM has conceptual categories rather than grade-level categorization. “Statistics and Probability” is one conceptual category including the content related to “interpreting categorical and quantitative data, making inferences, and justifying conclusions, conditional probability and the rules of probability, and using probabilities to make decisions” (CCSSI, 2010, p.80).

Franklin et al. (2015) also stated that the Standards for Mathematical Practices in the CCSSM document should be seen through a statistical lens by the teachers’ teaching statistics. While Franklin et al. (2015) consider CCSSM through a statistical lens, GAISE document (Franklin et al., 2007), Developing Essential Understanding of Statistics 6-8 (Kader et al., 2013), Developing Essential Understanding of Statistics 9-12 (Peck et al., 2013), The Challenge of Developing Statistical Literacy, Reasoning, and Thinking (Ben-Zvi & Garfield, 2004), and Statistical Thinking in Empirical Enquiry (Wild & Pfannkuch, 1999) were taken into account (as cited in Franklin et al., 2015, p. 9). Detailed statistical practices aligned with CCSSM are provided in the Statistics Education of Teachers (SET) document (Franklin et al., 2015, p. 10-12). Watson and English (2017c) list the eight standards for mathematics based on the SET document and call them the Standards of Statistics Practice (p.264):

- S1 Make sense of statistical questions posed and persevere in answering them.*
- S2 Reason abstractly in drawing informal inferences and reason quantitatively during numerical analyses of data.*
- S3 Construct viable arguments based on evidence from data and critique the reasoning of others.*
- S4 Model with statistics to find patterns within the variation in data.*
- S5 Use appropriate statistical tools strategically.*
- S6 Attend to precision in planning, collecting, and analysing data, and in the statistical language used.*
- S7 Look forward and make use of structure in data and in the process of the practice of statistics.*

S8 Look for and express regularity in repeating the practice of statistics in different contexts.

Furthermore, Groth and Bargagliotti (2012) link the GAISE and the CCSSM in a way that both documents have overlapped nature due to learning expectations from the students. GAISE document is prepared for the classroom implementations of the Data Analysis and Probability content in NCTM (2000) (as cited in Groth & Bargagliotti, 2012). Therefore, all documents emphasize the importance of statistical investigations and statistical content. The GAISE document is a supplementary document including “pedagogical approaches for statistics, meaningful statistical connections, developmental trajectories for students’ statistical learning, and enhancement of the curriculum prescribed by CCSSM” (Groth & Bargagliotti, 2012, p.39).

NCTM, CCSSM, and GAISE are curriculum documents in mathematics and statistics education (Shaughnessy, 2019b). NCTM (2014) differentiates the terms “standards” from “curriculum” (p.70):

Standards are statements of what students are expected to learn. Standards are the ends. A curriculum is the program used to help students meet the standards, including instructional materials, activities, tasks, units, lessons, and assessments. The curriculum is the means.

According to the NCTM (2014), a strong curriculum is essential for effective teaching and learning. It is also reflected in the objectives of statistics curricula in schools, which aim to promote statistical literacy, reasoning, and thinking among students (Batanero et al., 2011). The curriculum expectations for statistics education in K-12 schools seek to provide students with a thorough understanding of statistical concepts and techniques and to develop their critical thinking and problem-solving skills to apply statistical methods to real-world situations (CCSSI, 2010; NCTM, 2000).

Burrill (2011) discussed different statistics school curricula around the world in terms of their structure, the details of content, or if nationally mandated based on the chapters in the handbook (i.e., Batanero et al., 2011). For example, she stated that the United States has no national curricula, but the states have their own general standards aiming at statistical processes (i.e., Newton et al., 2011). However, Uganda has a mandatory curriculum, but their curricula mainly focus on computational aspects (i.e., Opolot-

Okurut & Opyene Eluk, 2011). While Brazil and South Africa primarily focus on statistical literacy as they see students as consumers of data (i.e., Campos et al., 2011; Wessels, 2011, respectively).

Watson and English (2017c) discuss the place of statistical investigations in New Zealand at every grade level as a subcategory of the Statistics component of the Mathematics and Statistics curriculum with the increase in complexity throughout grade levels. Watson (2014) describes the curriculum expectations for teaching statistics in the Australian curriculum. She emphasizes the place of statistics and probability strands with the names of “chance” and “data representation and interpretations” from years 1 to 10. Callingham et al. (2021) state that the Australian curriculum is updated around five big ideas in the statistics and probability strand of the curriculum. The mathematics curricula in New Zealand and Australia significantly focus on teaching statistics in the context of statistical problem-solving with a focus on big statistical ideas.

Furthermore, Watson and English (2017c) explain the integration of statistical investigations within the curriculum of New Zealand at every grade level as a sub-component of the Statistics strand of the Mathematics and Statistics curriculum, with a gradual increase in complexity as students’ progress through the grades. Watson (2014) also outlines the curriculum expectations for teaching statistics in the Australian curriculum, highlighting the emphasis placed on the statistics and probability strands, specifically the concepts of “chance” and “data representation and interpretation,” throughout the primary and secondary education years. Additionally, Callingham et al. (2021) note that the Australian curriculum has undergone revisions to incorporate a focus on five central concepts within the statistics and probability strand. Both the New Zealand and Australian mathematics curricula emphasize teaching statistics within the context of statistical problem-solving and understanding fundamental statistical concepts.

In the following section, I will explore the place of statistics in the Turkish curricula. It will include related literature examining the curriculum expectations, focus areas, and the emphasis placed on statistical concepts and problem-solving within the

curricula throughout the Turkish Republic's history. It will also include an analysis of any recent revisions or updates made to the curriculum regarding statistics.

2.1.4.1. Statistics in Turkish Curricula

Regarding the Turkish curriculum, statistics is seen as one strand in mathematics curricula. Ader (2016; 2018) made comprehensive document analyses based on Turkish mathematics curricula throughout the republican period. Ader (2016) analyzed nine primary school mathematics curricula regarding objectives related to “data handling.” For example, all curricula focused on the context except the 1926 curriculum. Again, graphs were present in all curricula except 1926. The focus on a problem situation and collecting data for solving the problem were the focus in addition to graphs-related content. In summary, earlier curricula included more graph-related issues like reading, constructing, and interpreting graphs. The latest curricula also emphasized statistical processes like problem posing, data collection, analysis, and interpretation of tables and graphs (see Table 2.7)

Table 2.7

Learning Outcomes Related to Statistics in the Primary School Mathematics Curriculum

Grade Level	Learning Outcomes
1	Data Collection and Interpretation - Reads simple tables with up to two sets of data.
2	Data Collection and Interpretation -Collects data, classifies it, and organizes it in a tree diagram, scoreboard, or frequency table by asking questions on any problem or topic; creates an object and shape graph.
3	Data Collection and Interpretation - Makes and interprets conversions from graph to scoreboard and frequency table by explaining the information shown in the shape and object graph. - Solves problems that require addition and subtraction operations by using the information given in graphs or by creating graphs. - Reads and interprets simple tables belonging to up to three data groups and edits the data obtained from the table.
4	Data Collection and Interpretation -Examines the bar graph and makes interpretations and predictions on the graph. -Creates bar graph. -Uses different representations to present the data collected. -Solves problems related to daily life by using the information shown by a bar graph, table, and other graphs.

Note: Adapted from *Mathematics Teaching Program (Primary and Middle School Grades 1, 2, 3, 4, 5, 6, 7 and 8)*, Ministry of National Education (MoNE), 2018.

Ader (2018) analyzed ten middle school mathematics curricula presented in the republican period. According to the document analysis, the first three curricula do not include data-related content. It first started as a part of arithmetic and algebra topics. The following three curricula (1977, 1990, and 1998) were structured as units; data-related content was one unit in those programs. Later, with the reform movement in 2005, mathematics curricula were divided into strands, and data-related content belonged to the “Probability and Statistics” strand within five strands. Data-related content in the last two programs (2017 and 2018) is under the “Data Processing” strand within the five strands. All curricula from 1949 cover the goals like developing skills for real-life connection, constructing tables and graphs, using data, and having research perspectives in problem situations, and interpreting data and making generalizations based on data.

When all curricula were examined regarding the statistical problem-solving process components, problem-posing started with the 2005 program. Although data collection, analysis, and interpretation were implicitly present in curricula until the 2005 program, the objectives related to those components have been a part of curricula after the 2005 program. As the content of the objectives was also investigated, bar graphs, circle graphs, and line graphs have been present in all programs after the 1977 program. In contrast, a tree diagram was only present in the 2013 program, and a histogram was only present in the 2005 and 2013 programs. Also, measures of central tendency (i.e., arithmetic mean, mode, median) have been in the programs since 1977. Additionally, measures of spread were integrated into the curriculum in 2005. However, the interquartile range and standard deviation were removed in the 2013 and 2017 programs.

In this study, I studied seventh-grade statistics instruction because all big ideas related to statistics in K-8 grade education are covered until seventh grade. Also, the seventh-grade mathematics curriculum includes statistics in the most comprehensive way throughout the K-8 grade curriculum. During the study, the 2017 program was covered by the teachers. However, the latest program (MoNE, 2017a) was updated in 2018; there were minor changes in the program in the data processing strand. The objectives in all grades have not changed, but the order of the objectives changed within a grade

level. For example, in the 7th-grade program, the first objective was related to the circle graph, followed by the line graph objectives in the 2017 program. Still, it was updated in the 2018 program as other objectives, including circle graph-related objectives, follow the line graph objective. Table 2.8 shows the learning outcomes related to statistics in the middle school mathematics curriculum (MoNE, 2018).

Table 2.8

Learning Outcomes Related to Statistics in The Middle School Mathematics Curriculum

Grade Level	Learning Outcomes
5	Data Collection and Interpretation -create research questions that require data collection - collect data on research questions and display it with a frequency table and bar graph. - solve problems for interpreting data represented by frequency tables or bar graphs.
6	Data Collection and Interpretation - create research questions that require comparing two sets of data and obtaining appropriate data. -Show two groups' data with a binary frequency table and bar graph. Data Analysis -calculate and interpret the range of a data set. -calculate and interpret the arithmetic mean of a data group. -use the arithmetic mean and range to compare and interpret data of two groups.
7	Data Analysis -create and interpret line graphs related to data. -calculate and interpret a data set's arithmetic mean, median, and mode. -create and interpret the circle graph for a data group. -show data on the bar, circle, or line graphs and make appropriate transformations between these representations.
8	Data Analysis -interpret line and bar graphs of up to three data groups. -show data on the bar, circle, or line graphs and make appropriate transformations between these representations.

Note: Adapted from Mathematics Teaching Program (Primary and Middle School Grades 1, 2, 3, 4, 5, 6, 7 and 8), Ministry of National Education (MoNE), 2018.

As displayed in Table 2.8, the data processing strand in the latest middle school mathematics program in Turkey has a significant focus on the big ideas of expectation in the form of the average, variation in the forms of range and distribution, and representation of data using graphs. Ader's (2018) document analysis supports this focus on the big ideas in the latest curriculum.

Furthermore, both primary and middle programs were present in the document. The document briefly introduced the strands under primary and middle school back-to-back. The focus on the statistical problem-solving process is present in the data processing sections. The primary section states that (MoNE, 2018, p.11):

The objectives are prepared within the framework of two dimensions. First, the objectives and grade levels are created by considering the prominent model in teaching data. Teaching data consists of four steps: generating researchable questions, collecting data, processing and analyzing data, and interpreting results. The data learning area must be carried out based on these steps. The context of research problems and the numbers used in those problems should be considered within the limitations of the classroom. The second dimension is creating and interpreting various tables and graphs using data. In addition, as the grade levels increase, the students can work with more data groups. Then the program describes the specific objectives by grades. The objectives are related to tables and graphs from first to fourth grade. In the fifth grade, students are expected to form the statistical investigative question and collect, process, and interpret the data based on the question. In later grades, students are introduced to measures of central tendency and range concepts for exploratory data analysis with more complex graphs like circle and line graphs.

The statement above summarizes the objectives outlined in the data processing strand for both primary and middle school mathematics curricula. However, it was explicitly mentioned in the primary school section. It can be observed that the focus is on the statistical processes from primary to middle school years. This continuity and progression of statistical concepts and statistical problem-solving skills throughout the primary and middle school curriculum need to ensure that students develop a strong foundation in statistical thinking.

In addition to the comprehensive document analyses and the latest program's details (Ader, 2016; 2018), research has been conducted on the curricular materials in the Turkish context (Batur et al., 2019; Özmen & Baki, 2019). First, Özmen and Baki (2019) investigated the middle school mathematics curriculum regarding statistical literacy. They examined two national curricula published in 2013 and updated in 2018 from 5th to 8th grade. The indicators for statistical literacy are determined as statistical processes, reasoning, basic concepts, and context. They found that the objectives in all grades consist of statistical processes to some extent. Therefore, all components of the statistical problem-solving process are covered as a part of statistical literacy in the

curriculum. However, each grade's content covered different steps in detail. For example, problem-posing was explicitly present in the 5th and 6th grades. Also, the objectives consisted of reasoning while covering selecting appropriate central tendency measures and appropriate graphs concerning data. However, interpreting the data and making inferences based on data were not explicitly present. In addition, the basic concepts were stated in each grade's statistics section, like in other contents. Nevertheless, the mathematical process skills were present without the adaptation to statistics content. Lastly, the focus on context was limited in the curriculum. The real-life context was emphasized in just 6th-grade objectives.

Based on the curriculum document, mathematics textbooks have been prepared by committees under the supervision of the Turkish Ministry of Education and distributed to students all over the country free of charge. In this context, as a part of curricular resources, Batur et al. (2019) examined Turkish mathematics textbooks for middle schoolers based on the principles presented in the GAISE 2005 report. They found that the textbooks do not mainly reflect the principles in the GAISE 2005 report even though they included example questions with real-life context, interpretation questions, conceptual and procedural understanding of statistical terms, and linking different statistical concepts. However, the questions in the textbooks lacked real data, statistical investigative processes, the use of technology, randomness, and making inferences from a sample to the population. Furthermore, Bütüner (2020) conducted a study on the coverage of the central tendency, one of the big ideas in Turkish mathematics textbooks. The study's results revealed that the fair-share and balance models and the use of multiple representations were not prevalent in the 6th-grade textbooks. Instead, many examples were found to utilize the add-and-divide algorithm. Additionally, the contextualized problems provided in the textbooks were found to lack a focus on the meaning of the representative value of a data set and, instead, focused solely on finding a single value without considering its relationship to the other data in the set.

The literature also points out the differences and similarities among the statistics contained in the mathematics curricula of different countries, including Turkey (Batur et al., 2021; Frischmeier et al., 2022): First, Batur et al. (2021) compared the statistics

section of different countries' curricula regarding the statistical problem-solving process using the GAISE I framework. Their comparison consisted of Turkish, Singapore, Korean, United States, and New Zealand statistics curricula in K-12 education. They used the GAISE framework, including four steps in statistical problem-solving with three developmental levels. The results indicate that the Turkish curriculum mainly included Level A learning outcomes and no Level C learning outcomes. The number of level B learning outcomes in the Turkish curriculum was less than in other curricula.

Similarly, the Korean curriculum included Level A and Level B learning outcomes but limited Level C learning outcomes. Unlikely, the other three curricula mainly included Level B learning outcomes and a relatively higher number of Level C learning outcomes than the Korean and Turkish curricula. Further analyses showed that all compared curricula mostly had learning outcomes in the data analysis component, especially in Level B. However, the Turkish curriculum's most frequent learning outcomes related to data analysis are at Level A, unlike in other countries.

Furthermore, Batur et al. (2021) also compare the overall number of learning outcomes and the content of those learning outcomes using the GAISE I framework. The results show that the curricula from Turkey and Korea have a relatively low number of learning outcomes than the other countries (6% and 5% of all objectives, respectively). However, New Zealand spares 30% of learning outcomes to statistics education, the highest percentage among all curricula compared in the study. When the statistical concepts addressed in the learning outcomes are examined, Turkey has very limited concepts compared to other countries. Korea is like Turkey in that aspect, but unlike Turkey, Korea includes stem and leaf plot and normal distribution to their curriculum. Turkey consists of a very limited range of graphs, no distribution concepts, correlation and regressions, and no parametric or nonparametric tests integrated into the curriculum. Those are primarily present in higher grades in other countries.

Frischemeier et al. (2022) also compared curricular documents of primary schools from Germany, Ireland, Cyprus, and Turkey in terms of early statistical thinking based on the developmental level A from the GAISE II framework. The analyses of the

different countries' primary mathematics curricula regarding Level A concepts, activities, and experiences regarding the statistical problem process due to the primary years mostly correspond to Level A in the GAISE II framework. First, all countries address posing investigative questions of interest to students. Nuance differences among countries about the issue are present. To illustrate, while the Turkish curriculum includes such questions in second and fourth grade expecting students to collect data from the classroom; Germany expects students to pose investigative questions requiring real-life data. None of the countries has the Level A concepts, activities, and activities, including “understand when a statistical investigation is appropriate, pose statistical investigative questions about one variable... and extend to include comparison and association questions, and experience different types of questions” (p.3). In the collect data/consider data step, all corresponding curricula expect students to understand that data is information and how to collect and record information from the group of interest. None consists of students interrogating the data set to understand the context of the variables and students understanding that data may contain errors or have missing values.

Furthermore, Frischemeier et al. (2022) also compared data analyses and interpretation steps of statistical problem-solving in Level A of the GAISE II framework. In the data analysis step, Turkey, unlike other countries, does not have a particular emphasis on distributions in primary years. Cyprus is similar to Turkey, but there is a limited emphasis on distribution. Both countries mostly focus on data representations in primary years, but Germany and Ireland focus on variability and distribution concepts in addition to representing data for the purpose of analysis. In the interpretation step, all countries focus on “the use of statistical evidence to answer the questions and communicate results” (p. 4). The focus on the interpretation is on later grades in Turkey with the reading and interpreting tables and graphs, without the focus on generalizability and comparisons of findings. Other countries also have a similar approach, but the scope might differ. Ireland's focus is on answering the statistical question through communicating based on graphs used in data analysis. Also, in the Cyprus curriculum, fourth-grade learning outcomes include two group comparisons based on range and mode.

In summary, the learning objectives in the Turkish mathematics curriculum are mainly consistent with those of other curricula around the world, with a greater emphasis on statistical problem-solving over time (Ader, 2016, 2018; Frischemeier et al., 2022; Özmen & Baki, 2019). However, the depth of these objectives is relatively shallow, as they do not fully encompass Level B and Level C objectives and activities as outlined in the GAISE frameworks (Batur et al., 2021), and the concepts addressed are limited in comparison to those of other countries (Batur et al., 2021). Despite efforts to incorporate current approaches in statistics education, such as integrating statistical problem-solving processes into the curriculum (Özmen & Baki, 2019), the instructional materials provided by the Ministry of National Education (MoNE) do not fully align with the principles outlined in the GAISE reports (Batur et al., 2019).

2.1.5. Related Literature on Statistical Investigations

Statistics is a complex field encompassing various elements and stages in statistical investigations. As such, numerous studies have focused on examining a specific aspect of statistics. For instance, many studies have emphasized the importance of posing a research question to students or collaborating with them to formulate a question (e.g., Allmond & Makar, 2010; Arnold, 2008; Arnold & Pfannkuch, 2018; Watson & English, 2017a). Furthermore, several studies have focused on data collection, with a prevalent emphasis on data collection tools and sampling (e.g., Fielding-Wells, 2010; Groth, 2003; Hedges et al., 2022; Hedges & Given, 2023). Furthermore, the studies may also encompass the analysis of data, which may involve statistical concepts such as creating tables and graphs, understanding measures of central tendency and variability, sampling, and distribution, or may even extend to decision-making with justification provided for the conclusions reached (e.g., Akkoç & Selçuk, 2017; Chick & Watson, 2002; Çatman Aksoy & Işıksal Bostan, 2020; Fitzallen & Watson, 2011; Makar, 2014; Makar & Confrey, 2005; Mokros & Russell, 1995; Santos & Ponte, 2013; Watson & Moritz, 2000). In addition, some studies examined the entire statistical problem-solving process from start to finish with a focus on students and teaching practices (e.g., Lavigne & Lajoie, 2007; Öz & Işıksal Bostan, 2021; Silva et al., 2022; Watson & English, 2015; Watson & English, 2017b; Watson & English, 2017c).

In this section, I will discuss studies examining the statistical problem-solving process, focusing on specific steps or concepts in statistical investigations, and studies examining the entire statistical problem-solving process. These studies aim to provide insight into the various elements and phases involved in statistical investigations and how they can be effectively taught and learned. The studies discussed will provide a comprehensive understanding of the statistical problem-solving process.

2.1.5.1. Studies on formulating statistical investigative questions

The section focuses on the studies on formulating statistical investigative questions (Allmond & Katar, 2010; Arnold, 2008; Arnold & Pfannkuch, 2018; Watson & English, 2017a). Table 2.9 summarizes key findings from the literature focusing on formulating statistical investigative questions.

Table 2.9

The Overview of The Studies on Formulating Statistical Investigative Questions

Study	Participants	Research Focus	Key Findings
Arnold (2008)	14-15-year-old students	Types of statistical investigative questions	Students improved their summary questions after a 10-lesson unit on the PPDAC cycle.
Allmond and Katar (2010)	25 nine-year-old students	Ability to generate purposeful statistical investigative questions	A specific focus on questioning improved the students' ability to pose statistical investigative questions.
Watson and English (2017a)	89 Grade 6 students	Statistical investigation process with a specific focus on problem posing	Students could complete the statistical problem-solving at least at the multi-structural level in all expected areas in the investigation
Arnold and Pfannkuch (2018)	27 14-year-old students	Ability to critique investigative questions posed by others	Students improved the quality of their critiques on statistical investigative questions from pre-test to post-test
English et al. (2017)	82 fourth-grade students	Developments in meta-questioning	Students could identify different types of questions and evaluate the quality of their own and others' questions.
Leavy and Frishemeier (2022)	158 prospective elementary mathematics teachers	Designing statistical investigative questions for group comparisons	Pre-service teachers improved their ability to create appropriate statistical investigative questions for primary school students that focused on making group comparisons through feedback from peers and experts.

First, Arnold (2008) conducted a study to examine the types of statistical investigative questions posed by 14-15-year-old students with multivariate data. The study was conducted as part of a teaching experiment in which a teacher and researcher collaborated to develop a ten-lesson unit on the PPDAC cycle. The study used pre-and post-test results from 12 students to analyze the types of questions posed at the beginning of the PPDAC cycle. In the pre-test, students were expected to pose a question, while students were expected to pose three summary questions and three comparison questions in the post-test. The results revealed that eight categories of summary and seven categories of comparison questions were identified. The pre-test results revealed that seven students posed non-summary questions, and four posed pre-summary questions. These questions are useful for data analysis but not for initiating statistical investigations. However, after the teaching experiment, most students improved their summary questions and could ask more appropriate questions, such as “how many” questions. Regarding comparison questions, most students could use the category of “A question that is answerable by the data.” Nonetheless, the researcher stated that only 18 questions posed in the post-test were suitable, and of the ten students, only one could create a fair comparison question.

Allmond and Katar (2010) conducted a study to examine the improvements in the ability of 25 nine-year-old students to pose purposeful statistical investigative questions in an inquiry-based classroom setting. The study was designed to evaluate the student's ability to generate statistically investigative questions by considering the characteristics of the questions, such as whether they were investigative or not, descriptive or inferential, and the type of data that could be collected based on the questions. The study was conducted throughout eight lessons, during which the students were asked to generate statistical investigative questions. The student-generated questions were then evaluated based on a hierarchy created by Chin and Kayalvizhi (2002) and Arnold (2009), which included categories such as irrelevant/off-topic, nonmathematical, non-investigative, closed, potential for investigation, investigative, and inquiry. The study included three groups of students: a main group focused on generating statistical questions, a comparison group working in an inquiry-based classroom without a focus on statistical questions, and a control group not in an inquiry-based classroom. Before the first group of students covered

the questioning unit, all groups were asked to write questions that needed investigation. The study results show that 4% of the main group created good questions, whereas 11% and 9% of the other groups created good questions. After completing the questioning unit, 60% of the main group created good statistical questions, whereas 11% and 0% of the other groups created good questions, respectively. The researchers concluded that a specific focus on questioning improved the students' ability to pose investigative questions in statistics rather than just providing an inquiry-based learning environment.

Watson and English (2017a) explored 89 Grade 6 students' statistical investigation process with a specific focus on problem posing. They used SOLO taxonomy for different phases of the study. They defined the levels of SOLO taxonomy for Grade 6 students; "(a) unistructural (Uni), where students use the elements of statistics and/or context in a singular fashion; (b) multi-structural (Multi), where students use several elements sequentially; and (c) relational (Rel), where students show evidence of interrelating elements in a meaningful way in the conclusion reached" (p.353). First, students formed statistically investigative questions based on context, and they suggested the types of data needed for the question. About 38% of students were at the unistructural level in which they focused only on one element, either question or data. The remaining students can focus on both the question and data; some provided more details for the question and data. Then they were given data and refined the question they posed, analyzed, interpreted the results, and re-interpreted them with additional context information. About 70% of students could complete the statistical problem-solving at least at the multi-structural level in all expected areas in the investigation.

Arnold and Pfannkuch (2018) conducted a study to evaluate the ability of 27 14-year-old students to critique investigative questions posed by others. The study analyzed the students' evaluations of five investigative questions using the SOLO taxonomy, which was based on descriptors of good statistical investigative questions and student responses. The SOLO taxonomy ranges from no response (0) to extended abstract (5) and evaluates the ability of students to identify both good and bad questions while also providing suggestions for improvement. The results show that students improved the quality of their critiques on statistical investigative questions from pre-test to post-test

due to the teaching experiment with learning activities that focused on students' understanding of statistical investigative questions. In the pre-test results, most of the students' answers were under the multi-structural level (3), where they could provide two pieces of evidence or related evidence to incorrect questions. In the post-test results, most answers were at the multi-structural level (3) or above. Additionally, 20% of the students reached the relational thinking level (4), where they could determine whether a question was good or bad and provide suggestions for improvement.

English et al. (2017) investigated 82 fourth-grade students' developments in meta-questioning in a statistical investigation through design-based research. They defined meta-questioning with three components; "(a) identifying the contextual issue to be investigated and posing initial questions, (b) interpreting and reviewing/critiquing initial questions, and (c) anticipating data collection and refining questions" (p.229). They looked at students' survey questions posed at the beginning and how and why the students revised their questions. The results show that the students could form questions like yes/no, open-ended, and multiple choice or selection questions. Through the revisions, the number of yes/no questions was declined due to students' realization of the limited answers provided by such questions. Then the students also tend to have more open-ended questions. Since the data from open-ended questions was difficult to handle, the researchers stated that multiple-choice questions were the most sophisticated, and a few students could manage to form such questions.

Leavy and Frishemeier (2022) investigated the problem-posing skills of 158 prospective elementary mathematics teachers in designing statistical investigative questions that required them to make group comparisons. The study used a three-phase design, beginning with the "Think" phase, in which the pre-service teachers were asked to design their initial statistical questions. In the second phase, called "Peer feedback," the pre-service teachers received feedback from their peers based on criteria developed from research by Arnold (2013) and Frischemeier and Biehler (2018). They were instructed to revise their statistical questions accordingly. In the final phase, called "Expert Feedback," the revised statistical questions were evaluated by experts using the same criteria. The study's results revealed that the quality of the statistical questions improved throughout the phases. The pre-service teachers were

able to develop statistical questions that were appropriate for primary school students and focused on making group comparisons.

Specifically, in the “look at the question” criterion, the pre-service teachers created meaningful and interesting statistical questions from the beginning, but they were not always clear. However, with the feedback provided by their peers and experts, the clarity of the questions improved from 62% to 85%. Additionally, the number of variables in the questions needed to compare groups improved from 52% to 88% with the feedback provided. The clarity of the population of interest also improved from 62% to 90% through the feedback provided. In the second component, “variables in the questions,” the clarity of the variables improved from 42% to 89%. The measurability of the variables was good from the beginning, but it improved even further throughout the feedback. In the third component, “relationship between data and question,” the pre-service teachers mainly selected reasonable question types, such as yes/no, nonbinary, and specific value questions. They developed the sophistication of their questions based on expert feedback. In the final component, “look at the data,” only certain participants' answers were used because some did not collect data based on their questions—however, those who did perform well in each sub-criterion.

To conclude, the results of the studies reviewed in this section point to significant improvements in students' ability to formulate and evaluate statistical investigative questions. The findings highlight the effectiveness of targeted learning activities in improving students' understanding of statistical questioning. Further findings indicate that using problem-posing and questioning in teaching statistics leads to more meaningful and appropriate questions posed by students.

2.1.5.2. Studies on Collecting/Considering Data

The studies discussed in this section focus on students' understanding and levels of thinking during the data collection phase of statistical investigation (Groth, 2003; Fielding-Wells, 2010; Hedges et al., 2022; Hedges & Given, 2023). Table 2.10 summarizes the main issues that emerged from the reviewed literature on collecting and considering data.

Table 2.10*The Overview of the Studies on Collecting/Considering Data*

Study	Participants	Research Focus	Key Findings
Groth (2003)	15 high school students	Levels of thinking in study design during the data collection phase of statistical investigation	Most students were at the uni-structural level, offered non-experimental methods, and did not understand representativeness.
Hedges et al. (2022)	20 seventh-grade students	Understanding of related concepts in the data collection phase of statistical investigation	Most students were at the uni-structural level and were aware of the need for collecting data but lacked an understanding of representativeness. Some students' ideas developed from uni-structural to relational levels during the lesson.
Hedges and Given (2023)	20 middle school students	Construction of statistical questions and study designs based on these questions	Students did not ensure the representativeness of their samples and reflected confirmation bias in their study designs.
Fielding-Wells (2010)	24 fifth-grade students	Planning of statistical investigations	Students improved their planning efficiency through explicit structure, group collaboration, and whole-class discussions.

First, Groth (2003) conducted a study to assess high school students and recent graduates' levels of thinking in terms of study design based on SOLO taxonomy. Fifteen students were interviewed and asked to design studies based on given sub-questions. The results indicate that only one student relied heavily on the internet or other sources rather than more formal strategies for collecting data, which placed her at a pre-structural level of thinking. Four students were at the uni-structural level and were aware of the need for collecting data but not of representativeness. Two students were at the multi-structural level and considered representativeness, but with a specific strategy. The remaining eight students were at the relational level and could propose one or more methods to obtain a representative sample. Five students at the uni-structural level offered a non-experimental method for data collection in the sub-question related to the experimental study. In comparison, five students at the multi-structural level could design an experimental study. Only four of these students went further and suggested control issues in the experimental research to address the trustworthiness issue.

Hedges et al. (2022) investigated students' understanding of related concepts in the statistical investigation data collection phase. They replicated Groth's (2003) research on study design with 20 seventh-grade students during a lesson. They categorized students' discourses based on the SOLO taxonomy adapted by Groth (2003). They found that most students were at the uni-structural level in which they were aware of the need for collecting data, but their comments did not include the concept of representativeness. Having no students at the pre-structural level limits the study due to the teacher's direction to collect data for the activity. Some students' ideas developed from uni-structural to relational levels during the lesson; they mentioned the concepts of sample and sample size but did not state the concept of stratification, even if they used the method. Hedges et al. (2022) also stated that their results align with Groth's (2003) findings. Groth (2003) and Hedges et al. (2022) used similar research methods, where students were given tasks to design studies based on given sub-questions and analyzed the results based on the SOLO taxonomy. Both studies found that most students were at the uni-structural level and were aware of the need for collecting data but lacked an understanding of representativeness.

Hedges and Given (2023) took this a step further by exploring middle school students' progress in constructing statistical questions and designing studies based on these questions as part of a classroom project. Twenty students were involved in the project and were tasked with generating research questions related to social issues and planning their studies using publicly available electronic databases. The results show that students adjusted their initial questions and variables based on online data. However, the student's final questions and study designs reflected confirmation bias. During the study design process, students did not ensure the representativeness of their samples. Instead, they selected samples aligned with their preconceived ideas, leading to results that supported their prior beliefs.

Fielding-Wells (2010) conducted a study to examine the planning of statistical investigations by 24 fifth-grade students through two tasks across six cycles (three cycles for each task). The student's performance on the planning tasks was assessed based on data collection, analysis, and conclusions. In data collection, the evaluation focused on identifying necessary data, determining the target population, proposing a

method, and considering the need to record the data. The data analysis aspect emphasized the organization and handling of the data to enhance its interpretability and usefulness. Finally, the conclusions involved identifying the intended audience and choosing an appropriate format for communicating the results. The considerations were scored on a scale of 1 to 4, with 1 representing inappropriate answers and 4 indicating substantial and relevant details provided. The results show that the students improved their planning through six cycles with explicit structure provided by the teacher, group collaboration, and whole-class discussions. In the first three cycles, students struggled with planning in all three aspects. However, they realized the importance of considering all three areas for effective planning. By cycles 4 to 5, the focus shifted to the conclusions before data collection and data analysis, which improved the students' planning efficiency of statistical investigation.

In conclusion, the studies discussed in this section indicate a need for improvements in students' understanding and levels of thinking in the data collection phase of statistical investigation. Groth (2003) and Hedges et al. (2022) found that most students were at the uni-structural level and were aware of the need for collecting data but lacked an understanding of representativeness. Hedges and Given (2023) revealed that students did not ensure the representativeness of their samples and reflected confirmation bias in their study designs. However, Fielding-Wells (2010) showed that students improved their planning efficiency of statistical investigation through explicit structure, group collaboration, and whole-class discussions.

2.1.5.3. Studies on Data Analysis

Many studies have been conducted on analyzing the data through the work on the concepts within exploratory and inferential data analyses. Data analysis is a highly studied step in teaching and learning statistics. In this section, the studies carried out on students' understanding of exploratory data analyses like average and variation will be discussed due to the scope of the current study (e.g., Çatman Aksoy & Işıksal Bostan, 2020; Makar, 2014; Makar & Confrey, 2005; Mokros & Russell, 1995; Watson & Moritz, 2000; Watson et al., 2022). Table 2.11 provides an overview of the selected studies on data analysis.

Table 2. 11*The Overview of The Studies on Data Analysis*

Study Title	Participants	Research Focus	Key Findings
Mokros and Russell (1995)	21 children (4th to 8th graders)	Understanding of average in terms of representativeness	Younger students struggled to comprehend representativeness in understanding the average concept. Most older students were able to conceptualize average using representativeness.
Watson and Moritz (2000)	94 students (grades 3 to 9)	Understanding of average over time	Six levels of understanding of average were identified, which increased with the higher grades. Some students did not improve their understanding of average over time.
Makar (2014)	26 year-3 children (age 8)	Conceptualization of average in an inquiry-based learning environment	Due to informal statistical inference, students conceptualized “typical” as the most common, middle, and medium, and representative.
Çatman Aksoy and Işıksal Bostan (2020)	164 seventh-grade students	Statistical literacy about average and variation concepts using bar and line graphs	The students mostly achieved Level 3 Inconsistent and Level 4 Consistent-noncritical in understanding average concepts while struggling with interpreting variation concepts.
Makar and Confrey (2005)	17 pre-service mathematics and science teachers	Understanding of variation using formal and informal language	Participants improved their use of standard statistical language, including the mention of standard deviation, and used nonstandard statistical language, such as spread and distribution chunks.

Mokros and Russell (1995) studied children's understanding of average in terms of representativeness. They interviewed 21 children using open-ended problems to construct and interpret. The results revealed two distinct groups: the first group did not recognize representativeness and believed average to be the mode or an algorithmic procedure, while the second group acknowledged representativeness and viewed average as the midpoint, mathematical point of balance, or what is reasonable. Participants in the study were fourth to eighth graders, and it was found that younger students struggled to comprehend representativeness in understanding the average concept. Still, most older students were able to conceptualize average using representativeness.

Watson and Moritz (2000) explored 94 students' understanding of average from grade 3 to 9 and how this understanding evolved over time. Follow-up studies were conducted with 22 students after three years and 21 students after four years. The results show six levels of understanding of average, ranging from preaverage to application of average in one or two complex tasks. The levels of understanding average increased with the higher grades. The longitudinal studies showed that 16 of the 22 students observed after three years had a higher level of understanding average. Similarly, 15 of the 21 students observed after four years were at a higher level of understanding. However, some students did not improve their understanding of average over time. Nonetheless, their understanding did not decrease either. These results are consistent with Mokros and Russell's (1995) study, which showed that students' understanding of average was better in higher grades. In lower grades, students struggled to conceptualize average and often used the concepts of “most” and “middle.”

Makar (2014) investigated the conceptualization of average among twenty-six year-3 children (age 8) in an inquiry-based learning environment to develop informal inferential statistics. This study was part of a long-term design-based research project, where the children were asked to investigate whether there is a typical height for year-3 students and what it is. The investigation took place over four weeks, beginning with a discussion of the research question and planning of data collection in the first week. Data were collected within the class in the second week, and typical height was estimated. During this week, students conceptualized “typical” as the most common. Data was collected from another class in the third week, and comparisons were made between the classes. As a result of this comparison, students conceptualized “typical” as middle and medium because of the variability across classes. In the final week, students concluded their studies. They conceptualized “typical” as representative due to informal statistical inference, allowing them to question the typical height of year-3 students worldwide.

Çatman Aksoy and Işıksal Bostan (2020) examined the statistical literacy of 164 seventh-grade students about average and variation concepts using bar and line graphs. Six levels of statistical literacy, adapted from Watson and Callingham (2003), were

employed in the study, ranging from Level 1 Idiosyncratic to Level 6 Critical Mathematical. The findings indicated that the students mostly achieved Level 3 Inconsistent and Level 4 Consistent-noncritical in their understanding of average concepts, being able to interpret them in the given context. Moreover, the students predominantly utilized the arithmetic mean when dealing with questions related to average concepts and could use both graphs to interpret the average concept. However, regarding variation concepts, the students generally reached only Level 2 Informal and struggled to interpret them. The students used bar graphs more than line graphs while interpreting variation.

Makar and Confrey (2005) examined the understanding of variation among 17 pre-service mathematics and science teachers. They investigated the use of formal and informal language by these teachers while comparing distributions in a statistical task during pre- and post-interviews. The study took a course focused on teaching assessment, data analysis, equity, and inquiry. The results show that most participants improved their use of standard statistical language across all categories, such as proportion, mean, maximum/minimum, sample size, outliers, range, shape, and standard deviation. In particular, the standard deviation was not mentioned during the pre-interviews, but two participants mentioned it during the post-interviews. The participants used two types of nonstandard statistical language: spread and distribution chunks, and the use of both categories increased similarly to the formal language. When discussing variation in nonstandard statistical language, they used words such as spread out, skewed left, even distribution, clump, scattered, dispersed, grouped, gathered, and clustered. The researchers found that the increase in standard language use was statistically significant compared to nonstandard language.

To conclude, studies have been conducted to investigate students' understanding of exploratory data analyses, particularly average and variation. Various studies showed that the concept of average could be challenging for younger students to grasp, but understanding improves with age and experience (Makar, 2014; Mokros & Russell, 1995; Watson & Moritz, 2000). Also, studies showed that children had difficulties conceptualizing variation (Çatman Aksoy & Işıksal Bostan, 2020; Makar & Confrey,

2005). These studies highlight the importance of teaching and learning data analysis in a way that promotes an understanding of key concepts like average and variation.

2.1.5.4. Studies on Interpreting the Results

The studies discussed in this section focus on students' understanding and levels of thinking during the interpretation phase of statistical investigation (Akkoç & Selçuk, 2017; Chick & Watson, 2002; Fitzallen & Watson, 2011; Santos & Ponte, 2013). Table 2.12 shows the overview of the studies selected from the literature on interpreting the results step of the statistical problem-solving process.

Table 2.12

The Overview of The Studies on Interpreting The Results

Study	Participants	Research Focus	Key Findings
Chick and Watson (2002)	27 fifth and sixth-grade students	Levels of interpretation and representation skills during a statistical investigation	Most students reached a relational level in interpreting cause-and-effect relationships among variables, while most were at the multi-structural level in representing variables through various graphs.
Fitzallen and Watson (2011)	12 fifth and sixth-grade students	Skills in creating and interpreting multiple attribute graphs	Students tended to create graphs with bins in computer-based interviews and focused on individual cases in their interpretation. The importance of context in graph interpretation was emphasized, with students introduced to it in three different forms.
Santos and Ponte (2013)	21 pre-service teachers	Interpretation of average concepts during a statistical investigation	Most teachers had difficulty choosing appropriate measures for specific contextualized variables and interpreting statistical concepts such as mean, median, and mode. Only a limited number of appropriate interpretations were made.
Akkoç and Selçuk, (2017)	34 9th-grade students	Informal statistical inferences regarding average and variation	Most students utilized evidence from data and employed probabilistic language while making informal inferences, with insufficient explanations mainly coming from the generalizing beyond the data.

First, Chick and Watson (2002) conducted a study to assess the hypotheses and representation skills of 27 fifth and sixth-grade students regarding given data. The study utilized the SOLO taxonomy to classify students' levels of interpretation and representation skills into three categories: uni-structural (U), multi-structural (M), and

relational (R). The results indicate that most students reached a relational level, where they could identify cause-and-effect relationships among variables. Only one student was at the uni-structural level, focusing solely on individual data points. In terms of representation, most students were at the multi-structural level. They could represent single variables through various graphs, including tables, pie charts, line graphs, and bar graphs. In addition, most students were at the relational level and could show relationships between two variables through bar graphs, pie charts, and scatter graphs. However, only three students were at or below the uni-structural level, where they could only show some or all data points in a table. The results revealed that the levels of interpretation and representation skills were interrelated.

Fitzallen and Watson (2011) conducted a study to assess the skills of 12 fifth and sixth-grade students in creating and interpreting multiple attribute graphs in a classroom setting and a one-to-one computer-based interview setting. The students used Tinkerplots to create six different graphs based on two attributes heart rate and gender in classroom settings and six attributes like weight, eye color, age, and favorite activity, considering whether to use bins or scaled axes in the creation process. The importance of context in graph interpretation was emphasized, with students introduced to it in three different forms in both settings: individual icons or data cards, trend-based descriptions, and story-based interpretations. The results show that students created 48 graphs in the classroom setting, with 15 focusing on single attributes and the others on relationships between two attributes. In terms of interpretation, three students based their interpretations on individual values, ten made general comments on trends, and the remaining used their experiences and possible causes to interpret the graphs. In the computer-based interview, students created 42 graphs, most of which used bins, and tended to interpret graphs based on individual cases more than in the classroom setting.

Santos and Ponte (2013) investigated pre-service teachers' interpretation of average concepts while conducting statistical investigations on recycling, food, water habits, and routines. Twenty-one pre-service teachers formed twelve groups to perform exploratory data analyses using Excel, and they calculated the arithmetic mean, median, and mode. While interpreting the mode, some pre-service teachers made

appropriate interpretations such as the value associated with the highest frequency, most of the sample, the most common value, or the value that appears more times. However, inappropriate interpretations, such as the highest value or a significant portion of the sample, were also made, and most teachers did not interpret the mode. Similar results were found for the mean and the median, with only a limited number of appropriate interpretations. For instance, for the mean, some interpretations were the balance point and the fair share, while for the median, the value divides the ordered data. However, most teachers did not interpret the mean and median or were confused about them. Some interpreted the median inappropriately, stating it as the value in the center or a maximum of 50% of respondents. Overall, the results indicate that the teachers had difficulty choosing appropriate measures for specific contextualized variables and interpreting statistical concepts.

Akkoç and Selçuk (2017) analyzed the informal statistical inferences of 34 9th-grade students regarding average and variation. The informal inference was considered a part of the interpretation phase. They utilized four open-ended questions based on Makar and Rubin (2009)'s informal statistical inference framework, which comprised three components: generalizing beyond data, evidence from data, and probabilistic language. The findings revealed that most students utilized evidence from data and employed probabilistic language while making informal inferences. Additionally, when generalizing beyond data, students primarily provided sufficient explanations, while insufficient explanations mainly came from this component.

Overall, the studies discussed in this section provide valuable insights into students' understanding and thinking levels during the statistical investigation interpretation phase. Chick and Watson's (2002) study showed that most students reached a relational level of interpretation and representation skills, while Fitzallen and Watson's (2011) study demonstrated the importance of context in graph interpretation. Santos and Ponte's (2013) study revealed that pre-service teachers had difficulty choosing appropriate measures for specific contextualized variables and interpreting statistical concepts. Finally, Akkoç and Selçuk's (2017) study highlighted those students utilized evidence from data and probabilistic language while making informal inferences but struggled with generalizing beyond data. Overall, these studies contribute to our

understanding of how students think and reason statistically in the interpretation phase of the statistical problem-solving process.

2.1.5.5. Studies Focus on The Entire Statistical Problem-Solving Process

This section focuses on studies that examine students' practices and understanding throughout the entire statistical investigation process (e.g., Lavigne & Lajoie, 2007; Öz & Işıksal Bostan, 2021; Silva et al., 2022; Watson & English, 2015; Watson & English, 2017b; Watson & English, 2017c). Table 2.13 provides an overview of studies to improve students' practices and understanding related to this issue by covering the entire statistical problem-solving process, including formulating investigative questions, collecting, analyzing, and interpreting data.

Table 2.13

The Overview of The Studies on The Entire Statistical Problem-Solving Process

Study	Participants	Research Focus	Key Findings
Watson and English (2015)	91 fifth-grade students	Evaluating statistics practices in the context of environmentally friendly habits	Most students recognized their class as a sample, set criteria for determining environmental friendliness, and made decisions about their class with a degree of certainty. However, they were not as confident in their conclusions regarding all fifth-grade students in Australia.
Watson and English (2017b)	96 fifth-grade students	Investigating different data collection methods	Students were engaged in the practice of statistics. They were careful in collecting and analyzing data, but there is room for improvement in their understanding of expressing reasons for limited confidence in their conclusions.
Watson and English (2017c)	85 Grade 6 students	Understanding average concepts and inferences	Students could conduct statistical investigations comparing two groups, had a good procedural understanding of average concepts, and could make inferences while considering within and between-group variations without being formally introduced to these concepts.
Lavigne and Lajoie (2007)	12 seventh-grade students	Exploring statistical reasoning during a statistical investigation	Students engaged in statistical reasoning and employed various reasoning types, including reasoning about samples, data, statistical measures, data representations, and alignment with the investigation question during different stages of the investigation.

Table 2.13(cont'd)

Silva et al. (2022)	20 first- grade students	Examining statistical investigations	Students developed their skills in statistical investigations, but their arguments were relatively weak due to limited knowledge of statistics and abilities in speaking and writing.
Öz and Işıksal Bostan (2021)	121 seventh- grade students	Assessing understanding of statistical problem- solving	Most students posed comparative questions but did not plan a data collection procedure effectively. Most students could create comparative graphs for data analysis but did not select the appropriate graph in all cases. Most students could interpret the graphs to some extent in the interpretation phase.

First, Watson and English (2017b) conducted a study on the practices of statistics among 96 fifth-grade students, focusing on different data collection methods and two datasets to answer the same research question: “What is the typical reaction time of Grade 5 students?” The study followed the statistical problem-solving process outlined in their previous study (Watson & English, 2015). In this study, the students collected data through two methods: the Ruler Drop activity, which measures reaction time through length, and the Reaction Timer Test from the ABS CensusAtSchool website, which measures reaction time by clicking a mouse. The researchers asked questions to understand the students' conclusions based on the data and their confidence in their judgments. The results indicate that the students were engaged in the practice of statistics and were careful in collecting and analyzing data. However, there is still room for improvement in their understanding of expressing reasons for limited confidence in their conclusions for a wider population of fifth-grade students.

Watson and English (2017c) conducted a study to examine the experiences of 85 Grade 6 students concerning statistical investigations that involved data from two groups of individuals. The investigation focused on students' reaction times based on their eye color. The study assessed the students' understanding of average concepts and their inferences based on various sample sizes. The results show that the students could conduct statistical investigations comparing two groups and had a good procedural understanding of average concepts. Furthermore, they could make inferences while considering within and between-group variations without being formally introduced to these concepts.

Lavigne and Lajoie (2007) conducted a case study that explored the statistical reasoning of seventh-grade students during a statistical investigation. The study involved two groups of six students who posed a statistical investigation question, collected data based on the question and analyzed and interpreted the collected data. The researchers employed a combination of survey and experimental methods to model the inquiry process. Prior to conducting their investigation, the teacher introduced the students to statistical concepts, providing them with a foundation for statistical thinking and an understanding of the investigative process. The study found that the students engaged in statistical reasoning and employed a variety of reasoning types, including reasoning about samples, data, statistical measures, data representations, and alignment with the investigation question, to varying degrees, during different stages of the investigation.

Silva et al. (2022) examined 20 first-grade students learning statistics through statistical investigations. The researchers conducted pre and post-tests to understand students' learning statistics with the statistical investigations. First, the students selected topics from the subject called Environmental Studies to formulate a statistical question. Despite challenges in the question-posing and planning phase, students gathered, arranged, and examined data with teacher assistance. In the conclusion phase, students can make sense of the data. However, their arguments are not relatively weak due to their limited knowledge of statistics and limited abilities in speaking and writing. However, pre-and-posttests results also showed that students developed their skills in statistical investigations.

Öz and Işıksal Bostan (2021) surveyed 121 seventh-grade students' understanding of statistical problem-solving. The survey comprised two questions with four sub-questions, each aligned with the four steps of the statistical problem-solving process. The results indicate that most students posed comparative questions in both questions, but most did not require data collection. Additionally, most students were unable to plan a data collection procedure effectively. Regarding data analysis, most students could create comparative graphs in the first question, but the appropriate graph was not selected in all cases. For the second question, most students were unable to create

graphs. In the interpretation phase, most students could interpret the graphs to some extent, showing a higher level of understanding by reading beyond the data.

The studies discussed until now highlight the research on the entire statistical problem-solving process, from formulating investigative questions to interpreting data and providing insights into students' statistical practices and their understanding of statistical concepts, such as sampling, data collection methods, and data analysis within the statistical problem-solving process. Also, certain studies have been conducted to understand pre-service and in-service teachers' understanding of and teaching practices on the whole statistical problem-solving process (e.g., Litke & Hill, 2020; Yılmaz & Yetkin-Özdemir, 2021). Table 2.14 includes an overview of studies related to teaching and teacher-related issues.

Table 2.14

The Overview of The Studies on The Entire Statistical Problem-Solving Process (Teacher-Focused)

Study	Participants	Research Focus	Key Findings
Yılmaz and Yetkin-Özdemir (2021)	Pre-service teachers	Understanding of statistics and its implications in teaching	A limited initial understanding developed into a more connected cycle of statistical investigations through lesson study
Litke & Hill (2020)	Fourth and fifth-grade teachers	Practices in statistics education	Lessons focused on calculating central tendency measures and range and creating graphs; emphasis on investigative questioning and interpretation was missing

Yılmaz and Yetkin-Özdemir (2021) explored also pre-service teachers' understanding of doing statistics and the implications of such understanding in teaching through a lesson study. They conducted three lesson studies in both college and real classrooms. The results show that pre-service teachers had an underdeveloped understanding of doing statistics at the beginning of lesson studies. However, through engaging in the lesson study and the implications of teaching statistics in both college and real classrooms, their understanding developed into a more connected cycle of statistical investigations; they could understand the importance of posing questions, followed by data collection, and the use of graphs for answering the question posed with a degree of struggle due to the challenges of transforming knowledge into practice.

Litke and Hill (2020) explored teachers' practices in statistics education by observing fourth and fifth-grade classroom instruction in terms of statistical investigation steps presented in GAISE. One hundred forty-four lessons were examined as part of a bigger study examining the quality of mathematics instruction. They created a coding protocol called "Transcript Coding Protocol Using GAISE Framework," considering four steps of the statistical problem-solving process. They found that the lessons consisted of three broader categories: calculating central tendency measures and range, graphs, and both categories. The mathematical focus is prominent. While teaching the statistical concepts, they did not use the cycle of statistical investigations. Later, they discussed findings based on four steps of the statistical problem-solving process. The teachers did not focus on formulating investigative questions mainly. Also, interpretation is missing part of the lessons covered. However, they concluded that the focus is not on the whole cycle of statistical investigations.

To sum up, this section examines studies that explore students' and teachers' practices related to statistical problem-solving processes. The studies aim to understand how students engage in investigations and use various reasoning types at different stages of the process. While the studies show that students can complete investigations from start to finish, some suggest that they still struggle with effectively planning data collection procedures and expressing reasons for limited confidence in their conclusions (Watson & English, 2015; Watson & English, 2017b).

The studies related to teaching practices reveal that although teachers may not focus on the entire problem-solving process in the classroom, they can improve their understanding and practices by engaging with them (Litke & Hill, 2020). Overall, when students are given statistical tasks that require the entire problem-solving process, their statistical thinking, reasoning, and literacy improve (Watson & English, 2017c). Therefore, teachers should provide opportunities for students to enhance their statistical skills.

2.2. Mathematics and Statistics Classroom Practices

Teachers play a crucial role in creating a positive learning environment that fosters student learning through effective teaching practices, as highlighted by various studies

(Copur-Gencturk, 2012; Conway IV et al., 2019; Charalambous & Kyriakides, 2017; Hill et al., 2008). They make important decisions regarding curriculum and pedagogy based on the unique needs of their students (Eichler, 2011; Lombaert et al., 2007; Stein et al., 2007). Numerous studies have focused on identifying the teaching practices for effective mathematics teaching (e.g., Charalambous et al., 2012; Schoenfeld, 2013; Stein & Kaufman, 2010). Understanding these practices is crucial for improving student learning outcomes (Boaler & Staples, 2008; Charalambous & Litke, 2018; Hill et al., 2008; Schoenfeld, 2013; Stein & Lane, 1996).

Eichler (2011) noted that while there is extensive research on mathematical teaching practices, there is a lack of research specifically on statistics teaching practices. Although curricular documents discuss mathematical practices (CCSSI, 2010; NCTM, 2014), the nature of statistics and mathematics is different, highlighting the need to understand statistical teaching practices better in the mathematics classroom (Eichler, 2011; Gerstenschlager, 2019; Shaughnessy, 2007). The SET document and Watson and English (2017c) discuss the eight standards for Mathematics Practice (CCSSI, 2010) and their corresponding statistical practices, as mentioned in the Curriculum and Curriculum Expectations subsection of the Learning and Teaching Statistics section. In addition, NCTM (2014) provides research-based Mathematics Teaching Practices that promote high-quality instruction, and Gerstenschlager (2019) provides corresponding practices in statistics teaching with subject-specific explanations in Table 2.15.

Gerstenschlager (2019) explains several teaching practices in statistics education, including setting goals for tasks, implementing tasks promoting reasoning and problem-solving, representations, and discourse issues (Practice 1 to 4 in Table 2.15 from <https://www.statisticteacher.org/2019/01/02/statistics-teaching-practices>). She notes that these practices differ from their counterparts in mathematical education due to the unique nature of statistics. For instance, while mathematical tasks often rely on algorithms and procedures, statistical tasks emphasize processes and statistical reasoning.

Furthermore, statistics requires the use of appropriate graphical representations based on the data, whereas mathematical representations are more prevalent in traditional mathematics education. Finally, discourse issues are essential in mathematics and statistics education, but skepticism is particularly important in statistics discussions. Nonetheless, some effective discourse patterns in mathematics classrooms can also be valid in statistics classrooms by integrating a skeptical approach.

Table 2.15

Mathematics Teaching Practices and Corresponding Statistics Teaching Practices

Mathematics Teaching Practices	Statistics Teaching Practices
1 <i>Establish mathematical goals to focus learning</i>	<i>Establish statistical goals to focus learning</i>
Effective teaching of mathematics establishes clear goals for the mathematics that students are learning, situates goals within learning progressions, and uses the goals to guide instructional decisions.	Effective teaching of statistics establishes clear goals for the mathematics that students are learning, situates goals within learning progressions, and uses the goals to guide instructional decisions.
2 <i>Implement tasks that promote reasoning and problem solving.</i>	<i>Implement statistical tasks that promote statistical reasoning and the statistical problem-solving process.</i>
Effective teaching of mathematics engages students in solving and discussing tasks that promote mathematical reasoning and problem solving and allow multiple entry points and varied solution strategies.	Effective teaching of statistics engages students in solving and discussing tasks that promote statistical and probabilistic reasoning and problem solving and allow multiple entry points and varied solution strategies.
3 <i>Use and connect mathematical representations.</i>	<i>Use and connect mathematical representations.</i>
Effective teaching of mathematics engages students in making connections among mathematical representations to deepen understanding of mathematics concepts and procedures and as tools for problem solving.	Effective teaching of statistics engages students in making connections among statistical representations to deepen understanding of variability inherent in the data and as tools for making decisions during the statistical problem-solving process.
4 <i>Facilitate meaningful mathematical discourse.</i>	<i>Facilitate meaningful statistical discourse.</i>
Effective teaching of mathematics facilitates discourse among students to build shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments.	Effective teaching of statistics facilitates discourse among students to build shared understanding of statistical ideas by analyzing and comparing student approaches and arguments.

Note. From *Re-Envisioning the Mathematics Teaching Practices as the Statistics Teaching Practices*, by N. E. Gerstenschlager, 2019 (<https://www.statisticsteacher.org/2019/01/02/statistics-teaching-practices/>). Copyright 2023 by American Statistical Association

Table 2.15 (cont'd)

Mathematics Teaching Practices	Statistics Teaching Practices
5 <i>Pose purposeful questions.</i>	<i>Pose purposeful statistical questions and develop students' ability to create statistical questions.</i>
Effective teaching of mathematics uses purposeful questions to assess and advance students' reasoning and sense making about important mathematical ideas and relationships.	Effective teaching of statistics uses purposeful questions to assess and advance students' statistical reasoning and sense-making about important statistical ideas and relationships. Additionally, effective teaching of statistics fosters students' ability to create and ask their own appropriate statistical questions to guide their statistical investigations
6 <i>Build procedural fluency from conceptual understanding</i>	<i>Build fluency with technology to conceptually understand the purpose of statistical and probabilistic procedures and use this technology to help make sound statistical decisions.</i>
Effective teaching of mathematics builds fluency with procedures on a foundation of conceptual understanding so that students, over time, become skillful in using procedures flexibly as they solve contextual and mathematical problems.	Effective teaching of statistics builds fluency with technology to perform statistical procedures and simulate probabilistic situations so students, over time, become skillful in using technology to make sound statistical decisions and develop conceptual understanding of the procedures.
7 <i>Support productive struggle in learning mathematics.</i>	<i>Support productive struggle in learning statistics</i>
Effective teaching of mathematics consistently provides students, individually and collectively, with opportunities and supports to engage in productive struggle as they grapple with mathematical ideas and relationships.	Effective teaching of statistics consistently provides students, individually and collectively, with opportunities and supports to engage in productive struggle as they grapple with statistical ideas and relationships.
8 <i>Elicit and use evidence of student thinking.</i>	<i>Elicit and use evidence of students' statistical thinking.</i>
Effective teaching of mathematics uses evidence of student thinking to assess progress toward mathematical understanding and to adjust instruction continually in ways that support and extend learning.	Effective teaching of statistics uses evidence of students' statistical thinking to assess progress toward statistical understanding and to adjust instruction continually in ways that support and extend learning.

Note. From *Re-Envisioning the Mathematics Teaching Practices as the Statistics Teaching Practices*, by N. E. Gerstenschlager, 2019 (<https://www.statisticsteacher.org/2019/01/02/statistics-teaching-practices/>). Copyright 2023 by American Statistical Association

Gerstenschlager (2019) discussed additional practices related to posing questions, building fluency, productive struggle, and using evidence in statistics teaching (Practice 5 to 8 in Table 2.15). Firstly, posing questions is a crucial aspect of the statistical problem-solving process, which differs from contributing to students'

mathematical reasoning and sense-making. Secondly, while mathematics and statistics rely on procedures that require fluency and conceptual understanding, being fluent in statistics means making sense of data through the procedures. Mathematical aspects can be handled through technology, and the focus is on making sound decisions, not just on procedures. Thirdly, the productive struggle should be supported in mathematics and statistics classrooms. Still, it is essential to understand that the struggle in statistics comes from the concept of variability at every step in statistical problem-solving. Finally, statistical thinking is distinct from mathematical thinking. While students may use mathematical thinking to calculate specific statistics, they must also think about what the results mean to make sound inferences. Gerstenschlager (2019) emphasized the importance of understanding the similarities and differences between teaching practices in mathematics and statistics to provide high-quality instruction, even though statistics is often covered in mathematics classrooms.

Eichler (2011) also compared the classroom practices of mathematics teachers and statistics teachers using the “four phases of the curriculum” framework by Stein et al. (2007). These four phases include the written curriculum, teachers' intended curriculum, teachers' enacted curriculum, and students' learning. The intended curriculum is the teachers' interpretation of the written curriculum, and the enacted curriculum refers to the interactions among teachers, students, and content. The enacted curriculum, influenced by teachers' beliefs and knowledge regarding student learning, transforms the intended curriculum. Eichler (2011) further emphasizes starting the study of classroom practice with the intended curriculum, influenced by teachers' beliefs and knowledge systems. However, teaching and learning statistics have different aspects than those in mathematics. Although the intended curriculum in statistics affects classroom practices, research in this area is limited. Therefore, to create classroom practices that align with the written curriculum, it is necessary to understand teachers' thinking and goals regarding teaching and learning statistics and to shape their beliefs, building on the research done in classroom practices in mathematics.

In conclusion, there is a growing need to understand and improve classroom practices in statistics education, as it differs from traditional mathematics education (Eichler,

2011; Gerstenschlager, 2019; Watson & English, 2017c). While both disciplines share some similarities, statistics education requires a different focus due to the nature of statistics as a field. In the following sections, previous research on teaching practices in mathematics education will be explored to shed light on teaching practices in statistics. Specifically, the current study will focus on examining the teaching practices of mathematics teachers during the statistics unit. This examination will consider the unique nature of statistics as well as the teaching practices in mathematics relevant to statistics.

2.2.1. Classroom Discourse

Vygotsky's theory asserts that active learning occurs within a cultural context, and knowledge is built through individual or collective idea exchange, leading to rational judgment (Hiebert & Grouws, 2007). Consequently, social experiences contribute to learners' existing knowledge base. Mathematics learning, for instance, necessitates the reconstruction of mathematical concepts through spoken and written language. This is why the National Council of Teachers of Mathematics (NCTM) standards emphasize communication as a crucial process for social learning. NCTM underscores the importance of communicating mathematical ideas for improved learning outcomes (Van de Walle et al., 2010). Besides explaining one's understanding through written or spoken mathematics, learners should analyze and evaluate others' ideas and clarify why an argument is either valid or invalid (NCTM, 2000). This learning environment is fostered through classroom discourse, where students engage in collective idea exchange to reconstruct existing knowledge under the guidance of teachers.

NCTM's communication standard places learners at the center of the learning process. In traditional classrooms, teachers control classroom communication through teacher talk (Franke et al., 2007). They explain mathematical concepts, strategies, or procedures while students interact minimally with teachers and peers. This style of classroom discourse is simple to manage for teachers because they control the communication flow through close-ended questions (Ponte & Quaresma, 2016). On the other hand, a student-centered learning environment is necessary to provide rich experiences for constructing knowledge through high-quality mathematics lessons

(Sherin, 2002). Classroom discourse plays a critical role in creating an effective learning environment for students (Otten & Soria, 2014). Teachers and students collaborate in classroom discourse to shape mathematical learning through mathematical communication practices (Guerreiro et al., 2013).

Classroom discourse patterns vary regarding teachers' and students' roles in literature. Traditional teacher-centered classroom discourse practices limit students' opportunities to explain their thinking or reach a consensus about mathematical ideas (Franke et al., 2007). Effective classroom discourse facilitating a student-centered learning environment is needed where students are the active constructors of knowledge (Drageset, 2014; Ponte & Quaresma, 2016; Stein et al., 2008).

One example of a traditional classroom interaction pattern is Initiation-Response-Evaluation (IRE) (Franke et al., 2007). In IRE, teachers ask close-ended questions with the expectation of one correct answer and evaluate the correctness of responses. The classroom discourse is dominated by teacher talk, and teachers do not guide students to construct multiple strategies for completing the task (Cazden, 1988). Drageset (2014) identified three types of classroom discourse actions: Redirecting, progressing, and focusing. Redirecting actions aim to change students' approaches and include practices such as advising new strategies or correcting questions. Progressing actions aim to move the process forward, including demonstration, simplification, closed progress details, and open progress initiatives. However, these actions are similar to IRE, as teachers hold power over classroom discussions. In contrast, focusing actions require active participation from students through justifying answers or requesting assessments from other students. Teachers expect students to be the main participants of the discussion in focusing actions.

Bredefur and Frykholm (2000) classified discourse patterns into unidirectional, contributive, reflective, and instructive communication. Unidirectional communication, like IRE, is the most common pattern in which teachers dominate classroom discourse through close-ended questions. In contributive communication, teachers provide opportunities for students to share and discuss their ideas, but in a corrective manner. Reflective and instructive communication both place students at

the center of classroom discussion. In reflective communication, students justify or refute conjectures, explore mathematics, and participate in classroom discussions. Instructive communication aims to modify instructional sequences to improve students' mathematical understanding.

Ponte and Quaresma (2016) proposed a framework for analyzing a teacher's actions in managing classroom discourse in mathematics. This framework consists of four actions: inviting, supporting/guiding, informing/suggesting, and challenging. The framework emphasizes the importance of analyzing teachers' actions in terms of mathematical processes, such as representing, interpreting, reasoning, and evaluating. NCTM also elaborated on the teacher's role in classroom discourse, which includes posing questions and tasks, listening to students' ideas, asking students to clarify and justify their ideas, deciding what to pursue in-depth, attaching mathematical notation and language to students' ideas, deciding when to provide information, clarifying issues, modeling, leading, and monitoring student participation (as cited in Silver & Smith, 1996).

Shaughnessy (2007) stated that the classroom discourse in mathematics education is considered in detail. Still, more research is necessary about classroom discourse in statistics education since the nature of statistics and mathematics differs. It is important to learn how students pose questions and make arguments as a community through high-quality statistical tasks encouraging students to use multiple representations, think, and critically analyze. Therefore, it is necessary to differentiate mathematical and statistical practices in the classroom to explore the discourse patterns in teaching and learning statistics.

2.2.2. Teaching Triad as the Conceptual Background

Many researchers mention the complexity of teaching (Charalambous & Kyriakides, 2017; Charalambous & Praetorius, 2018; Cohen et al., 2003; Cohen, 2011). Therefore, describing what is meant by teaching or instruction is crucial to understand good teaching practices (Charalambous & Praetorius, 2018; Hiebert & Grouws, 2007).

Teaching is a collection of practices, including pedagogy, learning, instructional design, and managing organization. There are more practitioners than teachers, more practices than pedagogy, and the environments of teaching and learning are implicated in the interactions (Cohen et al., 2003, p.124).

As understood from the teaching definition, the occupation of teaching differs from teaching practices (Cohen, 2011). Teaching practice might be considered more comprehensive and broader than the occupation of teaching and aims to set the relationship with learning to improve learning (Cohen et al., 2003; Cohen, 2011). Similarly, Thorndike (1906) defined teaching with student outcomes in which students achieve learning outcomes through teaching (as cited in Hiebert & Grouws, 2007).

Moreover, in a school context, Cohen et al. (2003) described teaching as not teachers do alone in the classroom but it is an interaction between teachers and students around the content. The interactions between teacher and student and student and student around content create complexity in the teaching or instruction (Charalambous & Praetorius, 2018). Hiebert and Grouws (2007) conceptualized teaching through “classroom interactions among teachers and students around content directed toward facilitating students’ achievement of learning goals” (p.372) by combining teaching definitions of Cohen et al. (2003) and Thorndike (1906). Figure 2.5 illustrates Cohen et al. (2003)’s instruction model called the teaching triad or instructional triangle, which is one of the conceptual frameworks of this study.

Teaching involves complex interactions between teachers and students around content (Charalambous & Praetorius, 2018; Cohen et al., 2003). While not tied to a specific discourse movement, these interactions take place in an environment. Cohen (2011) highlighted the importance of discourse in teaching and learning due to its ability to facilitate knowledge exchange and shared understanding.

Effective mathematics teaching requires various instructional phases that promote mathematical communication and discourse (Schoenfeld, 2014). Whole-class discussions are crucial for constructing meaning and developing students' understanding of mathematics (Hill et al., 2008; Van de Walle et al., 2010). To create a student-centered learning environment that fosters high-quality mathematics learning, teachers must anticipate, monitor, select, sequence, and connect student

responses. They must also attend to student thinking and demand mathematical explanations and justifications (Stein & Kaufman, 2010).

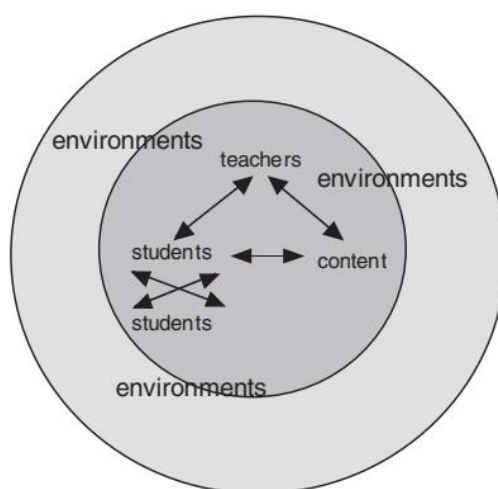


Figure 2.5 Instructional Triangle

Note: From “Resources, Instruction, and Research”, by D. K., Cohen, S. W. Raudenbush, and D. L. Ball, 2003, *Educational Evaluation and Policy Analysis*, 25(2), p.124
(<https://doi.org/10.3102/01623737025002119>)

Researchers have developed observation tools to identify key elements of effective mathematics instruction and conceptualize instruction quality. The following section will explore these components of good teaching in mathematics and how observation tools can facilitate research on mathematical and statistical teaching practices.

2.2.3. Instructional Quality in Mathematics Education

Effective mathematics teaching requires high-quality instruction, and teachers play a crucial role in creating a productive learning environment by managing various factors, including the learning environment, content, tools, and students (Franke et al., 2007). However, there is an ongoing debate in the literature regarding what constitutes good teaching practices and how to achieve them (Schoenfeld, 2014; Weiss & Pasley, 2004). The literature defines instructional quality as “teachers' classroom observable behavior” (Blömeke et al., 2022, p.2). While research has identified various components of instructional quality, the ultimate goal is to improve teaching and create opportunities for effective student learning (Hill & Grossman, 2013; LMT, 2011; Schlesinger &

Jentsch, 2016; Schoenfeld, 2014; Stein & Lane, 1996). Therefore, measuring and understanding teaching practices is necessary to achieve this goal (Bell et al., 2019).

Bostic et al. (2019) noted that mathematics education research involved classroom observations over the past 30 years to understand teaching practices. Over this period, research has been conducted to assess classroom implications in terms of attending to different levels of quality. First, classroom observations without attending to the levels of quality and implementation have been divided into three considerations: attending to specific measurable attributes such as the number of questions, the presence of a particular intervention like cooperative learning, and assessing teaching practices based on survey and self-report. In contrast, three types of research attend to the levels of quality and implementation through informal observations, formal observations with a holistic score, and formal observations with indicators.

The following section will discuss specific classroom schemes for analyzing teaching practices that attend to the levels of instructional quality through formal observations with a set of indicators defining the characteristics of high-quality mathematics instruction. It is important to note that literature uses various terms interchangeably, such as observational instruments, classroom schemes, rubrics, tools, schemes, and frameworks, to refer to different conceptualizations of instructional quality.

2.2.3.1. The Classification of Frameworks or Observational Instruments

Various frameworks are used in the literature to understand the complexity of instruction. These frameworks differ from each other in terms of the characteristics they include for good instruction (Charalambous & Kyriakides, 2017; Cohen et al., 2022; Praetorius & Charalambous, 2018; Senden et al., 2021). Praetorius and Charalambous (2018) stated that the frameworks differ based on the approaches taken during development, the purpose, conceptualizations, and measurement (see Figure 2.6). Similarly, Senden et al. (2021) support this idea, stating that the differences are due to conceptualizations, measurement, and operationalizations.

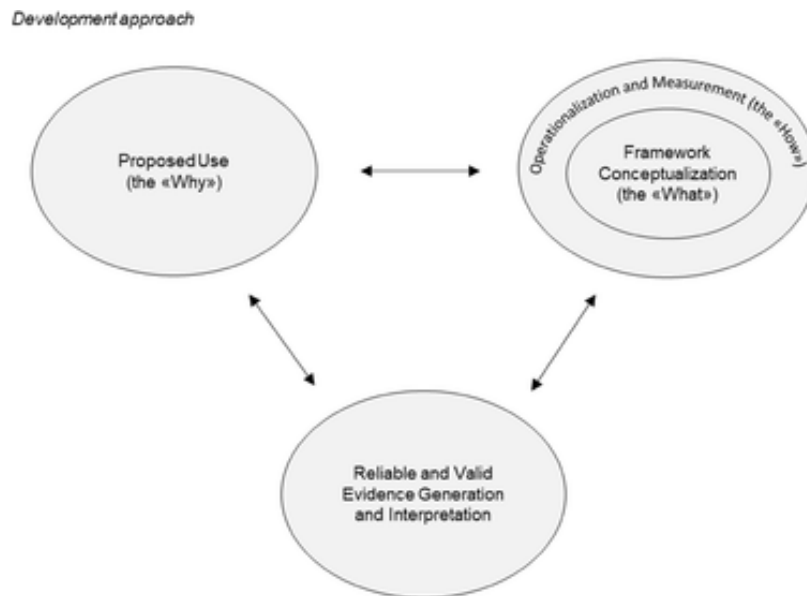


Figure 2.6 Prototypical Process in Developing and Using a Framework

Note: From “Classroom Observation Frameworks for Studying Instructional Quality: Looking Back and Looking Forward”, by A. K., Praetorius and C. Y. Charalambous, 2018, *ZDM*, 50, p.536 (<https://doi.org/10.1007/s11858-018-0946-0>)

It is essential to distinguish the focus of different frameworks or observation tools that address instructional quality in terms of “subject specificity” (Berlin & Cohen, 2018; Charalambous & Praetorius, 2018; Hill & Grossman, 2013). Charalambous and Praetorius (2018) represent the frameworks regarding their subject-specificity on a continuum (see Figure 2.7). They emphasize that one side of the continuum has generic frameworks, while content-specific frameworks are on the other end.



Figure 2.7 A Continuum of Classifying Frameworks and Observational Instruments

Note: From “Studying Mathematics Instruction Through Different Lenses: Setting the Ground for Understanding Instructional Quality More Comprehensively” by C. Y. Charalambous and A. K., Praetorius, 2018, *ZDM*, p. 356 , (<https://doi.org/10.1007/s11858-018-0914-8>)

All observation protocols emphasize different aspects of good teaching, with a particular focus on the theory of instruction (Cohen et al., 2022). Praetorius and Charalambous (2018) discussed twelve frameworks, including four generic, three subject-specific, and five hybrid frameworks, in terms of their development, purpose, conceptualization, and measurement. In this review, I consider ten of these frameworks and adapt the related tables from Appendix A and Appendix B1 of Praetorius and Charalambous' (2018) study based on this decision.

In addition to subject specificity, the purpose, theoretical background, and development process are also points of discussion in the literature when developing frameworks (Praetorius & Charalambous, 2018). Praetorius and Charalambous (2018) noted that researchers had used both top-down approaches, including literature-based, expert judgment, and existing frameworks and bottom-up approaches, including observing lessons and exploratory analyses, when developing frameworks related to instructional quality. Various theoretical foundations have been used, including educational effectiveness research, teaching theories/conceptualizations, learning theories/conceptualizations, motivational theories/conceptualizations, developmental theories/conceptualizations, social-interactive theories/conceptualizations, research on classroom discourse, research in mathematics education, instructional triangle, and conceptualizations on teacher knowledge and task potential in different frameworks. Details about the theoretical background and developmental process of specific example frameworks used in mathematics education are presented in Table 2.16, adapted from Praetorius and Charalambous (2018).

In addition, Praetorius and Charalambous (2018) provided characteristics of the example frameworks in terms of their operationalization and measurement. The frameworks use level 1 and level 2 classifications, which are named differently in each framework. The number of level 1 elements varies from three to nine, and level 2 elements range from eight to forty. Most frameworks focus on both the frequency and quality of teacher behavior while implicitly considering student behavior. Details on framework characteristics based on operationalization and measurement can be found in Table 2.17 (Adapted from Praetorius & Charalambous, 2018, Appendix B1). The following section will briefly explain generic, subject-specific, and hybrid frameworks.

2.2.3.1.1. Generic frameworks

Berlin and Cohen (2018) propose that instructional quality frameworks should consider context, content, and subject-specific practices. While generic frameworks measure teaching practices such as classroom management, climate, and instructional time, they do not account for subject-specific practices necessary for effective mathematics instruction (Hill & Grossman, 2013; Charalambous & Kyriakides, 2017; Schlesinger & Jentsch, 2016). Generic frameworks measure context and content without depending on the subject (Berlin & Cohen, 2018). Therefore, it is crucial to consider subject-specific aspects to improve teaching and enhance student learning (Charalambous & Praetorius, 2018).

Various generic frameworks exist in mathematics education literature, such as the Classroom Assessment Scoring System (CLASS) (Hamre et al., 2013), the Dynamic Model of Educational Effectiveness (DMEE) (Creemers & Kyriakides, 2008; Kyriakides et al., 2018), International System of Teacher Observation and Feedback (ISTOF) (Muijs et al., 2018), Three Basic Dimensions (TDB) (Klieme et al., 2009; Praetorius et al., 2018), and Framework for Teaching (FfT) (Danielson, 2013).

Charalambous and Kyriakides (2017) stated that these frameworks are also used in different subject areas to address general teaching practices but differ in scope, including instructional dimensions, as well as measuring characteristics, coding, and interpreting data collected from classrooms (see Table 2.16 and Table 2.17).

2.2.3.1.2. Subject Specific Frameworks

Subject-specific frameworks, also known as content-specific frameworks, focus on teaching practices related to the subject, such as mathematical errors and misconceptions, mathematical language, and mathematical content (Charalambous & Praetorius, 2018; Schlesinger & Jentsch, 2016). These frameworks provide valuable feedback on subject-specific aspects of instruction but may not cover general teaching practices.

Table 2.16

Process of Development and Theoretical Foundations of Example Frameworks

	Process of Development				Theoretical Foundations										
	Top-down		Bottom-up		Educational effectiveness research	Teaching theories/ conceptualizations	Learning theories/ conceptualizations	Motivational theories/ conceptualizations	Developmental theories/ conceptualizations	Social-interactive theories/ conceptualizations	Research on classroom discourse	Research in mathematics education	Instructional triangle	Conceptualizations on teacher knowledge	Conceptualizations of task potential
Generic frameworks															
CLASS	×	×	×	×	×	×	×	×	×	×	×				
DMEF	×		×	×	×	×	×								
ISTOF		×			×	×	×						×		
TBD	×			×	×	×	×	×		×					
Subject-specific frameworks															
IQA	×	×	×	×	×	×	×				×	×			×
MQI	×	×	×	×								×	×	×	×
M-Scan	×	×	×	×		×					×	×			×
Hybrid frameworks															
TEDS-Instruct	×	×	×	×	×		×	×				×		×	
TRU	×	×	×	×		×	×				×		×	×	
UTOP		×	×	×	×	×	×						×	×	×

a. CLASS: Classroom Assessment Scoring System; DMEF: Dynamic Model of Educational Effectiveness; ISTOF: International System of Teacher Observation and Feedback; TBD: Three Basic Dimensions; IQA: Instructional Quality Assessment; MQI: Mathematical Quality of Instruction; M-Scan: Mathematics Scan; TEDS-Instruct: Teacher Education Study in Mathematics-Instruct; TRU: Teaching for Robust Understanding; UTOP: UTeach Observation Protocol
 Note: Adapted from “Classroom Observation Frameworks for Studying Instructional Quality: Looking Back and Looking Forward”, by A. K., Praetorius and C. Y. Charalambous, 2018, *ZDM*, 50, Appendix A (<https://doi.org/10.1007/s11858-018-0946-0>)

Table 2.17
The Characteristics of Example Frameworks

Number of Framework Elements			Observational Foci				
Level-1 elements ^a	Level-2 elements (range per Level-1 element)	Types					Key Observed Subjects
		Frequency & Quality (intertwined)	Frequency	Quality	Other aspects	Teacher	
Generic frameworks							
CLASS	4	12 (1-5)	×			×	(×)
DMEE	8	40 (5) ^c		×	×	×	
ISTOF	7	21 (2-4)			×	×	
TBD	3	21 (4-10)		×	×	×	(×)
Mathematics-specific frameworks							
IQA	2	8 (3-5)	×			×	
MQI	4	20 (3-7)	×			×	(×)
M-Scan	9	22 (2-3)	×			×	
Hybrid frameworks							
TEDS-Instruct	5	27 (4-7)	×			×	
TRU	5	Variable ^d	×		×		×
UTOP	4	27 (6-8)	×			×	

Notes.

^a Different frameworks are using different names for their Level-1 classification (e.g., domains, dimensions, and components). Similarly, they use different names for the Level-2 classification (e.g., dimensions, sub-dimensions, indicators, items, rubrics, and codes).

^b An × in this column implies that students is a main focus for the framework; an (×) implies that the students are a key focus in this framework, but not the main one.

Note: Adapted from “Classroom Observation Frameworks for Studying Instructional Quality: Looking Back and Looking Forward”, by A. K., Praetorius and C. Y. Charalambous, 2018, *ZDM*, 50, Appendix B1 (<https://doi.org/10.1007/s11858-018-0946-0>)

In recent years, several subject-specific frameworks have been developed in the field of mathematics education, including the Mathematical Quality of Instruction (MQI) (Charalambous & Litke, 2018; LMT, 2011), Instructional Quality Assessment (IQA) (Boston & Candela, 2018), the Reformed Teaching Observation Protocol (RTOP) (Sawada et al., 2002), and the Mathematics Scan (M-Scan) (Walkowiak et al., 2018). These frameworks offer unique insights into the teaching practices and instructional strategies that are most effective in mathematics education.

Litke et al. (2021) compared MQI, IQA, and M-Scan and found that the frameworks use similar approaches to discuss instructional quality. However, nuanced differences exist in their use. For example, all three frameworks focus on “the importance of cognitive demand, implementation of tasks, students' use of representations, and presence of mathematical explanations” (p.9). However, the use of representations has slight differences among the frameworks. While IQA and M-Scan expect students and teachers to use representations, MQI only uses it as an indicator if there is a link among representations. There are also differences in the way the frameworks address explanations and justifications. IQA requires students to make general explanations and justifications, while M-Scan looks at the presence and depth of explanations provided by students and teachers. However, in M-Scan, the role of teachers' explanations is related to the frequency of asking what, how, and why questions and soliciting students' explanations and justifications (Walkowiak et al., 2018). Although student explanations or justifications are considered in IQA and M-Scan, the key observed subjects are teachers in both frameworks. In MQI, both students and teachers are expected to make explanations in different dimensions, and the key observed subjects are explicitly teachers and implicitly students in MQI.

Boston et al. (2015) also compared IQA, RTOP, and MQI, stating that RTOP is developed for science and mathematics education and can be adapted to other areas, while IQA, MQI, and M-Scan are mathematics-specific. The constructs of IQA are academic rigor and accountable talk, and the unit of analysis is mathematical tasks. In contrast, the construct in MQI is the rigor and richness of mathematics in a lesson and unit of analysis of the chunks of the lessons (7.5-minute segments) rather than observing different quality indicators around specific tasks. Furthermore, MQI is not

reform-based but curriculum-aligned, and M-Scan is a standards-based instrument like R-TOP, which aims to understand reform-oriented teaching practices. Therefore, MQI considers all kinds of teaching practices rather than focusing on just reform or standard-based practices. As another difference, the alternative activity structures are present in IQA, not in MQI, and MQI provides a more general view where alternative activity structures are not used.

Aside from the subject-specific frameworks, several domain-specific frameworks have been developed to evaluate the quality of instruction in specific areas of mathematics rather than focusing on broader teaching practices. For example, Litke (2020) developed a framework for evaluating contemporary algebra instruction, while Mesa et al. (2019) built on the work of Litke (2015) and the MQI to create a framework for assessing algebra instruction in post-secondary mathematics. Thurlings et al. (2019) developed The Quality of Instruction for Fraction Lessons (QIFL) framework based on MQI and general pedagogical instructional strategies to evaluate the quality of instruction for fractions.

All in all, the current study aims to explore the statistics practices of 7th-grade mathematics teachers. The Mathematical Quality of Instruction (MQI) framework was chosen to understand the quality of these practices within the broader context of mathematics instruction. While each framework has its own strengths and limitations, the MQI was selected for its comprehensive approach. In the next section, the relevant literature on the MQI will be discussed to describe the characteristics of the instrument derived from the framework and to justify its selection over other subject-specific frameworks, as outlined in Table 2.16 and Table 2.17. Before introducing the details regarding the MQI, the discussion of hybrid frameworks will also be briefly touched upon.

2.2.3.1.3. Hybrid Frameworks

Subject-specific frameworks have a limited focus and do not provide feedback on general teaching practices essential for effective instruction. However, while applicable across subjects, generic frameworks do not offer subject-specific feedback to improve

subject-related instruction (Hill & Grossman, 2013). To address this issue, researchers have combined both content-specific and content-generic teaching practices to capture instruction more comprehensively (Blazar et al., 2017; Kane & Staiger, 2012; Senden et al., 2021). Some studies have utilized both subject-specific and generic frameworks to capture both aspects of instruction (e.g., Blazar et al., 2017; Kane & Staiger, 2012). Hybrid frameworks, which incorporate both subject-specific and general teaching practices, have been developed and used in some studies (Schoenfeld et al., 2018; Senden et al., 2021). Hybrid frameworks have been developed to address single subject matters, like subject-specific frameworks, while also including general teaching practices (Senden et al., 2021). Combining both practices provides a comprehensive approach to understanding instructional quality and improving teaching practices for better student learning (Kane & Staiger, 2012; Praetorius & Charalambous, 2018). Examples of hybrid frameworks include Teaching for Robust Understanding (TRU) by Schoenfeld (2018), the UTeach Observation Protocol (UTOP) by Walkington and Marder (2018), and Teacher Education Study in Mathematics-Instruct (TEDS-Instruct) by Schlesinger et al. (2018).

In conclusion, the complexity of instruction can be understood through various frameworks, each with unique characteristics and approaches discussed in this section. Researchers identified the need for a comprehensive framework to understand better and improve instruction. As stated, the differences between these frameworks lie in their purpose, conceptualizations, measurement, and operationalizations. The frameworks are used in observational studies of mathematics education research for program evaluation, professional development, and teacher education. Such frameworks could potentially enhance the design and delivery of instructional practices, leading to improved student learning outcomes. In the next section, I will provide an in-depth description of the MQI, as it was used to understand statistical teaching practices within broader mathematics instruction for 7th-grade teachers in this study.

2.2.3.2. The Mathematical Quality of Instruction Framework

The MQI focuses on subject-specific practices in mathematics. The primary purpose of developing the MQI tool is to analyze the quality of mathematics lessons to improve

teaching and learning (Charalambous & Litke, 2018). In the developmental process, the MQI arises from the need for “more comprehensive pictures” related to instructional quality in mathematics (LMT, 2011, p. 30) and benefits from existing literature on effective mathematics teaching, including MKT, affordances, deficiencies, expert judgment, and observing lessons (Hill et al., 2008; LMT, 2011; Charalambous & Litke, 2018). Furthermore, the MQI is based on the instructional triangle that Cohen et al. (2003) developed, and the MQI considers teacher and student interactions around the content.

In the LMT (2011), the mathematical quality of instruction was defined as referring solely to the mathematical content presented during instruction. They clarified this by stating, “By the mathematical quality of instruction, we mean only the nature of the mathematical content available to students during instruction. This is distinct from a focus on instructional strategies or pedagogical style” (LMT, 2011, p. 30). The LMT also provided definitions for the terms “mathematical,” “quality,” and “instruction.” Specifically, they explained that the term “mathematical” referred to the disciplinary nature of the MQI, which was derived from the literature on mathematics knowledge for teaching. Additionally, it was noted that this concept is related to subject-specific practices in mathematics, but it encompasses more general mathematics than just specific topics.

While we include basic information about the mathematical topics being addressed, our focus is more on mathematical features closely related to the work of teaching, such as the presence of error and how it is treated, or precision in the use of mathematical language (LMT, 2011, p.31).

Then they describe what they mean by quality. Their conceptualization comprises “the distinctive character of something” (p.30). However, the quality considers “which features are worth considering and to examine constellations of features that constitute the overall character of particular instruction” (p.30). Ultimately, they conceptualized “instruction” from Cohen et al. (2003) by considering the instructional triangle through the interactions among teachers, students, and the content. They added that instruction “...includes not only what teachers say and do, but also what students say and do and

what a curriculum affords. The work of the teacher (namely teaching) is to manage these interactions” (p.30).

MQI's early works by Hill et al. (2008) focused on aspects of quality instruction in mathematics based on their expectations from the literature on mathematical knowledge for teaching (MKT). Hill et al. (2008) expected that “teachers with stronger MKT offering “students fewer of the ‘deficits’ and more of the ‘affordances’ seen in the literature and, perhaps, additional benefits as well” (p.437). The first aspect of quality instruction is the richness of mathematics, including “the use of multiple representations, linking representation, mathematical explanation and justification, and explicitness around mathematical practices such as proof and reasoning” (p.437). The second aspect is about connecting classroom practice to mathematics, ensuring teachers guide instruction while maintaining mathematical focus. The third aspect is attending to student thinking and responding appropriately, which requires teachers to be aware of students' mathematical thinking and respond accordingly. The fourth aspect is avoiding ineffective use of students' misunderstandings or responding inappropriately to student utterances. The fifth aspect is the proper use of mathematical language, including accurately conveying mathematical ideas and avoiding errors in computational, linguistic, and representational aspects.

Later, LMT (2011) presented a conceptual scheme that evolved through classroom observations and literature review, describing MQI with a set of constructs and codes. They emphasized that the study developed a theory on key mathematical classroom practices and an instrument to measure the presence of those characteristics. LMT categorized the constructs into two categories: “negative instances of classroom mathematics (e.g., errors, responding to students inappropriately) and positive instances of classroom mathematics (e.g., rich mathematics, proper use of mathematical language, promoting equity)” (p.43). These categories correspond to affordances and deficits in Hill et al.'s (2008) description of instructional quality in mathematics.

As the MQI is developed through an iterative process, the latest version of the MQI was presented in 2014 (Hill et al., 2014). The latest form has four main dimensions, including twenty codes and one separate code from the four dimensions for segment-

level coding. The detailed descriptions of the four dimensions and the twenty codes are presented in Table 2.18 (Charalambous & Litke, 2018, p.447-448). As seen in Table 2.18, the dimensions have varying codes; seven codes for the Richness of Mathematics dimension, three for the Working with Students dimension, four for Errors and Imprecision dimension, and six codes for the Common Core Aligned Practices. Furthermore, an overall code is present in all dimensions addressing the holistic perspective of the segments in terms of conceptualizing the dimensions, not the summary of the remaining codes.

As previously discussed, Charalambous and Litke (2018) stated how the instructional triangle functions in the MQI. First, the two dimensions, the Richness of Mathematics and Errors and Imprecision, look at the relationship between the teacher and the content. The Richness of Mathematics dimension looks at the quality and depth of the mathematical content teachers provide their students. In contrast, the Errors and Imprecision dimension examines the accuracy and precision of the instruction provided.

The Common Core Aligned Student Practices dimension focuses on the interaction between students and the content. It considers how well instruction supports the development of students' mathematical practices aligned with the Common Core State Standards. Turkish curriculum does not follow common core state standards even with certain commonalities. Therefore, in the following sections of the current study, this dimension is referred to as “Student Participation in Meaning-Making and Reasoning,” as per its previous name in LMT (2011), due to its broader use worldwide with different curricula.

Finally, the Working with Students and Mathematics dimension focuses on the relationship between the teacher and the students. It explores how the teacher facilitates students' interactions with mathematics, including how they manage the classroom, engage with students, and provide feedback to help support their mathematical learning. The additional code separate from the four dimensions is called Classroom Work is Connected to Mathematics, focusing on whether the mathematical content is the primary concern for the segment.

Furthermore, the latest version consists of ten whole lesson codes assigned to a lesson (Hill et al., 2014). The codes represent the overall quality of a lesson. The whole lesson codes are like a summary measure of the necessary components of a lesson. Nine lesson-level codes parallel the segment-level codes and four dimensions in the MQI. These include considering the amount of mathematical work done, rich mathematics covered in a lesson, and student engagement throughout the lesson (NCTE, n.d.). Nine codes are specific: *“Lesson Time is Used Efficiently,” Lesson is Mathematically Dense,” Students are Engaged,” “Lesson Contains Rich Mathematics,” “Teacher Attends to and Remediate Student Difficulty,” Teacher Uses Student Ideas,” “Mathematics is Clear and not Distorted,” “Tasks and Activities Develop Mathematics,” and “Lesson Contains Common Core Aligned Student Practices.”* The last code is *Whole-Lesson Mathematical Quality of Instruction*, giving an overall impression of the lesson. Like 'Overall' codes in the segment level codes, the code is not the average of the scores of other whole lesson codes (see the detailed explanation of each code in whole lesson codes in Hill et al., 2014).

2.2.3.2.1. Psychometric Properties of MQI Instrument

The MQI has strong technical properties making it a reliable and valid instrument for evaluating mathematics instruction. Charalambous and Litke (2018) emphasized the importance of scoring design in obtaining reliable estimates of instructional quality. To ensure reliability, the researchers conducted various analyses on different dimensions of MQI, including Inter-Rater Reliability-ICC2, Cronbach alpha's, and Relative (ρ)/Absolute (ϕ) coefficient for D study (see Praetorius & Charalambous, 2018, Appendix C, for all reliability scores). Furthermore, studies suggested different scoring designs for different dimensions. Hill et al. (2012a) discussed that three lessons with two raters are necessary to obtain reliable scores in the richness of mathematics, working with students and mathematics, and errors and imprecision at least $\rho=0.70$ level. For working with students and the mathematics dimension, four lessons with three raters or vice versa are necessary to obtain reliable results. Different combinations of lessons and raters are also calculated to provide reliable estimates for the MQI in different contexts (see Hill et al., 2012b, p.100, for a detailed comparison of reliability estimates).

Table 2.18

Description of the Four Main Dimensions of the Mathematical Quality of Instruction

Dimension	Description	Codes
Richness of the Mathematics	Captures the depth of the mathematics offered to students. Rich mathematics can be established by either focusing on the meaning of facts and procedures (e.g., through linking representations or providing explanations) or by focusing on key mathematical practices (e.g., working on multiple solutions, developing generalizations, using appropriate mathematical language).	a. Linking and connections: explores if explicit connections are drawn between representations of mathematical concepts. b. Explanations: examines whether mathematical explanations are offered by the teacher and/or the students c. Mathematical meaning and sense-making: captures the extent to which instruction is geared toward supporting students to make meaning of mathematical ideas. d. Multiple procedures or solution methods: investigates whether teacher/ students use(s) and discuss(es) multiple approaches to a given problem. e. Patterns and generalizations: examines the extent to which teacher/students identifies patterns and use(s) them to develop generalizations. f. Mathematical language: captures the depth and density of the mathematical language used during instruction. g. Overall code: a holistic code that captures the overall depth of the mathematics offered to students
Working with Students and Mathematics	Refers to teachers' ability to appropriately interpret and respond to students' mathematical ideas and errors, as well as their capability to capitalize on these ideas and errors during instruction.	a. Remediation of student errors and difficulties: examines the degree to which the teacher responds to student errors or misunderstandings and offers procedural/conceptual remediation. b. Teacher uses mathematical contributions: captures whether and how the teacher responds to and builds on students' mathematical productions to steer the lesson toward a mathematical goal. c. Overall code: a holistic code that provides an evaluation of the teacher–student interactions around the content
Errors and Imprecision	Explores the degree to which instruction is mathematically incorrect.	a. Mathematical content errors: captures instances in which instruction includes major mathematical errors. b. Imprecision in language and notation: includes instances of teachers' imprecise mathematical language or notation that interfere with the presentation of the content c. Lack of clarity in presentation of mathematical content: captures instances when the presentation of mathematical ideas is unclear or confusing, thus obscuring the mathematical content. d. Overall code: a holistic code that captures the overall presence of teacher errors in doing and talking about mathematics

Note: From "Studying instructional quality by using a content-specific lens: the case of the mathematical quality of instruction framework" C. Y., Charalambous & E. Litke, 2018, ZDM, 50, p. 447 (<https://doi.org/10.1007/s11858-018-0913-9>)

Table 2.18 (cont'd)

Dimension	Description	Codes
Common Core Aligned Student Practices (Student Participation in Meaning-Making and Reasoning)	Focuses on student involvement with the mathematical content through cognitively demanding activities	a. Students provide explanations: assesses students' explanations of ideas, procedures or solutions. b. Student mathematical questioning and reasoning: explores instances of students engaging in meaning-oriented mathematical practices, such as posing mathematically motivating questions, making hypotheses, offering examples and counter-examples, etc. c. Students communicate about the mathematics of the segment: captures the extent to which students contribute mathematically to the lesson. d. Task cognitive demand: examines student engagement in tasks in which they think deeply and reason about mathematics; this code refers to the enactment of the task, regardless of the initial demand of the curriculum/textbook task or how the teacher sets up the task for students. e. Students work with contextualized problems: captures the extent to which students work on problems situated in real-world contexts. f. Overall code: a holistic code capturing evidence of students' involvement in the mathematics of the lesson and the extent to which students participate in and contribute to meaning-making and reasoning.
Score points: 1 (element not present), 2-Low (brief or pro forma enactment of a given element), 3-Mid (more than brief instance(s) of a given element but without element-specific features that would indicate a score of high) 4-High (substantive and detailed enactment of a given element)		

Note: From "Studying instructional quality by using a content-specific lens: the case of the mathematical quality of instruction framework" C. Y., Charalambous & E. Litke, 2018, ZDM, 50, p. 448 (<https://doi.org/10.1007/s11858-018-0913-9>)

Several studies have been conducted to validate the instrument's validity. Content validity was discussed with various researchers in the LMT project and mathematics educators other than LMT Project (Bostic et al., 2021; Charalambous & Litke, 2018; LMT, 2011). As the construct validity, Blazar et al. (2017) found that MQI is structured under two factors: (1) ambitious instruction, with the codes from the dimensions of the richness of mathematics, working with students and mathematics, and common core-aligned practices (student participation in meaning making and reasoning), and (2) errors and imprecision. Later, Hill and Lovison (2022) described the MQI with ambitious instruction, including “the presence of disciplinary features (e.g., mathematical explanation, generalization, language precision), the role students play in the production of knowledge (e.g., student explanation, cognitive demand of student tasks) and classroom discourse (e.g., teacher uptake of student ideas)” (p.16).

Furthermore, concurrent and convergent validity were also established, and it was found that MQI is highly correlated with MKT (Hill et al., 2008, 2012b, 2012c, 2015). In addition, Kane and Staiger (2012) found that MQI has a higher correlation with subject-specific instruments like UTOP ($r=.85$) than generic frameworks like CLASS ($r=.69$). MKT is not correlated with CLASS, which is evidence of divergent validity (Blazar et al., 2017). Furthermore, studies showed that MQI and student learning are related (Blazar et al., 2016; Blazar & Kraft, 2017; Charalambous & Kyriakides, 2017), and the Measures of Effective Teaching (MET) Project results found a positive but small correlation between MQI and student scores (Hill et al., 2015; Kane & Staiger, 2012). Overall, MQI is a reliable and valid instrument.

2.2.3.2.2. Justification for Using the MQI Framework

The MQI has been chosen for this study because it considers the quality of instruction from a content-specific perspective, considering the interactions among the teacher, students, and content (Praetorius & Charalambous, 2018). While the observation primarily focuses on the teacher and their interactions with other components, there is also an implicit focus on student participation in meaning-making and reasoning. In addition, it does not consist of general practices like classroom climate, which is not within the scope of the current study.

Unlike other frameworks, the MQI aims to provide a comprehensive view of the mathematics classroom rather than just emphasizing specific practices, such as reform-oriented practices. Boston et al. (2015) noted that the MQI is “not biased toward any type of instruction” (p. 166). Therefore, MQI provides an understanding of subject-related practices of instructional quality in mathematics education (Charalambous & Litke, 2018). The study aims to understand statistics instruction within the middle school mathematics curriculum. As previously discussed, domain-specific practices, also known as discipline-related practices provided in GAISE are analyzed through “Transcript Coding Protocol Using GAISE Framework.” Therefore, MQI helps to understand the quality of statistics instruction in the broader context of mathematics instruction. Furthermore, looking at the statistics instruction with the specific components of MQI inform the practice for the improvement to provide a better learning environment for students (Charalambous & Litke, 2018)

Furthermore, while all subject-specific frameworks discussed in this section are considered valid and reliable, the reliability studies conducted on the MQI have been extensive, with G and D studies using a large amount of data from the MET study. The MQI has also met various types of validity, including content, predictive validity (both affective and cognitive), construct, and divergent validity. As Hill and Lovison (2022) stated, “well-established technical properties” are present in MQI (p.13). Although the scope of the current study is limited, it is still essential to use an instrument that has been thoroughly discussed in terms of its validity and reliability.

Another reason for choosing the MQI framework is its provision of online training for raters, which is free and available to potential researchers in the field. The availability of this training during the study period helps raters to produce more reliable scores and inferences regarding the quality of statistics instruction provided by teachers. Additionally, the MQI requires raters to have a level of Mathematical Knowledge for Teaching (MKT), which is an important factor contributing to the accuracy of the framework.

In conclusion, the MQI framework has been used in this study due to its content-specific perspective, comprehensiveness, and extensive validity and reliability studies.

The framework provides a comprehensive view of the mathematics classroom, focusing on curricular-aligned practices rather than just reform or standard-based practices. Its online training provision for raters and requirement for Mathematical Knowledge for Teaching (MKT) further contribute to the accuracy of the framework. Using the MQI framework in this study provides insights into the quality of statistics instruction in the broader context of mathematics instruction, with the potential for informing practice for improving student learning outcomes.

2.2.3.2.3. Reviews on Descriptive Studies Using the MQI Framework

MQI is one of several mathematics-specific frameworks for classroom observation, along with IQA, RTOP, and M-Scan, which aim to understand mathematical teaching practices. This section provides reviews on the studies to examine mathematics instruction through the MQI framework (e.g., Blazar, 2015; Blazar & Pollard., 2017; Hill et al., 2019; Hill & Lovison, 2022; Kane & Stigler, 2012; Lynch et al., 2017; Mantzicopoulos et al., 2019; Park & Kim, 2023). The overview of example studies is presented in Table 2.19.

Kane and Stigler (2012) conducted a comprehensive analysis of 7491 lessons from 1333 teachers in the US using a variety of generic and subject-specific frameworks to describe instructional quality. One of the instruments used was an adapted version of MQI, called MQI lite, for the MET study. Kane and Stigler (2012) found that most lessons (93.4%) contained mathematical work and that errors and imprecision were not present in most lessons. However, they observed that the teaching practices related to working with students and mathematics, the richness of mathematics, and student participation in meaning-making and reasoning were scored low in most segments (70%). In addition to the results related to the MQI, the study also found a relationship between instructional quality measured by four observation instruments (MQI, UTOP, CLASS, FFT) and student achievement gain scores obtained through a standardized test assessing conceptual understanding of mathematics.

Table 2.19***The Overview of the Studies conducted on the Mathematical Quality of Instruction***

Study	Participants	Research Focus	Key Findings
Kane and Stigler (2012)	1333 teachers	Instructional quality using generic and subject-specific frameworks, relationship with student achievement	Most lessons contained mathematical work, but teaching practices related to student participation, the richness of mathematics, and working with students scored low; instructional quality measured by four observation instruments related to student achievement gained scores
Hill et al. (2019)	329 teachers	Quality of teaching practices aligned with curriculum, common core student practices	Most teachers used direct instruction, with few giving controls to students; meaning-oriented practices were under teacher control, and the quality of disciplinary practices was low; higher quality of mathematical practices in districts with reform-oriented curriculum
Hill and Lovison (2022)	98 teachers	The extent of ambitious instruction and errors and imprecision, comparison with typical instruction in 1999	Most lessons were typical based on overall MQI, limited presence of positive practices, and low scores when present; no differences in errors between in 1999 and 2016, but significant differences in ambitious instruction
Park and Kim (2023)	Two beginning teachers	Instructional quality over three years	One teacher improved instructional quality over time, while the other did not improve; the second teacher had major errors and imprecision in the last observation, while the first had fewer errors and imprecision.
Blazar (2015)	111 teachers and 3203 students	Relationship between high-quality instruction and math achievement	Ambitious instruction was related to the quality indicators of CLASS and errors and imprecision in the MQI. the MQI scores on ambitious instruction were related to student achievement obtained through low-stakes assessment.
Lynch et al. (2017)	298 teachers and 6780 students	Relationship between the quality of instruction and student achievement	MQI scores are related to student achievement in districts with standardized tests more aligned with MQI and cognitive demand levels.
Blazar and Pollard (2017)	368 teachers	Relationship between test preparation and instructional quality	Test preparation activities were negatively related to ambitious instruction. They predict lower instructional quality.
Mantzicopoulos et al. (2019)	20 teachers and 285 students	Kindergarten math instruction quality and relationship with student achievement and motivation	Quality indicators in ambitious instruction practices were relatively low, but the teaching practices were aligned with state standards. The MQI with ambitious instruction and whole lesson codes was related to teacher-rated math achievement but not standardized tests. Whole lesson codes were positively associated with teacher-rated student motivation.

Later, Hill et al. (2019) investigated the quality of teaching practices of 329 fourth and fifth-grade mathematics teachers. The participants were from five districts, of which three adopted reform-oriented curricula, while the remaining were familiar with reform movements but did not engage in many reform-based practices. The researchers used the earliest versions of MQI to assess whether teaching practices were aligned with the curriculum. They looked at practices like mathematical explanations, sense-making, student questioning and reasoning, and teacher use of student thinking. The results revealed that most teachers used direct instruction as the instructional format, with only seven percent giving control to students.

Hill et al. (2019) also looked at the richness dimension and found that meaning-oriented practices were under the teacher's control, and the quality of disciplinary practices was mostly low. When looking at common core student practices, they found that students were involved in cognitively challenging work in 84% of the segments, but mainly at a low level. Lastly, they found that most lessons covered mathematics-related work, with major errors, language imprecision, and lack of clarity staying low in most lessons. Based on the results, Hill et al. (2019) concluded that teachers who participated in the study carried out mathematical practices in reform-oriented requirements but used teacher-led practices. They also noted that districts with reform-oriented curricula experienced higher-quality mathematical practices based on MQI. The results of this study are similar to the Measures of Effective Teaching study reported by Kane and Staiger (2012) in that most teachers used meaning-oriented and discipline-oriented practices at a low level when they used them. However, this study showed more reform-oriented practices than Kane and Staiger's (2012) research due to consisting of the districts with reform-oriented curricular materials.

Similarly, Hill and Lovison (2022) conducted a study to investigate the nature of mathematics instruction in the US, focusing on the extent to which it includes ambitious instruction and errors and imprecision and how it has changed from 1999 to 2016. The study involved 98 teachers, and 93 of their videos were analyzed using codes from the latest version of MQI. They found that most lessons received an overall MQI score of 3, considered typical or traditional, with a mix of affordances and deficiencies. When they looked at each code separately, they found that 58% of the

lessons had at least one positive incidence, but most segments lacked such practices, and when present, they were scored low in general. In terms of errors and imprecision, most segments did not contain errors, but when present, they were generally low-level.

When comparing the mathematics instruction in 2016 to typical instruction in 1999 in terms of ambitious instruction and errors, Hill and Lovison (2022) found significant differences in ambitious instruction but no differences in errors. These findings support the arguments made by Hill et al. (2019) and Kane and Stigler (2012) that lessons were primarily traditional and teacher-led, with less participation of students in classroom discourse. Overall, Hill and Lovison's (2022) study provides further evidence that mathematics instruction in the US is still largely traditional and teacher-led, with a limited engagement of students in discourse, and that more reform-oriented practices are needed to improve the quality of mathematics instruction.

In a more recent study, Park and Kim (2023) conducted a case study on the instructional quality of two beginning teachers in South Korea and tracked changes over three years. They found that one teacher improved the instructional quality over the years, with practices related to the richness of mathematics, working with students, and student participation in meaning-making and reasoning improving over time. However, the other teacher did not enhance their instruction. Their teaching practices related to all three dimensions remained around score 1, indicating a general lack of or low quality of practices. In addition, the second teacher had major errors and imprecision in the last observation, while the first had fewer errors and imprecision in their lessons.

Blazar (2015) examined the relationship between high-quality instruction and mathematics achievement with 111 teachers and their 3203 students using MQI and CLASS to analyze the quality of instruction from a maximum of three lessons. The results show that ambitious instruction was related to the quality indicators of CLASS and errors and imprecision in the MQI. Furthermore, student achievement was analyzed through an assessment tool developed by the researchers, coded in terms of cognitive demand levels, including non-routine problems. The results show a relationship between the MQI scores on ambitious instruction and student achievement

obtained through low-stakes assessment. This result is consistent with Kane and Stigler's (2012) findings in the MET study. However, the study found that CLASS scores, representing the generic instructional quality, were unrelated to student achievement. This result is not aligned with the MET study's results. Blazar (2015) explained that such differences might arise from the background variables controlled in his study or differences in the student achievement tests (high-stakes tests vs. low-stakes tests).

Lynch et al. (2017) examined the relationship between elementary mathematics teachers' quality of instruction and student achievement. The study included 298 teachers and their 6780 students from five districts in the US. The researchers analyzed the relationship between teacher observation scores obtained from the ambitious instruction section of the MQI in 1560 lessons (with an average of 5.23 lessons per teacher) and student achievement scores obtained through standardized tests and researcher-developed tests. The findings indicated that MQI scores are related to student achievement in the first two districts, where the standardized tests were mostly open-ended items and more aligned with common core standards. However, the remaining three districts had standardized tests with a mid/low level of alignment to MQI and cognitive demand, mainly consisting of multiple-choice items. Moreover, the researchers developed tests with the same cognitive demand level for all districts. They found a significant relationship between MQI scores, and student achievement scores based on these tests in district 1 and district 5.

In a subsequent study, Blazar and Pollard (2017) explained the relationship between test preparation and instructional quality. The study involved 368 teachers from five districts, and the researchers analyzed 73 lessons on test preparation and other activities from 60 teachers observed in Lynch et al.'s (2017) study. All observed lessons were analyzed using the ambitious instruction section of the MQI. The results revealed that test preparation activities were negatively related to ambitious instruction. Particularly in District 1, which had one of the highest MQI scores, the negative relationship was dominant, even though the test required more cognitive demand than the last three districts. The researchers concluded that test preparation activities predict lower instructional quality. They suggested that test preparation could

be distracting in the context of high-quality instruction, which might explain why the other districts did not show similar predictions.

Mantzicopoulos et al. (2019) focused on kindergarten mathematics instruction quality and its relationship with student achievement and motivation. The study involved 20 teachers and 285 students. The researchers used the latest version of MQI to analyze five lessons per teacher. Mathematics achievement was assessed through teacher ratings and standardized tests on mathematical reasoning, while motivation was assessed through teaching ratings and administering a motivation scale. The study found that quality indicators in ambitious instruction practices were relatively low, but the teaching practices were aligned with state standards. The second factor, errors and imprecision, yielded positive results since the teaching practices included low mathematical errors. When analyzing the whole lesson codes, most observed instruction was typical. The MQI with ambitious instruction and whole lesson codes was found to be related to teacher-rated mathematics achievement but not standardized tests. Ambitious instruction was not associated with student motivation, but whole lesson codes were positively related to teacher-rated student motivation. The study also found that the MQI scores were low, and the observed lessons were limited. Although some types of tests did not show a relationship with the MQI, the researchers obtained at least a small relationship with student achievement and motivation.

In conclusion, the studies discussed provide a comprehensive overview of the state of mathematics instruction, particularly in the US, due to the origin of the MQI framework and emphasize the need for high-quality instruction to improve student achievement and motivation. While traditional and teacher-led practices remain prevalent in classrooms, the studies suggest that reform-oriented practices that engage students in discourse and reasoning are needed to improve the quality of mathematics instruction. The findings from observational studies indicate that ambitious instruction practices are positively associated with student achievement. The studies also highlight the importance of aligning instruction and assessments to ensure accurate measurement of student achievement. Furthermore, high-quality instruction may require a different approach to preparing students for assessments, as test preparation activities are negatively associated with instructional quality. Overall, the studies offer

valuable insights into the current state of mathematics instruction. The following section will provide a brief review of factors related to instructional quality, specifically teacher knowledge, due to the scope of the current study.

2.2.3.3. Factors Related to Instructional Quality

Statistics education needs to be understood in depth through exploring teaching and learning statistics within a classroom context and factors affecting the teaching practices for a better learning environment and student achievement (Harrison, 2020). Koehler and Grouws (1992) proposed a research model on classroom processes to understand these dynamics better. This model tells the complexity of teaching by stating the interaction among teaching and learning mathematics as well as the teacher and student characteristics (see Figure 2.8).

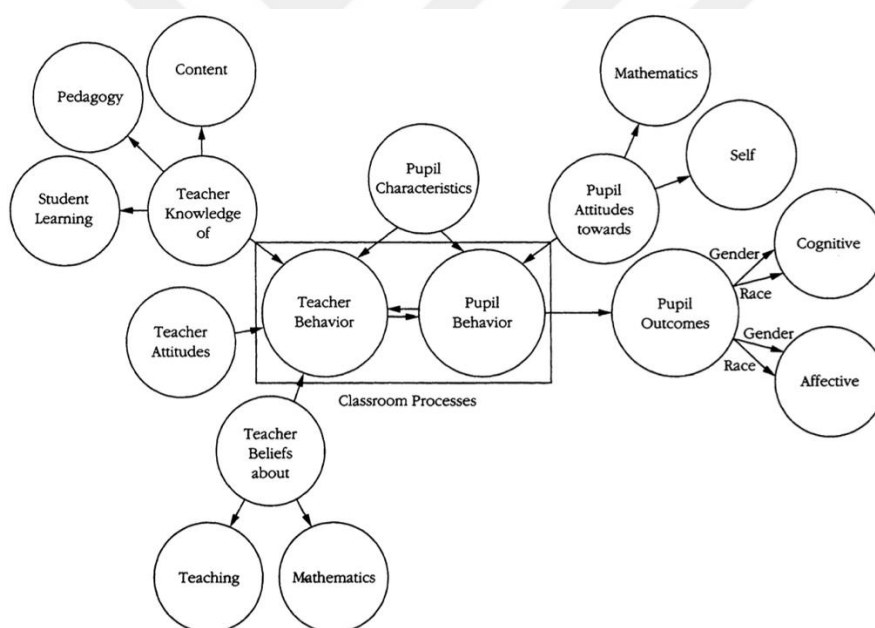


Figure 2.8 Research Model

Note: From “Mathematics teaching practices and their effects,” by M. S. Koehler & D. A. Grouws, 1992, *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*, p.118

Blömeke et al. (2020) defined teacher competence as a continuum encompassing teacher dispositions, situation-specific skills, and performance. This competence is

essential for delivering high-quality instruction because it includes cognitive characteristics (e.g., teacher knowledge), affect-motivation (e.g., teacher beliefs), and teacher skills (e.g., diagnostic skills) that influence teaching practices and ultimately impact student achievement. This study focuses on statistics teaching practices of mathematics teachers. Therefore, in the following section, I will discuss teacher knowledge and its relationship with instructional quality in more detail, as it is relevant to the MQI framework previously introduced in the literature review.

2.2.3.3.1. Teacher Knowledge

Teacher knowledge has an important place in teaching mathematics (Hill et al., 2008). In literature, there is an emphasis on the interrelations between student knowledge, teacher knowledge, and high-quality teaching (Kahan et al., 2003). Hill et al. (2008) further emphasized that instruction is a mediator between student learning and teacher knowledge. Although the significant role of teacher knowledge in instruction is mentioned in the literature, there is a lack of a connection between theory and practice that leads to not convincing results in policy and teacher education (Hill et al., 2008). First, it is essential to understand how teacher knowledge is conceptualized in the literature.

Shulman (1986) developed the conceptualization of teacher knowledge to stand against people assuming that teacher knowledge is merely related to subject matter knowledge or pedagogical knowledge. Shulman (1986) defined teacher knowledge into three categories: curricular knowledge, pedagogical content knowledge (PCK), and content knowledge.

Grossman (1995) detailed Shulman's conceptualization of teacher knowledge by categorizing it into six domains: knowledge of learners and learning, knowledge of content, knowledge of general pedagogy, knowledge of context, knowledge of curriculum, and knowledge of self. For example, although Shulman (1986) separated content knowledge and PCK, Grossman (1995) defined both into the same category as knowledge of content. Shulman's categorization of teacher knowledge differed is also seen through Grossman's categorization of PCK, including knowledge of students'

understanding, instructional strategies, purposes of teaching, and curriculum (as cited in Graeber & Tirosh, 2008).

Following Shulman (1986), researchers have defined, and detailed teacher knowledge as Grossman did (Ball et al., 2008; Graeber & Tirosh, 2008; Grossman, 1995; Hill et al., 2005). Although various definitions are provided based on Shulman's conceptualization of teacher knowledge with the missing paradigm as PCK that links content knowledge and practical knowledge, PCK could not develop much because of a variety of definitions and empirical bases (Ball et al., 2008). Researchers have discussed the difficulty of separating two knowledge categories: Subject-matter knowledge and PCK (Borko et al., 1992). The study also showed that PCK and subject-matter knowledge are sometimes inseparable (Krauss et al., 2008). In addition, the definitions also discuss whether teacher knowledge, especially PCK is static, as Shulman presented, or an active process (Graeber & Tirosh, 2008). Although such discussions about the teacher knowledge concept have continued, the concept of teacher knowledge is studied with respect to the subject matter, unlike Shulman's conceptualization of teacher knowledge applied to general education (Ball et al., 2008; Borko et al., 1992; Kahan et al., 2003).

2.2.3.3.1.1. Mathematical Knowledge for Teaching

In mathematics education research, mathematical knowledge for teaching (MKT) is developed onto Shulman's conceptualization and the developments in the teacher knowledge concept. For example, Fennema and Franke explained Shulman's conceptualization in terms of mathematics with five components, including "the knowledge of the content of mathematics, knowledge of pedagogy, knowledge of students' cognitions, context-specific knowledge, and teacher beliefs" (as cited in Mohr, 2006, p.219). Furthermore, Hill et al. (2005) conceptualized mathematical knowledge for teaching as follows:

By "mathematical knowledge for teaching," we mean the mathematical knowledge used to carry out the work of teaching mathematics. Examples of this "work of teaching" include explaining terms and concepts to students, interpreting students' statements and solutions, judging and correcting textbook treatments of particular topics, using representations accurately in the

classroom, and providing students with examples of mathematical concepts, algorithms, or proofs (p.373).

Unlike Fennema and Franke's teacher knowledge in mathematics, Hill et al. (2005) conceptualized mathematical knowledge for teaching apart from beliefs, just pedagogical concerns and subject-matter concerns. However, they focused on more specialized pedagogical content knowledge in Shulman's conceptualization. However, different from Shulman, they conceptualized MKT as an active process, not static, since they defined it as concerning the work of teaching.

Hill et al. (2008) proposed a conceptualization of mathematical knowledge for teaching (MKT), which consists of subject matter knowledge and pedagogical content knowledge. Subject matter knowledge is divided into two subdomains: common content knowledge and specialized content knowledge. Pedagogical content knowledge is also divided into two subdomains: knowledge of content and students and content and teaching. Groth and Meletiou-Mavrotheris (2018) further explained that the conceptualization of statistical knowledge for teaching (SKT) was mainly based on Hill et al.'s (2008) work on MKT.

2.2.3.3.1.2. Statistical Knowledge for Teaching

Several studies have discussed statistical knowledge for teaching (SKT) (e.g., Burgess, 2011; Groth, 2007; Groth, 2013; Leavy, 2015; Ponte, 2012). Groth (2013) developed a framework for SKT similar to Hill et al.'s (2008) conceptualization of MKT. The framework has two domains: Subject Matter Knowledge and Pedagogical Content Knowledge, each with three subdomains.

The Subject Matter Knowledge domain includes three subdomains: Common Statistical Knowledge, which contains knowledge of statistical concepts and procedures, including conceptualizing descriptive and inferential statistics and carrying out the procedures related to those concepts; Specialized Statistical Knowledge, which includes knowledge related to the specific age group of students being taught for instance, knowing the difficulties and needs of young children while interpreting a graph; and Horizon Knowledge, which relates to the broader context of statistics.

The Pedagogical Content Knowledge domain also includes three subdomains: knowledge of content and students, which involves understanding students' errors and misconceptions in statistics; knowledge of content and teaching, which requires knowledge of effective teaching strategies; and knowledge of curriculum, which involves understanding the curricular sequence that benefits student learning.

Burgess (2011) suggested that both domains of SKT are essential for teachers to effectively use statistical investigations for both student learning and their own learning of statistics. Lovett and Lee (2018) also emphasized the importance of developing Subject Matter Knowledge for developing Pedagogical Statistical Knowledge.

Ponte (2012) introduced a model of teacher knowledge called didactics knowledge, which comprises four components: knowledge of mathematics, knowledge of teaching practice, knowledge of students and their learning, and knowledge of the curriculum (as cited in Rodrigues & Ponte, 2022). In a later study, Rodrigues and Ponte (2022) presented an adapted version of the first two components of the framework for statistics education. They defined statistical knowledge as “the fundamental concepts and procedures of the area, in addition to knowledge about school statistics” (p. 227). This component was similar to statistical content knowledge described by Groth (2013) in SKT. They emphasized that this type of knowledge focuses on teachers' perceptions and the implications of those perceptions for practice, as Burgess (2011) also stated. The second component, knowledge of teaching practice, is described as “the ability to plan each class, organize the performance of tasks, and deal with the preparation, conduction, and evaluation of the teaching-learning process” (Rodrigues & Ponte, 2022, p. 227). As some statistical concepts, such as variability, are subject-specific, teachers must consider the nature of statistical concepts when designing the learning environment, including tasks and other processes.

Burgess (2011) presented a framework for teacher knowledge to teach statistical investigations, which differs from the work of Groth (2007, 2013) and Rodrigues and Ponte (2022) in that it emphasizes the need for statistics as a separate field. Burgess combined two frameworks to create the model. The first framework is statistical knowledge for teaching, divided into two domains: content knowledge (common

knowledge of content and specialized knowledge of content) and pedagogical content knowledge (knowledge of content and students and knowledge of content and teaching). This framework is based on the work of Ball et al. (2008). Also, this part of the model parallels Groth's (2007, 2013) work. The second framework is the statistical thinking model of Wild and Pfannkuch (1999), which includes statistical thinking with transnumeration, variation, reasoning with models, and integration of statistical and contextual, investigative cycle, and interrogative cycle. Through this model, Burgess aimed to understand what kind of knowledge is needed in teachers' classroom practices. The second framework was also in line with the knowledge of teaching practice proposed by Rodrigues and Ponte (2022).

Several researchers have discussed statistical knowledge for teaching in relation to the key concepts in the literature (Gonzalez et al., 2011; Harradine et al., 2011; Jacobbe & Carvalho, 2011; Reading & Canada, 2011; Sanchez et al., 2011). Among these studies, Gonzalez et al. (2011) focused on teachers' graphical competence, which refers to their knowledge of statistics related to graphs. They adopted Curcio's (1989) and Friel et al.'s (2001) model of graph comprehension, which has four components: reading the data, reading between the data, reading beyond the data, and reading behind the data. Additionally, they reviewed Wu's (2004) work on students' graph comprehension skills, such as reading, interpreting, building, and evaluating graphs. They also discussed Aoyama's levels of graph comprehension, based on Watson and Callingham's (2003) work, which includes the idiosyncratic level, basic graph reading level, rational/literal level, critical level, and hypothesizing and modeling level. Drawing from these studies, they concluded that graphical competence encompasses three elements:

- *The ability to extract data from different sorts of graphs and to interpret meanings from them by reading between, beyond, and behind the data displayed to form hypotheses about the phenomena represented in the graph;*
- *The capacity to select and create appropriate graphs for specific situations, with or without the support of technology; and*
- *The ability to critically evaluate graphs and to distinguish the relative strengths and limitations of graphical representations, recognising that creating a graph involves an interpretation of the original data. (Gonzalez et al., 2011, p.190).*

Numerous studies have explored the statistical content knowledge of teachers regarding graphs (e.g., Burgess, 2002; Espinel, 2009; Jacobbe & Horton, 2010; Kurt, 2015; Ulusoy & Çakıroğlu, 2013). Similarly, studies have investigated pedagogical content knowledge related to graphs (e.g., Gonzalez & Pinto, 2008; Özaltun-Çelik, 2022). In general, these studies have found that teachers require improvement in their statistical knowledge of graphs.

Furthermore, Jacobbe and Carvalho (2011) reviewed the limited studies on teachers' understanding of the average, with more extensive research on statistical content knowledge (e.g., Batanero et al., 1997; Groth & Bergner, 2006; Jacobbe, 2012; Leavy & O'Loughlin, 2006). For example, Groth and Bergner (2006) examined pre-service elementary school teachers' knowledge of average concepts, using open-ended questions, and analyzing their responses through SOLO taxonomy and PUFM framework. They found that most pre-service teachers were at the multi-structural/concrete symbolic level, where they could calculate all values but lacked a comprehensive understanding of what they referred to. Only a few were at the relational/concrete symbolic or extended abstract level, with deeper conceptual understanding. In Jacobbe's (2012) study, three elementary school teachers were asked to define and explain the differences between the mean and median. While all teachers could explain how to calculate them, only one could illustrate both concepts conceptually. The teachers were then presented with a statistical task called the catapult question to determine if they considered variability when determining the center. The results show that two teachers thought reducing variability, and all could determine the center by considering the balance point. Finally, the teachers were asked to sort cards with different distributions based on their mean and median values. All teachers accurately identified the position of the mean and median in the distributions. However, one teacher struggled to sort the cards based on their mean values. Jacobbe (2012) demonstrated that teachers' understanding of the mean and median could be connected to their knowledge of variability and distribution.

Sanchez et al. (2011) noted that research on teachers' understanding of variation was relatively scarce compared to research on teachers' knowledge of graphs and averages. Reading and Canada (2011) reviewed teachers' knowledge of distribution. It concluded

that teachers need to have a strong understanding of other concepts, such as graphs, averages, and variation, to develop knowledge of distribution. Building on the work of Reading and Canada (2011), Harradine et al. (2011) discussed teachers' knowledge of sampling and inference and pointed out the limited research conducted in this area.

Lovett and Lee (2018) studied the common statistical knowledge of pre-service mathematics teachers within the subject matter knowledge domain proposed by Groth (2013), in addition to their content knowledge on specific key ideas. They used the Levels of Conceptual Understanding of Statistics (LOCUS) assessment developed by Jacobbe et al. (2014). LOCUS includes multiple choice and constructed response items within four components of statistical problem-solving processes: Formulating Statistical Questions, Collecting Data, Analyzing Data, and Interpreting Results. The assessment is parallel to GAISE and CCSSM, and the percentages of each component are as follows: “10-15% on Level A/B and 15-20% on Level B/C for Formulating Statistical Questions, 25-30% on Level A/B and 20-25% on Level B/C for Collecting Data, 30-35% on Level A/B and 25-30% on Level B/C for Analyzing Data, and 25-30% on Level A/B and 30-35% on Level B/C for Interpreting Results” (Jacobbe et al., 2014, p.4).

Lovett and Lee (2018) administered a 30 multiple choice Intermediate/Advanced Statistical Literacy version of the LOCUS assessment to 217 pre-service mathematics teachers for their study. The results show that only a quarter of pre-service teachers had common statistical knowledge for teaching at the high school level. Pre-service teachers scored highest on the first two components but scored the lowest on the last two components. In the Data analysis component, even though the questions were at Level C, pre-service teachers seemed to have difficulty with the concept of variation. In the Interpretation component, the possible reason was that the questions were at Level B, where students need to compare distributions using knowledge related to center and variability. Additionally, students had difficulty with the concepts of p-value, statistical significance, and confidence interval in this component. Overall, Lovett and Lee (2018) concluded that pre-service teachers' common statistical knowledge had weaknesses in variability, distribution, p-value, and confidence interval concepts, indicating that most of them need improvements in their common statistical knowledge.

In conclusion, there is a clear need for further research and professional development opportunities to improve teachers' statistical content knowledge, particularly in areas such as graphs, average, variation, and distribution, as well as in sampling and inference and also within the context of statistical investigations and more in pedagogical content knowledge.

2.2.3.3.1.2. Teacher Knowledge and Instructional Quality

Teacher knowledge and instructional quality are two critical factors that are essential for effective teaching and learning outcomes. The link between the two is a topic of research in mathematics education—as discussed earlier Hill et. Al. (2008) classified the research in teacher knowledge and instruction into two approaches, “deficit and affordance approaches” (p.433). The deficit approach emphasizes the shortcomings or gaps in teacher knowledge, which might affect instructional quality, while the affordance approach highlights the strengths of teacher knowledge that enable effective instruction. In this section, I will provide the existing literature on the relationship between teacher knowledge and instructional quality, explicitly focusing on the deficit and affordance approaches (e.g., Borko et al., 1992; Charalambous et al., 2012; Charalambous & Hill, 2012; Hill & Charalambous, 2012b).

First, Borko et al. (1992) investigated a student teacher's unsuccessful instruction while teaching a standard division-of-fractions algorithm. To examine the instruction with respect to teacher knowledge, they used Shulman's conceptualization of teacher knowledge. It was a one-case study including a student-teacher whose knowledge related to mathematics was investigated based on the number of mathematics courses taken, the interview done including hypothetical cases related to classroom teaching, the questionnaire conducted on the day of teaching experience, and the written work done in mathematics methods courses. They called the instruction unsuccessful because the teacher could not respond to the student's demand for a conceptual explanation for the concept taught. Although the student teacher seemed to have good knowledge of mathematics based on various data sources, the limited repertoire of instructional representations for division of fractions, limited knowledge of students' understanding, and conceptual understanding of subject-matter knowledge were found to be the reason

behind unsuccessful lesson implementation. Although student-teacher knowledge was represented as good according to the number of courses she took when pedagogical content knowledge was concluded as weak based on the interviews, and subject matter knowledge was also not conceptual based, leading to low instructional quality that teacher could not respond to students appropriately because of her mathematics errors.

Table 2.20

The Overview of the Studies Conducted on the Relationship between Teacher Knowledge and Instructional Quality

Study	Participants	Research Focus	Key Findings
Borko et al. (1992)	A student-teacher	Investigation of unsuccessful instruction	A limited repertoire of instructional representations, limited knowledge students' understanding, conceptual understanding of subject matter knowledge were the reasons behind unsuccessful lesson implementation
Charalambous and Hill (2012)	Nine teachers with varying MKT	Relationship between teacher knowledge, curriculum materials, and instructional quality	High MKT provided teachers with high-quality lessons. The joint contribution of MKT and the curriculum to instructional quality was also positive. Low MKT did not support high-quality enactments in the lessons.
Hill and Charalambous (2012b)	Two teachers with high and low MKT	Relationship between MKT and instructional quality based on fractions teaching through the enactment of a pizza problem	Teacher with high MKT provided adaptations for the lesson when necessary for conceptual understanding of learners while directing learners' attention to main points through comparisons of ratios. A teacher with low MKT could not make much adaptation for the sake of conceptual understanding; she could not even articulate the main idea
Charalambous et al. (2012)	Three teachers characterized by low-MKT and mid-MKT	Relationship between instructional quality, curriculum materials, and MKT based on integers	Teachers with higher MKT used representations related to operations in integers more effectively. Limitations in both MKT and curriculum materials led to suffering instruction.

Charalambous and Hill (2012) investigated the relationship between teacher knowledge, the use of curriculum materials, and the quality of instruction. They used multi-case studies including nine teachers with varying MKT. They defined

instructional quality with respect to Hill et al. (2008). Likely Charalambous (2010), high MKT provided teachers with high-quality lessons. As a further finding, a joint contribution of MKT and the curriculum to instructional quality was also found positive.

The details of Charalambous and Hill (2012)'s study were provided in Hill and Charalambous (2012)'s study. Hill and Charalambous (2012) explained how instructional quality is contributed by MKT and curriculum materials. The results indicate that MKT and curriculum materials affected instructional quality by providing fewer errors, teachers' capacity for mathematical ideas, and engaging students in mathematical reasoning. Furthermore, they also indicated that the instruction did not go well for teachers with low MKT since they did not support high-quality enactments in the lessons.

Hill and Charalambous (2012b) conducted another research on the relationship between MKT and instructional quality based on a specific topic in mathematics as fractions teaching through the enactment of a pizza problem. Two teachers with high and low MKT were the participants. They used instructional quality themes provided by Hill and his colleagues (2008). They found that MKT contributes to instructional quality in a way that the teacher with high MKT provided adaptations for the lesson when it is necessary for conceptual understanding of learners while directing learners' attention to main points through comparisons of ratios. Furthermore, the teacher with low MKT could not make much adaptation for the sake of conceptual understanding; she could not even articulate the main idea of the lesson.

Likely the study of Hill and Charalambous (2012b), Charalambous et al. (2012) conducted research on the relationship between instructional quality, curriculum materials, and MKT based on a specific topic in mathematics as integers. They included three teachers in the study characterized by low-MKT and mid-MKT. They found that higher MKT provided teachers with attending to students' thinking while they were in the meaning-making process. When the teachers with mid-MKT and low MKT were compared, they found that teachers with higher MKT used representations related to

operations in integers more effectively. They concluded that when there were limitations in both MKT and curriculum materials, then instruction suffered.

Although limited research has been conducted on classroom practices in statistics education, to my knowledge, there has been no specific study into instructional quality. However, there is existing research on the relationship between teaching practices and statistical knowledge for teaching, which sheds light on the crucial role of statistical knowledge in improving instructional quality. The following section provides a brief overview of the existing literature on the relationship between teaching practices and statistical knowledge for teaching (e.g., Leavy, 2015; Rodrigues & Ponte, 2022).

Rodrigues and Ponte (2022) explored two teachers' knowledge development related to teaching statistics and statistics knowledge through a teacher education program. They collected data through questionnaires, classroom observations, and interviews. They considered didactic knowledge of statistical content with the subcategories of statistical representations and measures and the investigative cycle and knowledge of teaching practice with the subcategories of management of statistical tasks and conducting statistical investigations.

Table 2.21

The Overview of the Studies Conducted on the Relationship between Statistical Knowledge for Teaching and Teaching Practices

Study	Participants	Research Focus	Key Findings
Rodrigues and Ponte (2022)	Two teachers	Knowledge development related to teaching statistics and statistics knowledge through a teacher education program	Teachers had limited statistical content knowledge and teaching practice in the initial questionnaires and interviews. The teachers developed their statistical content knowledge in both subcategories during the teacher education program.
Leavy (2015)	73 pre-service teachers	Pre-service teachers' practices in teaching statistics	Knowledge needed for planning and teaching statistics is the subcategories of statistical content knowledge, knowledge of students' thinking and reasoning about data, and knowledge of the teaching of data handling.

The results show teachers' limited knowledge of statistical content and teaching practice in the initial questionnaires and interviews. They could not express the conceptual background of calculations or had a limited understanding of collecting data or sampling in the school context. However, during the teacher education program, the teachers developed their statistical content knowledge in both subcategories. While implementing the tasks requiring statistical investigations in the program, the teachers connected concepts to the contexts. In the last interviews, the teachers emphasized the need for statistical investigations in the school context and their lack of knowledge before the teacher education program. They stated that their prior statistics teaching included limited exercises without the active engagement of students.

Leavy (2015) conducted a lesson study that examined pre-service teachers' practices in teaching statistics to understand the knowledge needed for planning and teaching data handling. They used a practice-based theory of content knowledge for teaching as the framework of the study. They examined fourteen cases from the lesson studies consisting of 73 pre-service teachers throughout three years, focusing on various aspects of statistics. The results show that the knowledge needed for planning and teaching statistics are the subdomains of the pedagogical content knowledge as the knowledge of content and students and knowledge of content and teaching. The knowledge demanded from both subdomains was found to be related to each other. Furthermore, the students developed pedagogical content knowledge within the context of lesson study, and this led to contributing to the specialized content knowledge.

In conclusion, this review emphasizes the significant relationship between teacher knowledge and instructional quality contributing to student learning outcomes. It highlights the importance of equipping teachers with the knowledge needed to effectively design and implement subject-specific instructional strategies to promote student learning.

2.3. Summary of Literature Review

This chapter aims to provide background information on the issues discussed in the study by reviewing and analyzing relevant literature. The literature suggests that teaching and learning statistics should focus on developing statistical literacy, reasoning, and thinking skills, essential for students to be informed citizens (Gomez-Blancarte & Chavez-Aguilar, 2022). Big ideas or statistical concepts and processes are critical components of teaching and learning statistics (Garfield & Ben-Zvi, 2004; Garfield & Ben-Zvi, 2008; Shaughnessy, 2007; Shaughnessy, 2019a; Shaughnessy, 2019b; Watson et al., 2013), and students need to understand these ideas instead of just learning procedures (Callingham et al., 2021; Shaughnessy, 2019a; Watson et al., 2018).

Statistical problem-solving processes are essential for teaching and learning statistics (Wild & Pfannkuch, 1999). Students must engage in complete statistical investigations or specific steps within them to draw meaningful conclusions (Watson & Callingham, 2015). The GAISE reports as curricular documents provide recommendations for statistics education at the K-12 level to promote statistical literacy and informed decision-making based on data (Bargagliotti et al., 2020; Franklin et al., 2007). The role of statistical investigations in statistics education has been discussed in different curricula, including the Turkish curriculum (Ader, 2016; 2018). Several studies have emphasized various aspects of statistics, including posing research questions, data collection, data analysis, and the entire statistical problem-solving process. However, despite the importance discussed in curricula and literature, it is not reflected in classroom practices (Litke & Hill, 2020). Therefore, gathering evidence from practice to improve statistical instructional practices is important. The literature review emphasizes the crucial role of teachers in cultivating a positive learning environment that promotes effective teaching practices, particularly in mathematics and statistics education. As statistics education differs from traditional mathematics education, it demands distinct teaching approaches (Gerthenschlager, 2019). With the increasing demand for improving classroom practices in statistics education, it is essential to thoroughly consider statistics-specific practices, including the statistical problem-solving process and the big ideas covered. The literature review includes an in-depth discussion of the statistical problem-solving process and research conducted on it, along

with statistical literacy, reasoning, thinking, big ideas, and curricular approaches in statistics education.

The literature review underscores the importance of instructional quality and understanding teaching practices to enhance student learning in mathematics. Effective teaching necessitates high-quality instruction, which can be measured and comprehended through instructional quality frameworks and their components. One such framework is the Mathematical Quality of Instruction (MQI), which is the focus of this study due to its content-specific perspective, comprehensiveness, and curricular-aligned practices. The MQI is supported by validity and reliability studies, has an online training provision for raters, and requires Mathematical Knowledge for Teaching (MKT). These aspects improve the framework's accuracy and make it useful for examining statistics instruction within the broader context of mathematics instruction. The literature on the MQI is also discussed, which includes potential factors that affect instructional quality in mathematics, such as MKT. The study aims to gain insights into the quality of statistics instruction using the MQI framework to inform practice and improve student learning outcomes.

Overall, the literature review chapter comprehensively covers statistical literacy, reasoning, and thinking, highlighting the investigative approach's importance in teaching, and learning statistics and the role of statistical problem-solving in developing statistical skills. In addition, it provides insights into changes made in the Turkish mathematics curricula to incorporate statistical processes. Effective teaching practices are crucial for positive learning outcomes in mathematics and statistics education, and teachers play a significant role in this regard. Understanding the similarities and differences between teaching practices in these fields is necessary to provide high-quality instruction. Thus, measuring and comprehending teaching practices is essential to create opportunities for effective student learning. Considering the aim of the study, the literature discusses statistics-specific practices and their place in combination with mathematical practices in the school context. All in all, the literature review shows that while there has been a wealth of research on statistics education in recent years, there is still a lack of understanding of how statistical practices are integrated into mathematics instruction. Specifically, the quality of statistics instruction within the

broader context of mathematics instruction remains under investigation. Therefore, the current study will explore mathematics teachers' statistical practices in the classroom, using “Transcript Coding Protocol Using GAISE Framework” and the Mathematical Quality of Instruction framework to fill this gap in the literature and provide insights into effective statistics instruction.

The next section presents the method used to explore mathematics teachers' statistical practices in the classroom. This approach addresses research questions and fills the gap in the literature on integrating statistical practices into mathematics instruction. The method utilizes the “Transcript Coding Protocol Using GAISE Framework” and the Mathematical Quality of Instruction framework to measure teaching practices, comprehend the findings, and provide insights into effective statistics instruction. The following section outlines the research design, participants, data collection, and analysis to achieve the study's purposes.

CHAPTER 3

METHODOLOGY

This study aims to examine mathematics teachers' statistical practices in the classroom and the quality of their statistics instruction relating to their broader mathematics instruction. A qualitative case study design was employed, which allowed for an in-depth examination of the teaching practices within the real-life context of the classrooms. The methodology section of this study includes the following subtitles: research design, participants and the context of the study, data collection, and data analysis. The research design outlines the qualitative case study approach and its key features. At the same time, the participants' section describes the middle school mathematics teachers who participated in the study and the context of the study. The data collection section details the methods used to collect data, including observation, field notes, interviews, and classroom artifacts. The data analysis section outlines the strategies used to analyze and interpret the data collected. Overall, this methodology section aims to describe the methods used in this study and show the credibility of the findings.

3.1. Research Design

This research examines statistical teaching practices in terms of statistical content covered and how statistical investigations are incorporated into middle school mathematics instruction and evaluates the Mathematical Quality of Instruction (MQI) in these classrooms. A qualitative research design is necessary because it involves “an interpretive and naturalistic approach to the world” (Denzin & Lincoln, 2017, p.43) and allows for a “complex and detailed understanding of the issue” as well as “an exploration of the context and settings” (Creswell, 2013, p.41). In addition, qualitative research is appropriate for empowering individuals, developing theories, explaining

causal theories or models, and identifying issues with quantitative analysis (Creswell, 2013). Thus, qualitative research should be conducted in natural settings.

The case study is a commonly discussed qualitative research method in the literature (Creswell, 2007; 2013; Merriam, 2009; Stake, 2000; Yin, 2003; 2018). Yin (2018) define the case study as “an empirical method that investigates a contemporary phenomenon in depth and within its real-world context, particularly when the boundaries between the phenomenon and context are unclear” (p.45). Similarly, Creswell and Poth (2018) define the case study as a qualitative method that “explores a real-life, contemporary bounded system or multiple bounded systems over time through detailed, in-depth data collection from multiple sources (e.g., observations, interviews, audiovisual material, and documents and reports,” to provide an in-depth description of cases or case themes (p.153). Both definitions emphasize the importance of real-life context and a bounded system in the case study research method.

A qualitative case research design is selected to examine statistical teaching practices in real-life classroom contexts. This approach allows for a thorough examination and exploration of the teaching practices of two seventh-grade mathematics teachers, providing a detailed understanding of statistical teaching practices in the context of mathematics education. This approach is supported by the literature on qualitative research methods, emphasizing the importance of in-depth exploration and examination of real-life contexts (Creswell & Poth, 2018; Denzin & Lincoln, 2017; Merriam, 2009; Stake, 2000; Yin, 2018).

Depending on the research aim, case studies can be categorized as explanatory, exploratory, or descriptive (Yin, 2003). Explanatory case studies verify a hypothesis or explain a phenomenon, while exploratory case studies generate new insights or explore new topics. Descriptive case studies, like the one employed in this study, describe and explain the characteristics and features of a particular case or phenomenon within its context (Yin, 2003). This research can be categorized as a descriptive case study, as it aims to describe and explain the teaching practices within the context of the classrooms under examination.

Furthermore, case studies can be classified according to the type of case being studied. Stake (2000) proposes three types of cases: intrinsic, instrumental, and collective. Intrinsic case studies aim to understand a single case in depth, while instrumental case studies use a single case to illustrate a broader point or theory. Collective case studies examine multiple cases to compare and contrast experiences, perspectives, and processes (Stake, 2000). Yin (2013) introduces a four-part classification system for case studies based on whether they involve single or multiple cases and whether they use a holistic or embedded design. Embedded studies have subunits within single or multiple cases. The four design types are Type 1 - single-case (holistic) designs, Type 2 - single-case (embedded) designs, Type 3 - multiple-case (holistic) designs, and Type 4 - multiple-case (embedded) designs.

This study can be also classified as a collective case study examining two middle school mathematics teachers' statistics teaching practices, allowing for comparisons and contrasts between their experiences, perspectives, and processes (Creswell, 2007). In addition, it can be described as a multiple case embedded design since it analyzes the quality of instruction and statistical investigations within the context of each individual case, followed by cross-case analysis by comparing “the explanation from another case” (Creswell, 2013; Yin, 1981, p.108).

This research utilizes a descriptive collective case study design and a multiple case embedded design to obtain a comprehensive understanding of how two middle school mathematics teachers cover statistical content and integrate statistical investigations in their instruction and the quality of their statistics teaching practices. The cases under investigation are the statistical teaching practices of Ali and Esra. Through within-case analysis and cross-case analysis of multiple cases, this study aims to identify common themes and patterns and unique features and variations (Creswell, 2007; Yin, 1981;2018). According to Yin (2018), the pathway to conduct a multiple-case study involves using individual cases for literal replication to predict similar results, followed by a theoretical replication to recognize contrasting results with possible reasons. Figure 3.1 depicts the pathway Yin (2018) described for conducting a multiple-case study. I adopt a modified approach in this study.

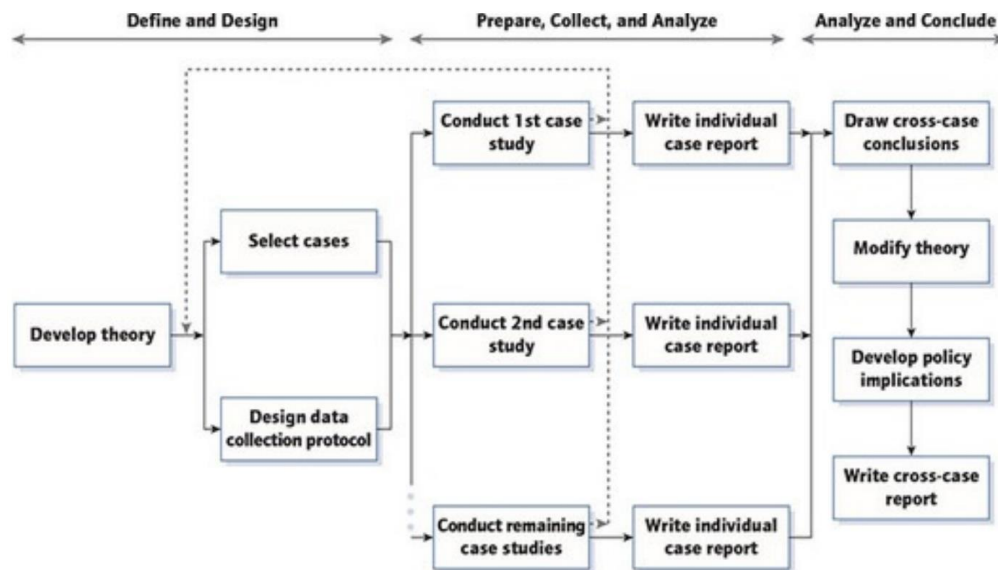


Figure 3.1 Multiple-Case Study Procedures

Note: Case study research and applications: Design and methods (6th Ed.)(p.95), by R. K. Yin, 2018, Sage Publications.

To begin with, I identify and select already existing theoretical frameworks to guide my study instead of developing a theory. These established theoretical frameworks provide a foundation for understanding key teaching practices relevant to my study in terms of discipline-related practices as the elements of the statistical problem-solving process and instructional quality. Using these existing theories, I analyze and interpret the data collected from the multiple cases I study.

In terms of literal replication, I examine each individual case as the teaching practices of each middle school mathematics teacher to gain a comprehensive understanding of the statistical teaching practices within the context of mathematics instruction. For the theoretical replication, I compare and contrast the findings from the individual cases to recognize if there are any differences between the teaching practices of the two teachers and explore the possible reasons behind these variations.

After analyzing each individual case and cross-case analysis, I write individual case reports and cross-case reports. While my study does not involve modifying existing theories, as they are already well-established, I explore the implications of my findings for the practice and further research.

Overall, by following Yin (2018)'s pathway, this study aims to understand the teaching practices of two middle school mathematics teachers in relation to statistical investigations and the quality of statistics instruction within a broader context of mathematics. To achieve this purpose, the background information of the selected cases, including the teachers and their teaching context, will be described in detail. This will provide a deeper understanding of the observed teaching practices and their context.

3.2. Participants and the Context of The Study

According to Creswell and Poth (2018), a case study is a qualitative method that provides a comprehensive and detailed description of cases and their context. Thus, this section provides detailed information about the participants, such as starting with the selection process, a description of the participants, including their background information, views on teaching and learning mathematics and statistics, understanding of statistical content, and reports on their typical instruction. Furthermore, the study's context, including the observed classroom context and teachers' use of curricular materials, is presented to provide a broader perspective for understanding the current study.

3.2.1. Selection of the Participants

Purposive sampling is a type of sampling strategy in which the researcher intentionally selects specific individuals or groups to participate in the study based on specific characteristics or criteria (Patton, 2002). In this study, purposive sampling is used to select two middle school mathematics teachers who met specific criteria, including teaching 7th grade, having an undergraduate degree in a middle school mathematics education program, having at most 12 years of experience in teaching, and working at the current school for at least two years.

First, I choose to study mathematics teachers teaching 7th grade because the statistical content in the 7th-grade mathematics curriculum is the most comprehensive compared to other grade levels. Until the end of 7th grade, students learn about all types of graphs used in middle school and measures of central tendency and spread. The emphasis is

placed on all steps of statistical investigations and the complete cycle by the end of 7th grade.

The second criterion of having at most 12 years of experience in teaching was chosen to ensure that the selected teachers had experience learning to teach and teaching with the reform-based curriculum, which has been in use in Turkey since 2005-2006 and emphasizes the development of skills such as problem-solving, reasoning, communication, and representation. The criterion for teachers to work at the same school for at least two years fosters a sense of familiarity and comfort with the learning environment.

Also, Creswell (2013) also stated that purposeful sampling is one sampling method that can be used in a case study through different perspectives like “purposeful maximal sampling” or just selecting “ordinary cases, accessible cases, or unusual cases” (p.100). Most qualitative research selects the maximum variation sampling method to “maximize differences at the beginning of the study” (Creswell & Poth, 2018, p.224) for different perspectives. However, this study focuses on ordinary and accessible cases, also called convenience cases, that “represent sites or individuals from which the researcher can access and easily collect data” (Creswell & Poth, 2018, p.224). The participants are not selected haphazardly but meet all the selection criteria and could provide relevant and valuable data for the study. In addition, they are accessible to the researcher.

I identified middle school mathematics teachers in the Sarıyer district who met the inclusion criteria and were accessible to the researcher with limited resources. Eleven teachers were contacted through gatekeepers, and four 7th-grade teachers from three different schools agreed to participate. These four teachers, who comprised seven classrooms, fully participated in the research process and met the selection criteria. However, they were relatively homogeneous, with just one teacher had an opportunity to use technology in their teaching practices, which could have created some variation. Unfortunately, a small fire in the school just before starting classroom observations resulted in electricity problems, preventing the use of technology. Therefore, the four teachers' practices remained homogeneous.

For this study, I selected two middle school mathematics teachers, Ali and Esra, as participants. I chose them because their teaching practices were observable in multiple classrooms, whereas the remaining two teachers only had one 7th-grade classroom each. Although I observed all their classes, I focused my analysis on one specific class for each teacher. This allowed me to examine how their teaching approaches varied within each teacher across different classrooms and identify any patterns or trends in the data. Understanding these differences was crucial to investigate how the teachers' practices differed in various learning environments. In addition, by analyzing the selected classroom contexts, I could better make sense of the teaching practices observed in the chosen classroom context. The next section will describe the data collection tools used to understand the cases.

3.2.2. Data Collection Tools for the Context of the Study

This section outlines the data collection tools used to understand the participants and the study's context comprehensively. Interviews were conducted to gather information about the participants' backgrounds, views on teaching and learning mathematics and statistics, understanding of statistical content, and typical instruction. Moreover, student tests provided insight into the classroom environment. Interviews with teachers were also conducted to explore their use of curricular materials and provide further context for the study. It should be noted that I will present the main data collection instruments of this study in the next section under the title of Data Collection.

3.2.2.1. Interviews

I conducted interviews with teachers to explore their characteristics, views on teaching and learning mathematics and statistics, and their understanding of statistical content (see Appendix B for the interviews). Before the observation period, two interviews were conducted to gather information about the teachers' backgrounds and their views on teaching and learning mathematics and statistics. The first interview focused on general background information and their views on teaching and learning mathematics and lasted about 16 minutes. The second interview centered on the statistics-specific background, their views on teaching and learning statistics, and the relationship between mathematics and statistics and lasted about 18 minutes.

A Statistical Content Interview was conducted to assess the teachers' understanding of statistical content taught in middle school and lasted about 19 minutes. This interview included questions about average concepts, range, outliers, data (adapted from Jacobbe, 2007), and data collection, handling, and analysis. After the observation period ended, additional interviews were conducted to reflect on the semester and the teachers' statistics teaching process and lasted about 18 minutes. All interviews were conducted in a quiet environment, mostly in the teachers' room during the teachers' free periods. All interviews were transcribed and analyzed through open coding. The data gathered from these interviews provided a deeper understanding of the cases studied.

3.2.2.2. Student Tests

Two Likert-type scales were used to measure students' attitudes toward mathematics and statistics. The reliability of these scales was established for the current study. After completing the topic/unit, a statistical literacy test was administered to students, and a holistic rubric was used to code the test results. Data collected from students in the classrooms were used to provide context for the cases but were not included in the findings section. However, these data inform the discussion of the findings and provide a more in-depth understanding of the cases. The psychometric characteristics of the student tests will be provided in the next section. Descriptive analyses were conducted on student tests to describe the teaching context.

3.2.2.2.1. Attitudes towards Mathematics

To measure the attitudes of students towards mathematics in the classroom context, the “Your Thoughts about Mathematics” scale (Matematikle İlgili Düşünceleriniz) was administered. This Likert scale consists of 20 items, with 12 positive items reflecting positive attitudes toward mathematics and eight negative items reflecting negative attitudes towards mathematics (reverse-coded). The scale is based on Erol's (1989) attitudes towards mathematics scale and was shortened by Nazlıçıçek and Erktin (2002). Nazlıçıçek and Erktin (2002) analyzed the factor structure and reliability of the current version of the scale, reporting a Cronbach's alpha of .84. For

the current study, I conducted a reliability analysis and found the Cronbach's alpha to be .90.

3.2.2.2.2. Attitudes towards Statistics

The Attitudes toward Statistics Questionnaire (ATSQ) used in this study was adapted from the Probability Attitude Survey Questionnaire by Bulut (1994) by changing the word 'probability' to 'statistics' (Yolcu, 2012). The questionnaire consisted of 27 items on a Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree), with 16 positive items reflecting positive attitudes towards statistics and 11 negative items reflecting negative attitudes towards statistics (reverse-coded). Yolcu (2012) established the validity and reliability of the ATSQ, reporting a Cronbach's alpha of .94 and validity in terms of content and construct validity. Similarly, for the current study, the reliability analysis showed a Cronbach's alpha of .94, consistent with the original study by Yolcu (2012).

3.2.2.2.3. Statistical Literacy Test

Yolcu (2012) developed the Statistical Literacy Test (SLT) to assess 8th-grade students' statistical literacy using Watson's (1997) three-tiered framework. The test comprises 17 questions with subquestions covering sample, average, graphs and tables, probability, inference, and variation, based on the Turkish curriculum context. Yolcu (2012) reported that the test had good content validity and reliability (Cronbach's $\alpha = .72$). However, due to changes in the curriculum, I decided to remove one question on standard deviation, which had not been introduced to the students. To give equal weight to the remaining 16 questions, I proportioned and evaluated the test overall as 16 points. I then conducted a reliability analysis, which yielded a Cronbach's α of .69, consistent with Yolcu's (2012) findings.

3.2.3. Teachers and Teaching Context

This section describes the participants and teaching context gathered through teacher interviews and student tests. It includes information on the participants' backgrounds, perspectives on teaching and learning mathematics and statistics, understanding of

statistical content, and typical instruction methods based on teacher interviews. In addition, the teaching context is described through the descriptive analysis of student test results and the background of students observed in the classrooms.

3.2.3.1. Teacher Ali

This section provides detailed information about Ali, the pseudonym, including his background and experiences with mathematics teaching practices and statistics education. It explores his perspectives on the relationship between mathematics and statistics and evidence of his understanding of statistical content. Also, the section describes the observed classroom context and curriculum to provide context for the study. By describing Ali, this section aims to contribute to a better understanding of the case of Ali's teaching practices.

3.2.3.1.1. The Background Information

Ali was a mathematics teacher who had been teaching for six years. He graduated from a public university in Ankara in 2012. Ali has worked at this school for 2.5 years, a religious middle school with gender-based classrooms for the 7th and 8th grades. He taught three 7th-grade and one 6th-grade class for 27 hours per week. He reported that the average number of students in those classes was 26-27. He also described the socio-economic status of his students as mid-level. In addition to teaching mathematics, he was also teaching an elective course about technology and design. However, he stated that he could not use technology while teaching mathematics or the elective course even if he wanted to use it. Thus, the school did not have technological opportunities. He stated that he had experience teaching mathematics with technology in the previous school because the school had smart boards in the classrooms.

According to Ali, he took content-specific courses such as geometry, statistics, and probability during his undergraduate studies. However, he did not remember the exact names of the courses related to his pre-service education. As for his pedagogical content knowledge, he reported that he had limited exposure to it, with only one course covering teaching mathematics and another course on general pedagogy related to

evaluation and assessment. Ali mentioned that he did not receive any in-service education related to teaching mathematics and statistics after becoming a teacher.

3.2.3.1.2. Ali's views on Learning and Teaching Mathematics

When asked about his resources for learning to teach a specific topic in mathematics, Ali stated that he had comprehensive knowledge of middle school topics. However, he also referred to teaching mathematics books to introduce new concepts. In addition, he mentioned that he often used the resources available on EBA, an online educational platform used in Turkish schools. He also stated that he tried to use real-life contexts depending on the topic and the student's interests. For example, he used football context for boys and soap operas or YouTube videos for girls. However, he acknowledged that not all subjects were appropriate for the real-life context and provided the example of fractions as a difficult topic to teach with real-life examples.

Regarding classroom discussion, he stated that he generally took students' opinions and did not scold them; he used humor when the students gave answers since students liked the humor. Then when asked whether there was a student-student interaction; he expressed that teacher-student interaction was present; he did not prefer the student-student interaction because especially boys' classroom would be noisy. The teacher also reported that the students were unfamiliar with working in groups because Ali did not use this method. During the classroom discussion, the students could say their ideas, but there was no direct interaction among students; the interaction was generally between the teacher and the student. The classroom observation was parallel with what the teacher reported about the interaction type used in the classrooms. The students can comment on other students' ideas, but the student-teacher interaction was more dominant during teaching statistics.

3.2.3.1.3. Ali's Views on Learning and Teaching Statistics.

Ali expressed his interest in statistics during the interview, particularly sports statistics. He mentioned that he enjoyed watching sports and wanted to learn how to calculate statistics related to football and basketball. He stated that the statistics he learned in undergraduate studies were theoretical and did not include sports statistics. As a

learner, he felt more confident about measures of central tendency but struggled with statistics that involved formulas, such as calculating z-scores. In daily life, he used arithmetic mean frequently to calculate the average wage or the number of people who visited a place. Overall, Ali's experiences with statistics were limited to undergraduate courses in probability and statistics. However, he expressed interest in learning more about the topic, mainly related to real-life contexts such as sports.

Furthermore, Ali did not take any courses on teaching statistics. When asked how he learned to teach statistics, he mentioned that the concepts he taught were not complex, as he knew them from his undergraduate education. However, he tried to give real-life examples as students could not make sense of abstract topics. He looked at the questions which were not complicated and tried to connect them with real-life scenarios. He reported that he learned how to teach statistics through experience over the years.

Ali expressed that he enjoyed teaching statistics, especially arithmetic mean, and felt confident teaching all middle school topics. However, he stated that he would not feel confident teaching high school statistics as there are more advanced concepts. He believed that statistics were taught for later academic purposes and to make sense of the news people follow, such as economy news.

3.2.3.1.4. Ali's Views on the Relationship between Mathematics and Statistics

Ali believed that mathematics and statistics were not distinct disciplines, as mathematics was utilized in applying statistics. He emphasized that statistical analysis required mathematical or quantitative reasoning, much like solving a problem with fractions. He gave an example of calculating the arithmetic mean, where one must comprehend the problem, analyze the information given, and perform numerical computations.

3.2.3.1.5. Ali's Understanding of Statistical Content

Statistical Content Interview was administered to understand the teacher's conceptualization of the statistical terms he covered in the 7th-grade mathematics

curriculum (Jacobbe, 2007). First, Ali described statistics as research about a problem situation. After defining statistics with participants, questions to ask the participants for data collection, data, and the interpretations of those data with particular methods, he concluded statistics as a process:

Researcher: You think it is a process.

Ali: Maybe it is not correct. Statistics might be the interpretation part at the end, but I think it is a process from determining the problem [situation] to finalizing and getting a result. I think as a whole.

Researcher: Problem

Ali: Problem or research question

Researcher: ... question, then you included data collection

Ali: I think also analyzing [data] and interpreting [data]

He defined data as the comments or numerical values gathered through sampling by asking questions. He explained the data collection process starting from identifying the research problem, choosing an appropriate sample, and asking proper questions. He mentioned frequency tables and graphs as ways of handling data and emphasized the importance of selecting appropriate graphs and measures of central tendency for data analysis. He stated that data analysis involves interpreting the information on frequency tables or graphs and that there were other methods besides graphs, such as measures of central tendency. He focused on the importance of using appropriate graphs and measures of central tendency.

Later, he defined average as a measure of central tendency and noted that students often associate it with the word “ortalama” in Turkish, which means “approximately even.” He explained that the average represents a data set's typical value or most noticeable characteristic.

Furthermore, Ali provided the computational meaning of arithmetic mean as “The type of average found by adding the data given to us and dividing by how many data there are is called the arithmetic mean.” In addition to using the arithmetic mean formula, he provided another way of finding the arithmetic mean; he mentioned leveling interpretation of the arithmetic mean. Then he conceptually defined the arithmetic mean as a typical value.

Regarding the concept of the median, he provided computational meaning as “the number in the middle when we sort the data ascending or descending order.” He mentioned the differences in finding median or middle points when the number of data was even or odd. He pointed out the conceptual meaning as similar to the arithmetic mean. When there were extreme values, he stated that the median was better because extreme values did not affect the median; the arithmetic mean might have given wrong ideas about a group.

Then he defined mode as the most repeated “number” in a data set. He stated that it was the most typical characteristic of the group studied. Finding the most favorite color, finding the most selected elective course, and finding the most sold shoe size were some examples he gave after he defined mode. His definition only covered numerical data, not categorical data. However, his examples contained categorical data.

Last, Ali's definition of the range was correct, but his definition of an outlier was incomplete and incorrect, as it did not include extreme values. He defined range as the difference between a data set's maximum and minimum value, which was accurate. However, his definition of an outlier as only the minimum and maximum values in a data set did not fully encompass the concept of outliers, which were extreme values significantly different from other values in the data set.

3.2.3.1.6. Ali's Typical Instruction

In the first interview with Ali, he described his typical teaching route as how he re-designed his teaching route not to spend time with nonmathematical work:

I was writing the questions on the blackboard in the past, and sometimes I distributed some worksheets. However, I was starting to think that this was a waste of time. Writing a problem on the blackboard and waiting for students to solve the problem took at least 10 minutes. Therefore, I wanted students to buy additional sources/books, and I started to solve problems in those books. I did not dictate the definitions in mathematics for students to take notes. When necessary, I want students to write brief notes with their own words on the book... Again I am distributing worksheets as well if there are suitable question types...

[Ali's First Interview 30.04.18]

Ali described how he changed his teaching style to prevent long transitions between problems (e.g., waiting for the teacher to write the question and copying the solution). Besides, he described how he shaped his instruction to prevent students from being off task:

In the past, I was writing a question [on board], and then 10 minutes passed; since there was only one question, I think that the children who grasped it very quickly are waiting for a while. Because he solves the problem that takes 30 seconds for the child, I think it is negative. So I changed it [the method of teaching]. I [now] tell [students] that we are solving the third question in the book; the one who finishes can continue to take the test and evaluate the gap in the meantime. I think the book has such an advantage... At least, that is how he uses the [waiting] time.

[Ali's First Interview 30.04.18]

Furthermore, Ali also shared that he frequently monitored his students during different tasks, and if he noticed an important point, he called on them to listen. This approach prevented students from wasting time waiting for others and ensured that those who followed his instructions worked through the problems effectively.

3.2.3.2. Teacher Esra

This section provides detailed information on Esra, the pseudonym, including her background and experiences with mathematics teaching practices and statistics education. It explores her views on teaching and learning mathematics and statistics and evidence of her understanding of statistical content. In addition, the section describes the observed classroom context and curriculum to provide context for the study. By describing Esra, this section aims to contribute to a better understanding of the case of Esra's teaching practices.

3.2.3.2.1. Background Information

Esra was a mathematics teacher who had been teaching for 11 years. She graduated from a public university in İstanbul in 2006. Esra has been working at her current school for five years. Previously, she worked as a paid/substitute teacher for one year and spent two to three years in the other two schools. She was currently teaching at a public middle school in the county town. The school did not have technological

facilities. She had four classes: two 7th-grade classrooms and two 8th-grade classrooms for 28 hours in total per week. She reported having about 24-28 students in 7th grade and 30 in 8th grade. She described the socio-economic background of her students as good compared to the other schools she worked.

When asked about her pre-service education, Esra could not recall the exact names of her courses. However, she remembered taking courses related to pedagogical content knowledge, such as teaching mathematics and science, assessment in mathematics teaching, and computer-assisted mathematics teaching. She did not take content-specific courses for teaching subjects such as geometry or statistics but mentioned taking a classroom management course. In informal conversations, she mentioned taking calculus courses, differential equations, and matrices. She also reported not having taken any in-service education related to teaching mathematics and statistics.

3.2.3.2.2. Esra's Views on Learning and Teaching Mathematics

Regarding her teaching practices for mathematics, Esra reported that she would first lecture on the content, followed by working through related tasks. When appropriate, she encouraged students to explore the concept before solving related tasks. She did not frequently use real-life examples unless the tasks themselves included them. Regarding classroom discussions, Esra said that students would discuss and come to a conclusion after exploring a topic. She stated that there were no specific rules for the classroom, only general expectations such as not talking unnecessarily and paying attention during lectures. She encouraged students to ask questions and would answer any questions they had. She also welcomed requests from students to solve problems and expressed her enjoyment in learning about their studies. Observations support that the teacher gave lectures, solved tasks, and allowed students to ask questions if they had confusion or difficulty with the concepts or tasks.

3.2.3.2.3. Esra's Views on Learning and Teaching Statistics

Esra stated that she took a statistics course but not a pedagogical one, but it was a content course. She did not mention the course details, she just remembered the parts

she taught as a teacher. She stated that she does not like statistics. As a learner, she stated that central tendency and variability concepts were the concepts she learned during her undergraduate studies. She was introduced to the concepts late. She also pointed out that she was good at graphs. She did not have difficulty understanding the topics. She did not feel confident about standard deviation as a learner.

Esra revealed that she had not taken a course in teaching statistics but learned the subject independently. She relied on course textbooks and additional books as resources to teach effectively. In summary, she described her process of learning to teach statistics as self-taught using course materials and supplementary books:

I pretend to tell myself first. How would I describe it if I told it to myself or a particular student? I first look at the subject as a summary; I take notes on certain basic things that are certain to be learned; I summarize aside, I explain the summary information to the student. For example, I do not want to talk about a topic for a long time, I give essential information that needs to be known, and it makes more sense to go through the rest of the examples. It is better to see and discover new information in every question rather than drowning it in information.

[Esra's First Interview- 02.04.18]

She described her approach to teaching statistics as traditional, focused on mathematical instruction without using hands-on activities. However, she indicated that she might begin with a questionnaire to collect information if relevant to the topic and then present the data in a table. According to her, the starting point for studying statistics is often a questionnaire to gather information.

When asked if she enjoyed teaching statistics, she stated that she did not like the subject but found the concepts of mode, median, and variability to be the most enjoyable. Due to the calculations involved, she felt confident in constructing graphs, especially circle graphs. Bar graphs were considered dull, as she preferred calculations.

Moreover, she did not give any specific concept she felt was insufficient but stated some points in the questions she knew the answer to. However, she did not know how to explain. She also believed they taught statistics because statistics was vital for students to think and interpret. She focused on the sense-making process. For example,

she stated that they explored the data's meaning after collecting it. The numbers were not meaningful if they could not interpret and make sense of them. When the students asked why they learned statistics, she gave an example that included numbers as collected data and explored what the data told them with statistics.

3.2.3.2.4. Esra's Views on the Relationship between Mathematics and Statistics

According to Esra, mathematics is at the heart of statistics because, without mathematics, there would be no statistics. She stated, "Mathematics is the root; statistics is a branch in the mathematics tree."

3.2.3.2.5. Esra's Understanding of Statistical Content

Esra's statistical knowledge for teaching was explored based on the Statistical Content Interview (adapted from Jacobbe, 2007). I aimed to understand the teacher's conceptualization of the statistical terms he covered in the 7th-grade mathematics curriculum.

Firstly, Esra described statistics as an extensive topic. She then described statistics as a topic consisting of collecting, enumerating, and representing data with graphs to facilitate interpreting the information. She reported that they could conduct analysis, reach a conclusion, and interpret all obtained information. She did not state statistics as a process or a cycle, but her answer to what statistics represented was a part of the statistical problem-solving process. She mentioned three steps in the statistical problem-solving process: collecting, analyzing, and interpreting data.

Then she described data as the information, the collected information in specific. She did not explain the types of data. When asked how and on whom to collect data, she stated that one could use different ways to collect data— surveys, interviews, and screening. Then she mentioned a sample but did not provide detailed explanations. However, she stated they chose a sample and collected information just before the data collection. The data handling section stated that the data were organized through tables. Then the tables were shown in graphs. She stated that they could use line, bar, and circle graphs. She correctly provided the purposes of the graphs. She especially

stated that not all table [data] could be used in all graphs. Just her explanation of the line graph's purpose was not complete. She stated that one could use a line graph for the increases or decreases. She stated that the data analysis was about reaching a conclusion and interpreting what was done. She also stated that constructing tables and graphs were used for conducting analysis and reaching a conclusion.

When asked about some statistical concepts she taught until 7th grade, Esra explained the average by an example with the number of questions solved in context. Then she listed three measures of central tendency as average types. Then she clarified the purpose of why one used the average with the same context from the class as follows:

For example, I solved questions for five days. The number of questions I solve every day is different from each other. If you say how many questions you solved in a day, I have to give five different answers. Instead of giving five different answers, we are already using the average by saying that I have solved almost as many questions every day.

[Esra's Statistical Content Interview 10.06.2018]

Esra recognized the need for a single value to represent the number of questions solved over five days. However, her language indicated that she perceived the average as an estimated value rather than a representative or typical value within the data set.

When asked what the arithmetic mean was and how to find it in different ways [if any], Esra stated that "arithmetic mean is directly a calculation. It is a number found by adding all data numbers and dividing by the counts of the numbers". She further added that there was only one formula and was taught one way, but she could calculate practically by estimating the differences among numbers. Moreover, she answered why using the arithmetic mean as similar purpose to the average "to tell how many questions were solved in a day.... Each day, I solved 12.5 questions on average". Then she defined arithmetic mean procedurally, not conceptually in general. Her conceptualizations of arithmetic mean were not detailed or precise. She also stated that the arithmetic mean was the most reliable for students because they could calculate it. She believed that the calculations could give definite results. This conceptualization was not about the reliability of the arithmetic mean itself but the use of formulas in

general. She also pointed out that arithmetic mean could be used when the numbers in the data sets were close to each other.

Furthermore, Esra provided both the computational and conceptual meaning of median. She stated that the median is found by sorting the data and identifying the middle value, and it gives a general idea of the average of the data set. However, the median provides less detailed information as compared to the arithmetic mean.

She defined mode as the most repeated “number.” She stated that the mode was also one type of average. She described where to use the mode as follows: “We are given such a data set that a number of repeats so much. For example, there are ten numbers there. If six of ten numbers are a repeating number, its average is likely to be close to that six numbers, so it makes sense to use the mode there.” Her explanation did not include the type of data. She focused on the closeness of the mode to the real center found by the arithmetic mean.

Last, the teacher defined the range as the difference between the minimum and maximum values in a data set and mentioned its importance in deciding which central tendency measure to use and determining the number interval. However, her definition of outliers was limited to only the minimum and maximum values in a data set, not including extreme values. She gave an example related to the number of children who visited a hospital. The minimum number was three, and the maximum number was ten. She stated that the minimum and maximum values could be found by looking at the outliers.

3.2.3.2.6. Esra’s Typical Instruction

In the initial interview, Esra described her typical lesson as follows:

Mathematics teaching varies from subject to subject. If there is information that needs to be given; that the student cannot discover, I will give the information beforehand. I do the lecture, and then I solve many questions about it. However, if it is a piece of information that the child can discover, I would explore it beforehand, raise a question mark, explore that subject together, and

then solve it together. However, I give importance to problem-solving. I want them to see as many questions as possible related to that topic.

[Esra's First Interview 02.04.18]

She described her teaching route as introducing the topic first and solving many questions. Her descriptions and teaching moves were based on direct instruction with Initiation-Response-Evaluation (IRE) instructional sequence. She also added that she prefers to keep the topic manageable; she gave the main concepts and expected students to discover concepts through examples.

3.2.3.3. Context of the Study

This section provides context for the study by describing the observed classroom settings in which Ali and Esra's teaching practices were analyzed. It includes descriptive statistics of student background variables, academic achievement, attitudes toward mathematics and statistics, and statistical literacy. Additionally, teachers' statements regarding their use of curricular materials are included.

3.2.3.3.1. Teaching Context in the Case of Ali's Teaching Practices

In the teaching context of Ali's case, I present his observed classroom context, including student background and descriptive statistics of their attitudes towards mathematics and statistics, academic achievement based on end-of-year grade point average scores, and statistical literacy. In addition, I will provide information on the curricular resources or materials that Ali used in his teaching practices.

3.2.3.3.1.1. Ali's Observed Classroom Context

One of three seventh-grade classrooms was used for examining Ali's teaching practices. The classroom consisted of 28 male students. As for the classroom organization, the arrangement of the desks was straight lines. There were three lines with seven desks in each line. The students were sitting in pairs; no individual desks were present. In addition, the teacher's desk was in front of the classroom next to the

blackboard. Table 3.1 shows the demographic background of the students in the classroom.

Table 3.1

Demographic Background of 7A Class

Demographics		N	%
Mother's education level	Primary school	5	18.52
	Middle school	6	22.22
	High school	6	22.22
	Undergraduate	6	22.22
	Graduate	1	3.70
	Do not know	3	11.11
Father's education level	Primary school	2	7.41
	Middle school	7	25.92
	High school	6	22.22
	Undergraduate	7	25.92
	Graduate	1	3.70
	Do not know	4	14.81
Number of siblings	1	2	7.41
	2	9	33.33
	3	13	48.15
	4	3	10.7
	5 and more	0	0
Number of books at home	0-10	3	11.11
	11-24	1	3.70
	25-100	9	33.33
	100-250	7	25.92
	250 and above	7	25.92
Computer at home	Yes	21	77.78
	No	6	22.22
Internet at home	Yes	23	85.19
	No	4	14.81

Table 3.1 consists of 27 students' background information because one student did not fill out demographic information parts of the instruments. Parents' education level, including both mother and fathers, the number of siblings, the number of books, and the presence of computer and internet connection at home were collected from students. The reports of the students showed that the education level of the mother and father were similar to each other; most parents did not attend or complete college education.

Furthermore, most students (n=13, 48.15%) were three siblings. The number of books reported was above 100 for 50% of the students. Most students had computer and internet connection at home (n=21, 77.78%; n=23, 85.19%, respectively). In addition, I collected students' mother and father job information. Twenty-two students filled in the mother job section; 20 of 22 reported that their mothers were not working; five did not state their mother's jobs. Two mothers were working as a teacher and a secretary. According to 23 students, the fathers worked at different types of jobs (e.g., driver, teacher, carpenter, cook, fisher, and accountant).

Table 3.2 shows the descriptive statistics of 7A students' academic achievement, attitudes towards mathematics, statistics, and statistical literacy variables. The arithmetic mean of students' end-of-year grade point averages described as academic achievement was 73.19 out of 100, with a standard deviation of 15.23.

Table 3.2

The Descriptive Statistics for Classroom Context Variables

Variable	N	M	SD	Range	Min-Max
1.Academic achievement	28	73.19	15.23	41.47-99.50	0-100
2.Attitudes toward mathematics	21	77.76	13.74	41-94	20-100
3.Attitudes toward statistics	22	94.32	17.41	71-134	27-135
4.Statistical literacy	24	6.56	2.74	2.67-11.17	0-16

Furthermore, the arithmetic means (with standard deviations in parentheses) of 7A students' attitudes toward mathematics and statistics were 77.76 out of 100 (13.74) and 94.32 out of 135 (17.41), respectively. These scores indicate that, on average, students had a positive attitude toward both mathematics and statistics. The arithmetic mean of students' statistical literacy was 6.56 out of 16, with a standard deviation of 2.74. This suggests that, on average, students had a moderate level of statistical literacy, with some variability in their scores after the statistics unit.

3.2.3.3.1.2. Curricular Materials in Ali's Teaching Context

Ali stated that he has not been using the textbooks from MEB as curricular materials. He preferred to follow an additional supplementary textbook. He expected all his students to buy the book. He did not use notebooks for students to take notes. He wanted students to take notes on the supplementary textbook. Furthermore, he stated that he was distributing additional tests to show and solve different problems on the board. His statements were parallel to the observed classroom.

3.2.3.3.2. Teaching Context of Esra's Case

In the teaching context of Esra's case, I present her observed classroom context, including student background and descriptive statistics of their attitudes towards mathematics and statistics, academic achievement based on end-of-year grade point average scores, and statistical literacy. In addition, I will provide information on the curricular resources or materials that Esra used in her teaching practices.

3.2.3.3.2.1. The Observed Classroom Context

One of two seventh-grade classrooms was used for examining Esra's teaching practices. It included 14 girls and ten boys. The organization of the classroom setting was traditional. The teacher's desk was next to the blackboard, and student desks were lined back-to-back with three lines. The students were sitting in pairs; no individual desks were present.

The demographic background of the students in the classroom is presented in Table 3.3. This table included 22 students' background information because two female students did not fill out the background information sections in the instruments. As the background information, parents' education level, including both mother and fathers, the number of siblings, the number of books, and the presence of a computer and internet connection at home were collected from students. The results show that the distributions of both parents' educational backgrounds were similar to each other; none of the parents completed the graduate studies. Most parents graduated from high school ($n=8$, 36.36% for mothers, $n=9$, 40.91% for fathers).

Table 3.3*Demographic Background of 7D Class*

Demographics		N	%
Mother's Education Level	Primary school	3	13.64
	Middle school	6	27.27
	High school	8	36.36
	Undergraduate	2	9.09
	Graduate	0	0
	Do not know	3	13.63
Father's Education Level	Primary school	2	9.09
	Middle school	6	27.27
	High school	9	40.91
	Undergraduate	3	13.63
	Graduate	0	0
	Do not know	2	9.09
Number of siblings	1	4	18.18
	2	7	31.82
	3	6	27.27
	4	4	18.18
	5 and more	1	4.55
Number of books at home	0-10	2	9.09
	11-24	8	36.36
	25-100	8	36.36
	100-250	1	4.55
	250 and above	3	13.63
Computers at home	Yes	18	81.82
	No	4	18.18
Internet at home	Yes	18	81.82
	No	4	18.18

Furthermore, the arithmetic means (with standard deviations in parentheses) of 7A students' attitudes toward mathematics and statistics were 77.76 out of 100 (13.74) and 94.32 out of 135 (17.41), respectively. These scores indicate that, on average, students had a positive attitude toward both mathematics and statistics. The arithmetic mean of students' statistical literacy was 6.56 out of 16, with a standard deviation of 2.74. This suggests that, on average, students had a moderate level of statistical literacy, with some variability in their scores.

Table 3.4 shows the descriptive statistics of 7D students' academic achievement, attitudes toward mathematics, attitudes toward statistics, and statistical literacy variables.

Table 3.4*The Descriptive Statistics and Correlations for Classroom Context Variables*

Scale	N	M	SD	Range	Min-Max
1. Academic Achievement	24	63.63	22.77	20.42-99.75	0-100
2. Attitudes toward Mathematics	23	74.91	15.35	43-97	20-100
3. Attitudes toward Statistics	20	102.30	19.86	44-125	27-135
4. Statistical Literacy	21	7.06	3.21	2.5-13.17	0-16

The arithmetic mean of students' end-of-year grade point averages was 63.63 out of 100, with a standard deviation of 22.77. Furthermore, the arithmetic means (with standard deviations in parenthesis) of 7D students' attitudes toward mathematics and statistics were 74.91 out of 100 (15.35) and 102.30 out of 135 (19.86), respectively. These scores show that, on average, the students had a positive attitude toward both mathematics and statistics. Also, the arithmetic mean of 7D students' statistical literacy was 7.05 out of 16, with a standard deviation of 3.21. This shows that, on average, the students had a moderate level of statistical literacy, with some variability in their individual scores.

3.2.3.3.2.2. Curricular Materials

Esra stated that she has been using an additional textbook for 7th grade. However, she was also using the textbook from MEB for 8th grade as curricular materials. In classroom observations, I observed that she used some tasks from the 7th grade MEB textbook, but she did not want students to bring their books; she took photocopies of the pages she used and distributed them to the students.

In this section, I describe the cases of Ali and Esra to highlight the complex nature of teaching practices in statistics and mathematics education. In the following section, I will comprehensively explain the data collection and analysis procedures used in this study. This includes the methods utilized for data collection, such as observations,

video recordings, and field notes, along with steps taken to ensure the validity and reliability of the data. Furthermore, I will discuss the data analysis process, including coding procedures, analysis tools, and strategies implemented to ensure the credibility of the findings.

3.3. Data Collection

This section provides an overview of the data collection procedures employed in the study. It outlines the steps taken to gain entry into the field, the researcher's role, and the data collection tools, such as classroom observations, videos, field notes, and student documents. In addition, informal interviews were conducted with the participating teachers to understand the teachers' perspectives on their observed teaching practices.

3.3.1. Data Collection Procedure

After obtaining ethical approval from the board, I identified middle school mathematics teachers in the Sarıyer district who met the criteria for inclusion in the study. I contacted 11 teachers, and four 7th-grade teachers from three schools agreed to participate. These four teachers, comprising a total of seven classrooms, fully participated in the research process and met the selection criteria.

Before the primary data collection period, I conducted a pilot study where I observed a 6th-grade mathematics teacher for a few hours. I informally analyzed their lessons using statistical investigations and the Mathematical Quality of Instruction framework. In addition to observations and video recordings, I interviewed the teacher using the same questions in the main study. I received feedback from the teacher and revised the interview questions accordingly.

The main study took place during the second semester of the 2017-2018 academic year, from March to June 2018. Data collection activities were conducted during this period, including observations, interviews, and student assessments. The observation period for the “circle” topic began in late March, as this topic was included in the annual plan for 7th-grade mathematics courses. I observed the “statistics” unit in April,

which consisted of approximately 15 40-minute lessons—the observation period for Ali's classrooms differed slightly as the order of units was changed, and observations took place in May. Table 3.5 presents the time schedule of the data collection process with a brief reference to the activity.

Table 3.5

Timeline of Data Collection Process

Timeline	Activity
Feb 2017	Obtained ethical approval
Sep-Dec 2017	Identified eligible teachers
Jan 2018	Contacted 11 teachers; received 4 positive responses
Feb 2018	Contacted 4 teachers from 3 schools for participation
Late Feb 2018	Conducted pilot study; observed and interviewed a six-grade teacher
Mar-Jun 2018	Main data collection period: observations, interviews, assessments
Apr 2018	Updated ethical approval
Late Mar 2018	Observed “circle” topic in classrooms
Apr 2018	Observed “statistics” unit in classrooms
May 2018	Observed units in Ali's classrooms
Mar-Jun 2018	Conducted interviews with participating teachers
Mar-Jun 2018	Conducted student assessments

Table 3.6 shows the weekly program for classroom observations. The first part of the table shows the period of March 26 to April 30 for observing three teachers’ teaching practices. The second part of the table shows the period of April 23 to May 15 for Ali teacher’s teaching practices. The details of when the observation is done during the week and days for the main classes of Ali and Esra can be seen in the Table 3.6.

I conducted a total of 104 hours of classroom observations with four teachers, including Esra, Teacher X, Teacher Y, and Ali. This comprised 13 hours for two of Esra's classrooms, 21 hours for Teacher X, 15 hours for Teacher Y, and 14 hours for three of Ali's classrooms. Each lesson was video-recorded and accompanied by descriptive and reflective field notes. These observations, along with the field notes and video recordings, served as the primary data sources for the study.

Table 3.6*Weekly Observation Period*

March 26 to April 30			
Day	Observation Periods		
Monday	08:00-08:40 (Esra -other class)	11:40-13:40 (Teacher X)	12:10-13:40 (Esra-main class)
Tuesday	09:40-11:10 (Esra-main class)	13:00-13:40 (Teacher X)	
Wednesday	08:00-09:30 (Teacher Y)	10:50-12:20 (Teacher X)	12:10-13:40 (Esra -other class)
Thursday	08:00-09:30 (Teacher Y)	10:50-12:20 (Teacher X)	
Friday	08:00-09:30 (Esra -other class)	09:40-11:10 (Esra-main class)	13:00-13:40 (Teacher Y)
April 23 to May 15			
Day	Observation Periods		
Monday	12:10-12:50 (Ali- other class 1)	13:40-14:20 (Ali-main class)	
Tuesday	08:50-09:30 (Ali- other class 1)	10:30-12:50 (Ali- other class 2)	13:40-15:05 (Ali-main class)
Wednesday	11:20-12:00 (Ali- other class 2)	12:10- 12:50 (Ali- other class 1)	
Thursday	10:30-12:00 (Ali- other class 2)		
Friday	08:50-10:30 (Ali-main class)	11:20-12:50 (Ali- other class 1)	

Additionally, interviews with the teachers were conducted both before and after the study. Student data was collected through attitude tests, achievement tests, and background information sheets to understand the cases under investigation thoroughly. The cases of the study were Ali and Esra's teaching practices, as discussed in the participant selection section. The following sections will provide a more detailed description of gaining entry to the field, including the ethical considerations and establishing rapport, and the role of the researcher.

3.3.2. Gaining Entry to the Field and Ethical Considerations

I followed several steps in order to gain access to the field. First, I applied to the Human Research Ethical Community board at Middle East Technical University and obtained permission for ethical considerations (See Appendix A). During the application process, I provided the community with detailed information about the study and filled

out the detailed application form. After receiving permission from the board, I applied for permission from the Ministry of National Education and gained access to the field.

Once I had obtained the necessary permissions for the study, I used gatekeepers to access teachers. According to Bogdan and Biklen (2007), it is necessary to present certain explanations to the gatekeepers, including the reasons for choosing the site/participants, the procedures that I would follow during the study, the duration of the study, the impact of my presence on the site, and the potential benefits for the participants and gatekeepers. Therefore, I contacted the gatekeeper and responded to these questions.

To establish rapport with the participants, I contacted and met with each participant individually before the start of the study. Creswell and Poth (2018) emphasized that a qualitative researcher needs to inform the participants so that no deception should be present about the nature of the study. Therefore, I explained the general purpose of the study and outlined the specific details of the data collection procedures (such as interviews, observations, and achievement tests). During the first meeting, I emphasized the importance of confidentiality, as the identities of the participants, the site, and the collected data would be kept confidential during any part of the study, including reporting the research findings. For this reason, the anonymity of the participants was considered, and I presented all participants with pseudonyms in the current document. I also informed them that they could drop out of the study at any point of the study.

After the explanations, the teachers agreed on participating to the study with the information given and filled out the consent forms. Teachers and school administrators informed the parents of the students about the study and obtained consent from them. Later, verbal consent from students was also obtained. Especially with the presence of the camera during the observations, students and teachers were informed that the video recordings could be watched by me and my supervisors and a small portion of the videos with a second-rater and would not be shared by any other people. Therefore, if they did not want to be in the video, the position of the camera would be positioned accordingly. None of the teachers and their students stated that they did not want to be

video recorded. As a result, informed consent from both teachers and students were obtained before the administration of the study.

During the study, I recognized the importance of establishing a positive relationship with the participating teachers to facilitate open communication and a comfortable research environment. To accomplish this, I spent prolonged periods of time with the teachers, observing classes and engaging in informal conversations during breaks or after school. I was responsive to their needs, assisting with tasks such as photocopying, solving mathematics questions, or suggesting mathematics activities upon request. By being present and attentive, I built trust and rapport with the teachers, creating a supportive and collaborative research environment. Additionally, I openly discussed my role as a researcher with the teachers. In the next section, I will elaborate on my role in the current study.

3.3.3. The Role of the Researcher

Denzin and Lincoln (2017) described the role of qualitative researchers as being in the field as natural settings of the case being studied and to understand and interpret the data collected from the field through the meanings of participating people. Therefore, as the qualitative researcher, I spent extensive time in the field to make sense of the meanings of the participating teachers. Establishing rapport with the participants helped me to understand and interpret the data by considering their meanings as well.

My role in this study is a nonparticipant observer during the classroom observations. This means that I did not actively participate in the activities or lessons being observed, but rather, I recorded my observations and collected data from a distance. My role is to objectively document the events and interactions taking place in the classroom without influencing them in any way. By remaining a nonparticipant observer, I maintained a neutral perspective and gathered unbiased data. Also, I took reflective field notes in order to prevent researcher bias (see details on the field notes section of data collection tools).

As stated earlier, I aimed to establish a positive relationship with the teachers by responding to their needs and building trust. I offered help during my waiting time or breaks since I requested their time to clarify issues or conduct interviews. I assured them that I would not interfere with their teaching and would not confirm or disconfirm any activities they brought to the classroom during observation. However, some teachers still sought my approval or feedback on their teaching, and it was difficult for me to remain neutral. I reminded them of the purpose of the study. Also, I promised to give feedback at the end of the semester. Despite these reminders, it was challenging to stay neutral during such interactions, but I noted each instance and how I responded. For example, one teacher prepared a task on the blackboard just before I began observing the classroom on the statistics unit. She scolded the children when she saw that the blackboard was empty, but later acknowledged that she wanted to have a good introduction because I was there. Although my presence might have had certain influences on the teachers, spending prolonged time in the field helped to decrease such incidents. After the study ended, we talked off-record about their teaching and mathematics and statistics education in general.

3.3.4. The Data Collection Tools

The data collection is described as extensive so that multiple sources of information are needed to be used in a case study (Creswell, 2013). Possible data collection tools that could be used in a case study are observations, including direct observation or participant observation, interviews, and documents, including physical artifacts or archival records (Creswell, 2013; Yin, 2018). In this section, I will describe the data collection tools utilized in the study. The primary data sources are the observations which are video recordings of the observations accompanied by field notes and classroom artifacts, and interviews (see Appendix B)

As discussed earlier, I also used data collection tools such as interviews and student tests to gather additional information on the cases and contextual factors relevant to the study. The results of these data sources are presented in detail in the “Participants and The Context of the Study” section of the method chapter.

3.3.4.1. Observations

Observation is used in qualitative research to understand the case within its natural environment (Creswell & Poth, 2018; Yin, 2018). Creswell and Poth (2018) stated that the observations might be carried out with “the physical setting, participants, activities, conversations, and your own behaviors” (p.232). Furthermore, they provided procedural steps for preparing and conducting observations. In Figure 3.2, the steps include selecting and gaining access to the field, choosing who and what to observe through gatekeepers, introducing oneself as an observer, using an observational protocol, taking rich and thick descriptive and reflective notes, and slowly withdrawing from the field when the observation is completed.

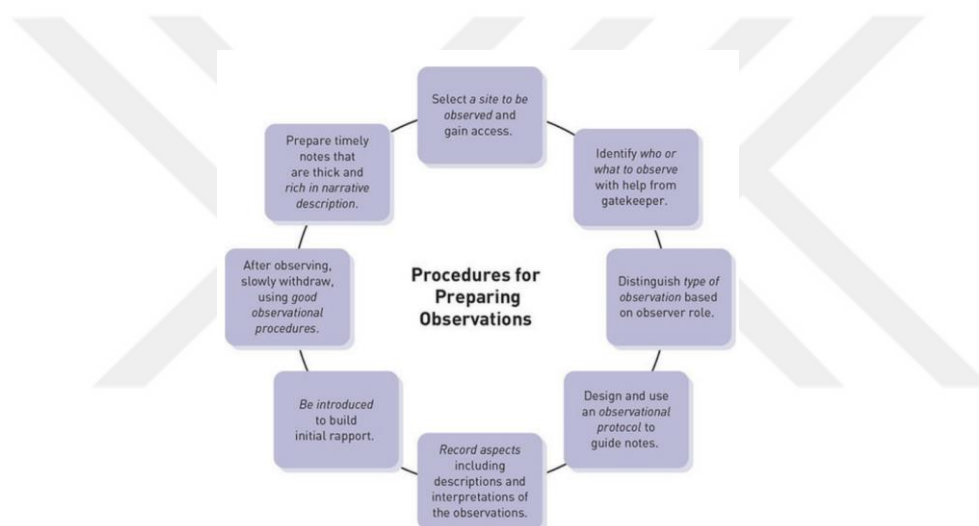


Figure 3.2 Procedures for Preparing Observations

Note: Qualitative inquiry and research design: Choosing among five approaches (p.235), by J. W., Creswell and C. N. Poth, 2018, Sage Publications.

Bostic et al. (2021) stated that classroom observations allow “direct measures of instructional quality” (p.6). That is to say, how teaching practices are implemented or what is going on in the classroom could be understood with classroom observations. In the current study, the purpose of the study is to understand seventh-grade mathematics teachers’ teaching practices in statistics instruction in terms of the statistical problem-solving process and the quality of instruction. Therefore, I observed statistics instruction during the whole statistics unit consisting of about fifteen 40-minute lessons. This involved observing content-student-teacher interactions in small

group and individual settings and whole class interactions. The lessons were videotaped, with a camera placed at the back of the classroom to minimize distraction for students and teachers.

Certain steps were followed in the observation process (Creswell & Poth, 2018; Fraenkel & Wallen, 2014). After obtaining the necessary permissions and gaining entry to the field through gatekeepers, I established rapport with the participants and defined my role as a non-participant observer. I observed the teacher from the back of the classroom as an outsider. Both the teachers and students were aware of my presence. The purpose of the study was generally explained to the teachers and students, and steps were taken to minimize observer effect by starting observations before the statistics unit began, during the “Circle” topic, which lasted for two weeks. Furthermore, I slowly left the field even after I completed the observations due to collecting other types of data and being accessible to both teachers and students. In the following section, I briefly explain field notes and video recordings as data collection tools during observations.

3.3.4.1.1. Field Notes

To minimize the potential observer bias, qualitative researchers need to take notes on what they “hear, see, experience, and think in the course of collecting and reflecting on the data” (Bogdan & Biklen, 2007, p.119). Therefore, I took detailed field notes during the observations, including descriptive and reflective elements. These notes provided a comprehensive record of the data and allowed for an in-depth analysis. To further mitigate observer bias, I also included my own perspectives and interpretations of the observations in the research report. This transparency allowed for a more thorough understanding of the data and the impact of the observer on the study. In addition to taking detailed field notes and being transparent about my biases, it was important to remain reflexive throughout the study and consider the potential influence of my assumptions and beliefs on the observations and interpretations. Engaging in reflexive practice, such as regularly reflecting on my role and biases, helped minimize observer bias in the study.

3.3.4.1.2. Video Recordings

Classroom videos are powerful tools for conducting classroom observation studies in recent years (Borko, 2016). They capture the complexity of the events happening in real classrooms for ready-to-use for later analysis that “focus ex-post facto on various aspects of the material under investigation” (Janík et al., 2009, p.7). That is to say, they captured the interactions among teachers, students, and content and provide rich data to analyze the teaching practices. Classroom videos provide researchers to analyze the data in hand multiple times to capture the richness of the complexity of the classrooms (Borko, 2016). In addition, reviewing and analyzing the data multiple times provide greater accuracy and reliability of the findings and also mitigate observer bias with an objective record of what happened in the classroom. In the current study, all 104 observations were video-recorded, and within all, 26 classroom videos were analyzed based on observational tools (see detailed explanations in the Data Analysis section).

3.3.4.3. Interviews

To gain insight into the teaching practices, I conducted pre-observation and post-observation interviews adapted from Miller (2000). The pre-observation interviews covered the teachers' lesson planning process and their expectations for student interaction. The post-observation interviews covered the teachers' thoughts on how the lesson went, whether the students learned the concept, and if they participated. However, the interviews were unsustainable in their original form as teachers were often unavailable during breaks, and some were occupied with solving math questions for students. Daily interviews were also offered, but teachers were not consistently available before and after school.

Therefore, I divided the interviews into three-time frames that were aligned with the objectives covered in the study to avoid disrupting the research process. These interviews were referred to as planning and evaluation interviews, and the aim was to gain insight into how the teachers planned and approached topics such as graphs, averages, and graph transformations. During the planning interviews, I asked the teachers about their teaching plans for graphs, as well as any possible errors or misconceptions that students might have. During the evaluation interviews, I asked if

the lesson went according to plan, if any errors or misconceptions occurred, and if the students learned the topic. I also followed up with additional questions to encourage elaboration on their responses. These interviews provided valuable information to supplement the analysis of teaching practices based on classroom videos.

3.3.4.4. Classroom artifacts

Before beginning classroom observations, I requested teachers to inform me about the curriculum materials they were using and obtained those books to follow the instructions. I also requested them to provide any additional resources they planned to use for each observed lesson. When teachers brought classroom photocopies of tests or activities, they usually gave me a copy. These artifacts helped me to take field notes more easily, as I could see in-hand what the students and teachers were working on. Additionally, teachers provided me with scans of the exam sheets of students taken after the unit was covered. This artifact enabled me to analyze the number and types of questions teachers asked on the statistics unit and how the students answered them.

In this section, I reviewed the data collection tools used in this study, which included observations with field notes and video recordings, interviews, and classroom artifacts. In the following section, I will describe the data analysis process employed during the study.

3.4. Data Analyses

In this case study, I analyzed multiple data sources, such as classroom videos with field notes, interviews, and classroom artifacts, to investigate the nature of statistics instruction in 7th-grade statistics classrooms. I described the statistical instructional practices of 7th-grade mathematics teachers in terms of the statistical problem-solving process and the instructional quality to examine how these practices aligned with the objectives in the curriculum, the Transcript Coding Protocol Using GAISE Framework, and the Mathematical Quality of Instruction Framework.

Data analysis in qualitative research is described as the meaning-making process of collected data (Flick, 2014). The qualitative data analysis starts with the data collection

process since the data collection, analysis, and reporting of the findings are interrelated (Creswell, 2007). Flick (2014) stated that the meaning-making process includes a rough and detailed examination of the data throughout the research process. Yin (2016) stated that qualitative data analysis does not have a strict structure but rather follows a series of interrelated steps, including compiling, disassembling, re-assembling, interpreting, and concluding. Similarly, Creswell (2007) outlined a specific set of steps for the analysis of qualitative data, including data management, reading and memoing, describing, classifying, and interpreting.

I followed the five-phased cycle of Yin (2016) for each case in this study. In the first phase, I compiled data by organizing data sources in order. This included organizing the videos and field notes, watching the classroom videos, summarizing the lessons, and re-reading the field notes. I transcribed the full six videotaped lessons and specific clips from the remaining lessons. This compiling phase was the same for both purposes of the study.

3.4.1. Schematic Diagram

To explore the nature of statistics instruction in terms of the statistical investigation process and the instructional quality, I prepared a list of statistical tasks, called the Schematic Diagram, by matching classroom artifacts and resources with segments of videos and field notes. This was used to identify the tasks used in the classroom. I then began the disassembling process, during which I reorganized the tasks based on their purpose. I rearranged the diagram according to Litke and Hill's (2020) definition of "episode" (2020). According to Litke and Hill (2020), episodes, which I defined as statistical tasks, refer to "the class' interactions with each unique set of data encountered by the class" (p.487). They provided examples of what they meant by a unique set of data: "if a class found the median and mode of a set of numbers and then collected and graphed data on their favorite pets, this counted as two episodes." (p.487). Based on this definition of uniqueness, I finalized the schematic diagram. All in all, the schematic diagram helped further understand the tasks' distinctiveness and purpose within the statistics instruction (see Figure 3.3)

	Topic/Concept	Task	Time period	Date
1	Introduction to statistics ----- Data collection ----- Graphs Circle graph	He asked students what they remembered from the previous year about graphs ----- He handed out 12 colored papers, the same-colored ones next to each other. I want you to form a circle. Do you think he started the circle graph like this? 2 yellow, 3 green, 4 blue, and 3 pink. ----- First, he placed the numbers. Do I know the angles here? What do we see here? What does it mean to me that this is 90 degrees? What does it mean to me that these 3 people correspond to 90 degrees? "One-fourth of a circle graph" What if I drew it as a bar graph? ... Do you think there is something I see in this one that I don't see in this one? Does the bar graph tell you anything when you look at it? Circle graph? Can you divide the circle by 180 if there are 180 people instead of 12? 	00:00-12:20	26.04

Figure 3.3 An Example from a Teacher's Schematic Diagram

To illustrate, in the schematic diagram, I first assigned codes to the statistical tasks which had unique data sets. For example, if the task involved a graphical representation, I assigned a “graph” code to the task. Then, I assigned a more specific code to the task, indicating the type of graphical representation used (e.g., a bar graph, a line graph, or a circle graph). Yin (2016) referred to these codes as Level 1 (types of graphs) and Level 2 (graph) codes. The main topics and contents covered in this analysis included various types of graphs (e.g., circle, bar, line), transformations of graphs, and measures of central tendency (e.g., mean, median, mode). With this coding process, I answered the research question of what statistical content mathematics teachers cover in their statistics instruction by providing the frequencies and percentages of each code appearing during statistical task analysis.

3.4.2. Video Analysis

After creating the schematic diagram, I employed the use of Computer-Assisted Qualitative Data Analysis (CAQDAS) for data analysis in this study. I chose to use video analysis because of my interest in classroom research and the desire to capture important information regarding the classroom environment and instructional aspects (Kuckartz & Rädiker, 2019). Knoblauch et al. (2014) defined “videography” as the

analysis of videos recorded by researchers in natural settings to explore social interactions. This method was appropriate for my research as the coding protocols and instruments aimed to examine the social interactions around the content in the classrooms, including the interactions between the teacher, students, and content. I recorded the lesson videos in a natural setting and then analyzed the recordings.

The video analysis process divided video clips into episodes and segments according to two frameworks. Therefore, I reorganized the video data in CAQDAS by constructing a document system that placed the videos in chronological order by the teacher. I had previously gathered information on which lessons contained similar topics for different teachers through the compiling and disassembling phases. I then followed a specific route to code all videotaped lessons focusing on the same topic or content. In some cases, teachers covered more than one topic in a single lesson, so I watched the video twice while following the coding route. I coded all lessons for both purposes alone twice, and also conducted a re-coding process by retracing the first codes after a period of time. The coding of classroom videos was considered the re-assembling phase of qualitative research. To further ensure the accuracy of the coding process, I followed a specific order in which I coded all of one teacher's classroom videos before moving on to the next teacher.

During the re-assembling phase of the study, I used the procedural method of protocol coding for data analysis in qualitative research. I used two different coding protocols to explore the different purposes of the study. These observational protocols contained pre-established codes developed by other researchers (Saldaña, 2013). This process is also referred to as standardized video analysis (Knoblauch et al., 2014). Both protocols were based on established theories and had already been tested in the field (Hill et al., 2008; 2014 for the MQI; Litke & Hill, 2020 for Transcript Coding Protocol Using GAISE Framework). The pre-established coding systems included specific themes and codes, which were described in detail. Additionally, both protocols utilized magnitude coding, as defined by Saldaña (2013) as the use of existing codes “to indicate its intensity, frequency, direction, presence, or evaluative content” (p. 265). For the Transcript Coding Protocol Using GAISE Framework, magnitude coding was used to indicate the presence and intensity of statistical problem solving processes. Similarly,

the MQI used magnitude coding to show the quality codes' presence, intensity, or direction.

The data was analyzed using content analysis based on two frameworks, and both within-case analysis and cross-case analysis were conducted due to the presence of multiple cases. Creswell (2013) suggested that “the typical format is to first provide a detailed description of each case and themes within the case, known as within-case analysis” (p.101). Thus, I provided a detailed description of each case using the Transcript Coding Protocol Using GAISE Framework and the MQI. Following the within-case analysis, I conducted “a thematic analysis across the cases” based on the frameworks used in the study, and I compared both cases (Creswell, 2013, p.101). Additionally, I provided examples from the qualitative data to describe the statistical instructional practices of the teachers using the Transcript Coding Protocol Using GAISE Framework and the MQI.

In the subsequent sections, I will describe the process of coding classroom videos in terms of statistical investigations using the Transcript Coding Protocol Using the GAISE Framework developed by Litke and Hill (2020). I will also describe the exploration of the nature of statistics instruction based on the quality of instruction using the Mathematical Quality of Instruction (MQI) 4-Point Version developed and revised by Hill et al. (2014).

3.4.2.1. Analyzing Classroom Videos in terms of Statistical Problem-Solving Process

To analyze statistics instruction in relation to statistical investigations, I utilized the Transcript Coding Protocol with the GAISE Framework, created by Litke and Hill (2020). This coding protocol was designed to comprehend statistical instructional practices according to GAISE standards, aligning with statistical investigations' components. The framework includes four elements with three categories. The four elements are Formulation of Questions, Data Collection, Data Analysis, and Interpretation. The categories are as follows:

- (1) the element was not observed;
 (2) the element was observed to some extent but without the depth specified in GAISE (e.g., finding measures of central tendency and range but not using that information to address distribution or spread); and
 (3) the element was more fully aligned with GAISE expectations, such that it was implemented with a deeper focus on statistical ideas or concepts (Litke & Hill, 2020, p.487)

Table 3.7 displays the levels for each element with detailed descriptions. The categories in the table are referred to as levels. For instance, at Level 1 for Formulation of Questions, there is no formulation of questions or statistical questions, according to the table's explanation. At Level 2, there is some observation of the formulation of a statistical question posed by the teacher or text. At Level 3, the element is more fully aligned with GAISE expectations, and students are expected to generate statistical questions.

Table 3.7

Transcript Coding Protocol Using GAISE Framework

	Level 1	Level 2	Level 3
Formulation of questions	Work with data is not rooted in a question	Statistical question posed by teacher or text	Students generate statistical question
Data collection	Data are provided	Data are collected and recorded systematically	Data are collected and recorded systematically and the class discusses issues such as variability, sampling, measurement, etc.
Data analysis	No data analysis occurs	Data analysis focuses on creating a display and/or calculating measures of central tendency, range; questions are descriptive (e.g., "How many students like chocolate ice cream?")	Data analysis focuses on creating a display and/or calculating measures of central tendency, range; and the class uses the display to describe a distribution, compare groups, or observe a relationship between variables
Interpretation	No interpretation occurs	Interpretation uses the graph or data to answer "within" data questions (e.g., "Which flavor of ice cream should we order for our class party? What type of ice cream does our class like the best?")	Interpretation uses the graph or data to make predictions, compare sample to population, compare across samples, discuss variability in the data, recognize limitation of scope of inference, or generalize

Note.—GAISE = *Guidelines for Assessment and Instruction in Statistics Education*.

Note: From "Exploring Central Tendencies: The Teaching of Data And Statistics in Elementary Mathematics Classrooms" E. Litke and by H. C. Hill, 2018, The Elementary School Journal, p. 487 , (<https://doi.org/10.1086/707007>)

A similar structure applies to the other elements. In Data Collection, Level 1 implies that data are already available, while in Level 2, data are collected and systematically recorded. At Level 3, in addition to Level 2, the class discusses issues such as variability, sampling, and measurement. In Data Analysis, the element is not present at Level 1, meaning no data analysis occurs. At Level 2, data analysis focuses on creating tables and graphs, calculating average scores and range, and answering descriptive questions. At Level 3, further emphasis is placed on Level 2 and using graphs to describe a distribution or compare groups or observe the relationship between variables. In Interpretation, the element is not present, meaning no interpretation occurs, similar to data analysis. Level 2 involves using data or graphs to answer within-data questions. At Level 3, using graphs and data to make predictions, compare samples to populations, discuss variability and recognition of the scope of inference, and generalize the findings.

I utilized the schematic diagram prepared during the disassembling phase of the data analysis process to assist in this process, as it included episodes with statistical tasks dealing with unique data sets. Therefore, I used “episodes” as the unit of analysis in order to answer the research question, “How do mathematics teachers approach the four-step model of statistical problem-solving processes in their instruction?” To answer this question, I coded a total of 78 episodes, including 47 episodes from Ali's statistics instruction and 31 episodes from Esra's statistics instruction.

In the first episode, the teacher discussed the concept of circle graphs with the students. The teacher asked the students if they remembered any graphs from last year and if they know what a circle graph was. After some discussion, the teacher asked 12 students to draw a color from a set of paper prepared by the teacher. Then, the teacher asked the students to form a circle and stand side by side based on the color they drew. Next, the teacher drew an empty circle on the blackboard and asked how many people there were in total. The students responded that there were 12 people. The teacher noted the number of each color. Then, the teacher asked the students how to divide the circle on the blackboard based on the number of people with each color. The students suggested dividing the circle based on the largest number, but one suggested dividing it based on the threes because threes refer to the quarter of 12. The teacher then divided

the circle according to the number of students with each color from the answers of the students. After constructing the circle graph based on the data, the teacher also constructed a bar graph based on the data they had and asked students about the differences among these graphs (The first 7.5-minute segment ended here). Figure 3.4 demonstrates how I coded a video lesson using the protocol in CAQDAS. The statistical task presented in Figure 3.4 was coded as the first episode of this lesson.

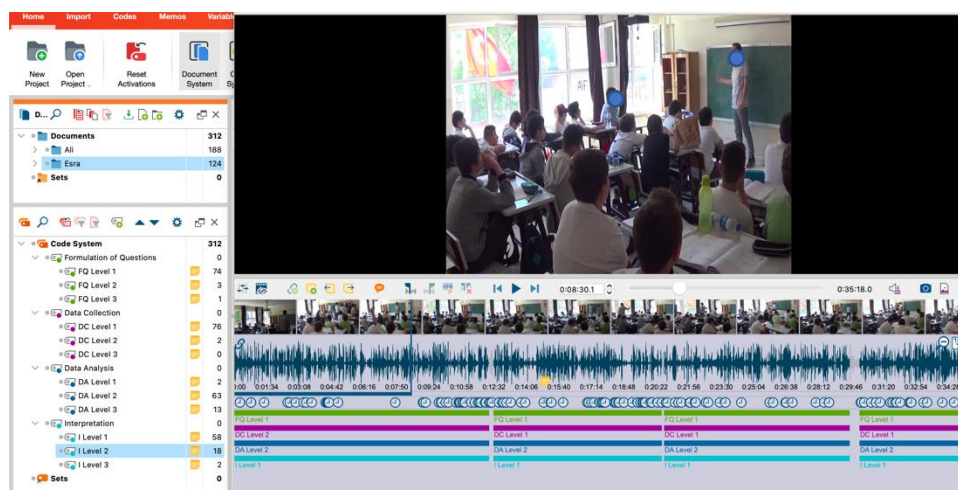


Figure 3.4 Coding Example of Episodes in a Lesson on Transcript Coding Protocol
Using GAISE Framework

The students asked questions to clarify their understanding of the graphs. At one point, they discussed the advantages of using a circle graph, which allowed them to see the entire picture and how many shares each part had. The teacher then moved on to discussing the angles in the circle graph and asked the students how they found them. One student suggested dividing the circle into equal parts and coloring some of them, and the teacher confirmed that this was a valid method. They also discussed how to find the angles when the total number of items was not divisible by 360. Then they discussed finding the angles in a specific example. The students suggested different methods, such as dividing the circle into quarters and halves or using proportions to find the angles. The teacher encouraged the students to explore different methods and think critically about their approaches. They also discussed the importance of using correct terminology and concepts, such as proportions and ratios (See Appendix C for

the transcripts of this episode). Table 3.8 displays the coding decisions made using the Transcript Coding Protocol Using GAISE Framework for the episode described above.

Table 3.8

Coding Decisions for the Episode with the Transcript Coding Protocol using GAISE Framework

Elements	Categories	Descriptions
Formulation of Questions	Level 1	Statistical investigative questions were not posed.
Data Collection	Level 2	The teacher collected data by asking 12 students to draw a color from a set of paper, then recorded the number of each color. However, there was no discussion of variability, sampling, or measurement, which are typically addressed in Level 3 of data collection.
Data Analysis	Level 2	The teacher and students created graphs (specifically a circle graph and a bar graph) to represent the collected data. The focus was mainly on graph construction rather than using graphs to describe a distribution or compare groups or observe the relationship between variables, which are typically addressed in Level 3 of data analysis. Therefore, the data analysis in the first episode can be categorized as Level 2.
Interpretation	Level 1	No interpretation occurs and no discussion of answering questions or making inferences based on the collected data or graphs.

The episode was coded according to the elements in the statistical problem-solving process described in the GAISE documents: Formulating Questions, Collecting Data, Analyzing Data, and Interpreting Results, with three categories for each element. When an element was not observed, such as in this episode where Formulating Questions and Interpreting Results were not discussed, it was scored as Level 1. When an element was observed to some extent, as in the case of Data Collection and Data Analysis in this episode, it was scored as Level 2. Since this episode did not contain any element that fully aligned with GAISE expectations, none of the elements were scored as Level 3. Specific descriptions of the episode's coding decisions can be found in Table 3.8.

In the following section, I will describe the Mathematical Quality of Instruction (MQI) framework as a standardized instrument to assess the quality of instruction in mathematics and statistics.

3.4.2.2. Analyzing Classroom Videos in terms of Mathematical Quality of Instruction (MQI) Framework

In this study, I utilized the Mathematical Quality of Instruction (MQI) framework as a standardized observational instrument to answer the research question, “How does the quality of mathematics teachers' statistics instruction within their broader mathematics instruction?”. As comprehensively discussed in the literature review chapter, the MQI focuses on subject-specific practices in mathematics education and is designed to evaluate “the depth of mathematics offered to students, teachers' response to students' mathematical ideas and misconceptions, teacher errors, and students' participation in mathematical meaning-making and cognitively demanding activities” (Charalambous & Litke, 2018; Litke & Hill, 2020, p.485).

In this study, both mathematical and statistical occurrences were considered, as statistics is a distinct field that is closely related to mathematics, and both fields benefit from each other. The MQI was used for all domains in school mathematics. Still, it may be more appropriate to refer to the codes as mathematical and statistical or mathematics and statistics, given their distinctiveness. Furthermore, I employed the latest version of MQI (Hill, 2014), which comprises segment-level codes and whole lesson codes. I utilized the segment-level codes to analyze the quality of statistics instruction and the “Whole-Lesson Mathematical Quality of Instruction” code to capture the overall quality of the instruction.

The segment-level codes of the MQI encompass five main themes/dimensions: (1) Classroom Work Connected to Mathematics and Statistics, (2) Richness of Mathematics and Statistics, (3) Working with Students and Mathematics and Statistics, (4) Errors and Imprecision, and (5) Student Participation in Meaning-Making and Reasoning (also known as Common Core Aligned Student Practices). Except for Classroom Work Connected to Mathematics and Statistics, the dimensions of the MQI

include segment-level codes that vary in quantity, depending on the theoretical framework of the MQI. Specifically, there are 17 distinct codes assigned to each segment, and four overall segment-level codes, such as Overall Richness of Mathematics and Statistics, are categorized as holistic codes and not a composite of other codes within the dimension (Hill et al., 2014). As a result, a total of 21 segment-level codes are allocated to each segment of a lesson.

Consistent with the guidelines outlined in the Mathematics Quality of Instruction (MQI) instrument (Hill et al., 2014; Hill et al., 2018; Hill & Lovison, 2022), I divided the lessons into 7.5-minute segments for coding purposes. Previous versions of the MQI utilized 5-minute segments (Hill et al., 2008). However, this resulted in a large number of segments that impeded the efficiency of the analysis. Although utilizing 15 to 20-minute segments was considered, it was considered inadequate as it did not clearly define the quality indicators due to the high frequency of mathematical occurrences (LMT, 2011). The latest version of the MQI (Hill et al., 2014) employs 7.5-minute segments, as researchers found this to be an appropriate length that did not result in lost time for scoring (Hill et al., 2018).

In this study, I used the latest version of the scoring unit to score 7.5-minute segments of observed lessons. The average lesson time was 32 minutes and 61 seconds, which resulted in leftover minutes regardless of whether 5 or 7.5-minute segments were used. However, during the first round of observations and video analysis, I found that using 7.5-minute segments provided a more accurate capture of mathematical and statistical occurrences. To score the segments, I used segment-level codes, observing 7.5-minute clips. For the last segment of a lesson that was less than 7.5 minutes, I coded 3- to 4-minute segments, as I observed that mathematical and statistical occurrences often happened in the final minutes. Depending on the lesson's length, the number of segments per videotaped lesson varied, ranging from one to five segments.

The segment-level codes use 4-point scales, ranging from Not Present (1), Low (2), Mid (3), and High (4), with the exception of the “Classroom Work is Connected to Mathematics and Statistics” dimension, which uses a 2-point scale of No (1) and Yes (2). The rating of Not Present (1) means that the segment is not qualified based on the

given codes, while High (4) indicates that the segment is highly qualified for a specific code. In the Errors and Imprecision dimension, Not Present (1) means that the segment is free from errors and imprecision, whereas High (4) indicates that the segment contains a significant amount of errors and imprecision.

The Richness of Mathematics and Statistics dimension of the MQI instrument is designed to evaluate the depth of the mathematical and statistical content being taught in the lesson (Hill et al., 2014). The dimension is divided into two categories: meaning of facts and procedures and key mathematical practices, with a total of seven codes assigned to these categories. The “meaning-oriented” codes focus on the three codes: *Linking Between Representations*, *Explanations*, and *Mathematical and Statistical Sense-Making*. The “practice-oriented” codes focus on three codes: *Multiple Procedures or Solution Methods*, *Patterns and Generalizations*, and *Mathematical and Statistical Language*. The final code in this dimension is the Overall Richness of Mathematics, which assesses the lesson's overall mathematical and statistical depth and complexity.

The Working with Students and Mathematics and Statistics dimension aims to explore to what extent the teacher attends to students' mathematical errors and contributions to address students' difficulties and contributions to the development of the content (Hill et al., 2014). There are three codes in total. *Remediation of Student Errors and Difficulties* and *Teacher Uses Student Mathematical and Statistical Contributions* are two specific codes for this dimension. There is also an overall code called *Overall Working with Students and Mathematics and Statistics*.

The Errors and Imprecision dimension aims to determine the mathematical errors made by the teachers and the imprecision in language and notation while covering the content (Hill et al., 2014). *Mathematical and Statistical Content Errors*, *Imprecision in Language or Notation*, *Lack of Clarity in the Presentation of Mathematical and Statistical Content*, and *Overall Errors and Imprecision* are the codes for the dimension. This dimension only considers the errors not corrected during the segment.

The Common Core Aligned Student Practices dimension is the last dimension for the segment-level codes. Common Core Aligned Practices are based on Common Core State Standards for Mathematics in the USA. Hill et al. (2014) stated that all codes are not directly connected with each standard but are highly related. This dimension aims to determine to what extent students are associated with mathematics and statistics, how they contribute to the mathematical and statistical content, and if they reason, ask questions, and participate in mathematical and statistical practices. *Student Provide Explanations, Student Mathematical and Statistical Questioning, and Reasoning (SMQR), Students Communicate about the Mathematics and Statistics of the Segment, Task Cognitive Demand, Students Work with Contextualized Problems, and Overall Common Core Aligned Student Practices* are the codes for this dimension. In the previous versions of MQI, this dimension is called Student Participation in Meaning-Making and Reasoning. As stated before, I prefer to call this dimension “Student Participation in Meaning-Making and Reasoning” due to the study's context.

After scoring the segment-level codes, I conducted a frequency analysis and presented the results in tables and bar graphs grouped by dimensions and case levels. Additionally, I calculated the mean and standard deviation of the scores. However, while frequency analysis provides a quantitative representation of instructional quality, it may not offer a comprehensive picture of the complex nature of classroom instruction. Thus, to supplement the quantitative data, I also provided qualitative evidence to support the scores assigned to the segments. This approach enabled a more nuanced interpretation of the instructional quality and provided a more holistic understanding of classroom teaching practices.

Figure 3.5 demonstrates how I used the MQI in CAQDAS to code a video lesson, which is the same as the one used in the Transcript Coding Protocol. However, unlike the protocol, the video lesson is divided into 7.5-minute segments, with each segment assigned 21 codes. The color coding of the codes in the video lesson represents their corresponding score levels. Green indicates a high level, yellow represents a mid-level, orange signifies a low level, and red indicates that the code is not present in that segment.

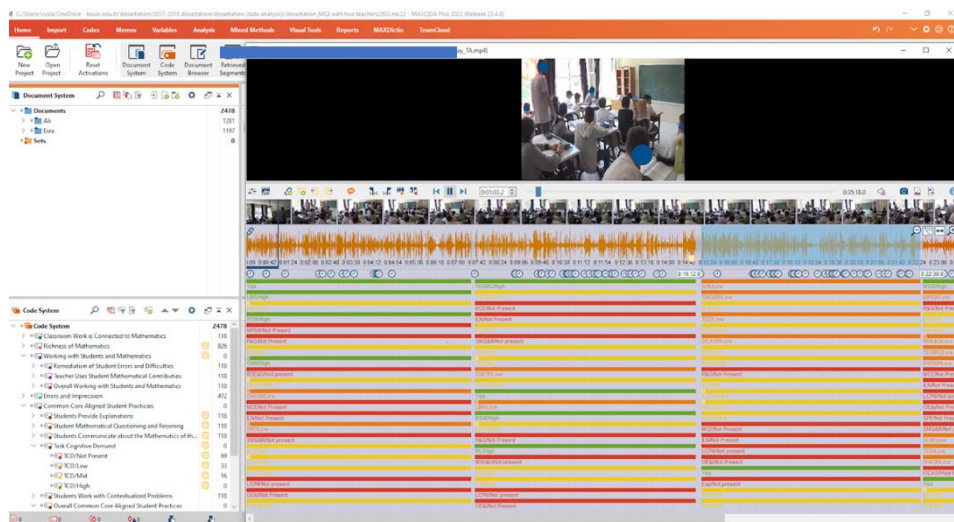


Figure 3.5 Coding Example of Segments in A Lesson on The MQI

In the first segment of the lesson, the teacher discussed the concept of circle graphs with the students. The first segment is a part of the episode described in the previous section. It was noted where to cut out the 7.5-minute segment. As part of this discussion, I assigned scores to the codes in the dimensions of classroom work connected to mathematics and statistics, the richness of mathematics and statistics, working with students and mathematics and statistics, errors and imprecision, and student participation in the meaning-making process. These codes were used to evaluate the quality of the lesson and the extent to which the students were engaged in the mathematical and statistical content. Table 3.9 summarizes the evaluations made for each code in four dimensions.

In addition to segment-level codes, whole lesson codes are assigned to measure the overall quality of a lesson (Hill et. al., 2014). Unlike segment-level codes, the coding unit for whole lesson codes is the entire lesson. These codes summarize the necessary components of a lesson or instruction. To analyze the overall quality of a lesson, I used the last code, Whole-Lesson Mathematical Quality of Instruction, which provides a “summary impression of the quality of instruction for the entire lesson” (NCTE, n.d.). It is important to note that the whole lesson MQI is not the average of the scores of other whole lesson codes. Rather, it is a combination of all MQI dimensions (NCTE, n.d.).

Table 3.9*Example of Scoring Decisions of Codes in MQI Corresponded to a Segment*

Codes	Score	Descriptions
1. Classroom Work is Connected to Mathematics and Statistics	Yes	Focus is on mathematical and statistical content for the majority of the segment
Richness of Mathematics and Statistics		
2. Linking Between Representations	High	The explicit and detailed link between two representations as students' being circled and the circle graph construction on the blackboard
3. Explanations	Mid	Brief two explanations; why to start with three while constructing the graph (math) and why to use a circle graph rather than a bar graph (stats).
4. Mathematical and Statistical Sense-Making	High	Teacher and students focus on the sense-making of circle graphs in both mathematical and statistical ways in a sustained way.
5. Multiple Procedures or Solution Methods	NP	No multiple procedures or solution methods present
6. Patterns and Generalizations	NP	No patterns and generalization present
7. Mathematical and Statistical Language	Mid	Moderate density: circle graph, bar graph, statistics, quarter, area, degree but not pressing students to use terms.
8. Overall Richness of the Mathematics and Statistics	High	A combination of elements that together saturate the segment with rich mathematics and statistics
Working with Students and Mathematics and Statistics		
9. Remediation of Student Errors and Difficulties	NP	No student misunderstandings or difficulties with the content observed.
10. Teacher Uses Student Mathematical and Statistical Contributions	Mid	Teacher uses students' contributions to construct the graph (e.g., using 3s to decide the half circle). However, instruction generally proceeds without the use of students mathematical ideas.
11. Overall Working with Students and Mathematics	Low	Interaction between teacher and students was present but with limited student input and no remediation.
Errors and Imprecision		
12. Mathematical Content Errors	NP	No mathematical or statistical errors observed
13. Imprecision in Language or Notation	NP	No imprecision in language or notation observed
14. Lack of Clarity in Presentation of Mathematical and Statistical Content	NP	The presentation of mathematical and statistical content is clear
15. Overall Errors and Imprecision	NP	No errors occurred

Table 3.9 (cont'd)

Student Participation in Meaning-Making Process		
16. Students Provide Explanations	Low	One or two brief student explanations were present (e.g., 3s make half for why to start with 3s question).
17. Student Mathematical and Statistical Questioning and Reasoning	NP	No instances of student mathematical or statistical questioning or reasoning are present.
18. Students Communicate about the Mathematics and Statistics of the Segment	Mid	Substantive student contributions like answering teachers' questions, offering explanations, and commenting on others were present but did not characterize the segment.
19. Task Cognitive Demand	Mid	Segment features mix of demanding and undemanding tasks and activities (e.g., constructing graphs started open-ended and was demanding; but the direct instruction with some student contributions at certain points lower the demand)
20. Students Work with Contextualized Problems	Mid	The contextualized problem is present and students play some role in deciding how to construct the graph. However, teacher mostly hints the ways to construct the graph based on the data they had.
21. Student Participation in Meaning Making and Reasoning	Mid	Students engage with content at mixed level. Few student explanations but substantive student contributions with middling density of task cognitive demand make the student participation in meaning making and reasoning in mid-level.

The whole lesson codes are scored on a 5-point Likert scale ranging from 1, "Not at all true of this lesson," to 5, "Very true of this lesson." According to Charalambous and Litke (2018), a score of 1 indicates "poor and problematic instruction," while a score of 5 represents "high-quality instruction in terms of combinations of characteristics from different dimensions" (p.450). A score of 3 is considered the "Default score," assigned when there is no evidence for low or highest quality. The MQI protocol provides descriptions for scores of 1, 3, and 5, but not for scores of 2 and 4. However, scorers are instructed to assign scores between these levels when necessary. Table 3.10 outlines the scoring criteria for the whole-lesson MQI code. As there were 26 videotaped lessons, a total of 26 whole-lesson MQI codes were assigned.

In conclusion, analyzing multiple data sources allowed for a detailed investigation into the nature of statistics instruction in 7th-grade mathematics classrooms. This section provided the data analysis process and tools. The following section will address the validity and reliability issues of the current study.

Table 3.10*The Scoring Criteria for the Whole-Lesson MQI Code*

Whole-Lesson Mathematical Quality of Instruction				
This code is intended to capture the overall mathematical quality of instruction for the lesson (MQI) as suggested by the teacher's work during the lesson.				
Low 1	Low/Mid 2	Mid 3	Mid/High 4	High 5
Instruction is characterized by one or more of the following: • Systematic teacher errors, imprecision, or lack of clarity around the mathematics • Major teacher conceptual error in a significant portion of the lesson • Unproductive teacher-student interactions around the content (e.g., teacher cannot effectively remediate) • Lack of directionality/unsystematic exploration • Lack of connection of classroom activities to mathematical content		Instruction does not have characteristics of Low and is mostly error-free, but it lacks the mathematical richness, appropriate use and discussion of procedures, and the sharp mathematical focus detailed under High. Examples: • Mostly error-free instruction, perhaps with occasional but not consistent elements of richness • Mainly pro forma interactions with students (inquiry, response, evaluation-type discussion)		Instruction is error-free (save for minor slips) and characterized by combinations of the following: • Mathematical richness in terms of explanations and linking between representations • Focus on mathematical practices (developed generalizations, mathematical efficiency) that is sustained and detailed • Instruction has a clear and sharp mathematical focus and directionality that allows students to develop the important mathematical ideas under consideration • Instruction is characterized by at least some productive teacher-student interactions around the content (either working with student ideas/errors OR Common Core aligned student practices)

Note: Mathematical Quality of Instruction (MQI): 4-Point Version, Hill et al., 2014, p.37
<https://hu.sharepoint.com/sites/GSE-CEPR/MQI-Training/Shared%20Documents/MQI%204-Point.pdf>

3.5. Validity and Reliability

Validity and reliability are essential issues in qualitative research (Creswell, 2009). Validity and reliability need to be considered in a study, from collecting and analyzing data to interpretation of the findings (Merriam & Tisdell, 2016; Yin, 2018). Merriam and Tisdell (2016) stated that providing validity and reliability is an ethical concern of research because the research findings need to be trustworthy due to their practical outcomes directly intervening in people's lives. In literature, four issues are considered in the validity and reliability of qualitative research: internal validity, external validity, reliability, and objectivity (Merriam & Tisdell, 2016; Yin, 2018). The terms credibility, transferability, dependability, and confirmability correspond to the four issues, respectively, and are primarily used in qualitative research (Merriam & Tisdell, 2016). I will use the corresponding terms while providing details regarding the validity and reliability issues for the current study.

First, Creswell and Poth (2018) described validity as “a distinct strength of qualitative research” due to the process of validity requiring “extensive time spent in the field, the thick, detailed description, and the closeness of the researcher to participants in the study” (p.338). Such a process increases the value of a qualitative study and makes the

findings “accurate from the standpoint of the researcher, the participant, or readers’ account” (Creswell, 2009; p. 22). Validity is not about verification but “trustworthiness and authenticity” (Creswell & Poth, 2018, p.338). Therefore, validity in qualitative research is called trustworthiness (Creswell, 2007). Creswell (2009) suggested eight primary strategies in the validity: triangulation, member-checking, rich and thick description, clarification of researcher’s bias, providing negative information, spending prolonged time in the field, peer debriefing, and external auditor. I consider the strategies while discussing the current study's credibility and transferability.

Credibility is interested in answering the question, “how do research findings match reality?” (Merriam & Tisdell, 2016, p.242). Like Creswell (2009), Merriam and Tisdell (2016) provided four strategies for establishing credibility. First, I employed a triangulation strategy by using multiple data collection methods, including three methods of data collection- classroom observations, interviews, and classroom artifacts. I used interviews and classroom artifacts against what I observed in the classroom. This allowed for a more comprehensive understanding of the statistical practices being analyzed. In addition, I used a member-checking strategy, stated as respondent validation (Merriam & Tisdell, 2016). After the data collection period ended, I reviewed the specific descriptions with the teachers to check if the preliminary results were accurate according to the participants’ perspectives. I also spent a prolonged time in the field; I engaged adequately in the data collection during the whole statistics unit covered and before and after the observation period. Furthermore, reflexivity was practiced throughout the study by taking reflective notes helped to reflect on my own biases and assumptions. Also, I noted the possible influences of my presence in the field. Lastly, I used peer examination to ensure credibility by having a Ph.D. in Mathematics Education review the data analysis process and the results of the study.

Transferability aims to increase “the possibility of the results of a qualitative study ‘transferring’ to another setting” (Merriam & Tisdell, 2016, p.256). The most common strategy to ensure transferability is a rich and thick description. In the current study, I provided a rich and thick description of the study, including the participants and the context of the study, to understand the cases. Also, the data collection and analysis processes were described in depth. Also, Merriam and Tisdell (2016) stated that

“purposefully select a typical or modal sample” similar to the current study’s selection of participants to increase the possibility of transferability due to the comparisons that the readers can make.

Confirmability refers to the results of a study based on the collected data rather than the researcher’s bias (Merriam & Tisdell, 2016). Creswell and Poth (2018) emphasized that qualitative researchers do not seek objectivity; instead, they seek confirmability of the findings. The audit trail is one way to ensure confirmability (Creswell & Poth, 2018; Merriam & Tisdell, 2016). The audit trail is interested in the detailed description of data collection, processing, and analysis. Therefore, research journals or memos are necessary to keep track of the whole process with emerging issues and ideas (Merriam & Tisdell, 2016). For this study, I had research journals and notebooks with coding decisions that helped me keep track of my process and my stand during the study to decrease the possible researcher bias and errors. Also, triangulation using multiple methods of data collection and reflexivity are the strategies adopted to establish confirmability.

Lastly, dependability is related to the reliability of the research findings and aims to decrease possible errors and researcher bias in a qualitative study (Merriam & Tisdell, 2016; Yin, 2018). Similar to confirmability, auditing of the research process could be used to ensure dependability (Creswell & Poth, 2018; Yin, 2018). In addition, triangulation using multiple methods of collecting data, peer examination, and the role of the researcher from the scope of reflexivity are other strategies used in the current study to ensure dependability.

Furthermore, different from the peer examination in the validation process, peer examination was employed to review the data analysis process and findings to ensure that the scorings done by the Transcript Coding Protocol Using GAISE Framework and the MQI are reliable. For this reason, I conducted intercoder agreement analysis for both tools with two different coders as my peers. In the intercoder agreement, I used MAXQDA with a minimum code overlapping rate of X% at the segment level analysis in which “per coded segment, it is checked whether the other person has assigned the same code to the segment” (MAXQDA Manual, n.d., p.764). In addition,

the kappa coefficients as chance-corrected coefficients were also calculated along with the percentage of agreement. MAXQDA Manual stated the kappa coefficient calculated based on Brennan and Prediger's (1981) definition. Walker (2010) described kappa values ranging between 0 to 1, and the values in between were interpreted based on Landis and Koch's (1977) guide. Walker (2010) provided the categorization of interrater agreement values " as: < 0.0 = Poor 0.0 to 0.20 = Slight 0.21 to 0.40 = Fair 0.41 to 0.60 = Moderate 0.61 to 0.80 = Substantial 0.81 to 1.00 = Almost Perfect" (p.40).

For the first protocol, the second rater held a Ph.D. in Mathematics Education and was familiar with the GAISE framework. Before scoring the lessons based on the protocol, we reviewed the protocol together and discussed possible examples for each element and category. Later we separately coded four lessons with 16 episodes in total. Each episode was assigned four elements with three levels of scoring. The agreement levels and kappa values of each element were calculated in MAXQDA. First, both the formulation of questions and data collection elements show perfect agreement between coders (100% for agreement level, $\kappa=1$). The other elements, data analysis, and interpretation, show lower agreement levels ranging from 93.75% and 87.50% for agreement levels and .88 and .75 for kappa levels, respectively. Still, the intercoder agreement is substantial to almost perfect in the remaining elements. After the intercoder agreement analyses, we compared our results and reached a complete agreement.

Furthermore, intercoder agreement analyses were conducted for the MQI instrument in MAXQDA with a co-coder, a Ph.D. student with a mathematics education background. Initially, we went over the MQI instrument together and discussed any uncertainties. Later, I used a video and its transcription to train the coder on analyzing a video lesson based on MQI. Subsequently, we coded a lesson using the video and its transcription together. Then we coded five lessons, dividing each into 7.5-minute segments with 21 codes separately.

The first agreement levels for each dimension were as follows: "Classroom Work is Connected to Mathematics and Statistics" had a 100% agreement level, while "The Richness of the Mathematics and Statistics," "Working with Students and the

Mathematics and Statistics,” “Errors and Imprecision,” and “Student Participation in Meaning-making and Reasoning” had agreement levels of 57.14%, 57.89%, 64.47%, and 64.04%, respectively. In addition, kappa values were calculated for each dimension. The kappa values ranged from 0.60 to 1.00, with “Classroom Work is Connected to Mathematics” having a perfect kappa of 1.00. The kappa values of other dimensions are .60, .59, .62, and .61, respectively, indicating moderate to substantial interrater reliability.

The results show that some dimensions had lower agreement levels and kappa values and needed improvement in additional training of coders. After conducting intercoder agreement analyses, we compared our results and discussed any discrepancies. Then we agreed on our disagreements at the end. However, since the intercoder agreement analyses corresponded to each dimension's lower agreement levels with moderate kappa values, we decided to review three more lessons I have coded. I provided the scoring decisions for three more lessons to the co-coder, and she decided whether the assigned scores to the code were accurate.

This section provided a detailed description of the method used in the current study. The following chapter will present the findings aligned with the research questions.

CHAPTER 4

FINDINGS

This chapter presents the main findings of the study, which aimed to investigate middle school mathematics teachers' statistical teaching practices in mathematics classrooms. The following research questions guided the study:

1. What are middle school mathematics teachers' statistical practices in the classroom?
 - 1.1 What statistical content do mathematics teachers cover in their instruction?
 - 1.2 How do mathematics teachers approach the four-step model of statistical problem-solving processes in their instruction?
2. What is the quality of middle school mathematics teachers' statistics instruction within their broader mathematics instruction?
 - 2.1 To what extent is classroom instruction connected to both mathematics and statistics?
 - 2.2 How rich is the mathematics and statistics content covered in the lessons?
 - 2.3 How do mathematics teachers work with students to support their understanding of mathematics and statistics?
 - 2.4 What are the mathematical and statistical errors and imprecision demonstrated by mathematics teachers?
 - 2.5 How do students participate in meaning-making and reasoning during statistics instruction within the broader mathematics context?

The findings are presented in two sections. The first section provides a within-case analysis of two cases, the teaching practices of Ali and Esra, regarding the alignment of statistical content with curricular expectations, the statistical problem-solving

process, and the quality of mathematics and statistics instruction. The second section presents a cross-case analysis that compares Ali and Esra in the same aspects.

4.1. Within Case Analyses

This section presents the findings of the within-case analysis, where I examine each case separately within the context of a multiple-case study. The first case in this study is Ali's teaching practices, and I provide a detailed analysis of his statistics instruction. Specifically, I explore the alignment of statistical content with curricular expectations, the statistical problem-solving process, and the quality of mathematics and statistics instruction. Next, I examine the statistics instruction of the second case, Esra's teaching practices, and provide a similar analysis of her teaching practices. This section provides an in-depth understanding of the teaching practices of mathematics teachers in teaching statistics by examining each case in detail.

4.1.1. Case 1: Ali's Teaching Practices

This section presents the findings of the within-case analysis for the first case, Ali's teaching practices. The analysis focuses on the alignment of statistical content with curricular expectations, the statistical problem-solving process, and the quality of mathematics and statistics instruction in Ali's classroom. The section is divided into three subsections: Ali's statistical content and problem-solving process, the mathematical and statistical quality of instruction at the segment level, and the whole-lesson mathematical and statistical quality of instruction.

The first subsection provides a detailed analysis of Ali's statistical content, including the topics covered in the classroom and how they align with curricular expectations. In addition, Ali's practices are examined in terms of the four-step model of statistical problem-solving processes. The second subsection focuses on the mathematical quality of instruction at the segment level. Specifically, I analyze the quality of instruction regarding mathematics and statistics in individual segments of Ali's instruction. The third subsection examines the mathematical and statistical quality of instruction at the lesson level involving the whole-lesson mathematical and statistical

quality of instruction, focusing on an entire lesson based on a combination of all dimensions in the MQI.

4.1.1.1. Ali's Statistical Content and Statistical Problem-Solving Process

This section describes Ali's statistics instruction in two main subsections. First, I will describe the instruction based on the statistical content aligned with curricular expectations. I aim to explore which concepts are covered in Ali's statistics instruction and the distribution of the number of episodes allocated for each concept. As the second subdomain, the statistical problem-solving process will be examined. I aim to uncover Ali's instructional practices regarding the goals of statistics instruction in the GAISE Framework (Litke & Hill, 2020), which includes four elements. The four elements will be examined separately and as a whole process.

4.1.1.1.1 Statistical Content in Ali's Instruction

To examine statistical content in Ali's instruction, I used the schematic diagram developed to understand the statistical tasks' distinctiveness and purpose in terms of the objectives covered in the Turkish 7th-grade mathematics curriculum. From the schematic diagram, it can be seen that Ali's statistics instruction consisted of 49 episodes as a unique set of data the class worked on (Litke & Hill, 2020). I used the episodes regarding statistical content and the statistical problem-solving process. As statistical content, the episodes consisted of graphs and measures of central tendency (see Table 4.1).

The episodes with graphs included 28 (57.14%) episodes. In the two (4.08%) episodes, the teacher and students worked on a practice worksheet consisting of all types of graphs. Furthermore, the remaining graph episodes consisted of a circle graph ($n=16$, 32.65%), a line graph ($n=7$, 14.29%), a bar graph ($n=1$, 2.04%), and graph transformation ($n=2$, %4.08). The measures of central tendency were presented in 20 (40.82%) episodes. In eight (16.32%) episodes, students worked on a practice worksheet consisting of all measures of central tendency or specific examples in which they decided which types of averages they needed to use for a given context. The

remaining episodes consisted of the arithmetic mean (n=6, 12.24%), median (n=1, 2.04%), and mode (n=5, 10.20%).

Table 4.1

Statistical Content in Ali's Instruction

Level 1	Level 2	n	%
Graphs		28	57.14
	Circle Graph	16	32.65
	Line Graph	7	14.29
	Bar Graph	1	2.04
	Graph Transformation	2	4.08
Measures of Central Tendency		20	40.82
	Arithmetic mean	6	12.24
	Median	1	2.04
	Mode	5	10.20
Other		1	2.04

Note: N=49 (the total number of episodes)

Seventh-grade Turkish Middle School Mathematics curriculum consists of four objectives. These objectives are to construct and interpret line and circle graphs, calculate, and interpret central tendency measures, and display data in bar, line, and circle graphs and transform data between these types of graphs when appropriate. The results indicate that the statistical content was aligned with the objectives of a 7th-grade Turkish mathematics curriculum.

Ali covered at least one episode for each objective in the curriculum. In one of the episodes, the teacher presented a bar graph task to the students as a reminder of other graph types they had learned in the fifth and sixth-grade mathematics curricula. These graph types were standalone objectives in those years. However, this particular episode aligned with the current curriculum's last objective, which aimed to teach students how to display data using different types of graphs. Lastly, one episode (2.04%) was assigned the “other” code because it briefly covered the sample concept by defining what it means and determining the sample of an example teacher provided at that moment.

4.1.1.1.2. Statistical Problem-Solving Process in Ali's Instruction

I utilized the Transcript Coding Protocol Using the GAISE Framework by Litke and Hill (2020) to analyze Ali's statistics instruction concerning statistical investigations. The coding protocol is designed to understand statistical instructional practices based on GAISE standards, aligned with statistical investigations' components. The framework comprises four elements, namely, Formulation of Questions, Data Collection, Data Analysis, and Interpretation, each with three categories. The categories are:

Level 1: The element was not observed.

Level 2: The element was observed to some extent but without the depth specified in GAISE. For instance, finding measures of central tendency and range without using that information to address distribution or spread.

Level 3: The element was more fully aligned with GAISE expectations and implemented with a more in-depth focus on statistical ideas or concepts (Litke & Hill, 2020, p.487).

I coded each episode concerning the (a) formulation of questions, (b) data collection, (c) data analysis, and (d) interpretation in terms of the three categories presented above. Table 4.2 shows the episode frequencies of each code considering the level of presence of each code (through Level 1 to Level 3).

Table 4.2

Ali's Statistical Problem-Solving Process: The Occurrences of the Level of Elements

	Level 1		Level 2		Level 3	
	n	%	n	%	n	%
Formulation of Questions	46	93.88	2	4.08	1	2.04
Data Collection	48	97.96	1	2.04	0	0
Data Analysis	1	2.04	44	89.80	4	8.16
Interpretation	40	81.63	8	16.33	1	2.04

Note: N=49 (the total number of episodes)

Table 4.2 shows that the three steps for the statistical problem-solving process were in Level 1, so they were not primarily observed in the episodes of Ali's instruction, namely formulations of questions (n=46, 93.88%), data collection (n=48, 97.96%), and interpretation (n=40, 81.63%). Data analysis was observed to some extent, but most episodes did not meet all expectations provided in GAISE (n=44, 89.80%); that was about the conceptual understanding of statistical processes.

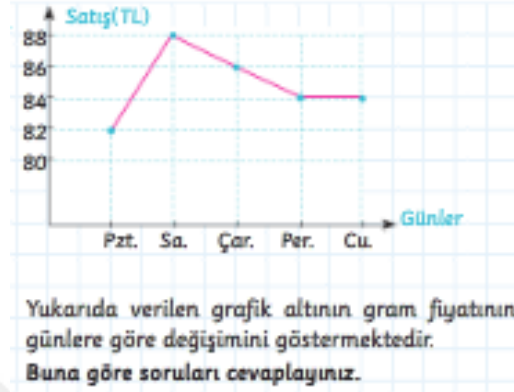
None of the episodes included data collection at Level 3; no discussion on issues about collected data. The frequencies of Level 2 and Level 3 at all steps except data analysis were low (below 16%). In the following section, I will describe each step of the statistical problem-solving process in Ali's instruction. Ultimately, I will present the co-occurrences of the steps in the statistical problem-solving process.

4.1.1.1.2.1. Formulation of Questions

First, Ali did not pose statistical investigative questions or create an environment for students to formulate statistical questions in most episodes (n=46, 93.88%). Therefore, those episodes were coded as Level 1, where the data worked in those episodes was not based on a statistical question, so it did not lead to a statistical investigation (see Figure 4.1). The teacher generally provided tasks with data sets for graphing, reading graphs, or calculating the measures of central tendency. There were questions in the tasks, but they could not be identified as statistical questions. The reason is that the questions did not demand students to inquire, and they did not require students to anticipate variability. They were generally for exploring and analyzing the data, not posing a statistical investigative question.

Two episodes (4.08%) included statistical investigative questions posed by Ali and described at Level 2. However, the questions posed by the teacher did not lead to a statistical investigation. First, one question was not in a question form. However, to some extent, it was a statistical investigative question, in which Ali introduced two data sets about the parents' income of two classes to discuss which average concept should be used for summarizing the data. Before providing the data and discussing the average, Ali stated, "I am wondering this, friends, I want to calculate. I wonder and

learn about the monthly income of the students' parents. Let the data obtained be as follows.” Then he wrote the data on the board. The data was a secondary one. It was second hand as described in the updated GAISE (Bargagliotti et al., 2020).



The graph above shows the change in the gram price of gold according to the days. Answer the questions accordingly.

- Which day is the highest price day?
- Which day is the day with the lowest price?
- On which day was the price unchanged?
- Which day did the price increase?
- On which days did the price decrease?
- How much TL is the difference between the highest and lowest prices?

Figure 4.1 Example Task of “Formulation of Questions” Element at Level 1 (Varışlı, 2017, p.153)

One episode (2.04%) consisted of student-generated statistical questions (Level 3), but they were not again led to statistical investigations. In the episode, the teacher asked students to form a question to ask the class or school while introducing the line graph topic. The students came up with example questions to be asked to students in the school as a survey: “How well do you play ball?”, “How much do you like math?”, “How much do you like school?”, “Do you like to sleep?”. The student-generated statistical questions were only partially accurate. They were survey questions, not statistical investigative questions. Some of them were deterministic and not leading to statistical investigations. After the examples, the teacher did not discuss students' example questions or focus on changing survey questions to statistical investigative questions that could lead to an investigation. However, I scored this episode as Level 3 since it involved questions generated by students.

In the following episode, Ali posed a question while introducing a line graph:

T: What if I choose one and ask a question? Let me ask Yiğit; we have wondered about your math averages since 1st grade. I looked in your file; what is Yigit's math grade? Yiğit?

Multiple Students: 30-20

T: Let us say he was 65, he dropped a little in the 2nd grade 60 ... in the third grade

S: Let it be ascension period, my teacher; let it be 80

T: He said, let it be the ascension period, let it be 80... In the 4th grade, let it be 85...

[Ali 30.04.2018-2 10:30-26:00]

His question was not statistical and somewhat similar to students' questions. The question was about one individual or group. Variability was present because the student's math grades over five years were given. However, the context was limited for anticipating variability. In addition, the data for the questions formed were not authentic or realistic in all examples above. The teacher came up with the data spontaneously during the class discussion, and although it appeared to be secondary data, it was essentially fabricated.

Overall, the results show that Ali did not formulate statistical questions in most episodes. The episodes with the formulation of questions included questions that did not require collecting data from a group or a population; they were deterministic and followed up on secondary data in hand or made up.

4.1.1.1.2.2. Data Collection

Table 4.2. shows that 48 episodes (97.96%) did not include a data collection process. In those episodes, the data were provided. Some episodes had specific tasks with predetermined data sets. The students were expected to work on the data provided (i.e., calculating measures of central tendency). In other episodes, the data were given in the graphs or tables that students were expected to read (i.e., finding a missing value in the graph). The episodes with no data collection also did not consider the data experiences (i.e., not considering the type of variables).

Only one episode (2.04%) included data collection at Level 2. In the first lesson of statistics instruction, Ali prepared 12 colored papers (two yellow, three green, four blue, and three pink). He did not mention the number of each color; he distributed them to the 12 students and asked them to be in a circle with the same colors next to each other. The teacher asked how many students had yellow, green, blue, and pink. Then they recorded the number of colors and constructed a circle graph based on the collected data. The task differed from other tasks involving predetermined data sets in data collection. However, it did not consider sampling and measurement or acknowledge variability and did not connect with a statistical investigative question to determine if the data fit the question.

To conclude, Ali's instruction did not involve a data collection process except for one episode. The episode with data collection was at Level 2. The data collection was not connected with a statistical investigative question. Therefore, the students did not have experience collecting data by considering sampling, measurement, and variability in Ali's statistics instruction.

4.1.1.1.2.3. Data Analysis

Ali's statistics instruction included a significant amount of data analysis. In 48 (97.96%) episodes, Ali presented data analysis tasks. Only one episode (2.04%) did not include data analysis. One of those episodes introduced a line graph (i.e., asking where they saw a line graph) and formulating questions (at Level 3). They did not work on the analysis of the data.

Forty-four episodes (89.80%) at Level 2 focused on constructing a graph, finding a value in a graph, and calculating the average. Therefore, they were described as Level 2 episodes: They did not include the description of distribution, comparisons of the groups, and examining the association of the groups. For example, students worked on the information (data) given in the episodes with graphs. Students found the total and the part in quantities, percentages, and degrees. They constructed the graphs based on those calculations or given graphs. For example, the students constructed a line graph or answered the questions like "What is the difference between the highest price and

the lowest price?"; "How many Liras were put in the piggy bank in March?" Those questions required calculation. Although students had experiences with data analysis with such examples, the experiences with most graphs (given or constructed) were procedural. They did not involve some graph elements to make sense of them (e.g., the title). In average examples, students worked with contextualized and noncontextualized tasks. They calculated the arithmetic mean, median, and mode of the data set. Mostly they did not use those calculations for further analysis; to understand the distribution.

Four episodes (8.16%) included data analysis at Level 3. Those episodes were also about the measures of central tendency. The episodes consisted of tasks discussing what happened to the measures of central tendency if a value changed in a small data set. The discussion led students to explore and describe the distribution. They explored how the center changed or did not change with the changes in the distribution. In addition, some tasks asked which central tendency measures best describe the data set. In such examples, the students needed to understand and analyze the distribution. The students needed to consider the data type, the outliers, and the variance in the data set for such discussion, even though the concepts were not mentioned in the episodes explicitly except "outlier." Furthermore, I found that the episodes in which the teacher instructed students to compare the parts and the whole in a circle graph or compare data from different periods in a line or bar graph often resulted in the level of data analysis classified as Level 3. This met the GAISE expectation that data analysis should focus on using the display to describe the distribution and compare groups or observe a relationship between variables.

As a result, Ali's statistics instruction covered data analysis except for two episodes. The majority of the episodes were at Level 2 in data analysis. The graph construction, descriptive analysis based on graphs, and the calculations of the measures of central tendency were present in general. In addition, there were some episodes at Level 3 where the students had opportunities to discuss distributions and compare groups in graph and average tasks.

4.1.1.1.2.4. Interpretation

The results show that 40 episodes (81.63%) did not include interpretation. The teacher did not ask students to make statements about the results of the data analysis. Eight of 49 episodes (16.33%) were at Level 2; the students read the graph and used the graph or data to answer the questions.

For example, in the line graph task about ‘the change in the gold prices based on days,’ the teacher asked, “Which day has the largest increase/decrease?” In the line graph task about the change in the plant-height task, the teacher asked, “In which interval did the greatest increase in height occur?” In the line graph task about the monthly change in the money in the piggy bank, the teacher asked, “In which month does the amount of money remain unchanged?”. These example questions were based on reading the data and considering the relationship among the data. They were not about reading beyond the data. In the other episode, the students made statements about the context based on data. The teacher provided a task about finding the arithmetic mean of mathematics grades of 10 students in 7A class. Then he asked students why they found the arithmetic mean for the group. The students stated that they determined which classroom was better. The episode was not at Level 3 because the students explained why the classroom was better by examining the data within samples. However, they did not compare classes for these episodes.

One episode (2.04%) was at Level 3, discussing the variability, making predictions, and comparing groups. The two groups were compared in the episode regarding their parents’ income. They discussed that if they used arithmetic mean in the data set with an outlier, they could not reach the correct conclusion about the parents' income. Hence, they discussed the variability in both data sets to summarize the data soundly.

Most episodes in Ali’s statistics instruction did not consist of the interpretation of analyzed data. Five episodes were at Level 2, and one was at Level 3. In those episodes, the students and the teacher interpreted the graph or the typical value to calculate an average. Furthermore, one episode consisted of a looking beyond data example; the students and the teacher made inferences about the analyzed data of two groups for comparison.

4.1.1.1.2.5. Overview of Ali's Statistical Problem-Solving Process.

The results show that Ali's statistics instruction did not focus on three elements in the statistical problem-solving process: posing statistical questions, collecting data, and interpreting data. However, there was a focus on data analysis. Most episodes included data analysis in the descriptive analysis of steps in the statistical problem-solving process. As a whole process, I can conclude that Ali's instruction did not focus on the statistical investigation.

To gain a more comprehensive understanding of the statistical problem-solving process as a whole, I determined the frequency with which different categories of elements appeared together in each episode by using MAXQDA. Table 4.3 outlines the frequencies of these co-occurrences, providing insights into the interaction between different elements' categories of statistical problem-solving process.

Table 4.3

Ali's Statistical Problem-Solving Process: The Co-occurrences of the Level of Elements

Formulation of Questions	Data Collection	Data Analysis	Interpretation	Episodes (n)	Percent (%)
FQ Level 1	DC Level 1	DA Level 2	I Level 1	34	69.39
FQ Level 1	DC Level 1	DA Level 2	I Level 2	8	16.33
FQ Level 1	DC Level 1	DA Level 3	I Level 1	3	6.12
FQ Level 1	DC Level 2	DA Level 2	I Level 1	1	2.04
FQ Level 3	DC Level 1	DA Level 1	I Level 1	1	2.04
FQ Level 2	DC Level 1	DA Level 2	I Level 1	1	2.04
FQ Level 2	DC Level 1	DA Level 3	I Level 3	1	2.04

Note: N=49 (the total number of episodes) FQ = Formulation of Questions, DC= Data Collection, DA= Data Analysis, I= Interpretation

Table 4.3 displays that in 34 out of 49 episodes (69.39%), the students solely analyzed the data at Level 2. This indicates that they constructed graphs, and calculated graph-related questions and measures of central tendency but did not pose statistical investigative questions, collect data, interpret data, or demonstrate a deeper

understanding of the concepts. Only three episodes (6.12%) showed students performing data analysis at Level 3, which included comparing groups and describing the distribution. Nevertheless, these episodes lacked posing questions, collecting data, and interpreting the results of the analysis.

Eight episodes (16.33%) with data analysis with interpretation at Level 2 co-occurred, where students analyzed the data descriptively and answered “within” data questions. Again, these episodes lacked posing questions and collecting data. In addition, data analysis and interpretation elements co-occurred at Level 3 in only one episode. However, this episode did not cover the whole process.

Furthermore, two episodes (4.08%) started with posing statistical questions at Level 2, with data provided not collected. In one of the two episodes, data analysis at Level 3 appeared with interpretation at Level 3 together. The other one scored data analysis at Level 2 with no interpretation. Furthermore, the formulation of questions at Level 3 did not follow data collection, analysis, and interpretation in one episode (2.04%). Last, in only one episode, data collection was present at Level 2 and co-occurred with data analysis at Level 2 without posing questions and interpretations.

These results also supported the separate findings on statistical problem-solving. The main focus was on data analysis. Other elements formulating the question, data collection, and interpretation co-occurred minimally. There was not even an episode starting with posing a statistical investigative question, followed by data collection and the analysis of the collected data and interpretation of those analyses.

4.1.1.2 Mathematical Quality of Instruction (MQI): Segment Level Analysis of Ali's Statistics Instruction

This section describes Ali's instructional quality in statistics instruction, which includes a segment-level analysis with 21 codes and a lesson-level analysis with one code. The data sources used for this section were videotaped lessons, field notes, lesson summaries, and interviews with the teacher. The MQI was employed to analyze the

instructional quality of Ali's statistics instruction by dividing the lessons into 7.5-minute segments.

Ali conducted 14 lessons with varying durations, resulting in various segments. Five lessons had five segments: four 7.5-minute segments and one 3-minute or longer segment. The remaining nine lessons had four segments. In total, I evaluated 61 segments using 21 segment-level codes.

This section presents the segment-level analysis divided into four main dimensions. Each dimension consisted of subsections related to the segment-level codes. First, I present the summary statistics of segment-level codes in each dimension. Then I describe Ali's instruction in each segment-level code under the dimensions. Next, an overview of each dimension is provided.

The following section starts with the first code: Classroom Work is Connected to Mathematics and Statistics. After the results of the segment-level analysis, I describe Ali's statistics instruction with the whole-lesson quality of instruction code from the lesson-level analysis.

4.1.1.2.1. Classroom Work is Connected to Mathematics and Statistics.

The results show that classroom work in all segments (%100) of the Ali's lessons was connected to mathematics and statistics. I observed that Ali's instruction consisted of reviewing the prior content, introducing new concepts, working on tasks as a whole class, or working on tasks individually. Although there were some administrative issues or spending time distributing the materials (the worksheets), such activities did not take most of the segments. In such interruptions, Ali wanted his students to work on the tasks. For example, in a lesson segment, the school principal checked the electricity leakage from a cable and talked to the teacher. Ali warned students to work on what they were doing.

I observed that Ali followed the workbook, but before the lesson, he chose which task he would cover with students and which questions students could continue to work on

if they were ahead. He continually checked students' work and tried to answer students' questions individually. He arranged most of the class time to work on problems individually. When he wanted to discuss the task as a whole class, he asked all students to look at the blackboard. He wanted students to work individually when distributing worksheets, including multiple-choice questions. However, when more than two students had difficulty with the specific questions, he called on those students at the teacher's or student's desk to solve the problem.

4.1.1.2.2. The Richness of the Mathematics and Statistics

This section aims to investigate the richness of the mathematics and statistics of Ali's statistics instruction. In light of this aim, I examined the richness dimension with seven codes in the MQI. Table 4.4 shows the results regarding the richness dimension in the MQI.

According to Table 4.4, the richness of the mathematics and statistics in Ali's teaching practices can be characterized by low to mid-quality. Particular codes had a small number of high-quality codes (e.g., Explanations). Furthermore, in most lesson segments, Ali's statistics instruction did not involve codes regarding Patterns and Generalizations (n=59, 96.72% absent/not present), Linking between Representations (n=49, 80.33% absent/not present), and Multiple Procedures or Solution Methods (n=47, 77.05% absent/not present).

The codes that had relatively higher presence than others were Mathematical and Statistical Sense-Making (n=30, 49.18% mid; n=15, 24.59% low), Mathematical and Statistical Language (n=45; 73.77% mid; n=12, 19.67% low), and Explanations (n=11, 18.03% mid; n=13, 21.31% low).

Furthermore, all lesson segments include codes regarding Mathematical and Statistical Language and Mathematical and Statistical Sense-Making (both 0% absent). Besides, none of the segments could be identified as 'high' quality in the codes of Patterns and Generalizations and Multiple Procedures or Solution Methods (both 0% high).

Table 4.4*“The Richness of the Mathematics and Statistics ” Dimension Scores of Ali’s Instructional Segments*

	1									
	Not Present		Low		Mid		High		M	SD
	n	%	n	%	n	%	n	%		
Linking Between Representations	49	80.33	8	13.11	3	4.92	1	1.64	1.28	0.64
Explanations	34	55.74	13	21.31	11	18.03	3	4.92	1.72	0.93
Mathematical and Statistical Sense-Making	0	0.00	16	26.23	30	49.18	15	24.59	2.98	0.72
Multiple Procedures or Solution Methods	47	77.05	9	14.75	5	8.20	0	0.00	1.31	0.62
Patterns and Generalizations	59	96.72	2	3.28	0	0.00	0	0.00	1.03	0.18
Mathematical and Statistical Language	0	0.00	12	19.67	45	73.77	4	6.56	2.87	0.50
Overall Richness of the Mathematics and Statistics	1	1.64	40	65.57	17	27.87	3	4.92	2.38	0.61

Note. N=61 (the total number of lesson segments)

Figure 4.2 shows that mathematical and statistical sense-making, mathematical and statistical language, and the overall richness of mathematics and statistics were observed in Ali's statistics instruction. Especially mathematical and statistical language and mathematical and statistical sense-making were present in all segments, and about three-fourths of the segments were mid or high-quality.

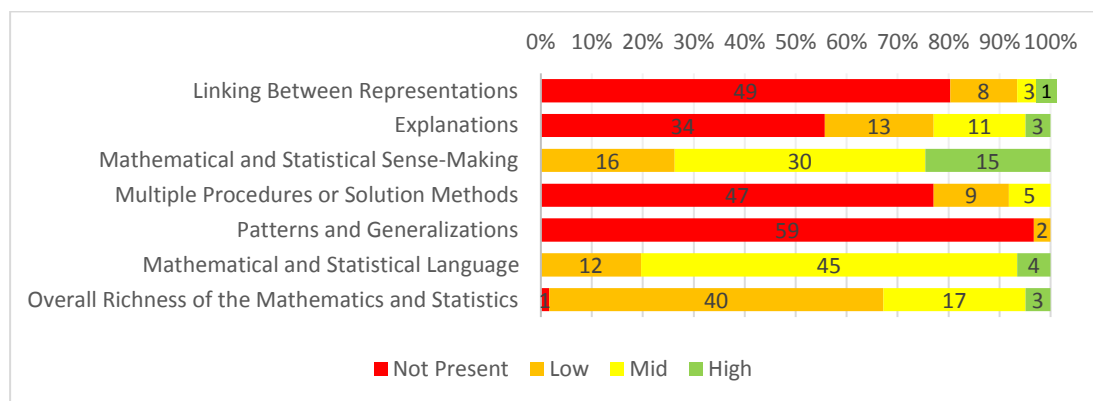


Figure 4.2 “The Richness of the Mathematics and Statistics” Dimension in Ali's Instruction

Note. N=61 (the total number of lesson segments)

However, linking between representations, multiple procedures or solution methods, and patterns and generalizations codes indicated the lack of richness in Ali's instruction while teaching statistics. Those codes were absent in half of the segments above three fourth of the lesson segments. In addition, explanations were not present more than half of the segments. The details of the codes in the richness dimension will be presented in the following sections.

4.1.1.2.2.1. Linking Between Representations

Figure 4.3 shows the distribution of scores assigned to lesson segments for the “Linking Between Representations” element, describing the explicit links between representations from different representational families. Figure 4.3 indicated that Ali did not link different representational families in 49 (80.33%) lesson segments. In 12 (19.67%) lesson segments, Ali linked representations from low to high levels. Ali connected representations while constructing graphs (geometrical structure of a graph,

especially a circle graph) and less often during teaching measures of central tendency. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.28 out of 4, with a standard deviation of 0.64. This section provides the descriptions and classroom examples of each score assigned to the Linking between Representations code in Ali's statistics instruction.

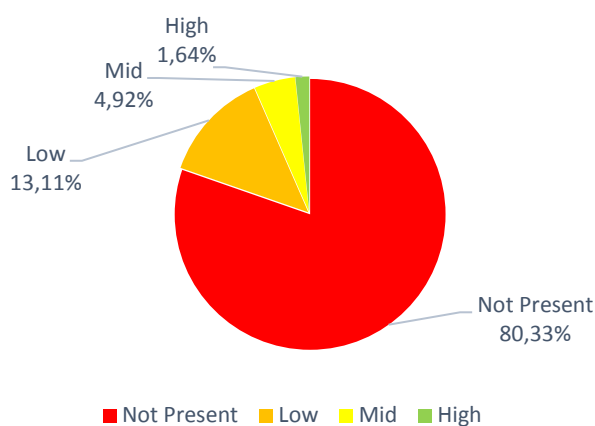


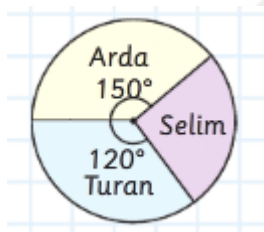
Figure 4.3 The Distribution of Scores Assigned to Lesson Segments for the “Linking between Representation” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

I assigned a “Not Present” score to 49 (80.33%) lesson segments for linking between representations because there were no links in those segments. There were no representations in specific segments, including individual work periods or introduction to the statistical concepts (e.g., a particular graph). Besides, many segments included representations without connections or explicit links (e.g., symbolic representations and graphs), or only one representation was present (e.g., only graphs or symbolic representations). For example, Ali used average as an algorithm definition. He did not introduce average as a mathematical balance point. He would have opportunities to use and connect among representations if he did. Also, while covering the objective related to transforming among different types of graphs when appropriate, the data were represented explicitly and in detail in at least two different types of graphs. However, I assigned a score of Not Present in those segments because the rule regarding the connections must be among “different representational families” (Hill et

al., 2014, p.5). Even though they were different graphs, they were from the same representational families.

In eight (13.11%) lesson segments, Ali linked the representations at a low level because he did not explicitly explain the connections or just pointed out the links without the presence of one of the representations. In one case, the task asked to find the central angle of a sector for representing a class with 30 students, including 20 girls. While solving the task, he described the sector by drawing a circle graph. While drawing the circle graph, he stated which sector represented girls and boys, emphasizing the central angle of the sector. The calculated angles and their placements in the circle graphs were mentioned briefly. The teacher did not elaborate on the relationship between the circle sector and the context of a classroom's gender distribution. Furthermore, similarly, in the following case, he introduced the following graph, which included three context-related sub-questions requiring proportions (see Figure 4.4).



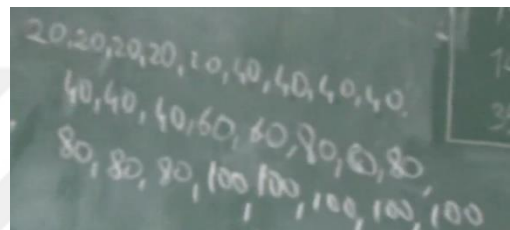
The Question Part:

The circle graph shows the ratio of the votes of candidates for class representative. When Turan got 24 of the votes;

- What is the class size?
- How many votes did Selim get?
- How many votes did Arda get?

Figure 4.4 Graph Task 1 (Varışlı, 2017, p.150)

First, Ali drew the graph on the blackboard. While answering how many votes Selim got, Ali and his class discussed that the remaining part -Selim- is 90° and 90° means quarter, and quarter means one-fourth, so that they could find Selim's votes by dividing class size by four. Although the fraction is not visually present, the links between 90° and the sector were connected through fraction explanations. In short, symbolic representation and graphical representation were linked. If both presentations were visually present, the segment would have been scored “mid” because of the details and elaboration about the connections among multiple representations, but not high because it was brief and not the focus of that segment.



Open form (Symbolic representation)

with the question of how many colored papers there are in total and then discussed as follows:

T: 12, isn't it? There are 12 people. Twelve cards, 12 people in total. Well, I will share these 12 people in the circle graph, right? According to the numbers here. So, where do you think, we should start sharing?

S1: Let us start from the middle

T: Okay, we will start in the middle, and how do we divide it?

S1: From the greatest.

T: Should we start with the biggest?

S2: We divide by four

T: You tell me

S3: Let us start with three

T: 3s. Why threes? [asking the student to speak]

S3: It makes half of it ... three plus three makes six

[Ali 26.04.2018- 00:00-07:30]

They discussed how they relate numbers to the graphical representation while constructing a circle graph. After putting the numbers based on their fraction knowledge, they described them with their understanding of the percent concept.

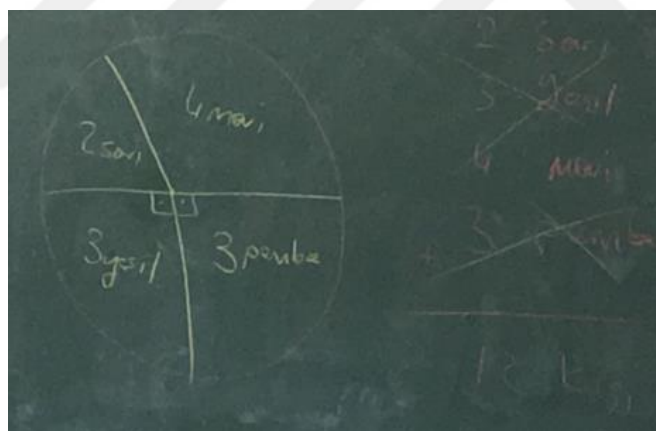


Figure 4.6 Graph Task 3

T: Now let us see, what did we see here? What does this 90° mean to me? What does it mean that these three people have such a response or that these three people have such a response?

S: Is it its quarter?

S: The graph?

T: Tell me, son

S: Teacher, one-fourth of the circle graph.

[Ali 26.04.2018- 00:00-07:30]

During the lesson segment, students were introduced to various numerical representations, such as numbers, degrees, and fractions, which were visually present along with a graph. They were asked to arrange the sectors of the circle for corresponding numbers, 90 degrees, or quarters while constructing the graph. Multiple representations (e.g., symbolic, and graphical) were used in this process, and students were encouraged to link the representations in detail. The segment focused on linking different representations, particularly the numerical and visual ones.

4.1.1.2.2.2. Explanations

Figure 4.7 shows the distribution of scores assigned to lesson segments for Ali's mathematical and statistical explanations based on the “why” question answers. Most lesson segments were assigned a score of Not Present for providing “explanations” ($n=34$, 55.74%). Ali provided explanations from low (a brief explanation) to a high level (explanations as to the segment's focus) in the remaining lesson segments ($n=27$, 44.26%). Ali provided explanations regarding why to use or not to use a specific solution method and why to use a graph or a measure of central tendency. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.72 out of 4, with a standard deviation of 0.93. This section provides the descriptions and classroom examples of each score assigned to the Explanation code in Ali's statistics instruction.

I assigned a score of Not Present to 34 (55.74%) lesson segments for explanations because these segments included descriptions of solution methods, the clarifications of the tasks, and the definitions of graphs and central tendency measures. For example, a circle graph task in one of the segments consisted of the number of people in three tourist groups. The question asked for the sector's angle corresponding to a tourist group in the circle graph. This was one of the first tasks solved in the circle graph topic and students' experiences with the connections between the circle and the context. He simply described how to form a direct proportion. He did not explain why the total given in the context relates to the sector's angle or why they used direct proportion to find the number of tourists in a tourist group.

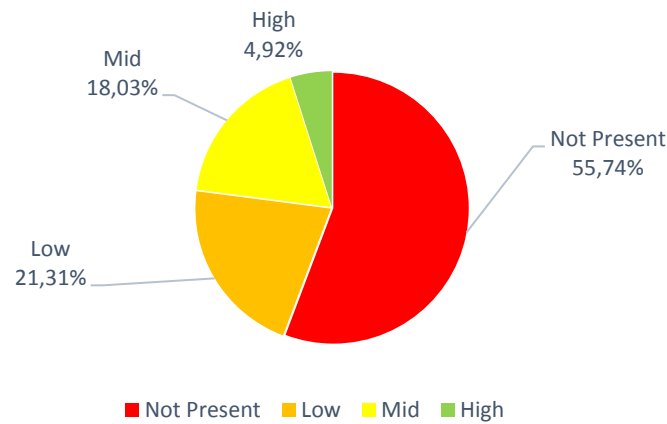


Figure 4.7 The Distribution of Scores Assigned to Lesson Segments for
“Explanation” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

I assigned a Low score to 13 (21.31%) lesson segments for providing explanations in Ali's statistics instruction. The reason is that Ali's explanations were briefly present for those segments. He briefly explained why a procedure works or does not generally work. Those explanations were not the focus of the segment. For instance, two segments after the previous example provided in the “Not Present” part, the class worked on a similar example (see Figure 4.4 Graph Task 1 in the LBR part). In this example, Ali solved the sub-questions about the circle graph showing the ratio of candidates' votes for class representative. First, he described the solution steps of the sub question of the class size when Turan got 24 votes. Second, when Ali and his students tried to find the second sub-question on Selim's number of votes, a student explained that 90 was one-fourth of 360 and needed to divide 72 into four. The teacher repeated the student's explanations to the class that “90 is a quarter of 360, so it is a quarter since it is divided into four ... 72 should be divided by 4, or a quarter of 72 should be the answer.” This explanation was an example of a brief explanation of why dividing by four is necessary for finding Selim's number of votes. If they worked on similar reasoning throughout the segment, not just a brief part of a segment, explaining why the solution steps needed to be taken, the segment would have been scored Mid or above Mid for providing explanations.

I assigned a Mid score to 11 (18.03%) lesson segments for providing explanations in Ali's statistics instruction. In the MQI, it is pointed out that explanations are scored as Mid when present more than briefly or there are one or two explanations, but they were not scored as high because the explanations are not the focus of instruction (Hill et al., 2014). In Ali's instruction, there were segments with one or two explanations, or the explanations were present more than briefly. For example, I scored Mid for explanations for the segment coded as High in LBR (See extract Ali 26.04.2018-00:00-07:30). In the segment, Ali and his students also discussed why they started using threes (three green and three blue cards) while discussing how to construct the graph and connect symbolic and graphical representations. A student explained that three and three made six, and six was half of 12. Then the teacher contributed to the explanation that threes were the quarters of 12. The discussion on using three while constructing a graph was more than brief, so I scored Mid.

I scored High for only three (4.92%) lesson segments in explanations. After a few lessons introducing average concepts and practicing computational parts, Ali wanted students to solve tasks requiring students to determine which average concept should be used for a given context. In a segment, Ali devised his example to differentiate which average concept should be used for the given context. He stated that he was conducting two independent research, and the numbers given are the incomes of five parents in 7A and 7B. They discussed why they used a specific average for each class:

T: If we wanted to interpret students' incomes in 7A and 7B, which average concept should we use for 7A and 7B? I want to comment on them; I want to conclude. I want to hear your ideas... Which one would you prefer to use for 7A?

S1: 7A median

T: You said median for 7A, and which one for 7B?

S2: Mode

T: Okay, why? Did you prefer the ones you like, or just guess?... Median, okay, do you have a reason?....

[Ali 14.05.2018- 22:30-30:00]

The student remained silent. After the student could not explain his reasoning, Ali decided to hear about other students' answers and chose each example without explanations. Three students said both were arithmetic mean; two students said both were mode, and two students said both were median. Then students started to give

different answers for each example. After taking the answers, Ali gave a clue to students on how to look at the given data:

T: Okay, everyone gave different answers. Look at two distinct groups. One group had values closer to each other. The other group had one value that was different from the others. I wanted to make interpretations about 7A and 7B. Let us think. Firstly, we need to prefer the arithmetic mean for 7A because the values are closer. Okay, when we calculated the arithmetic mean for 7A, does it give an idea for 7A for us?

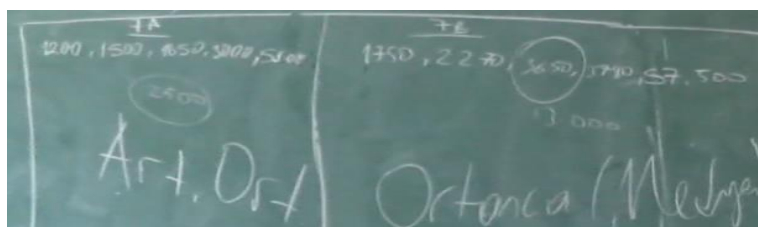
Ss: Yes, it does.

T: ... Okay, here, does the richest parent make the arithmetic mean larger? ... Don't you think this [value] makes me incorrectly interpret 7B? When a person [parent] is richer or apart from other values, makes arithmetic mean?

Ss: [inaudible]

T: [it] prevents me; I make incorrect interpretations. This makes me think that their income is very good to get donations from them for the school. However, it does not give the truth. I can then say that the first thing I should consider using is the arithmetic mean, but if there are values a way bigger numbers than the group or smaller numbers than the group, you do not prefer to use the arithmetic mean. Why? Because one value makes the average smaller, the larger value makes it there [showing high with his hand], it prevents me from making correct interpretations....

[Ali 14.05.2018- 22:30-30:00]



7A	7B
1200 1500 1650 3000 5100	1750 2270 3650 3780 57500
Arithmetic Mean	Median

Figure 4.8 Average Task 1

Ali explained why they chose median over arithmetic mean for the 7B context throughout the segment. He also explained why he used arithmetic mean for the 7A context. Overall, the teacher's explanations were the segment's focus and were substantive and detailed.

4.1.1.2.2.3. Mathematical and Statistical Sense-Making

Figure 4.9 shows the distribution of scores assigned to lesson segments for mathematical and statistical sense-making to uncover how Ali and his students focused on mathematical and statistical sense-making. Figure 4.9 shows that 16 (26.23%) lesson segments were assigned a score of at least Low. Furthermore, I assigned a Mid score to most lesson segments ($n=30$, %49.18). The remaining segments ($n=15$, 24.59%) were scored as High. Ali focused on the meaning while constructing graphs, using direct proportion to solve graph tasks, connect different graphs, and describe a data set with central tendency measures. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.98 out of 4, with a standard deviation of 0.72. This section provides the descriptions and classroom examples of each score assigned to the Mathematical and Statistical Sense-Making code in Ali's statistics instruction.

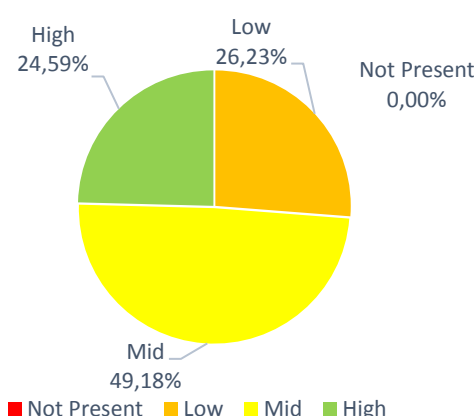


Figure 4.9 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Sense-Making” Code in Ali's Instruction

Note. $N=61$ (the total number of lesson segments)

I did not give any segments a “Not Present” score since no segments included incorrect or no sense-making processes in Ali's instruction. Ali's instruction included some incorrect or incomplete parts described in the Errors and Imprecisions section. Still, as the MQI instrument suggested, I scored only the correct portion of the segments for this code (Hill et al., 2014).

Furthermore, I gave a Low score for mathematical and statistical sense-making to 16 (26.23%) of the lesson segments because Ali focused on the meaning briefly during those segments. Most of those segments were part of a practice session where students practiced and worked individually. There were a few instances where the teacher or students talked to the class during practice. There were brief instances when the teacher observed students' work and worked on specific questions individually or in groups. Sometimes he directed the questions struggled to the class. The teacher interacted with students individually, in groups, and with the whole class during individual work time. Those brief interactions were mainly based on making sense of procedures or discussing the solution methods. In one of the active instruction parts, the teacher solved a task on the board and asked which average concept should be used for a given context (e.g., the favorite course in the school); arithmetic mean, median, or mode. Ali described that they need to consider arithmetic mean, then median, and lastly, mode. He mentioned they could not use the arithmetic mean because "it does not always work." For "the favorite course in the school" example, he said it is better to use mode.

T: Which course has a typical characteristic here?

Multiple students: Mathematics

T: Mathematics, right? Why? It is highly preferred... 127 people answered this [mathematics]. Then what do you use here? This [mode].

[Ali 14.05.2018- 15:00-22:30]

While Ali explained why use mode for a given context, he focused on the meaning of "average" and found a typical value. He briefly mentioned they needed to find a typical value but did not discuss using arithmetic mean or median. Ali did not focus on the reasonableness of using mode for finding the typical value of a given context rather than using the arithmetic mean or median. Therefore, I scored Low for the segment.

I gave a Mid score to 30 (49.18%) of the lesson segments for mathematical and statistical sense-making. All lessons included at least one segment the teacher and students focused on as Mid-level or above mathematical and statistical sense-making. Ali helped students make sense of statistical concepts such as graphs and averages more than briefly. For example, he connected the context and operations while

focusing on average concepts' meanings. He also encouraged students to read the graph and understand the relationship among values during constructing circle graphs (e.g., focusing on the multiplicative relationship between degrees (or percentages) and quantity). I did not score high because most segments included several instances that were not sustained or a long instance that was a part of practicing a related concept or procedure.

I assigned a score of High to 15 (24.59%) of the lesson segments for mathematical and statistical sense-making. In the segments, Ali and the students made sense of statistical concepts or procedures in a sustained way. All segments with a score of High for explanations (4.92% of lesson segments) were also scored High in sense-making for Ali's instruction. For example, when introducing average concepts, Ali explained why to use the arithmetic mean in general; he tried to make sense of the arithmetic mean with a conceptual focus. This instance was sustained and substantial for both explanation and mathematical and statistical sense-making codes. Moreover, I scored High for all segments in one lesson for sense-making. Ali used tasks that connected the context and the average concepts (mostly arithmetic mean) in that lesson. Also, he focused on the characteristics of the arithmetic mean by changing the context.

[on the board: 40,80,85,100, 50,75,20,80]

T: ... We have eight students. We will find the average of their exam results. You know the arithmetic mean... You added them all and divided them by the number of results. Okay, I will say something. Let us say I am distributing [exam] papers to 8 people. I made some mistakes with one student. Let us say one student's grade would be changed... It has changed from 85 to 90. It is decreased or increased. What do you think about whether the arithmetic mean, median or mode changes after the grade changes? Let us say this grade [85] changed to 90. I graded the answer sheet of this student wrong. Eighty-five will become 90. What do you think about whether the arithmetic mean would change? If so, why? Or would the median change? If so, why? The mode would change, and if so, why?

[Ali 10.05.2018- 00:00-07:30]

The discussion started with the extract above. The students and the teacher discussed whether three measures of central tendency changed if one value had changed. First, they changed one value slightly (85 to 90), and the number of participants has not changed. They tried to make sense of what changes (for median) or not (for median or mode). Then Ali asked what would change if the student who got 20 was sent to the

other class because of their low grade. He changed the number of students for this example. They discussed what happened to the arithmetic mean. In the following segment, he introduced a task where givens were one student's grade above the arithmetic mean, the number of students, and the arithmetic mean of the student's grades. Considering all examples above, the focus on average concepts' mathematical and statistical meaning, especially arithmetic mean, was substantial enough to score High on those segments.

4.1.1.2.2.4. Multiple Procedures or Solution Methods

Figure 4.10 shows the distribution of scores assigned to lesson segments for Multiple Procedures or Solution Methods (MPSM) to identify whether multiple procedures and solution methods are used and discussed during a segment. Ali did not use multiple procedures or solution methods for 47 (77.05%) lesson segments. In nine (14.75%) lesson segments, Ali and his students briefly mentioned a second method but did not discuss it. They discussed a second method at length but did not link or compare both methods in 5 (8.20%) lesson segments.

None of the segments were scored as High for MPSM because Ali's instruction did not include the unique features of MPSM (e.g., explicit connection and comparison among solution methods) described in the MQI (Hill et al., 2014). Ali used multiple procedures and solution methods to solve a circle graph task or construct a circle graph. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.31 out of 4, with a standard deviation of 0.62. This section provides the descriptions and classroom examples of each score assigned to the Multiple Procedures or Solution Methods code in Ali's statistics instruction.

I scored Not Present of 47 (77.05%) lesson segments for MPSM because Ali did not use multiple procedures and solution methods. Especially in parts where he discussed average concepts, he did not use multiple procedures and solution methods to find the average. Only one segment included multiple procedures for finding the median. While finding the arithmetic mean of a data set, he only used the standard algorithm (e.g., adding all values and dividing by the number of values) to find a data set's

arithmetic mean. In the “statistical content interview, “ he briefly mentioned a second procedure derived from “leveling interpretation.” When asked if there is another way to find the arithmetic mean, he said, “For example, I calculated grade average, grades are 70 85 90, ... you can decrease one grade by five from your mind...” Although he knew a second solution method, he did not introduce it. In the interview, he said not many students were not familiar with this method; 99% of students used the first one he mentioned.

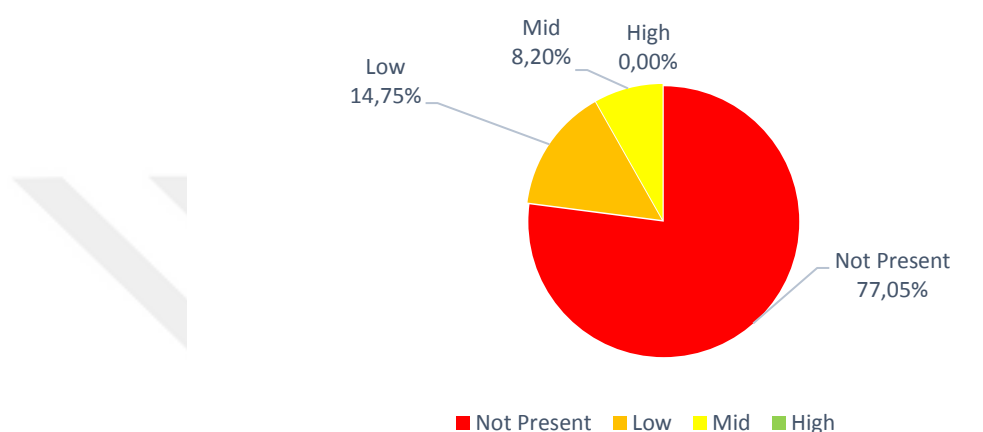


Figure 4.10 The Distribution of Scores Assigned to Lesson Segments for “Multiple Procedures or Solution Methods” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

I assigned a Low score for nine (14.75%) lesson segments for MPSM. In those segments, Ali briefly mentioned a second method but did not use it explicitly. For instance, in the circle graph task in Figure 4.4 (Graph Task 1) discussed in “linking between representation” and “explanation,” they found the class size and the number of Selim's votes. A student (Mehmet Ali) found the number of Arda's votes by subtracting the class size and Selim's votes on the board. In the end, Ali mentioned a second method briefly by saying that they could use multiple solution methods but used direct proportion (cross-product approach) in general. I did not score Mid for such segments because they did not use the second method during the segment.

I gave Mid to 5 (8.20%) lesson segments for MPSM. The teacher explicitly used more than one method in these segments but did not compare those methods or the

comparison brief. For example, while he was solving a second method, he stated that the second method was more straightforward than the other. Furthermore, he used multiple solution methods to find the components of a circle graph (e.g., the quantity, the angle of a sector, or the percentage). He used between ratio or within ratio and cross-product approach. As mentioned earlier, while finding the median of a data set with an even number of values (20,40,50,75,80,80,85,100), he discussed adding the two values in the middle and dividing by two. Then he asked students if there was another way to do it. They discussed the difference between two numbers and getting closer with the same amount in the middle. Ali stated that “the difference [between 80 and 75] is 5, it gets closer by 2.5 from there [75] since it is half of 5 and it gets closer by 2.5 from there [80]. [showing 77.5] what is the median, then?”. Then he continued to change a value in the data set to show whether the median changed. I scored Mid for the segment because he used two methods to find the median, but I did not score High because he did not compare and connect both methods.

4.1.1.2.2.5. Patterns and Generalizations.

Figure 4.11 shows the distribution of scores assigned to lesson segments for Patterns and Generalization code to describe to what extent the teacher and students develop generalization by examining patterns or examples. 59 (96.72%) lesson segments did not include a mathematical and statistical generalization development. Furthermore, the students did not examine patterns or construct a definition in the segments.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.03 out of 4, with a standard deviation of 0.18. This section provides the descriptions and classroom examples of each score assigned to the Patterns and Generalizations code in Ali's statistics instruction.

Only two (3.28%) lesson segments included working on a generalization or examining a pattern briefly. Those segments included patterns and generalization, mostly based on giving two or more examples and saying that the procedure used in those examples was valid for others. For example, in one segment, the teacher stated that they needed to do the same thing for all examples [for the circle graph examples]; they needed to

find the total and proportion [the whole and the part]. This situation lasted very briefly and was not the focus of the segment.

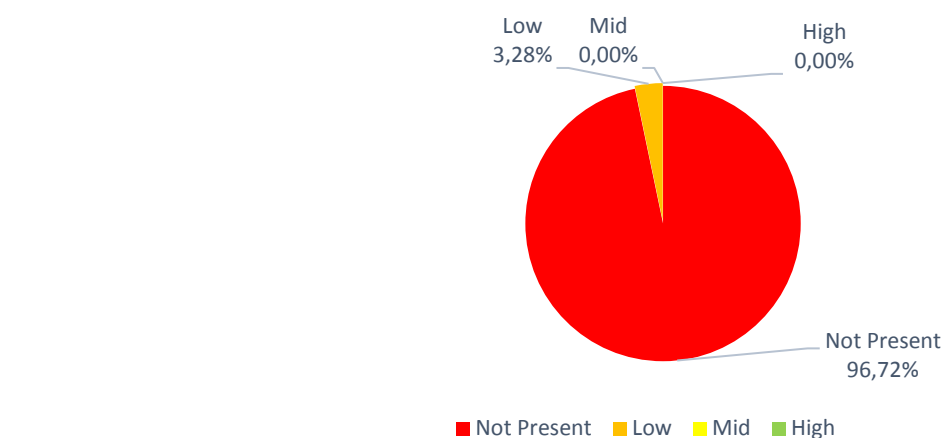


Figure 4.11 The Distribution of Scores Assigned to Lesson Segments for “Patterns and Generalization” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

Similarly, he discussed if the median or mode of a data set changed when a value changed with more than one example. In the end, he asked students, “Can I say that when a value has changed, the median does not ever change,” then, they concluded that it depended on the situation and discussed which condition could change the median. I also scored Low for this segment because this was a short work, although it contributed to constructing a statistical and mathematical definition based on examples and non-examples.

4.1.1.2.2.6. Mathematical and Statistical Language

Figure 4.12 shows the distribution of scores assigned to lesson segments for the language code to summarize Ali's use of mathematical or statistical language fluently and his support for students' fluent use of mathematical language. Ali's instruction included at least a low mathematical and statistical language level. In 12 (19.67%) lesson segments, a low mathematical and statistical language density was present. Four (6.56%) lesson segments had a high mathematical and statistical language density where Ali encouraged students to use it correctly and fluently. Lastly, Ali used

mathematical and statistical language at Mid density for 45 (73.77%) lesson segments because there were no unique characteristics mentioned in high density, but the teacher used mathematical language fluently.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.87 out of 4, with a standard deviation of 0.50. This section provides the descriptions and classroom examples of each score assigned to the Mathematical and Statistical Language code in Ali's statistics instruction.

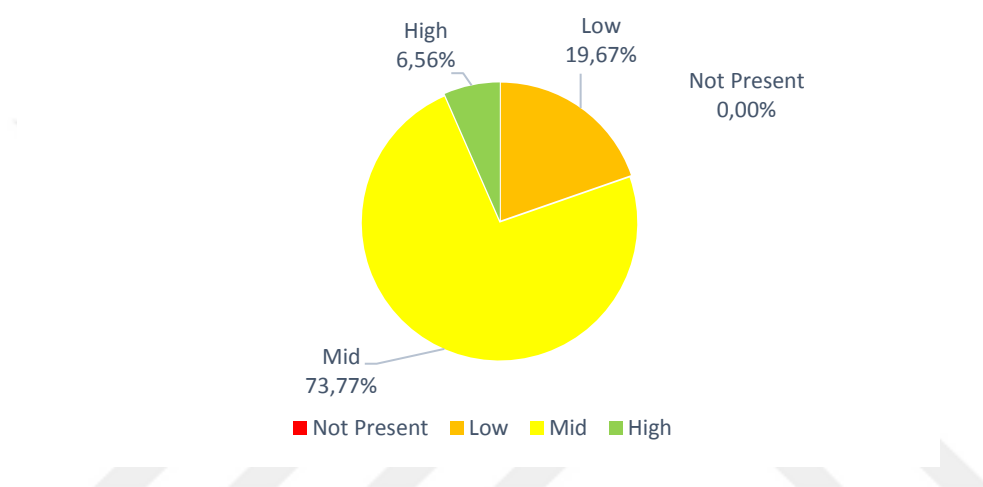


Figure 4.12 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Language” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

I scored Low on 12 (19.67%) lesson segments for mathematical and statistical language. Ali primarily introduced or clarified a task to the class or checked students' work during individual work time in those segments. The mathematical and statistical language density was low, or he used the same terms during those segments. None of the segments were scored as not present because all segments included mathematical and statistical terms, although there was some sloppiness in the language.

I scored Mid to 45 (73.77%) lesson segments in mathematical and statistical language because those segments included the middling density of mathematical and statistical language. Ali used mathematical and statistical terms correctly. The segments covered various statistical and mathematical terminology like the terms ratio, proportion, angle,

circle, circle graph, bar graph, line graph, sector of an angle, the total, cross-product method, degree, arithmetic mean, median, and mode. Especially those segments that did not include sloppiness in the language.

I scored High for only four (6.56%) lesson segments in which the mathematical and statistical language density was at least moderate. Those segments included the teachers' encouragement of students' appropriate mathematical and statistical language use. The main difference between the segments scored as Mid and High was teachers' support for using accurate mathematical and statistical language (Hill et al., 2014).

Furthermore, in one example, the teacher discussed the sample concept with students and wanted to hear the word “sample” from them. He encouraged students to use the correct terminology—one student defined the sample as the example selected “randomly” for a research study during the segment. Then the teacher wanted the student to repeat the definition and emphasized the importance of “randomness” by stating that “the word [randomness] is significant.”

4.1.1.2.2.7. Overall Richness of Mathematics and Statistics

This code describes the overall richness of mathematics and statistics of the instruction. With this code, the richness of most lesson segments in Ali's instruction can be regarded as low level ($n=40$, 65.57%) (see Figure 4.13). Richness elements were not observed in only one (1.64%) lesson segment within all lessons because students just started to work individually. The interaction among students, the teacher, and the content was observed minimally.

The richness codes were present with middling density in 17 (27.87%) lesson segments. Although those segments included High scores for some codes, most codes were at a moderate level, or the high score was considered isolated. Rich mathematics and statistics were present in three (4.92%) lesson segments, including the introduction to a circle graph and the characteristics of average concepts.

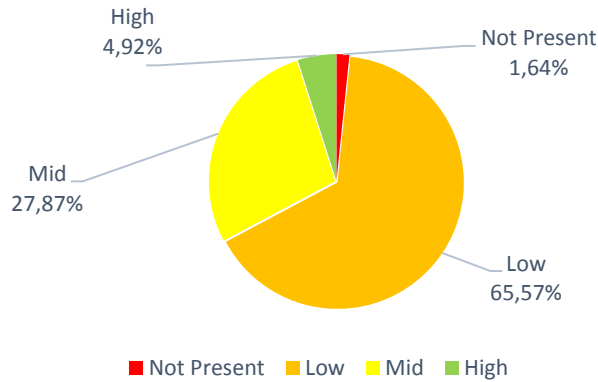


Figure 4.13 The Distribution of Scores Assigned to Lesson Segments for “Overall Richness of Mathematics and Statistics” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

The segments included High scores for explanations, mathematical and statistical sense-making, and mathematical and statistical language in common. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.38 out of 4, with a standard deviation of 0.61.

4.1.1.2.2.8. Overview of “Richness of Mathematics and Statistics” Dimension

To provide an overview of the Richness of Mathematics and Statistics dimension, a MAXQDA map was created to show the occurrences and co-occurrences of scores assigned to codes in a segment (see Figure 4.14). The MAXQDA Manual describes how this map is generated using the classic multidimensional scaling method to position codes based on the intersections in a segment (MAXQDA Manual, n.d.).

The distance between codes on the MAXQDA map is determined by a similarity matrix converted into a distance matrix. The MAXQDA Manual explains that the distance between codes is based on their co-occurrences in the data. When two codes always occur together, their distance is 0. When they frequently occur together, they are visually closer on the map. Conversely, when two codes never occur together, their distance is at its maximum. The analysis is based on the frequency of co-occurrences between codes in a segment.

The MAXQDA map shows the size of code symbols based on the frequency of codes in a segment. Codes with higher frequencies have larger symbols than those with lower frequencies. Overall, the MAXQDA map provides a clear visualization of the relationships between the MQI codes with their scores based on their co-occurrences in the data. The MAXQDA Manual explains the method used to generate the map in detail. It is important to note that MAXQDA maps created for the current study consisted of segment-level codes in the MQI with their assigned scores.

The MAXQDA map of the Richness dimension shows that the richness of mathematics and statistics of Ali's instruction could be characterized by low-quality to mid-quality (see Figure 4.14). The frequencies of Not Present scores assigned to Linking between Representations (LBR), Explanations (Exp), Multiple Procedures and Solution Methods (MPSM), and Patterns and Generalization (P&G) codes were the highest in the dimension. In addition, those scores assigned to the codes co-occurred with the highest frequency. In addition, Mid scores in Mathematical and Statistical Sense-Making and Mathematical and Statistical Language and the Low scores assigned to overall code were also present with the highest frequency in the segments. In addition, they co-occurred in the same segments and with the Not Present scores mentioned above.

The occurrences and co-occurrences indicate that in most segments, Ali focused on meaning more than briefly; he also used mathematical and statistical language at middling density. However, there were three issues in the richness dimension: linking between representations, multiple procedures or solution methods, and patterns and generalizations. Ali minimally linked between representations; most links stemmed from graph tasks. However, the majority of the links were not explicit or detailed. The nature of statistics instruction requires the teacher to use or construct graphs. If the teacher had used the statistical investigation process from the start and allowed students to collect or consider the data from a table and then construct the graph, the teacher might have had more opportunities to link the representations. In addition, if the teacher had used representations for summarizing data and distributions to show the center of the data, more links could have been connected.

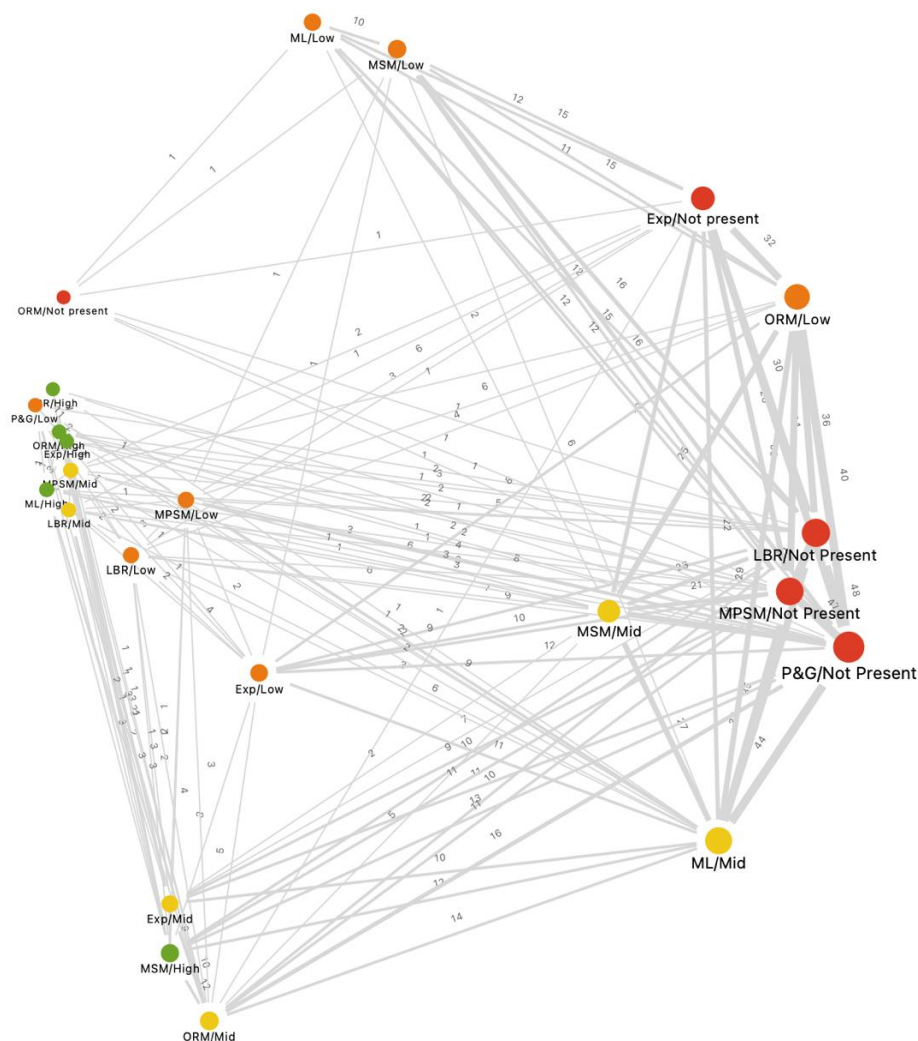


Figure 4.14 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “The Richness of Mathematics and Statistics” Dimension in Ali's Statistics Instruction

Note: LBR- Linking between Representations, Exp- Explanations, MSM- Mathematical and Statistical Sense Making, MPSM- Multiple Procedures and Solution Methods, ML-Mathematical and Statistical Language, P&G- Patterns and Generalizations, ORM- Overall Richness of Mathematics and Statistics

The second missing code was multiple procedures and solution methods. Although the teacher provided multiple procedures and solution methods in circle graph tasks, the remaining tasks did not include multiple procedures and solution methods. Especially in describing a data set, the teacher focused on one way to calculate central tendencies. The teacher asked for other ways to find the arithmetic mean in the Statistical Content Interview. He stated one more method, but the method was not well-formed and only

applicable to small data sets with normal distribution. Even if he worked on small data sets in the classroom, he did not provide other methods for calculating the arithmetic mean. Third, patterns and generalizations were absent in Ali's instruction. Statistics instruction, especially reading beyond data experiences, requires students to find the pattern in the data, and based on the given information, they are expected to make inferences. However, looking beyond the data was missing in the instruction.

Furthermore, the frequencies of high scores were low, the high scores assigned to richness codes generally interacted with each other, and the Mid scores assigned to the codes. The Low scores at LBR and MPSM co-occurred with mid and high scores in other codes. Those codes were mainly lacking in Ali's instruction. They interacted with Mid and High scores assigned to other codes when they were present. Therefore, in minor incidents, the teacher focused on meaning in a sustained way, and he pressed students to use mathematical and statistical language correctly. He explained why to use a procedure, choose a specific graph, or use a measure of central tendency based on context. However, the number of mathematical and statistical explanations varied in the segments. However, with a few segments, the explanations were the focus of the instruction. The teacher compared multiple procedures or solution methods in the segments and connected different representational families. Overall, the richness map shows that richness codes interacted at similar levels.

4.1.1.2.3. Working with Students and Mathematics and Statistics

In this section, the purpose is to investigate Ali's work with students over the content. Considering this aim, I examined this dimension with three codes in the MQI. Table 4.5 shows the results regarding working with students and the mathematics and statistics dimension in the MQI. Ali's statistics instruction can be characterized by low to mid-quality for understanding and responding to students' mathematical and statistical contributions.

Table 4.5 shows that Ali did not remediate student errors and difficulties in 20 (32.79%) segments, but he did not use student mathematical and statistical contributions in 13 (21.31%) segments. However, the Overall Working with Students

and Mathematics and Statistics code was scored as Not Present for only one (1.64%) segment and Low for 43 (70.49%) segments.

Table 4.5

“Working with Students and Mathematics and Statistics” Dimension Scores in Ali's Instructional Segments

	1		2		3		4		M	SD
	Not Present		Low		Mid		High			
	n	%	n	%	n	%	n	%		
Remediation of Student Errors and Difficulties	20	32.79	27	44.26	12	19.67	2	3.28	1.93	0.81
Teacher Uses Student Mathematical and Statistical Contributions	13	21.31	29	47.54	13	21.31	6	9.84	2.20	0.89
Overall Working with Students and Mathematics and statistics	1	1.64	43	70.49	16	26.23	1	1.64	2.28	0.52

Note. N=61 (the total number of lesson segments)

Furthermore, the majority of the segments were assigned a score of Low; respectively, Remediation of Student Errors and Difficulties (n=27, 44.26%), Teacher Uses Student Mathematical and Statistical Contributions (n=29, 47.54%), and Overall Working with Students and Mathematics and Statistics (n=43, 70.49%). All three codes included a small amount of the characteristics of high scores, but at least 20 percent of the lesson segments included mid-level quality for all codes in this dimension.

Figure 4.15 shows the remediation of student errors and difficulties, and the teacher uses student mathematical and statistical contributions. Overall, working with students and mathematics and statistics were primarily present in Ali's instruction.

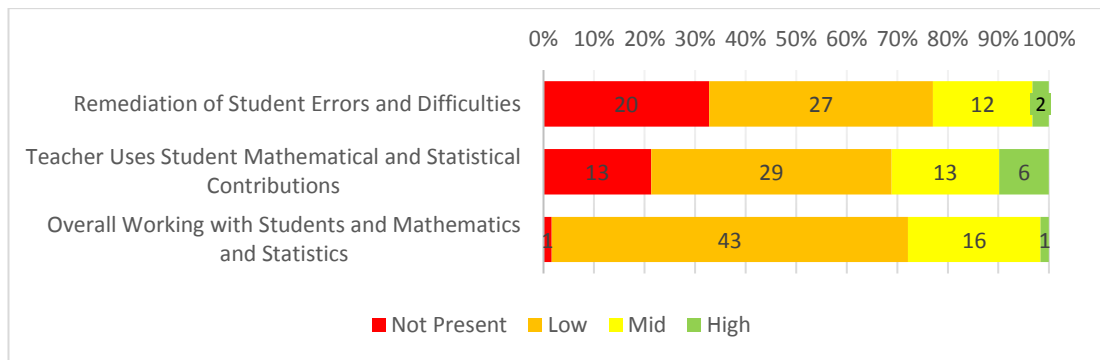


Figure 4.15 “Working with Students and Mathematics and Statistics” Dimension in Ali's Instruction

Note. N=61 (the total number of lesson segments)

The percentages or frequencies of low and mid scores for the remediation of student errors and difficulties code, and the teacher uses student mathematical and statistical contributions code are similar. However, the presences of high-quality instructional practices in terms of these codes were minimal. The details of each code in the Working with Students and Mathematics and Statistics dimension are presented in the following sections.

4.1.1.2.3.1. Remediation of Student Errors and Difficulties

Figure 4.16 shows the distribution of scores assigned to lesson segments for the Remediation of Student Errors and Difficulties code to explore to what extent the teacher remediated student errors and difficulties. Also, Ali's instruction did not include remediation in 20 (32.79%) lesson segments because the segments did not include students' errors or difficulties, or the teacher did not choose to remediate procedurally or conceptually. In addition, most segments included procedural or conceptual remediation briefly (n=27, 44.26%) or at a moderate level (n=12, 19.67%). However, only two (3.28%) segments covered systematic and at-length remediation of student errors and difficulties.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.93 out of 4, with a standard deviation of 0.81. This

section provides the descriptions and classroom examples of each score assigned to the Remediation of Student Errors and Difficulties code in Ali's statistics instruction.

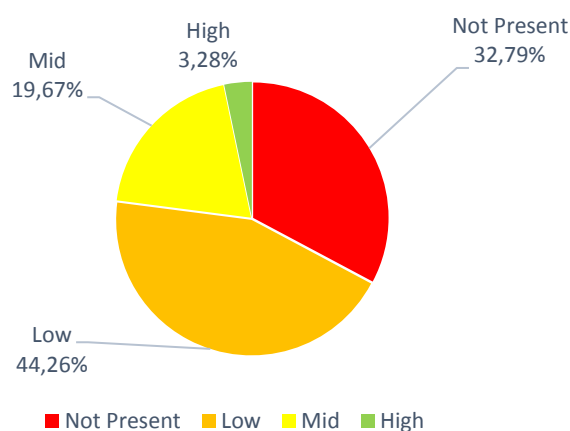


Figure 4.16 The Distribution of Scores Assigned to Lesson Segments for “Remediation of Student Errors and Difficulties” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

I assigned a score of Not Present for 20 (32.79%) lesson segments for remediation of student errors or difficulties. In some segments, student misconceptions were not observed because the teacher talk was dominant, or the students worked individually and silently. Some examples included that the teacher only said what was correct or incorrect, especially left the incorrect ones in the classroom discussion section, which was in the following section. In a specific example, the students solved a graph task asking the proportion of students in two classes based on the circle graph. The students stated they had difficulties because the proportion concept was hard to understand. They also said there were no numbers to solve the task. The teacher repeated what the students said, gave them time to solve it, and did not interfere with their work for a while. At the end of the segment, he asked students why the task did not include numbers, and when students could not give correct or complete answers, he clarified why the task did not include the “number” and why they did not need to use numbers. However, the remediation part was in the following segment, and it was procedural. The teacher did not choose to remediate immediately, so I scored not present for remediation for this segment.

I gave a Low score to 27 (44.26%) lesson segments for the remediation code. Ali mostly remediated students' difficulties or misunderstandings procedurally, especially during individual work time. While students worked on the questions individually, he told them to ask the questions they had difficulty with or observed a student struggling. He corrected students' procedures and explained the procedures step by step but did not mainly focus on the conceptual aspect. While discussing why and where they used average concepts, some students gave examples not suitable for arithmetic mean:

S: [We are using arithmetic mean] while giving Islamic charity fee (Fitre)

T: How do you calculate the arithmetic mean there?

S: You are dividing a number by the number of [individuals in] family

T: Is there no money per individual? You are a 4-individual family, and then you will multiply, for example, 50 lira, 50 lira, 50 lira, and 50 lira. It is not average in that there is a donation per individual.

[10.05.18 Ali 00:00-07:30]

The extract shows that it was a brief but conceptual instance. Independent from how the authority determined the donation amount, the individuals donate money based on the number of individuals in the family and the donation amount. The example from the extract was simply a calculation of the total amount of money. Ali explained the concept of donation. However, if the teacher explained how the average differed from calculating the total amount given, the remediation would be scored as mid or above for this and the following segments.

I scored Mid to 12 (19.67%) lesson segments for remediation. There was moderate level conceptual or extensive procedural remediation in some segments. For example, he warned students to consider the total number while working on a circle graph task individually, which was short conceptual remediation. In addition, he did extensive procedural remediation for the same task. He solved the task on board with a student for those who could not solve it while solving a question asking to show the distribution of a mixture for making a jar into a circle graph. A student drew the circle graph by starting with the quantities [i.e., 8 kg strawberry, 3 kg sugar, and 1 kg water]. In this segment, pre-remediation was present. The teacher warned students about plausible mistakes or missing parts. For example, the student did not consider the angles of the sectors corresponding to the given data; he tried to estimate the size of a sector with given numbers and totals. Then the teacher anticipated that other students

might have made the same mistake, so he immediately warned the class not to make similar errors.

Moreover, in Graph Task 2 (see Figure 4.5), he asked students to find the mode of the given data set on the bar graph and warned students by saying that “Do not say 100; the mode does not mean the highest number”. He anticipated possible confusion due to the translation of the mode concept in Turkish and warned students about it. According to MQI, when pre-remediation occurs, the minimum score given for the segment is Mid for remediation. The pre-remediation parts were not extensive, so I assigned a Mid rather than a High score.

I scored high for only two (3.28%) segments for the remediation code. Those segments included the teacher's actions to understand the source of student difficulties, errors, and misconceptions systematically and extensively. In one segment, the teacher anticipated a possible misunderstanding among the students regarding a line graph task, even though the students selected the correct answer (see Figure 4.17). The teacher asked follow-up questions to ensure the students understood the reasoning behind the correct answer, which they could not explain.

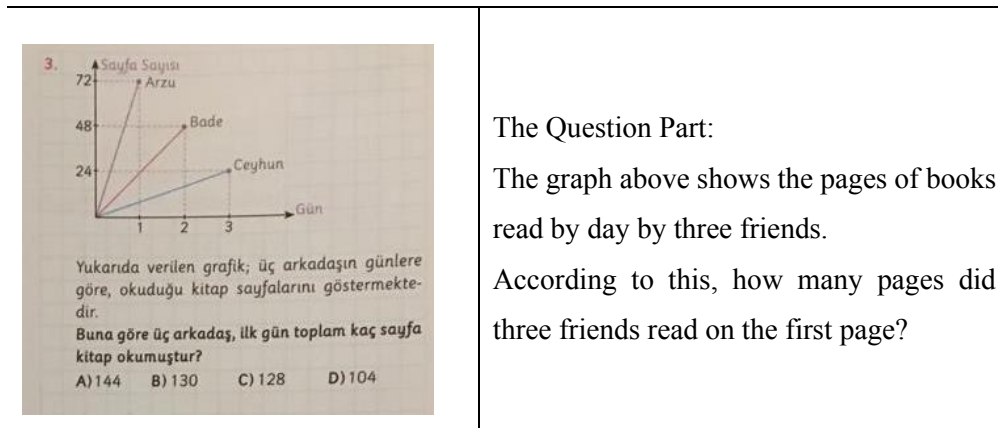


Figure 4.17 Graph Task 4 (Varışlı, 2017, p.155)

The problem involves calculating the number of pages read by a student named Ceyhun on three consecutive days. Initially, the students assume that Ceyhun reads eight pages each day. However, Ali claimed it did not have to be eight pages each day; the first could be 15 pages, and the second could be something else. To demonstrate

this, a student provided different numbers that added up to the exact total of 24. However, some students still assumed that the number of pages read each day must be equal, while another student pointed out that the increase in pages read is not linear and could not be eight pages, but they could assume that it could be eight pages.

The teacher's remediation approach included pre-remediation and extensive conceptual explanation to help the students understand the problem better, with a particular focus on the concept of linear graphs. This approach demonstrates Ali's commitment to ensuring students understood the concepts thoroughly and addressing potential misunderstandings before they became more significant problems. Specifically, the teacher's actions suggest a systematic and extensive effort to understand the source of and remediate student difficulties, errors, and misconceptions.

4.1.1.2.3.2. Teacher Uses Student Mathematical and Statistical Contributions

Figure 4.18 shows the distribution of scores to lesson segments for the “Teacher Uses Student Mathematical and Statistical Contributions” code, describing to what extent Ali used student mathematical and statistical contribution. In 13 (21.31%) lesson segments, the teacher did not use student mathematical and statistical contributions or student responses. Teacher’s limited use of student contributions was present in 29 (47.54%) lesson segments. In addition, 13 (21.31%) lesson segments were scored as Mid because the teacher responded to students' contribution to moving the instruction forward with some characteristics of high-quality indicators. Lastly, six (9.84%) lesson segments included teacher actions of hearing student contributions and helping students to make more contributions through elaborating on student ideas.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.20 out of 4, with a standard deviation of 0.89. This section provides the descriptions and classroom examples of each score assigned to the Teacher Uses Student Mathematical and Statistical Contributions code in Ali's statistics instruction.

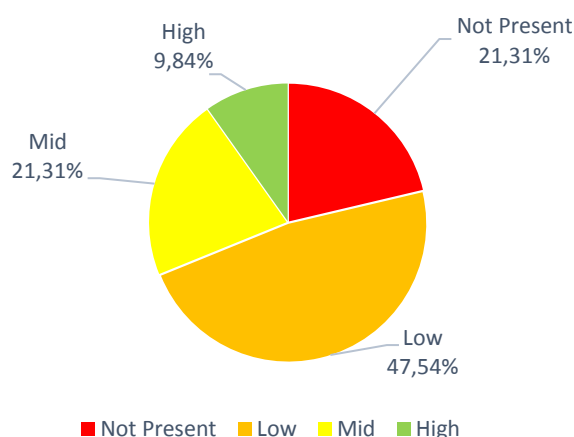


Figure 4.18 The Distribution of Scores Assigned to Lesson Segments for “Teacher Uses Student Mathematical and Statistical Contributions” in Ali's Instruction

Note. N=61 (the total number of lesson segments)

I gave a score of Not Present for 13 (21.31%) lesson segments for teachers using student mathematical and statistical contributions. Those lesson segments did not include student responses because the students worked individually on tasks. As described earlier, some individual work time included the teacher's interaction with students individually. However, the teacher helped students solve or clarify the questions to the students. The one-to-one sessions were dominated by teacher talk. Additionally, in some segments, Ali reviewed or introduced a topic or a task to the students, so he did not give opportunities to the students to contribute mathematically.

I scored Low on 29 (47.54%) lesson segments because students contributed, but Ali used student mathematical and statistical contributions to develop students' mathematical and statistical ideas in a proforma way. He expected students to give one or two-word answers when he clarified a question or a solution path. For example, the circle graph task, which was asking to find the central angle of a sector for representing a class with 30 students, including 20 girls” [presented in “linking between representation” code section]. He used student contribution to some extent:

T: How many people in total?
Multiple students: 30 people
T: How many girls?
Multiple students: 20

S: Teacher with the cross-product method...

T: Boys?

Multiple students: 10

T: According to this, what did I say just before? There are 30 people, right? Which degree is equal to a total of 30 people?

S: 360

T: Then, what would be the number of people if it is like that?

The teacher presented a way to solve the graph task with student contributions. After two to three minutes of giving time for students to work on the solution individually, he asked a student to solve it on board, then the student solved the task on board. However, the contribution of the student who solved the question remained nonverbal. While solving the task, the teacher asked the whole class the steps:

T: I found the total. How many people [are there in total]?

Multiple students: 30

T: The total is 30 people. To what degree corresponds to 30 people?

Multiple students: 360

T: All of them, right? 30 [corresponds to] 360 degrees; what do I want to find? What to find first?

Multiple students: Boys

T: Let us find boys, few people, okay guys, ... there is a relationship between these two, divided by 3, put division sign, divide by 3, and okay, divide 360 by 3, okay? X equals 3. To what degree [the number of] do boys correspond?

Multiple students: 120

[Ali 30.04.18-1 07:30-15:00]

The teacher asked answer-bound questions to the class and heavily scaffolded the student on board. He did not verbally let the student on board contribute to the mathematics and statistics. All in all, Ali incorporated student contributions into the instruction, but his responses did not fully engage with the student's ideas. Rather, he was primarily focused on his own agenda and used the student contributions more as a matter of formality than as a means of enhancing the learning experience.

The use of student contribution to the development of mathematics and statistics can be characterized as Mid for 13 (21.31%) lesson segments. Especially discussing the tasks where he introduced a new concept or topic, he tried to explore students' ideas; thus, students made contributions mathematically and statistically. He moved the instruction based on students' contributions in these segments. In some parts, he took

a misunderstanding or an erroneous idea to develop the mathematics and statistics of the lesson. For example, Ali asked the meaning of data analysis to the whole class, and the students described it with one word like controlling, reviewing, observing, and investigating. Ultimately, the teacher summarized what data analysis meant with students' contributions. If the teacher asked a student to comment on or expand the given answers or gave credit to students while explaining the definition, the segment might have scored High.

Furthermore, in another example, the teacher asked students to transform a line graph into a circle graph. A student drew the circle graph on the board. While constructing, the teacher asked the whole class about the relationship between degrees and quantities. The students gave answers, and the instruction moved forward with student contributions. The student on board answered at some point, finding a portion of the circle graph. The teacher heard what Buğra said, directing the class by stating, "Buğra said we need to multiply all of them by 2,... 55 will be multiplied by 2, 35 will also be... Why? What? What? What is the reason behind it?". However, he did not give students much time to think. He explained the reason to the class. Therefore, I assigned a Mid rather than a High score for such an instance.

I scored a High for 6 (9.84%) lesson segments because Ali took up student contributions and developed mathematics and statistics based on those contributions. For example, when introducing average concepts, he gave a small data set and asked students why he wonders about the arithmetic mean of the data set. Students provided a wide range of responses, and the teacher repeated some and elaborated on other answers. Then he acknowledged Hamza's answers by directing them to the classroom, "Hamza said that we could find the average of three classrooms, and we will be looking at which classroom's grade is better [than others]." He identified key ideas in Hamza's statement and asked a follow-up question.

Furthermore, in the following lesson segment, Ali introduced the median concept, and a student said, "There are five numbers, and the median is the number in the middle." During the discussion, a student asked, "What if there are six numbers?" Ali highlighted student questions by saying, "Mustafa asked a very nice question." In the

following part, he asked students to find the median of six numbers by referring to Mustafa's questions. These segments included Ali's use of student ideas to develop mathematics and statistics extensively and at length.

4.1.1.2.3.3. Overall Working with Students and Mathematics and Statistics

Figure 4.19 shows the distribution of the scores assigned to lesson segments for the Overall Working with Students and Mathematics and Statistics code. Ali's instruction can be characterized as a Low level of working with students and mathematics and statistics because 43 (70.49%) lesson segments included at least a little take-up of students' mathematical and statistical contributions or short remediation. In addition, only 1.64% of lesson segments, one segment in all lessons, did not include working with students, and mathematics and statistics because the teacher's use of student contributions was off-track.

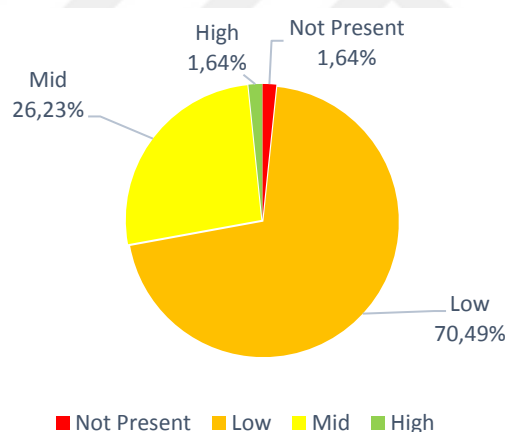


Figure 4.19 The Distribution of Scores Assigned to Lesson Segments for “Overall Working with Students and Mathematics and Statistics” in Ali's Instruction

Note. N=61 (the total number of lesson segments)

The codes describing working with students and mathematics and statistics were present at the Mid-level in 16 (26.23%) lesson segments (see Figure 4.19). Those segments included extensive teacher and student interaction over content, but Ali remediated students' misunderstanding conceptually at a moderate level or procedurally at an extended level. Only one (1.64%) segment was given a High score

because the teacher interweaved student ideas by discussing the meaning of randomness and conceptually remediated students' misconceptions of randomness. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.28 out of 4, with a standard deviation of 0.52.

4.1.1.2.3.4. Overview of “Working with Students and Mathematics and Statistics” Dimension

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the frequencies of occurrences and co-occurrences of Working with Students and Mathematics and Statistics codes with their scores as visualized on the MAXQDA map (see Figure 4.20)

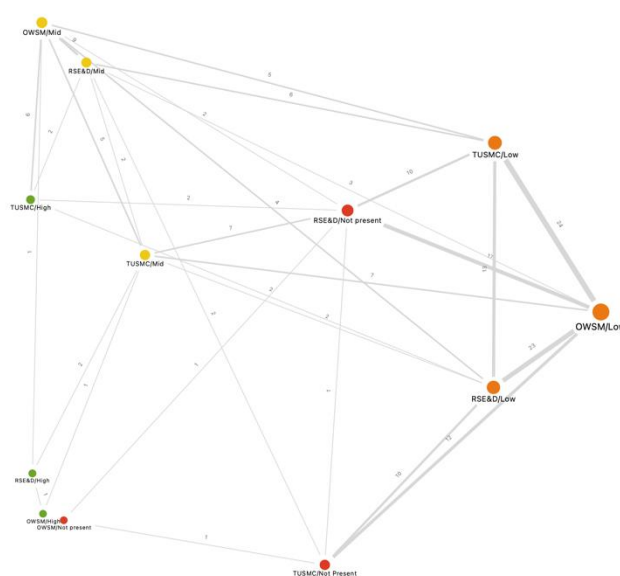


Figure 4.20 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “The Working with Students and Mathematics and Statistics”

Dimension in Ali’s Statistics Instruction

Note: RSE&D- Remediation of Student Errors and Difficulties, TUSMC- Teacher Uses Student Mathematical and Statistical Contributions, OWM- Overall Working with Students and Mathematics and Statistics

Ali’s work with the student and the mathematical and statistical content could be characterized by low quality. The frequencies (i.e., the size of codes) of Low scores

assigned to Remediation of Student Errors and Difficulties (RSE&D), Teacher Uses Student Mathematical and Statistical Contributions (TUSMC), and Overall Working with Students and Mathematics and Statistics (OWM) codes were the highest in the dimension. They also co-occurred in the same segments with the highest frequencies (i.e., the width of the connected lines). I observed that the segments included dominant teacher talk or individual work time. Student contributions were brief and taken up proforma, and teacher remediation remained brief and procedural in general. Therefore, Ali's statistics instruction was described as Low quality in working with students over mathematical and statistical content.

The interactions between the codes with Low scores and those with Not Present scores were also present (see Figure 4.20). In those segments, the teacher used student contribution with no remediation. Also, there were student contributions, but the teacher did not use those contributions to develop mathematics and statistics but briefly remediated student ideas. Moreover, the codes with Low scores also interacted with the codes with Mid or High scores; but those co-occurrences were less than those with Low and Not Present scores. In those incidents, the students contributed mathematics and statistics to the lesson; conceptual remediation was present. However, the teacher take-up student contributions were present briefly. Alternately, those segments included teacher remediation procedurally, but there was extensive use of student contributions; the teacher used student contributions to develop mathematics and statistics.

Furthermore, the frequencies of the codes with Mid and High scores on the codes of Working with Students and Mathematics and Statistics dimension were less. Thus, the codes with High scores occurred together less frequently; and the prevalence of the co-occurrences of the codes with Mid scores was also fewer. However, the codes with High scores and Mid scores co-occurred in general. Therefore, they clustered together. Infrequent incidences were assigned the codes with High scores that occurred with the codes with Low or Not Present scores. Especially, a segment with a High score in TUSCM co-occurred with a Not Present score in RSE&D. The reason was that there was no student difficulty observed during the segments. As a result, the map shows

that working with students and mathematics and statistics codes co-occurred at similar levels in most segments.

4.1.1.2.4. Errors and Imprecision

This section aims to investigate the problematic parts of Ali's statistics instruction, mainly his mathematical and statistical errors and imprecision. I examined this dimension with four codes in the MQI. Table 4.6 shows the Errors and Imprecision dimension results in the MQI. It is necessary to note that higher scores mean more problematic aspects in the instruction, and low scores are less problematic or error-free than other dimensions. I reversely assigned points to the scores while calculating the arithmetic mean and standard deviation to align with other dimensions. According to Table 4.6, most of Ali's statistics instruction can be characterized as "error-free."

Table 4.6

The "Errors and Imprecision" Dimension Scores in Ali's Instructional Segments

	4		3		2		1		M	SD
	Not Present		Low		Mid		High			
	n	%	n	%	n	%	n	%		
Mathematical and Statistical Content Errors	49	80.33	7	11.48	3	4.92	2	3.28	3.69	0.72
Imprecision in Language or Notation	54	88.52	6	9.84	1	1.64	0	0.00	3.87	0.39
Lack of Clarity in Presentation of Mathematical and Statistical Content	49	80.33	8	13.11	2	3.28	2	3.28	3.70	0.69
Overall Errors and Imprecision	36	59.02	17	27.87	5	8.20	3	4.91	3.41	0.84

Note. N=61 (the total number of lesson segments)

Table 4.6 shows that Ali made no mathematical and statistical content errors in 49 (80.33%) lesson segments; he did not use imprecise language or notation in 54 (88.52%) lesson segments. Furthermore, Ali's presentation was clear for 49 (80.33%) lesson segments. Errors and imprecision were absent for 36 (59.02%) lesson segments.

Mathematical and Statistical content errors were present in 12 (19.67%) lesson segments, but the distortion of mathematics and statistics ranged from low to high. Similarly, only seven (11.48%) lesson segments included imprecision in teacher language or notation. However, none of the segments included a high imprecision in language. In addition, 12 (19.67%) lesson segments lacked clarity in the teacher's presentation from low to high. Since the overall code considered errors or imprecision in the first three codes in this dimension, 25 (40.98%) lesson segments were assigned at least a Low score for errors and imprecision.

Figure 4.21 shows that the first three codes had similar frequencies for being free of errors, imprecision, and lack of clarity in Ali's instruction. Major errors, imprecision, or lack of clarity (above Mid) were also minimally present.

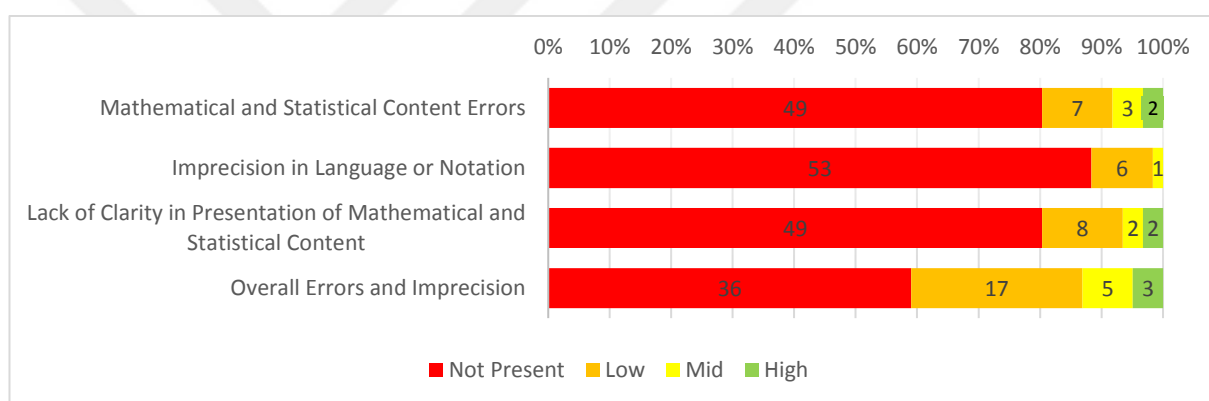


Figure 4.21 “Errors and Imprecision” Dimension in Ali's Instruction

Note. N=61 (the total number of lesson segments)

According to the Overall Errors and Imprecision code, some segments included at least one Low or above score. Not present scores were lower for overall code than other codes in this dimension. Still, 59.02% of lesson segments did not include any errors, imprecision, or lack of clarity. Therefore, Ali's statistics instruction was mainly error-free, precise, and clear. The details of the Errors and Imprecision dimension codes will be presented in the following sections.

4.1.1.2.4.1. Mathematical and Statistical Content Errors

Figure 4.22 shows the distribution of scores assigned to lesson segments for the Mathematical and Statistical Errors code, describing to what extent Ali made content errors while teaching statistics. Ali did not make content errors in 49 (80.33%) lesson segments. Seven (11.48%) lesson segments included a brief content error. In addition, three (4.92%) lesson segments contained mathematically incorrect parts, so I scored Mid. Ali made content errors in most or all two (3.28%) lesson segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.69 out of 4, with a standard deviation of 0.72. This section provides the descriptions and classroom examples of each score assigned to the Mathematical and Statistical Content Errors code in Ali's statistics instruction.

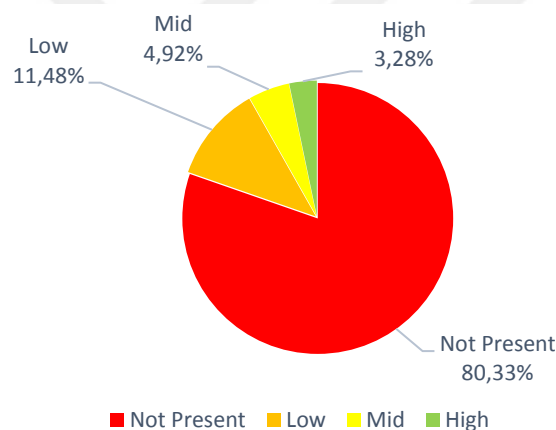


Figure 4.22 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Content Errors” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

I scored Low on seven (11.48%) lesson segments because Ali made brief content errors. Ali consistently defined the concept of mode as “the most frequent number” in different lessons or lesson segments. In the definition, Ali forgot a key condition in a definition (Hill et al., 2014). Although the definition was correct for the given contexts, he also provided examples with nominal or ordinal data sets following those contexts. When he defined with the focus of “number,” not “value,” then his definition of mode would be incomplete and somewhat incorrect. For example, Ali asked students to

answer a question on a worksheet asking which graph they should use to compare two classes' math and science courses' grade average. Then they briefly said that since there was a comparison, they needed to use a bar graph. However, using a bar graph to compare frequencies, not averages, would be more appropriate. Therefore, this segment included an incorrect but brief instance that did not obscure the mathematics and statistics of the segment.

I assigned a score of Mid to three (4.92%) lesson segments for mathematical and statistical content errors. Those segments included tasks requiring students to draw a line graph of a given context. The teacher introduced the line graph with the examples in the textbook, which were not wholly correct representations of the given data.

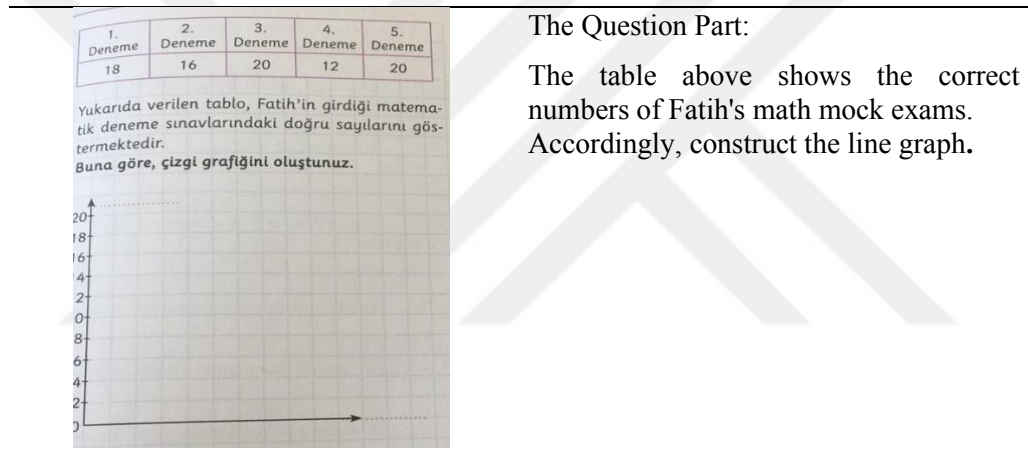


Figure 4.23 Graph Task 5 (Varışlı, 2017, p.162)

For example, the task in Figure 4.23 asked students to construct a line graph of Fatih's number of corrects in math in the mock exams according to the given table. The table in the question included the number of corrects in math for the first exam to the fifth exam. While introducing the line graph, the teacher emphasized the change in time for line graphs. However, in the example, there was no explicit time order. Although the mock exams were stated as ordinal numbers, the example gave no information about the time interval. The slope of the line drawn among two consecutive correct points by mock exams might have differed, so that this situation might lead to misinterpretation of the graph. We scored Mid rather than high because statistically incorrect tasks

obscured the mathematics and statistics of only a part of the segment. In the following parts of the segment, the teacher correctly used a line graph for another context.

I scored High in two (3.28%) lesson segments for mathematical and statistical content errors. While constructing a line graph in a segment, the teacher devised an example that included a student's (called Yiğit) grade point average in mathematics by years. The context was not appropriate for drawing a line graph. Even if the purpose was to observe the change in grade averages in time, the time is not continuous for the context. Therefore, when a point was selected in a line between two data points, the point did not yield meaningful data from the x and y axes. I scored High because the content error occurred during the whole segment. Furthermore, the second segment included an example from a textbook. The teacher asked students to find the mode of the given data "4,4,4,4,4,4". The students provided some answers to the teacher:

T: What is the mode of this?

S1: 4

S2: 0

S3: No [mode]

S: [Does the question] want to say if there is no number other than 4, teacher?

T: Does anyone say 4? Does anyone say that this has a peak value? There is. Okay. Anything else from 4?

S5: No [mode]

T: Why not?

S5: [because] no numbers other than 4

S6: I do not think so

T: Why not?

S5: Teacher, there is no number other than 4

T: What do you think?

S6: I do not think so. For example, if four were four, six, and five, there would be four directly.

T: You say it would. You tell

S7: I think 20 teachers

T: Why 20?

S7: I collected them all because

[Ali 10.05.02-2 30:00-36:30]

After discussing the mode, the teacher did not approve or disapprove of the student's ideas, and he left the incorrect answer about adding all the values unexplained. In the following part of the segment, he provided a context for the data set. According to the data set, he asked students the mode of the data set by constructing a bar graph based

on the context (see Figure 4.24). In the end, he concluded that there was no mode in the data set.

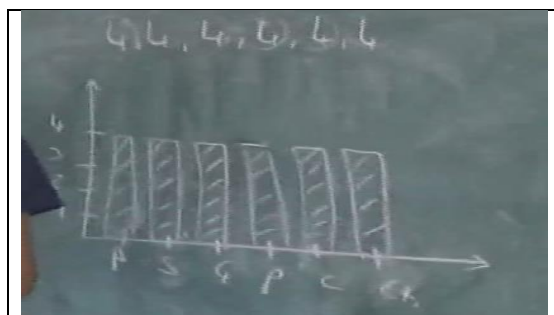


Figure 4.24 Teacher Construction of a Graph of a Data Set “4, 4, 4, 4, 4, 4”

T: ... Well, let us do something like this. Let us see what happens now on the graph. Let us say these are the number of students who come to ask me questions. Monday 4, Tuesday 4, Wednesday 4, Thursday 4, Friday Saturday. 4 students came every day. These are the number of students (y-axis), who will be here Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday (x-axis)

S1: no Sunday

S2: Sunday is off.

T: Let us take a look at the chart. What is the peak value, i.e., mode, in this graph?

S3: No [mode]

S4: How could it be, teacher?

T: Are they not all the same? Did someone make a difference?...

[Ali 10.05.02-2 30:00-36:30]

According to the definition of mode, the “no mode” conclusion was incorrect. Furthermore, considering the context and the graph, the data set did not include numbers (4s) but the name of days. The mode of the data set shown in the graph could be all the days mentioned because they all occurred at the same frequency. The graph had a kind of multimodal distribution. Therefore, the extract and the graph showed that the data contextualization and visualization were also incorrect and obscured the mathematics and statistics of the segment.

4.1.1.2.4.2. Imprecision in Language or Notation

Figure 4.25 displays the distribution of scores assigned to lesson segments for Ali’s use of inappropriate mathematical and statistical language or notation. In 54 (88.52%) lesson segments, Ali did not have language or notation imprecision. Six (9.84%) lesson

segments included low-level imprecision in which there was a brief instance in mathematical language, but this instance did not obscure the mathematics and statistics of the segment. Only one (1.64%) lesson segment included imprecisions in language or notation for a part of the segment. None of the lesson segments was assigned a High score. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.87 out of 4, with a standard deviation of 0.39. This section provides the descriptions and classroom examples of each score assigned to the Imprecision in Language or Notation code in Ali's statistics instruction.

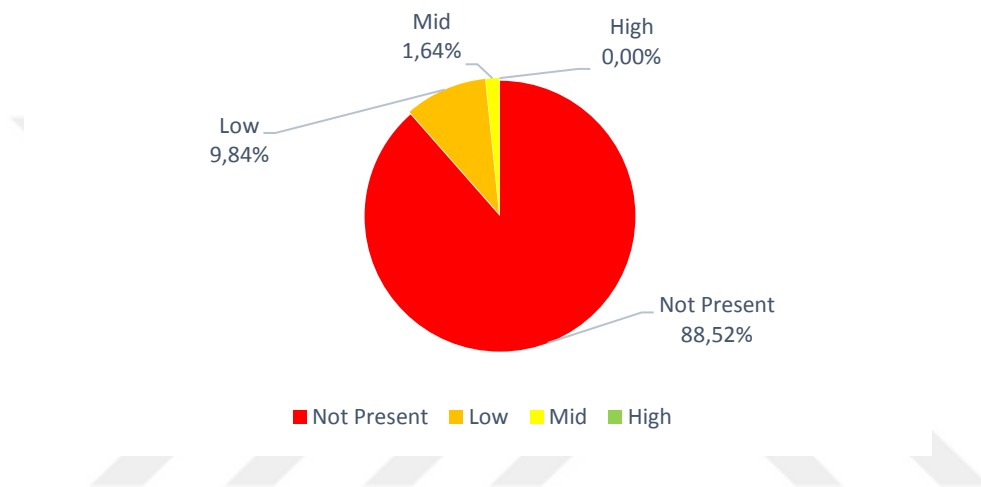


Figure 4.25 The Distribution of Scores Assigned to Lesson Segments for “Imprecision in Language or Notation” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

The lesson segments were scored Low (n=6, 9.84%) and Mid (n=1, 1.64%), including Ali’s use of imprecise language covering average concepts. Ali explained three kinds of averages: arithmetic mean, median, and mode. He also explained the average meaning to the students, like finding a typical value. However, sometimes Ali asked students to find the average instead of the arithmetic mean. This error was about the Turkish language; average means “ortalama” and “arithmetic mean” also means “aritmetik ortalama.” The arithmetic mean is generally called mean “ortalama” as well. Mathematical and statistical language imprecision occurred when the teacher called both concepts with the exact wording. However, in most lesson segments, imprecise language was a brief instance that did not obscure the mathematics and statistics of the segment.

Only one segment with such an instance was scored as Mid because the task asked students to find the mathematics average of a class in the segment. Before the task, they mentioned three kinds of averages as well. The teacher emphasized that it was asked to find the arithmetic mean. The imprecision obscured the mathematics and statistics of a part of the segment, so the teacher clarified which one to use for students.

4.1.1.2.4.3. Lack of Clarity in Presentation of Mathematical and Statistical Content

Figure 4.26 shows the distribution of scores assigned to lesson segments for Lack of Clarity in the Presentation of Mathematical and Statistical Content code, describing the teacher's mathematical and statistical practices as unclear for students. Ali's practices did not include a lack of clarity in 49 (80.33%) lesson segments. Eight (13.11%) lesson segments included a brief lack of clarity. In addition, Ali's teaching was confusing for a part of two (3.28%) lesson segments and most of the segments of two (3.28%) lesson segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.70 out of 4, with a standard deviation of 0.69. This section provides the descriptions and classroom examples of each score assigned to the Lack of Clarity in the Presentation of Mathematical and Statistical Content code in Ali's statistics instruction.

I assigned a Low score for eight (13.11%) lesson segments that included a brief lack of clarity, not making it difficult to understand mathematics and statistics. For example, Ali asked students to find the mode of a data set. The data set included one mode. While the students found the mode of the data set, Ali repeated the definition of mode. One student asked if more than one value had the same frequency. Another student claimed they would add both and divide by two [like median]. The teacher asked the mode of the data set: "1, 1, 3, 3, 7" Students gave answers like 3, 1, 2, and "1 and 3". The teacher concluded, "Look, here there are two modes, 1 and 3, okay?". There were a variety of correct and incorrect students' answers. Some students found the mathematical and statistical points confusing until the teacher stated the correct solution. Then students were engaged with such examples of distributions with no mode. The teacher pointed out that the distributions could have one mode, no mode,

or more than one mode. The mathematical and statistical points remained clear with those examples.

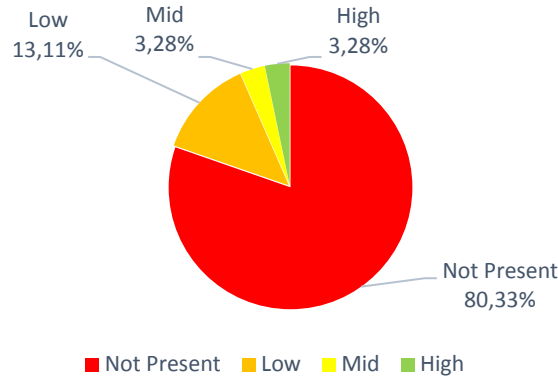


Figure 4.26 The Distribution of Scores Assigned to Lesson Segments for “Lack of Clarity in the Presentation of Mathematical and Statistical Content” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

I scored Mid to two (3.28%) lesson segments for lack of clarity in presenting mathematical and statistical content. One lesson segment asked about the “mathematical average” discussed in the Imprecision in Language or Notation section. In the lesson segment, Ali did not use clear mathematical and statistical language, so the mathematical and statistical point was muddled; it became difficult for students to understand. In the other segment, Ali introduced a task asking which average could be used to get information about the number of students' shoes in a school. The task stated that the data was collected from students by asking, “How many shoes do you have” and a frequency table was provided (see Figure 4.27).

Ayakkabı Sayısı	1	2	3	4
Öğrenci Sayısı	41	96	23	12

Bir okulun öğrencilerine “Kaç tane ayakkabınız var?” sorusu sorulmuş ve yukarıdaki tabloda verilen sonuçlar elde edilmiştir. Buna göre, okulda ayakkabı sayısı hakkında bilgi edinmek için hangi ortalama çeşidine bakmamız gerekir?

The Question Part:

The question “How many shoes do you have?” was asked of the students at a school, and the results were given in the table above. Accordingly, what kind of average should we use to learn about the number of shoes in school?

Figure 4.27 Average Example: The Number of Shoes (Varışlı, 2017, p.162)

Ali asked students which average concept should be used. Most of the students said “mode.” Nevertheless, one student said median. Ali did not ask for an explanation of the student’s answers. However, he made an explanation for choosing “mode.”

T: Everyone says mode. See when we use mode. I am trying to identify one distinct, typical feature within a group. Here is the one that shows me the most [circling 96 and 2 parts]. There is this. For example, you called it median, right? What does the median give you here?

S: Does not give anything

T: It does not give anything. We will use the median in a moment. We will use the median when there is no arithmetic mean, okay? Do you understand this? When did you use the mode? Those with two shoes for a particular feature, or those with the following shoe size ...

[Ali 14.05.2018- 15:00-22:30]

The extract above made the statistical point in Ali’s explanation confusing. Ali defined that the purpose of the average concept or measures of central tendency was to find a typical value of the data set in previous lessons. However, while determining which average should be used for the given context, he said mode because the aim was to find a typical feature within a group. With this explanation, the median or arithmetic mean was also appropriate. Therefore, this section of the lesson segment lacked clarity regarding why to use mode.

Two (3.28%) lesson segments included a high-level lack of clarity. Both segments were consecutive and included discussions on the same tasks. In the example, the teacher said he constructed a circle graph about the computer games Ömer played for a week (see Figure 4.28). He wrote the names of the games and stated that he divided the graph and wrote the degrees of sectors. Then he asked students to draw a line graph. First, a student tried to create a line graph. The teacher and the students then discussed constructing the line graph with information on a circle graph.

The teacher asked students for ideas about what to use to find the values for the line graph. Some students tried to match days with games. A student suggested that they needed to use “time.” Then they discussed which time concept they should use; 24 hours or days. Some students tried to form a direct proportion by using 130 and 360 degrees. The teacher asked if there was another number other than degrees. At the end

of the confusing discussion, the teacher asked why they constructed a line graph in general. A student answered that to observe the change in time. A student said that there was no time in the example. Then he asked if there was a possibility that they could not construct the line graph.

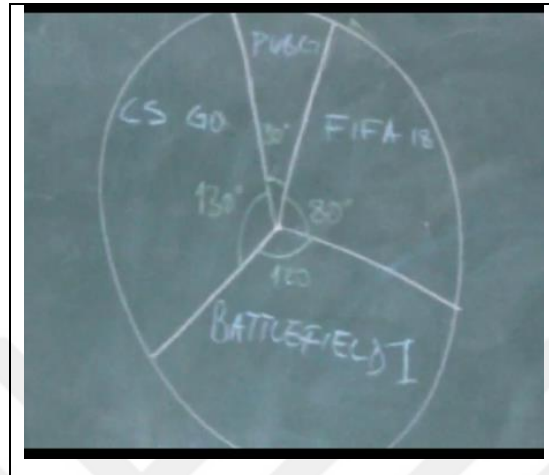


Figure 4.28 A Circle Graph Example of Ömer's Games [Ali 07.05.2018-1 00:00-15:00]

Furthermore, other students tried to develop ideas about time, like multiplying 7 and 24. The teacher asked if Ömer could play seven days and 24 hours per day. The teacher again asked if it was possible to construct a line graph based on the circle graph. The discussion lasted like 15 minutes. The next segment also included some parts of the discussion. However, the students concluded they could not construct the graph because they did not know what the whole represents in the next segment. The mathematical and statistical point was confusing and distorted for the two lesson segments. The students were confused about how to construct a line graph and came up with ideas that did not work to complete the task. Ali did not guide students to the solution of the task with a degree of clarity.

4.1.1.2.4.4. Overall Errors and Imprecision

Figure 4.29 shows the distribution of the scores assigned to lesson segments for the Overall Errors and Imprecisions in Ali's instruction. In 36 (59.02%) lesson segments, all the codes in this dimension were scored as Not Present. Ali's instruction included

minor errors which did not obscure the mathematics and statistics of the segments in 17 (27.87%) segments. In five (8.20%) lesson segments, at least one of the codes in the dimension was scored as Mid. Ali made errors in content, language, or a lack of clarity in certain parts of those segments, making it difficult for students to understand.

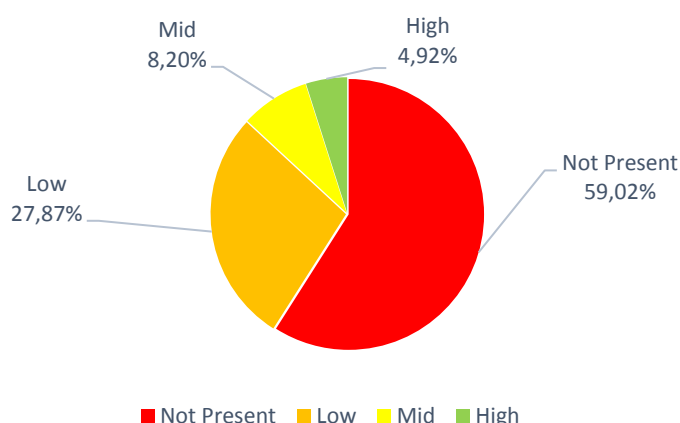


Figure 4.29 The Distribution of Scores Assigned to Lesson Segments for “Overall Errors and Imprecision” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

Ali made errors in most parts of three (4.91%) lesson segments. Those errors distorted the mathematics and statistics of the segments. The segments were the ones including the example that was converting the circle graph to a line graph (see Figure 4.23) and the example with the incorrect visualization of the data set “4, 4, 4, 4, 4, 4” (see Figure 4.24). The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.41 out of 4, with a standard deviation of 0.84. Based on the results regarding the Overall Errors and Imprecisions code, I concluded that most of Ali’s instruction was error-free in content, language, and clarity.

4.1.1.2.4.5. Overview of “Errors and Imprecisions” Dimension

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the frequencies of occurrences and co-occurrences of the Errors and Imprecision codes with their scores as visualized on the MAXQDA map (see Figure 4.30). Most of Ali’s

statistics instruction can be characterized as error-free in terms of content, language and notation, and precisions. Figure 4.30 shows the co-occurrence of the scores assigned to a segment's Errors and Imprecision Dimension codes.

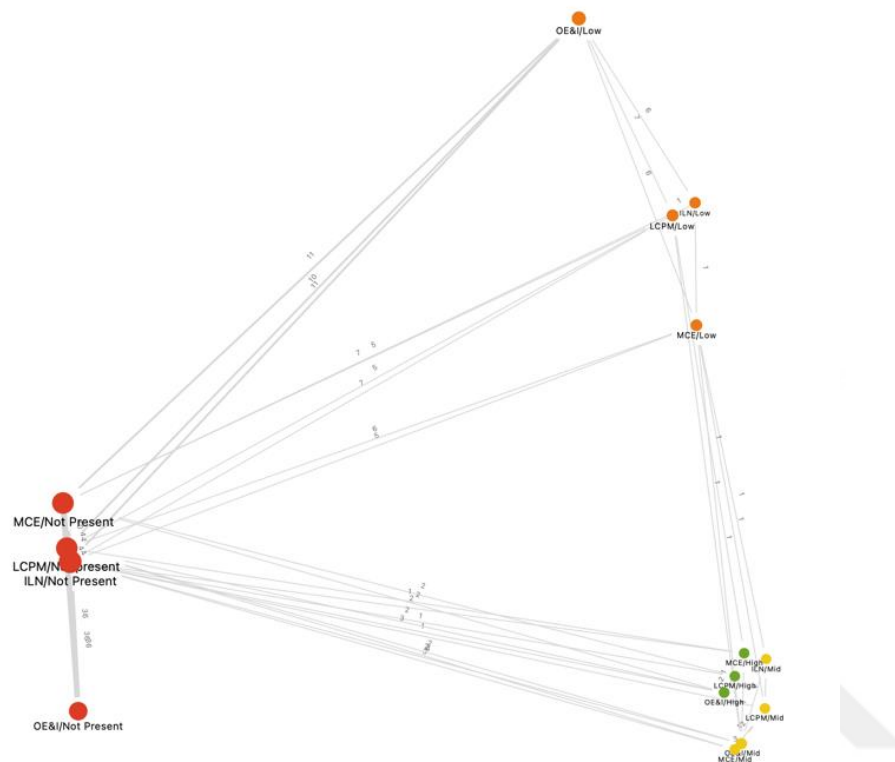


Figure 4.30 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in The” Errors and Imprecision” Dimension in Ali’s Instruction

Note: MCE- Mathematical and Statistical Content Errors, ILN- Imprecision in Language or Notion, LCPM- Lack of Clarity in Presentation of Mathematical and Statistical Content, OE&I-Overall Errors and Imprecision

The results indicate that Not Present scores were assigned to codes with the highest frequency. Furthermore, the codes with Not Present scores gathered around and co-occurred in the same segments with the highest frequency. The overall code with a Not Present score only interacted with other codes with Not Present Scores.

Figure 4.30 also presents that the codes with Low, Mid, and High scores less frequently occurred during the Ali’s instruction. Therefore, the segments with errors in Ali’s instruction were less frequent. First, the codes with Low scores were overlapped and gathered around the map. Accordingly, the codes with Low scores occurred in similar

segments. Those segments included brief errors in terms of content, language, notation, or clear presentation; those errors did not affect the mathematical and statistical points of the segments. In addition, the codes with Low scores also interacted with the codes with other scores. However, those codes were distant, so they were not present more likely in similar segments.

The codes with Mid and High scores were also less frequent and clustered. The Mid and High scores cluster showed that Ali's instruction consisted of errors and imprecision. A significant number of errors and imprecision co-occurred and obscured the mathematics and statistics of the segments. Similarly, the codes with Mid and High scores also interacted with those with Not Present scores. For example, content errors were present with middling density at some segments; other codes were not present except overall code with a Mid score, which was close to mid-level content error.

All in all, the map indicates that errors and imprecision codes co-occurred with each other at similar levels in most segments. In addition, it also confirms that the majority of lesson segments were error-free in Ali's instruction. The content errors at any level were less frequent and belonged to similar lesson segments.

4.1.1.2.5. Student Participation in Meaning-Making and Reasoning

This section shows the interaction between students and the content; how students are engaged in mathematics and statistics in Ali's instruction. For this reason, I examined this dimension with six codes in the MQI. Student participation in meaning-making and reasoning in Ali's statistics instruction can be characterized as Not Present to Low-quality (See Table 4.7). On the other hand, students talked about the mathematics and statistics of the segment (SCM), and students provided explanations (SPE) in some high-quality incidents. Table 4.7 shows that students did not participate in meaning-making and reasoning above 50% of the lesson segments in terms of Students Provide Explanation (SPE) code (n=37, 60.66% absent/not present) and Student Mathematical and statistical Questioning and Reasoning (SMQR) (n=50, 81.97% absent/not present).

Table 4.7

“Student Participation in Meaning Making and Reasoning” Dimension Scores in Ali’s Instructional Segments

	1		2		3		4	
	Not Present		Low		Mid		High	
	n	%	n	%	n	%	n	%
Students Provide Explanations	37	60.66	11	18.03	11	18.03	2	3.28
Student Mathematical and Statistical Questioning and Reasoning (SMQR)	50	81.97	10	16.39	1	1.64	0	0.00
Students Communicate about the Mathematics and Statistics of the Segment	13	21.31	12	19.67	27	44.26	9	14.75
Task Cognitive Demand	27	44.26	24	39.34	10	16.39	0	0.00
Students Work with Contextualized Problems	9	14.75	29	47.54	23	37.70	0	0.00
Overall Student Participation in Meaning Making and Reasoning	28	45.90	23	37.70	10	16.39	0	0.00

Note. N=61 (the total number of lesson segments)

Additionally, there was no high-quality lesson segment in SMQR, SCM, Task Cognitive Demand (TCD), Students Work with Contextualized Problems (SWCP), or overall code. The codes relatively present than other scores were SCM (n=27, 44.26% mid; n=12, 19.67% low), SWCP (n=23; 37.70% mid; n=29, 47.54% low), and TCD (n=10, 16.39% mid; n=24, 39.34% low). Furthermore, the overall code was also present at Mid (n=10, 16.39%) and Low (n=23, 37.70%) levels.

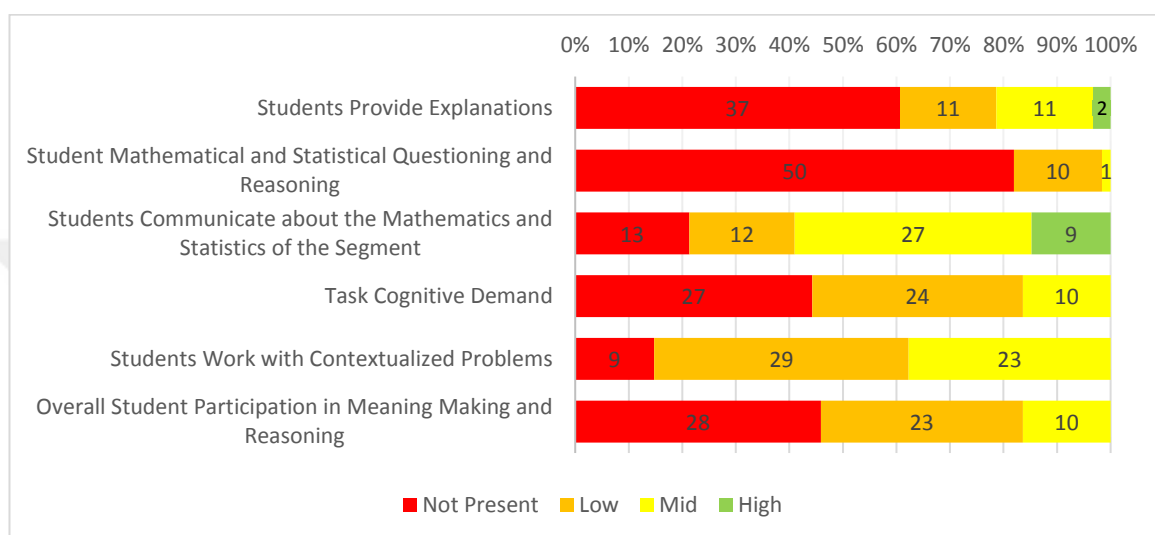


Figure 4.31 “Student Participation in Meaning Making and Reasoning” Dimension in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

As shown in Figure 4.31, Ali’s instruction lacked student and mathematical and statistical content interaction for many codes to some extent (e.g., SPE and, remarkably, SMQR). Even though there were interactions among teachers and students, they could be characterized as low quality in most lesson segments (e.g., TCD, SWCP, and Overall) and mid-level (e.g., SCM) for some lesson segments but not high for many codes in this dimension. The following sections present the details of the codes of this dimension.

4.1.1.2.5.1. Students Provide Explanations

Figure 4.32 shows the distribution of scores assigned to lesson segments for students’ mathematical and statistical explanations based on the answers to the “why” question

in Ali's instruction. Students did not provide explanations in most lesson segments ($n=37$, 60.66%). Students provided brief explanations in 11 (18.03%) lesson segments. 11 (18.03%) lesson segments included student explanations in a more sustained way. However, only two segments (3.28%) were mainly characterized by student explanations. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.64 out of 4, with a standard deviation of 0.90. This section provides the descriptions and classroom examples of each score assigned to the Students Provide Explanations code in Ali's statistics instruction.

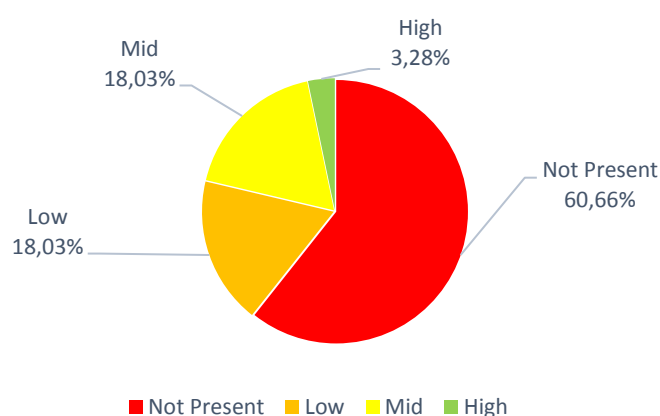


Figure 4.32 The Distribution of Scores Assigned to Lesson Segments for “Students Provide Explanations” Code in Ali’s Instruction

Note. $N=61$ (the total number of lesson segments)

I assigned a score of Not Present to 37 (60.66%) lesson segments for SPE. Student explanations were not observed during individual or group work in these segments. Although the interaction occurred between the teacher and students over content, students’ responses did not include any reasoning for those segments. The questions posed by the teacher did not require students to provide explanations for their reasoning. Instead, students were primarily expected to provide descriptions of solution steps or procedures, often in response to “what” or “how” questions rather than “why” questions. For example, when working on graph tasks, students were asked to read the graph at either an elementary level, which involved extracting information from the graph, or an intermediate level, which involved finding relationships in the

data through procedural steps. To clarify, the term “elementary level” refers to the most basic level of reading a graph, where students are expected to extract information directly from the visual representation of the data. This could include reading off values on the axes or identifying specific data points on the graph. Additionally, reading the graph at the intermediate level involves understanding the relationship between data points and the overall trend of the data. For example, students were expected to use the information presented in the graph to answer questions such as “What was the trend in the data over the given period?” (Friel et al., 2001)

I assigned a Low score to 11 (18.03%) of the lesson segments for SPE because there were one or two brief explanations of students based on teachers’ demand from students to answer ‘why’ questions. The teacher asked follow-up questions in some tasks, and a few responses were received from students, some of which were incomplete or incorrect explanations. For example, Ali asked, “Why do we construct a line graph? What does it tell us?” in a graph transformation task (see Figure 4.28 in Lack of Clarity in Presentation in Mathematical and Statistical Content). Some students gave incomplete answers. A student explained why to use a line graph “with the increasing or decreasing in a number.” Another student answered, “A change in terms of time.” In the same task, the teacher asked if they transformed the circle graph (with missing information) into a bar graph, and a student said no, and the teacher asked why they did not transform the circle [graph] into [bar graph]. There were two brief explanations after the why question. The student with a “no” answer said it was not time dependent. A student said yes and added that he could “put degrees on the y-axis.”

The Mid-level lesson segments (n=11, 18.03%) included student explanations of the meaning behind a procedure, why a solution method was appropriate, or why use a specific graph or a measure of central tendency. For example, the extract in the Explanation code included a student explanation regarding dividing 360 degrees into 12 [the number of people] to find how many degrees corresponded to a person to find the angle. In addition, the example in the Remediation code included a student's incomplete explanation regarding constructing a circle graph without knowing angles [the jar mixture example]. The student tried to estimate the size of a sector by thinking

about the relationship among numbers. He said they needed to start with eight because six was half of 12, and eight was bigger than the half. Those student explanations were sustained. Other segments included more frequent student explanations. Especially in selecting an appropriate average concept or transforming a graph into other types of graphs, the teacher asked ‘why’ questions more, and students gave explanations; some were incomplete, correct, or detailed. However, more frequent explanations were received from students.

I scored High in two (3.28%) segments for SPE because student explanations characterized those parts. The teacher and students constructed the line graph in Graph Task 5 example (See Figure 4.23). After constructing the line graph, the teacher asked why they used line graphs and their advantages. The students explained why to use line graphs with examples. One student gave an appropriate example for a line graph “Let us think about a firm, whether its income increased or decreased between 2007 and 2017, and in which years there was a big increase or in which years there was a decrease?” After the student’s explanation, other students gave a brief, incomplete, and incorrect explanations, like seeing the increase in terms of years and the rate of number and time. However, the teacher used the first student example to emphasize using a specific graph.

The second segment included the average task presented in the Mathematical and Statistical Sense-Making code. The task included eight students’ exam results, and the teacher asked students what would happen to the arithmetic mean, median, and mode when he changed a student’s grade. Students explained how the median did not change or when the median changed. Furthermore, students explained why the mode did not change and why the mode was 80. Student explanations shaped the instruction in the segment.

4.1.1.2.5.2. Student Mathematical and Statistical Questioning and Reasoning

Figure 4.33 shows the distribution of scores assigned to lesson segments for the Student Mathematical and Statistical Questioning and Reasoning code. Most lesson segments (n=50, 81.97%) did not include student mathematical and statistical

questioning. Students questioned and reasoned briefly in 10 (16.39%) lesson segments. In only one lesson segment, student mathematical and statistical questioning or reasoning was more frequent but not characteristics of the segment. The high quality of SMQR was not present in none of the segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.20 out of 4, with a standard deviation of 0.44. This section provides the descriptions and classroom examples of each score assigned to the Student Mathematical and Statistical Questioning and Reasoning code in Ali's statistics instruction.

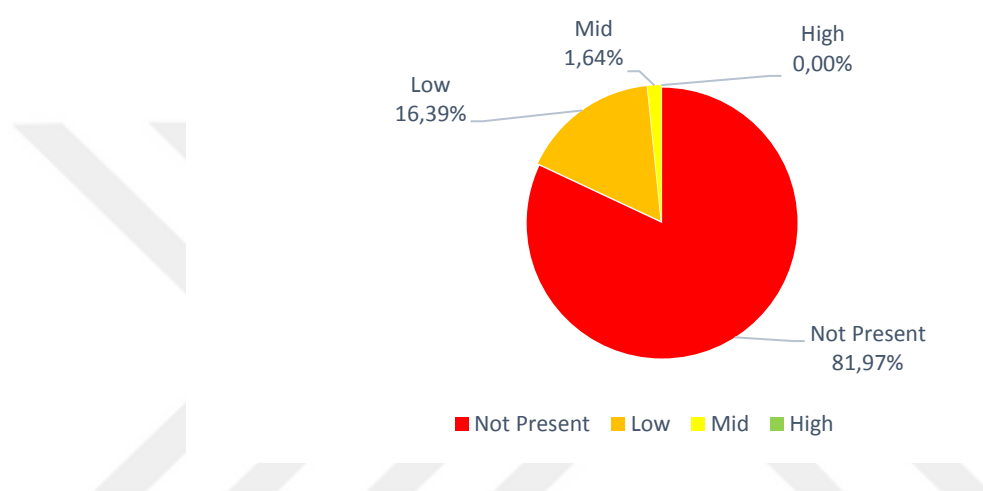


Figure 4.33 The Distribution of Scores Assigned to Lesson Segments for “Student Mathematical and Statistical Questioning and Reasoning” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

Lesson segments scored Low (n=10, 16.39%) included only one or two brief student mathematical and statistical questioning or reasoning. Four segments included graph tasks. The students tried to make sense of a line graph about the length of a tree. The teacher asked students if the tree grew at most; he wanted to choose between the first and second years. A student gave a counter-answer by not choosing two choices; he said one year and six months. In another example, the monthly change in Kerem’s money was given in the line graph. The students were asked to find the month Kerem’s money did not change. The teacher stated that the month at the end of the horizontal line was the month [May] his money did not change. However, a student questioned if the answer was not both months [April and May]. Those graph examples were

considered mathematical and statistical questioning and reasoning because the students came up with counterclaims other than what the teacher asked or stated.

Furthermore, six-lesson segments included average tasks. During average tasks, some instances of mathematically motivated questions were observed. For example, when the teacher presented the first examples of finding the median of an odd number of data sets, a student asked what would happen to the median if there were six numbers. Another student asked what would happen if they had values with a half for the median. In addition, a student asked what would happen to the mode if two values appeared most frequently or there were no repeating values. However, such instances were infrequent and brief, as the teacher provided general answers to those questions.

Only one (1.64%) segment with a score of Mid had more frequent student mathematical and statistical questioning and reasoning. The segment included a controversial issue about finding the mode of a data set with the same values. The example is also presented in the Mathematical and Statistical Content Error code. Some students claimed no mode, and others said the mode was four. A student claimed it would be four if there were one more value. A student stated that there was no mode because there was no different number. The students made conjectures about why there was no mode with their experience in mode definition. Moreover, after the teacher gave the wrong context for the given data (see Mathematical and Statistical Content Error code) and confirmed no mode in the data set, a student again asked why the data set did not include a mode. The student asked a motivating statistical question to the teacher for demanding explanations.

4.1.1.2.5.3. Students Communicate about the Mathematics and Statistics of the Segment

Figure 4.34 shows the distribution of scores assigned to lesson segments for students' communication about the mathematics and statistics of the segment in Ali's instruction. In 13 (21.31%) lesson segments, students did not communicate about the mathematics and statistics of the segment. 12 (19.67%) lesson segments included brief student contributions. Substantive student contributions did not characterize 27

(44.26%) lesson segments while they characterized nine (14.75%) lesson segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.52 out of 4, with a standard deviation of 0.99. This section provides the descriptions and classroom examples of each score assigned to the Students Communicate about the Mathematics and statistics of the Segment code in Ali's statistics instruction.

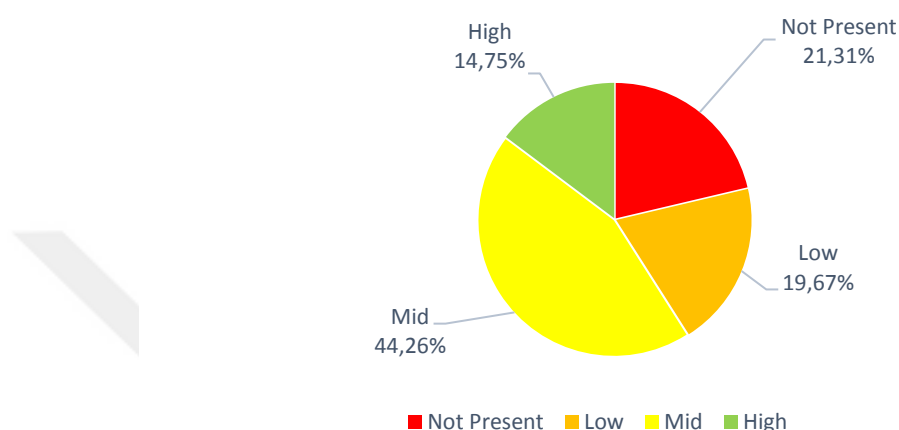


Figure 4.34 The Distribution of Scores Assigned to Lesson Segments for the “Students Communicate about the Mathematics and Statistics of the Segment” Code in Ali's Instruction

Note. N=61 (the total number of lesson segments)

Student contribution was not present in 13 (21.31%) lesson segments. Those segments could be characterized as individual work time where students solved tests individually and asked for help from the teacher when needed. The teacher sometimes clarified or solved the tasks for individual or group students. When the researcher heard, the teacher dominated the talk in those segments. However, in some segments, students remained silent, or the student talk was inaudible. However, all segments scored Not Present were a part of the individual-work period.

Students contributed briefly, providing one- or two-word explanations/answers to the questions in 12 (19.67%) lesson segments. Those segments were dominated by teacher talk, but there were some brief one-or-two-word answers from the students. For example, the teacher selected one task to solve on board during individual work time since most children could not solve it during individual work time. While the teacher

solved the task on board, he asked questions like the total number, angle, or data point in a line graph; students gave one-or-two-word answers; described the solution steps shortly. Similarly, in whole class sessions, although the teacher called a student on board in some segments, he did not allow students to contribute to mathematics and statistics; he used the students as a writer of his solution steps.

Students contributed more than briefly, but the contributions were not the segment's focus in 27 (44.26%) of the lesson segments. Mid segments included substantive student contributions like solving the task on board; some students solved the tasks without expressing them; the teacher expressed their solution steps. In addition, some students solved the tasks on board and expressed the solution process with the teacher's contribution. When the teacher asked what the sample meant in other segments, some students had substantive contributions to the definition, even if some were incorrect or complete. The teacher emphasized the “randomness” word with student definitions while selecting a sample. However, the segment included other concepts for the discussion. The sample and randomness concepts were not the segment's focus, so the student contribution did not characterize the segment.

I scored nine (14.75%) lesson segments High for SCM because the student contributions moved the instruction. High scores on SPE and SMQR were also scored as High for SCM based on the description in MQI. Some examples of High SCM were student questioning and reasoning about the median or explaining why arithmetic mean, median, and mode changed or did not change. Those segments were characterized by student contribution.

4.1.1.2.5.4. Task Cognitive Demand

Figure 4.35 shows the distribution of scores assigned to lesson segments for task cognitive demand during the task enactment phase in Ali's instruction. 27 (44.26%) lesson segments did not include students' engagement in cognitively demanding activities, while students were briefly engaged in cognitively demanding activities in 24 (39.34%) lesson segments. Furthermore, students engaged in cognitively demanding activities at Mid-level (n=10, 16.39%) described as a transition between or

a mixture of cognitively demanding and undemanding activities. None of the segments included students' engagement in a high level of cognitive demand. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.72 out of 4, with a standard deviation of 0.73. This section provides the descriptions and classroom examples of each score assigned to the Task Cognitive Demand code in Ali's statistics instruction.

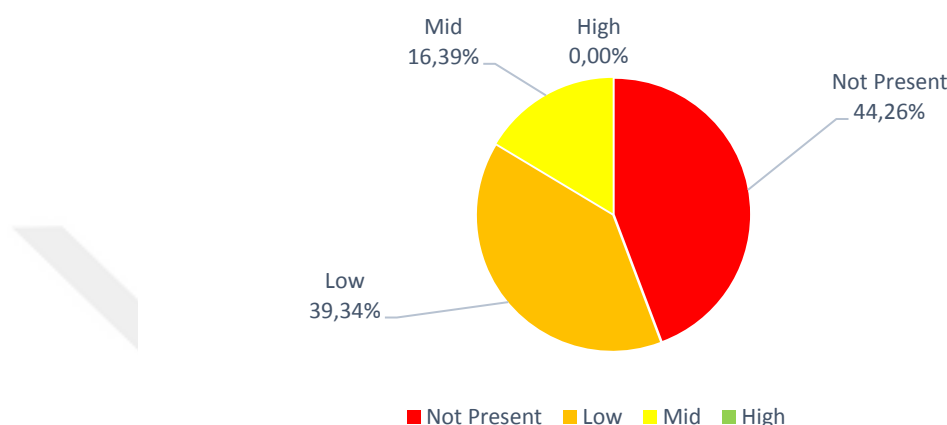


Figure 4.35 The Distribution of Scores Assigned to Lesson Segments for “Task Cognitive Demand” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

I scored Not Present for Task Cognitive Demand for 27 (44.26%) lesson segments. Those segments' tasks demanded students to apply the procedures or formulas during solving circle graph tasks or average tasks, especially in individual work time. Teacher dominated talk in whole class session, and unsystematic exploration. For example, the segments contained the task of transforming the circle graph about Ömer's games into a line graph (see Figure 4.28). As mentioned before, the context was inappropriate for creating a line graph. However, the teacher asked students to create one. In the end, they discussed why they could not form a line graph and what they needed for a line graph. However, the students dealt with the unsystematic exploration in the segments by discussing how to form a line graph.

Low lesson segments (n=24, 39.34%) contained examples of short cognitively demanding work where student explanations, questioning, or reasoning about the

solution process or the task were present briefly. For example, the student questioning why not starting at 0 in a graph made the task enactment momentarily cognitively demanding. Another example was constructing a circle graph where students tried to relate the context and the structure of a circle. Initially, the tasks seemed not cognitively demanding as curricular material or set-up; they seemed to apply well-established procedures. However, momentary cognitive demand appeared when students engaged with them; for example, considering a part-whole relationship and using proportional reasoning at different levels.

Mid-level lesson segments (n=10, 16.39%) had high cognitive demand task characteristics, but the teacher's presence and direction were also evident during these segments. The teacher often directed the instruction with his agenda and responses to students' reasoning, questions, or explanations. For instance, segments involving the introduction to graphs and reasoning about the change in averages with a change in the data set required high-level cognitive demand. However, the teacher's direction and guidance remained noticeable during these segments. In another example, students attempted to connect given information with circle characteristics by introducing the circle graph, and they also tried to understand the features of average concepts. Still, while students were actively engaged, the teacher-directed instruction was prominent, making it more teacher-centered, resulting in a middling density of cognitively demanding activities.

All segments scored Not Present, Low, and Mid included some tasks with similar characteristics at the curricular level. However, the Mid-score segments included making sense of a statistical concept with a statistical task. The teacher allowed students to explain or reason about the concepts through the enactment phase. After reasoning about the concepts, students engaged in similar tasks at a low level because there was room to discuss the concepts. Then there was student engagement with cognitively undemanding tasks. At that point, students were practicing what they learned, calculating the procedures for finding a value in a graph or finding the angle of a sector in a circle graph, or calculating the arithmetic mean or median or mode with or without context tasks.

4.1.1.2.5.5. Students Work with Contextualized Problems

Figure 4.36 shows the distribution of scores assigned to lesson segments for students' work with contextualized problems in Ali's instruction. I assigned Not Present to Nine (14.75%) lesson segments for SWCP. In 29 (47.54%) lesson segments, contextualized problems were routine exercises. Students worked with contextualized problems with some assistance from the teacher ($n=23$, 37.70%). None of the segments included high-level student reasoning or worked on contextualized problems. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.23 out of 4, with a standard deviation of 0.69. This section provides the descriptions and classroom examples of each score assigned to the Students Work with Contextualized Problems code in Ali's statistics instruction.

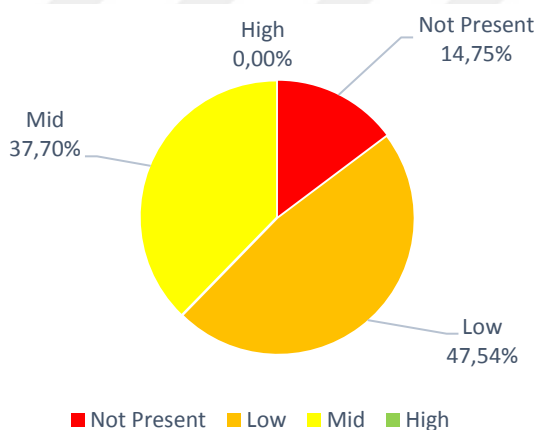


Figure 4.36 The Distribution of Scores Assigned to Lesson Segments for 'Students Work with Contextualized Problems' Code in Ali's Instruction

Note. $N=61$ (the total number of lesson segments)

Nine (14.75%) lesson segments did not include contextualized problems, or students did not work on contextualized problems. Students were expected to calculate the arithmetic mean, mode, and median for a data set in many segments. Those data sets were not contextualized. Therefore, students worked on exercises. In addition, two segments included contextualized problems, but the students did not work on them; the teacher briefly mentioned the answer or solved the task himself, not the students. There were no problems or tasks; the teacher recalled and revealed students' knowledge about the data analysis by asking about its definition in a segment.

The contextualized problems were present but solved as a routine exercise and heavily scaffolded in 29 (47.54%) lesson segments. The teacher mostly told students what to do. For example, while calculating the angle corresponding to the categories of data provided for constructing a circle graph, the teacher helped students reason multiplicatively; he heavily guided students in solving proportions. In addition, students worked individually on worksheets that consisted of contextualized problems. The teacher asked students to recall what they learned and solve the questions individually with the worksheets. Therefore, the segments with individual work time included routine exercises; when students did not know how to solve the contextualized problems, the teacher executed the operations for students.

Students reasoned about the contextualized problems in 23 (37.7%) lesson segments, but some teacher guidance was present. For example, the contextualized problems/tasks were about constructing a graph or reading a whole class in the first sections of introducing the topic. While the graphs were transformed, students were more engaged with context to decide if the graph could be transformed or how to transform a type of graph into another type of graph. The task about Ömer's games lacked clarity; it engaged students based on context. However, in the end, the teacher gave hints to students about missing information in the circle graph. In addition, some average tasks also started more open-ended for students. Students were expected to decide which average concept should be used for a given context. Those problems were non-routine but guided by the teacher. During student descriptions or explanations of their work on contextualized problems, the teacher provided guidance at some points.

4.1.1.2.5.6. Overall Student Participation in Meaning-Making and Reasoning

Figure 4.37 shows the distribution of the scores assigned to lesson segments for the Overall Student Participation in Meaning-Making and Reasoning in Ali's instruction. Student participation in cognitively demanding activities was absent, or the activities were procedural or heavily guided by the teacher in 28 (45.90%) lesson segments. Students were involved in mathematical and statistical activities, but the activities were primarily procedural in 23 (37.70%) lesson segments.

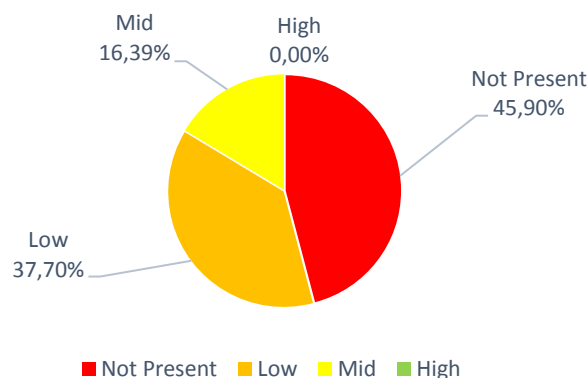


Figure 4.37 The Distribution of Scores Assigned to Lesson Segments for “Overall Student Participation in Meaning-Making and Participation” Code in Ali’s Instruction

Note. N=61 (the total number of lesson segments)

Students engaged in mathematical and statistical activities at a mixed level. There were some high and low-quality indicators for participation in meaning-making and reasoning in 10 (16.39%) lesson segments. There were no high-quality indicators for the overall code. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.70 out of 4, with a standard deviation of 0.74. Students’ participation did not include good communication for building mathematics and statistics, substantial contributions, a high level of reasoning, and asking good questions with cognitively demanding activities.

4.1.1.2.5.7. Overview of “Student Participation in Meaning-Making and Reasoning” Dimension.

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the frequencies of occurrences and co-occurrences of the Student Participation in Meaning-Making and Reasoning codes with their scores as visualized on the MAXQDA map (see Figure 4.38).

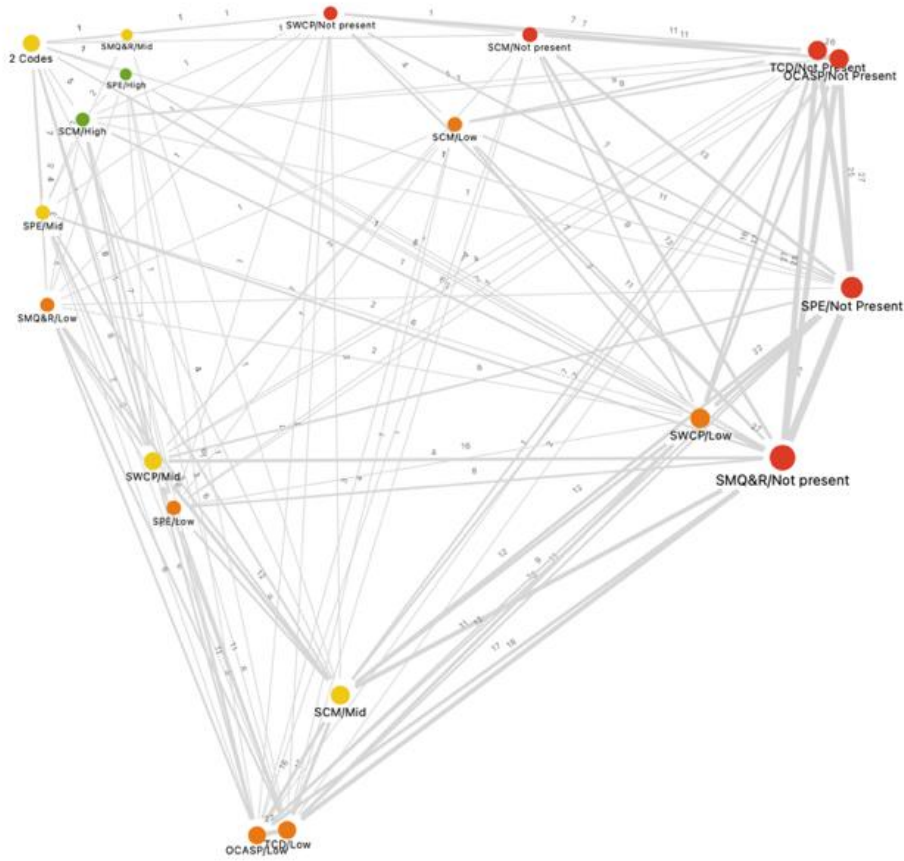


Figure 4.38 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Student Participation in Meaning-Making and Reasoning”

Dimension in Ali’s Instruction

Note: SPE- Students Provide Explanation, SMQ&R- Student Mathematical and Statistical Questioning and Reasoning, SCM- Students Communicate about the Mathematics and Statistics of the Segment, TCD- Task Cognitive Demand, SWCP- Students Work with Contextualized Problems, OCASP- Overall Student Participation in Meaning-Making and Reasoning

Student participation in meaning-making and reasoning in Ali’s statistics instruction can be described as not present to low quality. Figure 4.38 indicates the occurrences and co-occurrences of the scores assigned to the codes. The results show that SPE, SMQR, and TCD were frequently absent. In addition, the overall code was also not present most frequently in Ali’s instruction. They co-occurred most frequently in similar segments. The length of the lines among the codes showed the co-occurrences in similar segments. The width of the connected lines indicated the frequency of their co-occurrences. The overall code with Not Present score also co-occurred with the codes with Not Present scores more frequently.

Furthermore, the codes mentioned above with Not Present scores also co-occurred with the SWCP and SCM codes with all scores without a High score. In addition, they also interacted with the overall code and TCD code with Low scores. However, they were less frequently co-occurred than the codes with Not Present scores. Therefore, students did not provide explanations or question, or reason mathematically and statistically in many segments. The students did not also engage in cognitively demanding activities. In some segments, the students did not even contribute to the development of the mathematics and statistics of the segment. However, the student contributions were generally present briefly or at length during working on the contextualized problems related to graphs and average concepts in many segments. The teacher scaffolded students' ideas to varying levels. The level of cognitive demand was low during the enactment phase because the tasks mainly included brief non-routine aspects with some students' meaning-making and reasoning.

The codes with Low, Mid, and High scores were less frequent than codes with a Not Present Score. Therefore, the student participated in meaning-making and reasoning less frequently in Ali's instruction. Figure 4.38 shows that Mid scores assigned to SPE, TCD, SMQR, and overall code, and High scores assigned to SPE and SCM were present in Ali's statistics instruction with the lowest frequencies. The distinct codes with different scores (Mid score or High score) also co-occurred in similar segments according to the length of lines among the codes and the width of the connected lines.

The codes of SPE and SCM with High scores co-occurred with the other codes assigned Mid scores because Ali's statistics instruction did not have the characteristics of High scores in other codes. In addition, SMQR with Low scores also co-occurred with the other codes with Mid and High scores. The results show that the students engaged in cognitively demanding activities at a middling density. The student contributions were substantive and characterized the segments at some points. The students reasoned about statistical concepts and explained their reasoning or solution methods during the segments with the graph and average tasks. However, the high-quality practices of student participation in meaning-making and reasoning were limited in Ali's instruction.

Consequently, the map indicates that the codes in the Student Participation in Meaning Making and Reasoning dimension co-occurred at similar levels in most segments. In addition, the map also shows that most lesson segments did not include student participation in meaning-making and reasoning in Ali's instruction or the participation was present at a low level.

4.1.1.3. Whole Lesson Mathematical and Statistical Quality of Instruction (MQI): A Lesson Level Analysis in Ali's Statistics Instruction

This section aims to describe Ali's instruction as a whole based on one code that assesses the lesson's overall mathematical and statistical quality. As previously mentioned, Ali's statistics instruction consisted of 14 lessons. I coded each lesson using this one code on a 5-point scale.

The results show that the overall mathematical and statistical quality of Ali's instruction can be characterized as Mid quality (n=7, 50%) and Mid/High quality (n=7, 50%), with a mean of 3.50 and a standard deviation of 0.52. The Mid quality lessons did not exhibit characteristics of Low and High quality. However, they were generally free from errors and imprecision. Most lessons observed were typical and did not include special features, as described in the MQI. The Mid/High-quality lessons were scored higher because they had richer mathematics and statistics, more substantive student contribution and participation, and responsive teacher-student interaction. However, they were not scored as High quality because not all high-quality codes were present and combined in the same lesson.

In summary, Ali's instruction demonstrated an overall Mid to Mid/High quality, with most lessons exhibiting typical but error-free and well-organized instruction. The Mid/High-quality lessons demonstrated more engagement and interaction between teacher and students, with richer mathematical and statistical content.

4.1.2. Case 2: Esra's Teaching Practices

This section presents findings regarding Esra's teaching practices. The within-case analysis will focus on aligning statistical content with curricular expectations, the

statistical problem-solving process, and the mathematics and statistics instruction quality in Esra's classroom. The section is divided into three subsections: Esra's statistical content and problem-solving process, the mathematical and statistical quality of instruction at the segment level, and the whole-lesson mathematical and statistical quality of instruction.

The first subsection provides a detailed analysis of Esra's statistical content, including the topics covered in the classroom and how they align with curricular expectations. In addition, I examine Esra's teaching practices in terms of the four-step model of statistical problem-solving processes. The second subsection focuses on the mathematical quality of instruction at the segment level. Specifically, I analyze the quality of instruction regarding mathematics and statistics in individual segments of Esra's instruction. The third subsection analyzes the mathematical and statistical quality of instruction at the lesson level. This involves the whole-lesson mathematical and statistical quality of instruction, focusing on an entire lesson based on a combination of all dimensions in the MQI.

4.1.2.1. Esra's Statistical Content and Statistical Problem-Solving Process

This section describes Esra's statistics instruction in terms of two subdomains. First, I describe the instruction based on statistical content. I explore which concepts were covered in Esra's statistics instruction and the distribution of the number of episodes allocated for each concept. Then I describe the statistical problem-solving process with the protocol based on the goals of statistics instruction in the GAISE Framework (Litke & Hill, 2020). This subsection includes four elements with three categories for each element. The four elements will be examined separately and as a whole process. The data sources used for this section are videotaped lessons, field notes, lesson summaries, and interviews with the teacher.

4.1.2.1.1. Statistical Content in Esra's Instruction

To examine statistical content in Esra's instruction, I utilized a similar approach as in the case of Ali's teaching practices. I used a schematic diagram to understand the statistical tasks' distinctiveness and purpose in terms of the objectives covered in the

Turkish 7th-grade mathematics curriculum. This analysis shows that Esra's statistics instruction consisted of 32 episodes as a unique data set the class worked on (Litke & Hill, 2020). I utilized the episodes regarding statistical content and the statistical problem-solving process to gain insight into Esra's teaching practices.

There were episodes with graphs and measures of central tendency. As shown in Table 4.8, the episodes with graphs consisted of circle graphs (n=9, 28.13%), line graphs (n=5, 15.63%), bar graphs (n=4, 12.50%), and graph transformations (n=1, 3.13%).

Table 4.8

Statistical Content in Esra's Instruction

Level 1	Level 2	n	%
Graphs		20	62.50
	Circle Graph	9	28.13
	Line Graph	5	15.63
	Bar Graph	4	12.50
	Graph	1	3.13
	Transformation		
Measures of Central Tendency		12	37.50
	Arithmetic mean	2	6.25
	Median	0	0
	Mode	1	3.13

Note. N=32 (the total number of episodes)

The measures of central tendency were presented in 12 episodes (37.50%). In nine (28.13%) episodes, students worked on specific examples in which they decided the types of averages needed to use for a given context. Not a single episode just consisted of the median. The remaining episodes consisted of merely the arithmetic mean (n=2, 6.25%) and mode (n=1, 3.13%).

The analysis shows that the statistical content in Esra's teaching practices was aligned with the objectives of a 7th-grade mathematics curriculum. However, bar graphs are not the primary focus in the 7th-grade mathematics curriculum except in the objective related to graph transformation. Like circle and line graph objectives, constructing and reading bar graphs objectives are present in the fifth and sixth-grade mathematics curricula. In Esra's teaching practices, four episodes were about bar graphs because

the book the teacher followed included bar graph tasks. Additionally, based on students' statements, the class had not covered the sixth-grade statistics unit. The objective related to graph transformation required prior knowledge of the bar graph. Therefore, the teacher followed the book content for teaching bar graphs to fulfill the graph transformation objectives or due to the student's lack of knowledge from the previous year as stated by the teacher.

4.1.2.1.2. Statistical Problem-Solving Process in Esra's Instruction

This section aims to analyze Esra's statistics instruction regarding statistical investigations using the Transcript Coding Protocol Using the GAISE Framework, also used in Ali's section. The framework includes four elements, each with three categories, designed to understand statistical instructional practices based on GAISE standards. These categories help identify the extent to which statistical investigations are implemented, focusing on statistical ideas or concepts. The categories reflect varying levels of alignment with GAISE expectations, with Level 3 indicating the highest level of alignment.

I described Esra's statistics instruction based on statistical problem-solving processes. I assigned codes to each episode in terms of the phases of the statistical problem-solving; (a) formulation of questions, (b) data collection, (c) data analysis, and (d) interpretation with three categories. Table 4.9 shows each element's episode frequencies regarding the level of presence (through Level 1 to Level 3).

The results show that the first two elements of the statistical problem-solving process were mainly in Level 1, indicating that they were not observed in the episodes, precisely formulating questions ($n=31$, 96.87%) and data collection ($n=31$, 96.87%). Only one episode included the formulation of questions and data collection. Furthermore, interpretation was not present in 19 episodes (59.37%), but 12 episodes (37.50%) included interpretation at Level 2, indicating that interpretation was somewhat present.

Table 4.9*Esra's Statistical Problem-Solving Process: The Occurrences of the Level of Elements*

	Level 1		Level 2		Level 3	
	n	%	n	%	n	%
Formulation of Questions	31	96.87	1	3.13	0	0.00
Data Collection	31	96.87	1	3.13	0	0.00
Data Analysis	0	0.00	23	70.87	9	28.13
Interpretation	19	59.37	12	37.50	1	3.13

Note. N=32 (the total number of episodes)

None of the episodes covered the first three elements fully expected in GAISE. Besides, data analysis was observed in all episodes. Twenty-three episodes (70.87%) included data analysis to some extent but did not meet all expectations provided in GAISE. Nine episodes more fully covered data analysis according to the expectations of GAISE. The following section will describe each element of Esra's instruction's statistical problem-solving process. In the end, I will present the co-occurrences of the steps in the statistical problem-solving process.

4.1.2.1.2.1. Formulation of Questions

Table 4.9 shows that Esra's instruction did not include the formulation of questions (n=31, 96.87%). Therefore, those episodes were scored as Level 1. Those episodes included the work with data without focusing on the statistical investigative question. The episodes included graphs or measures of central tendency tasks not rooted in statistical investigative questions. The questions in the tasks did not consider variability; they were primarily for data analysis and interpretation. Therefore, they could not be described as a part of the formulation of the questions element. The teacher posed a statistical investigative question in one episode (3.13%). The episode was coded as Level 2 because the teacher posed the question. The question was about "the most favorite color in 7D class". Then she collected the data from students and summarized the situation.

What have I done? I wondered. I asked, what is the 7D class's favorite color? I wondered, a question mark appeared in my head, and what did I do to satisfy my curiosity? I asked you a question. I asked what your favorite color was.

[Esra 06.04.2018 07:30-26:40]

The teacher asked one statistical investigative question. The question anticipated variability, the teacher determined the options beforehand, but while deciding what to investigate, a student suggested that the question might be about the favorite “meals.” The teacher said that there would be many options for such a question. Then she proposed to ask about the favorite colors and determined the four-to-five options for leading to the data collection process. The statistical investigative question was not deterministic but about the whole group. It included a group of interests and also defined the variable clearly. This episode was assigned at Level 2 because students did not generate the question. However, with all characteristics, the question leads to statistical investigations. The details regarding the investigation will be provided in the Overview of Esra’s statistical problem-solving process section.

The results show that the formulation of questions was not present in most episodes of Esra’s instruction. However, the single episode with the formulation of the question included a statistical investigative question that carried most properties of a good question and led to statistical investigations.

4.1.2.1.2.2. Data Collection

Table 4.9 shows that Esra’s instruction did not consist of data collection in most episodes ($n=31$, 96.87%). Therefore, those episodes were scored as Level 1. In those episodes, the data were provided. In some episodes, tasks included pre-determined data sets. The students were expected to work on the data provided. For example, they calculated the measures of central tendency. In other episodes, the data were presented in graphs or tables. For instance, the tasks required finding a missing value in the graph while reading the graph.

Only one episode (3.13%) included data collection. The episode was the same as the one including the statistical investigative question. For the data collection, the teacher asked students their favorite color. As mentioned, the teacher formed options and

warned students that they could only choose one option. She did not collect the data individually and read aloud the options to the whole class. If the option was the students' choice, they raised their hands. The data collected was recorded systematically in the table on the blackboard. The data collected was first-hand and necessary for answering the statistical investigative question.

As a result, most of the episodes in Esra's instruction did not include data collection. The episode led to a statistical investigation with data collection at Level 2. The episode was more fully aligned with GAISE's expectations regarding data collection because the class did not discuss the variability, sampling, and measurement issues. Solely the data was collected and recorded systematically.

4.1.2.1.2.3. Data Analysis

In Esra's statistics instruction, all episodes included a data analysis element (see Table 4.9). In 23 episodes (70.87%), the data analysis occurred to some extent. The episodes focused on creating circle, line, and bar graphs. The students were expected to calculate the total quantity corresponding to the whole circle and a part of the quantity by using the circle sector characteristics (i.e., percentages and degrees). In short, they calculated the value asked in the task. In addition, they also answered descriptive questions based on the graphs. For example, in a bar graph example, the students were expected to find the answer to the descriptive question: "What is the total profit-loss amount in four years?". Other episodes focused solely on calculating central tendency measures, including tasks with or without context. Most data analysis episodes at Level 2 were descriptive and procedural. They did not include graphs for describing distributions, comparing groups, and examining the relationships among variables.

Nine episodes (28.13%) were at Level 3 in the data analysis. At Level 3, it was expected to describe the distributions, compare groups, and examine the relationships along with the criteria at Level 2. In Level 3 episodes, the teacher and students used graphs to describe the distributions or compare data sets. For example, when the students examined the circle graph they constructed, they stated that the ones who liked green were one-fourth of the whole group. In a line graph example, the example

contained a graph with two lines stating the correct answers over mock exams (based on ordinality) of two students. The question asked in the task was, “Arda’s total number of correct answers was higher than Turan’s total number of correct answers.” The students were expected to examine two lines, calculate the total number of correct answers for each student, and compare both. It was a kind of comparison of groups example, at Level 3.

In the other episodes, there were some examples regarding the measures of central tendency. The example tasks demanded students to find the most appropriate central tendency measure to find the typical value for describing the distribution. The teacher expected students to calculate the arithmetic mean, mode, and median in those examples. When the calculated value for all measures was close to each other, the teacher mentioned the even distribution of the data set, the distribution when there was an extreme value, and why they should not have used arithmetic mean for describing distribution. Furthermore, the teacher also brought up an example demanding students compare two data sets using their knowledge of the measures of central tendency. Therefore, the students compared groups based on their calculations.

To conclude, all episodes in Esra’s statistics instruction covered data analysis. Most episodes included data analysis at Level 2, constructing graphs, answering descriptive questions based on graphs, and calculating measures of central tendency. Furthermore, there were episodes at Level 3, including further data analysis, focusing on the distribution concept and comparisons of the groups based on graphs and measures of central tendency.

4.1.2.1.2.4. Interpretation

The interpretation element in Esra’s statistics instruction was absent in 19 episodes (59.37%) (see Table 4.9). Those episodes included the construction of a circle graph and a line graph. When they constructed the graphs, they did not make any interpretations based on the graphs. Furthermore, there were some calculations with central tendency measures for a data set without contexts. They were just practices on the calculations. Therefore, no interpretation occurred during practice time. There

were also some examples of data sets with context. However, even if the teacher had an opportunity to interpret calculated average means based on the context, the teacher just demanded students calculate the arithmetic mean, median, and mode; and decide which one to use based on the presence of an “outlier.”

In 12 episodes (37.50%), the teacher and students interpreted the data based on graphs and the measures of central tendency to answer “within” data questions. Those episodes were scored at Level 2. For example, in introducing the circle graph tasks, the teacher asked students to interpret the circle graph they constructed. The teacher did not have a set of questions for students to answer. She said, “Let us look at the graph and make interpretations. What can you say by looking at the graph?”. The students answered the open-ended questions about what they saw in the graph. The students' responses were “the least favorite color of 7D was red” and “the most favorite color of 7D was blue.” In addition, when the teacher introduced a line graph task from the book, she characterized the task as an “interpretation task” to the students. She also added that they would focus on the questions related to “the increase or the decrease” in the line graphs. They responded on which day the selling price of the gold decreased and increased or did not change. They also determined on which day the selling price of the gold was the highest or the lowest.

There were similar within-data questions based on bar graph tasks too. In addition, some average tasks used interpretations at Level 2. Based on the measures of central tendency, they commented on the context. For example, the teacher said, “Let us assume that Elif got 20, 30, and 100 [in math exams]. If I want to talk about Elif’s success, what can I say?... What was the average?”. The class said 50, and the teacher compared 50 with each grade point and asked students to decide on Elif’s success based on [arithmetic] mean. Then the teacher concluded that Elif was not successful. Those episodes were based on reading the data, not considering the relationship among data or reading beyond data.

Moreover, one episode (3.13%) was at Level 3, discussing the variability and comparing groups. The students were expected to determine the more successful class according to students' exam results in both classes. The students and the teacher mainly

discussed which central tendency measures were appropriate. Then they decided not to use the arithmetic mean because one class had outliers, so the arithmetic mean did not represent the class' success. It was not healthy to compare both groups based on the arithmetic mean. After all, they compared groups based on median and mode and decided more successful class. This example helped students consider the variability in both data sets to make inferences about the success of the classrooms.

All in all, most episodes in Esra's statistics instruction did not use interpretation. Only one episode was at Level 3; the students and the teacher made inferences about two samples and compared their mathematics exam results to decide which was more successful. The episode included comparisons across samples. The remaining episodes were about making statements based on the data; they were at Level 2 and included example answers to data questions.

4.1.2.1.2.5. Overview of Esra's Statistical Problem-Solving Process

To understand Esra's statistics instruction in terms of statistical problem-solving processes, I analyzed the co-occurrence of different categories of elements in each teaching episode using MAXQDA. Table 4.10 presents the frequencies of these co-occurrences, revealing the interplay between various categories of statistical problem-solving elements. This analysis provides valuable insights into Esra's instructional practices regarding teaching statistical investigations

Table 4.10 shows that most of Esra's statistics instruction did not focus on the statistical problem-solving process. The focus on posing statistical questions and data collection was absent except for one episode. The data analysis was present for all episodes. In addition, most episodes with data analysis were followed by the interpretation step. As a whole process, I can conclude that Esra's instruction did not focus on the statistical investigation. Table 4.10 shows the frequencies of the co-occurrences of the steps in the statistical problem-solving process.

In 15 episodes (46.88%), only data analysis occurred at Level 2 (i.e., graph constructions, calculating descriptive questions, and calculating average) with the

other elements at Level 1; no formulation of questions, data were provided, and no interpretation occurred. Similarly, the data analysis at Level 3 and other elements at Level 2 co-occurred in three episodes (9.38%). On the other hand, in 8 episodes (25%), the data analysis at Level 2 co-occurred with interpretation at Level 2. However, the episodes did not cover the formulation of questions and data collection. In four episodes (12.50%), data analysis at Level 3 co-occurred with interpretation at Level 2. One episode also included both data analysis and interpretation steps at Level 3.

Table 4.10

Esra's Statistical Problem-Solving Process: The Co-occurrences of the Level of Elements

Formulation of Questions	Data Collection	Data Analysis	Interpretation	Episodes (n)	Percent (%)
FQ Level 1	DC Level 1	DA Level 2	I Level 1	15	46.88
FQ Level 1	DC Level 1	DA Level 2	I Level 2	8	25.00
FQ Level 1	DC Level 1	DA Level 3	I Level 1	4	12.50
FQ Level 1	DC Level 1	DA Level 3	I Level 2	3	9.38
FQ Level 2	DC Level 2	DA Level 3	I Level 2	1	3.13
FQ Level 1	DC Level 1	DA Level 3	I Level 3	1	3.13

Note: N=32 (the total number of episodes) FQ = Formulation of Questions, DC= Data Collection, DA= Data Analysis, I= Interpretation

There was only one episode that covered the whole problem-solving process. This episode was the first episode of Esra's statistics instruction. Before the first lesson, I asked Esra what she planned to do; she said she would start asking what statistics was and look for what they remembered from previous years. Esra introduced the topic in the lesson and mentioned what the students had covered in the previous years. Then students stated that they did not cover the topic in the previous year because the mathematics teacher prepared a project and skipped it. The teacher started the topic from scratch:

Okay, we do not know what the statistics are, so let us start from the beginning... Statistics is the interpretation and evaluation of the data obtained from statistical research with the help of mathematics. It does not mean anything right when you say it as a definition. However, imagine there is a subject you are researching and what we should do to get information about this subject. There are some ways to gather information; there are some

methods; for example, one of them is a survey; we create a survey and distribute it to people. They answer the result, and we take it back and examine what they wrote. Let us do something simple ourselves. Shall we say the most favorite color?

[Esra 06.04.2018 7D (2) 00:00-15:00]

Then the teacher determined four colors and told students they could only choose one option. The teacher asked the whole class some questions to collect data from students. Then she organized the collected data on a table. She described what statistics was with an emphasis on “data,” “data collection,” “questions asked for data collection,” and research questions. Then she analyzed collected data and interpreted them:

What did I do now? I wondered. I said, what is the 7D class's favorite color? I was curious, a question mark appeared in my head; and I asked you what I did to satisfy this curiosity. I asked what your favorite color was. I just created a table. Our table is the color that the 7D class loves. Looking at this table, I learned that three students like red, 15 like blue, and six like green. Is it true? Now I can show you this in different ways. The way it is displayed is called a graph. We will learn three graphs. The first is the circle chart, the second is the column chart, and the third is the line chart. I can visualize these results with graphs. So I can make you see more clearly there. However, I was curious; I asked you a question. Can I come to a conclusion based on the information I have obtained? What can I say? So, can I say that most of the students in the 7D class like the color blue the most? Can I say that the least favorite color [in 7D class] is red? Or can I say that the number of students who love red is three less than those who love green? See, I can get different results. However, what do I need this data to achieve this result?

[06.04.2018 Esra 7D (2) 00:00-15:00]

Afterward, she introduced and constructed a circle graph with students. Then she wanted students to interpret what they saw in the circle graph. The episode with the statistical problem-solving process was an example set up and applied. Therefore, the flow was not like starting with a statistical question; the teacher later emphasized the research question. In addition, it was a significant emphasis on “wonder” while starting a statistical investigation. The episode started with a statistical investigative question, followed by data collection and then data analysis and interpretation.

The results show that many episodes focused on data analysis. It was followed by data analysis and interpretation of the analyses. Only one episode consisted of all steps in

the problem-solving process; the sole episode included formulating statistical questions and data collection. After about two months, I left the field; I conducted the last interview with the teacher. In the last interview with Esra, when asked what would make her statistics teaching better [than the observed one]. She responded with an emphasis on students' engagement in the statistical problem-solving process:

We would be going to start something ourselves so that we would say, let us start investigation; we would have the child collect information from their environment from daily life. We had to have him do that work exactly. They would bring the data to the class. For example, we could give the same subject to the class. We would create the table immediately on the board. Then [we would discuss] what kind of graph we can draw suitable for that table, [we would say] let us draw all the appropriate ones, let us interpret now, we could do something like this, that is, we could [bring] everyday life to the classroom.

[Esra's Last Interview 10.06.2018]

Even if she did not allow the whole class to experience the statistical problem-solving process individually and collectively, she repeated the need for statistical investigations when asked what was needed to improve statistics instruction. Furthermore, she stated that the students did not ask questions about statistics after the observation ended. Still, she gave a project requiring students to research three topics they were interested in. Based on her description, she expected students to select a research topic and investigate the topic by collecting data and representing the data using appropriate visualization. She gave importance to statistical investigations. When asked what she would want from policymakers for better statistics education, she demanded fewer students, technology integration, and time for starting the statistics topic with statistical investigations. She said, "bringing data from our daily life."

All in all, only one episode followed the complete cycle of statistical problem-solving, from formulating the question to data collection, data analysis, and interpretation in Esra's teaching practices. The focus was mainly on data analysis, with minimal co-occurrence of other elements such as formulating the question, data collection, and interpretation. These findings suggest a need for greater emphasis on the complete cycle of statistical problem-solving in Esra's teaching practices parallel with what Esra emphasized in the last interview.

4.1.2.2. Mathematical and Statistical Quality of Instruction (MQI): Segment Level Analysis of Esra's Statistics Instruction

This section describes the quality of instruction in Esra's statistics instruction. This subsection includes the segment-level analysis with 21 codes under four dimensions and the lesson-level analysis with one code from the MQI instrument. According to MQI, the lessons were divided into 7.5-minute segments. Esra spared 13 lessons for statistics instruction. I analyzed 12 lessons with different segments according to MQI because one lesson consisted of students taking a quiz about the concepts they covered in statistics instruction. Nine out of 12 lessons comprised five segments with four 7.5-minute segments and one three minutes or more except one with five 7.5-minute segments. Four lessons consisted of four 7.5-minute segments. In total, I scored 57 segments with 21 segment-level codes. I scored 12 lessons with a whole lesson code. The following sections describe the quality of Esra's statistics instruction based on segment-level analyses and a lesson-level analysis.

This section presents the results of the segment-level analysis regarding the MQI of Esra's statistics instruction. The section was divided into four main dimensions. Each dimension consisted of subsections related to the segment-level codes. First, the summary statistics of each dimension's segment-level codes are presented. Next, I will qualitatively describe Esra's instruction in each segment-level code. Then I will provide an overview of each dimension in MQI. The following section starts with the first code, as Classroom Work is Connected to Mathematics and Statistics. After the results of the segment-level analysis, I describe Esra's statistics instruction with the whole-lesson quality of instruction code from the lesson-level analysis.

4.1.2.2.1. Classroom Work is Connected to Mathematics and Statistics

Esra's instruction connected classroom work to mathematics and statistics ($n=56$, 98.25%). Those segments focused on mathematical and statistical content over 3.75 minutes of 7.5-minute segments. Esra spent time writing questions on board in some segments, but the time she spent did not last the half-length of the segment (3.75 minutes). In addition, Esra expected students to read and think about the task while

she was writing on the board. Therefore, those segments also focused on mathematical and statistical content. In the segments that scored Yes for Classroom Work is Connected to Mathematics dimension, Esra introduced new concepts (e.g., circle graph, line graph, the measures of central tendency). She reviewed the content from previous lessons (especially in the transformation of graphs). She practiced the content and solved tasks on the board. She discussed content with the students. Even if there were nonmathematical incidents in those segments (e.g., taking attendance), it did not last more than half of the segment. She focused on mathematics and statistics; she briefly described the topic and solved as many questions as possible on board.

Only in one segment she spent approximately the whole time checking on students' homework. She expected students to open the homework page and wait for her. She did not give directions to work on a task for students whose homework was checked. Therefore, I scored "No" for Classroom Work is Connected to Mathematics and Statistics dimension for the segment. In the following segment, the teacher asked students to solve the problems on the worksheet, which were assigned as homework for themselves. Therefore, the segment which did not focus on the content shaped the instruction for the lesson.

4.1.2.2.2. The Richness of the Mathematics and Statistics

This section explains the richness of the mathematics and statistics of Esra's statistics instruction. I examined the richness dimension with seven codes in the MQI. Table 4.11 shows the results regarding the richness dimension in the MQI. According to Table 4.11, the richness of Esra's statistics instruction can be characterized as Not Present and Low-quality for most codes in the richness dimension. Some codes had a small number of high-quality incidences (e.g., mathematical and statistical sense-making). Furthermore, Esra's statistics instruction did not include Patterns and Generalizations at all lesson segments (n=57, 100% absent/not present). Linking between Representations (n=51, 89.47%) and Multiple Procedures or Solution Methods (n=49, 85.96%) were absent in many segments. In addition, the explanation was absent in 35 (61.40%) segments.

Table 4.11

“The Richness of the Mathematics and Statistics” Dimension Scores in Esra’s Instruction

	1		2		3		4	
	Not Present		Low		Mid		High	
	n	%	n	%	n	%	n	%
Linking Between Representations	51	89.47	5	8.77	1	1.75	0	0.00
Explanations	35	61.40	11	19.30	10	17.54	1	1.75
Mathematical and Statistical Sense-Making	4	7.02	8	14.04	33	57.89	12	21.05
Multiple Procedures or Solution Methods	49	85.96	5	8.77	3	5.26	0	0.00
Patterns and Generalizations	57	100	0	0.00	0	0.00	0	0.00
Mathematical and Statistical Language	1	1.75	10	17.54	43	75.44	3	5.26
Overall Richness of the Mathematics and Statistics	6	10.53	31	54.39	19	33.33	1	1.75

Note. N=57 (the total number of lesson segments)

Furthermore, there was no ‘high’ quality in the Linking between Representations, Patterns and Generalizations, and Multiple Procedures or Solution Methods (0% high for all). Mathematical and Statistical Sense-Making and Mathematical and Statistical Language were present in Esra’s instruction.

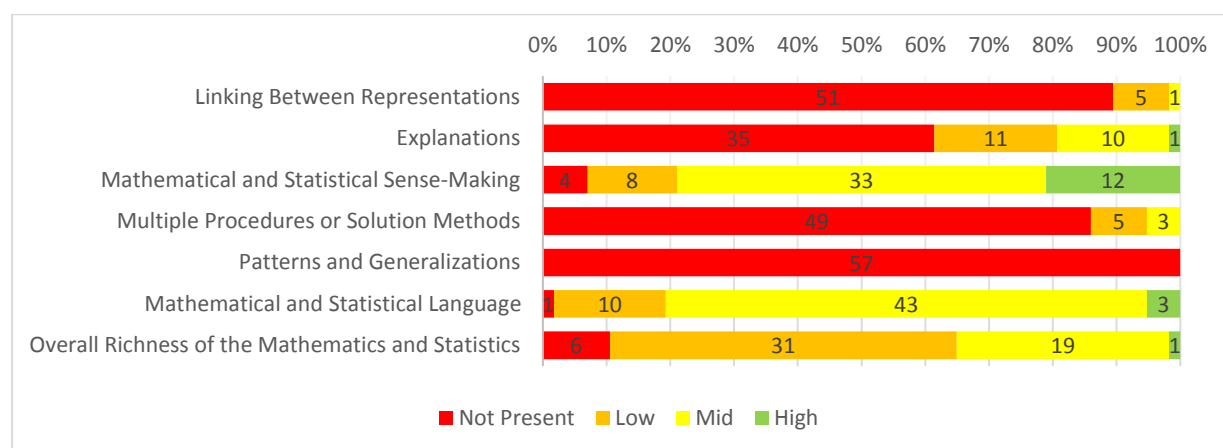


Figure 4.39 “The Richness of the Mathematics and Statistics” Dimension in Esra’s Instruction

The richness of mathematics and statistics in Esra’s instruction was partially present in mathematical sense-making, mathematical language, and the overall richness of mathematics (see Figure 4.39). However, explanations, linking between representations, multiple procedures or solution methods, and patterns and generalizations codes indicated the lack of richness in Esra’s statistics instruction. In the following sections, I will provide codes in the richness dimension of Esra’s statistics instruction.

4.1.2.2.2.1. Linking between Representations

Figure 4.40 shows the distribution of scores assigned to lesson segments for the Linking Between Representations code, describing the explicit links between representations from different representational families. Esra’s statistics instruction lacked the link between different representation families in 51 (89.47%) lesson segments. Five (8.77%) lesson segments included links, but those links were not explicit or detailed. Esra linked different representational families in one (1.75%)

lesson segment, but her emphasis on the connections was not the segment's focus. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the linking between representations code was 1.12 out of 4, with a standard deviation of 0.38. This section provides the descriptions and classroom examples of each score assigned to the Linking Between Representations code in Esra's statistics instruction.

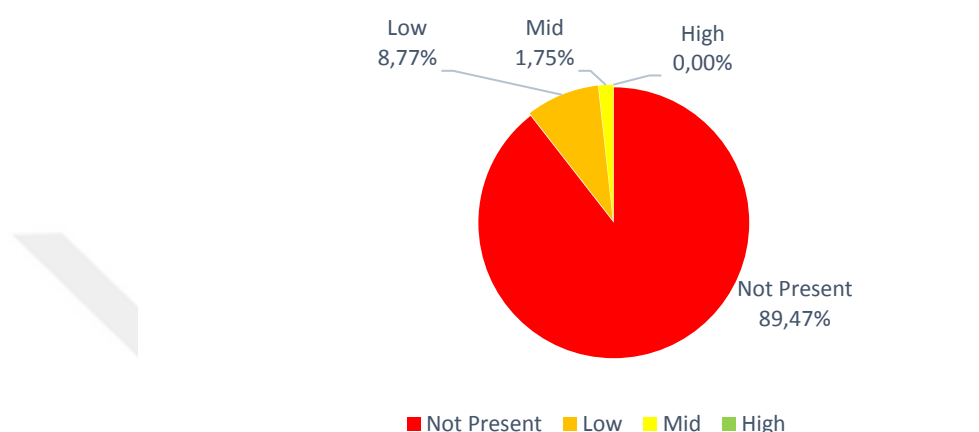


Figure 4. 40 The Distribution of Scores Assigned to Lesson Segments for “Linking Between Representation” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I assigned a “Not Present” score to 51 (89.47%) lesson segments for linking between representations since those segments did not include connections between different representations of mathematical and statistical concepts. Because of the nature of the objectives in statistics instruction, many segments included representations (i.e., visual and symbolic). For example, the segments included constructing graphs or working on existing graphs. While constructing the graph, there were some symbolic representations. Therefore, although some segments included different representations, the connections among the representations were not observed. In some segments, graph transformation was covered, and different graphs were used for one context. However, I did not observe the connections among different representations in those segments, although there were explicit links among the same representational family (Hill et al., 2014, p.5). In addition, some circle graph examples did not require students to draw the graph but found the values related to the sector of the graph. Only

symbolic representation was present. No graphical representation was present and connected. While covering average objectives, just one example consisted of linked representations. However, the teacher did not use possible different representations for the average concept; she did not know possible interpretations; and representations of the average based on the statistical content interview.

Esra connected different types of representations at a low level in five (8.77%) lesson segments. One lesson segment included teacher gestures related to the connections between the circle graph and the angle. While emphasizing the 360-degree angle corresponding to the whole circle, she drew the whole circle in the air with her finger repeatedly. Three of those lesson segments included the construction of a circle graph. During constructing a circle graph, the teacher and students calculated the angle for each category corresponding to the given data. In all examples, one category corresponded to a 90-degree angle. The teacher started with the sector at a 90-degree angle. In one example, she drew the sector by dividing the whole circle into half and then dividing half of the circle into half. She said the 90-degree angle corresponded to one-fourth of the circle.

Then she drew the sector with a 45-degree angle by dividing the remaining half-circle (the fourth part) into half. While she drew the sectors, she described what she did. The other two examples included drawing the sectors corresponding to a 90-degree angle with similar explanations. In constructing the circle graph related to the ingredients needed for making strawberry jam [i.e., 8 kg strawberry, 3 kg sugar, and 1 kg water], the students found that 1 kg water corresponded to the sector with a 30-degree angle. While drawing the graph, the teacher emphasized that 30 degrees were one-third of 90 degrees, divided the 90-degree angle into three, and labeled the sector as a water ingredient of strawberry jam. All examples related to constructing the circle graph were scored Low because the links between graphical and symbolic representation were unclear and not detailed. The links were relatively brief. In addition, the teacher mentioned one-fourth and one-third of the 90-degree angle in the examples. However, she did not write the fractions on the board or explicitly link the fractions with the graph.

Furthermore, I only scored one (1.75%) segment as Mid. The segment also included constructing a circle graph task. The data was given in the frequency table. The teacher found the angle for representing a person and found the angles for each category. While drawing, she purposefully selected two categories that made a total of 180 degrees (60 degrees plus 120 degrees). Then, she drew the sector at a 96-degree angle by connecting it with the sector's size at a 90-degree angle. Then she labeled the remaining sector. In the process of the construction of the circle graph, explicit connections between the symbolic representation and with graphical representation were present. However, the segment was not high because the teacher did not elaborate on the relationship between different representations. She used proportional reasoning while finding the angles corresponding to the categories of the data provided. Then she constructed the graph based on the calculations. However, she did not elaborate on the proportionality of the size of the sectors. She did not transfer her reasoning behind her calculations to the circle she drew.

Specifically, 51 lesson segments (89.47%) were scored as Not Present, 5 lesson segments (8.77%) were scored as Low, and 1 lesson segment (1.75%) was scored as Mid. No lesson segments were scored as High. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the linking between representation code was 1.12 out of 4, with a standard deviation of 0.38. This suggests that, on average, Esra's statistics instruction was lacking in explicit connections between different representational families.

4.1.2.2.2. Explanations

Figure 4.41 shows the distribution of scores assigned to lesson segments for Esra's mathematical and statistical explanations based on the answers to the "why" question. In many lesson segments, the teacher or students did not offer mathematical and statistical explanations ($n=35$, 61.40%). Esra provided explanations as an isolated occasion at 11 (19.30%) lesson segments. In addition, 10 (17.54%) lesson segments included mathematical and statistical explanations at the mid-level. Only one segment consisted of mathematical and statistical explanations as to the focus of the instruction (1.75%). In general, mathematical and statistical explanations in Esra's statistics

instruction were about why to use a specific method while calculating angle, percentages, or quantities in a circle graph task, why a procedure work or does not work, or why to use a specific graph or a measure of central tendency. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the explanations code was 1.60 out of 4, with a standard deviation of 0.84. This section provides the descriptions and classroom examples of each score assigned to the Explanation code in Esra's statistics instruction.

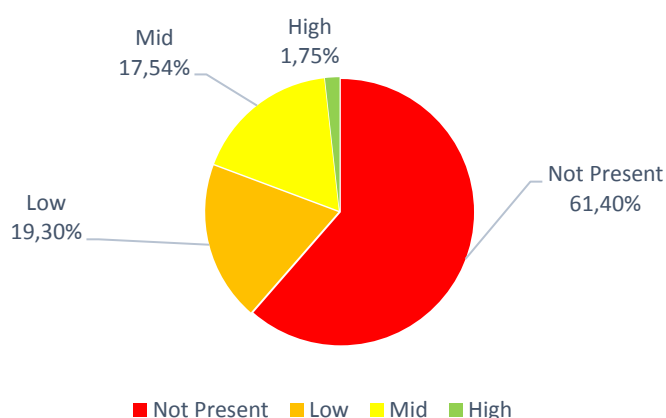


Figure 4.41 The Distribution of Scores Assigned to Lesson Segments for “Explanations” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I scored Not Present to 11 (19.30%) lesson segments for providing mathematical and statistical explanations because those segments did not include the answers to why questions. In those segments, the teacher mainly described the steps of the calculations for constructing a circle graph or finding a value in the circle graph, line graph, or bar graph. Moreover, she described how to draw a line graph in the first line graph task. She did not discuss why to use a line graph for representing a specific set of data. She only discussed when to use a line graph. In addition, none of the bar graph tasks included mathematical and statistical explanations. She introduced the central tendency measures without focusing on why; she described how to calculate the arithmetic mean, median, and mode of a data set in the introduction lesson. However, she later explained why choosing a specific measurement for a given data set.

I scored Low to 11 (19.30%) lesson segments for the Explanations code in Esra's statistics instruction. Mathematical and statistical explanations were about why a particular graph would be used, why a line graph started at 0, why a procedure related to proportions worked, or why the answer was invalid. At the beginning lesson of statistics instruction, the teacher collected data from the students and recorded the data on a frequency table. Then she stated that the purpose of learning graphs was to see results more clearly and visually. This explanation was necessary for students to understand why they learned graphs, but the explanation occurred briefly.

Furthermore, in a circle graph task, the circle showed the distribution of goals from the penalty. There were two categories; scored goals and missed goals. The number of total goals was given. Students were expected to find the number of missed goals with the corresponding angle. The teacher asked a student how to solve it. The student described the solution as subtracting 144 degrees (the angle of missed goals) from 360. Then the teacher stopped him and explained why his solution steps were not correct:

You want to calculate the rest. Now the rest, look, the missed goals is 144 degrees. Why are you trying to calculate the rest? Why are you trying to calculate the goals scored? The question did not ask us about the goals scored; what did it ask us? It asked for the missing ones. So, finding the angle of the remaining circle segment there would not mean anything to you. You act on what is given and what is asked.

[Esra 09.04.2018 7D 07:30-15:00]

Esra explained why subtracting 144 degrees from 360 degrees was incorrect, what the result of subtraction referred to, and what the student should consider while solving the task. However, it was a brief explanation. Therefore, I scored Low. In the following parts, the teacher asked another student how to find the value of the missing goals. A student provided the descriptions of solution steps.

I assigned a score of Mid to 10 (17.54%) lesson segments for explanations code in Esra's statistics instruction because those segments included two or more brief mathematical and statistical explanations, or the explanation was present more than briefly. However, the explanations were not the focus of the instruction. Mid scores were given to the segments that involved mathematical and statistical tasks related to

measures of central tendency. These segments provided a detailed explanation of why arithmetic mean, median, or mode should be chosen for a given data set instead of a brief one.

For example, in an average task, the question was asked to find the “mathematical average” (“matematik ortalaması” in Turkish) of a class. A student calculated the arithmetic mean to find the mathematical average of the class. Then the teacher asked why they calculated the arithmetic mean for finding the average since the question did not include the statement about finding “arithmetic mean.” Then they calculated the median and mode of the given data set. Then the teacher asked students to compare each value. All measures of central tendency had similar values for the example task (arithmetic mean 3.6; median and mode 4). She asked if they could use all values because they were close to each other. One student stated that she would instead use arithmetic mean. Then the teacher explained that they could use the arithmetic mean because it gave more detailed information about the data set. In the same segment, after she explained why to use the arithmetic mean for the given context, she asked students why to use the arithmetic mean; one student gave incomplete and incorrect explanations; the following segment included a correct brief explanation of a student.

Similarly, in another example, the teacher introduced the following task to the classroom for a part of the segment. In the example, the students were expected to find the more appropriate type of average for selecting course type for the summer school (see Figure 4.42). The teacher asked students which one to choose. A student answered, “Mode.” The teacher explained why the mode was appropriate. She also explained why not calculate the arithmetic mean but the mode. She provided mathematical and statistical explanations more than briefly. She explained that she could not use the arithmetic mean because the data were not numerical. They were categorical, so they could not do the calculations. She said that the data was like football, football, football, so they could not use other measures of central tendency other than mode. Therefore, I scored Mid for the segment, but the explanation was not the focus of the instruction.

Course Types	Football	Basketball	Volleyball	Handball
Number of people	120	50	40	60

"Which course would you like to open on the weekend?" question was asked to the students at a school. The results given in the table above were obtained. Accordingly, which type of average would be more appropriate to use to obtain information about which course will be offered at the school?

Figure 4.42 Example Task 1: Course Types (Varışlı, 2017, p.162)

In the following segment, the teacher introduced a similar task from the book. She asked students which average concept should be used for the task with another context. The context was about the number of babies visiting the emergency service (Figure 4.43).

ÖRNEK

Bir hastanenin acil servisine son 11 günde getirilen bebeklerin sayısı aşağıdaki gibidir.
8, 9, 6, 5, 5, 7, 10, 60, 8, 5, 9

Buna göre acil servise getirilen bebek sayılarını temsil eden en uygun merkezi eğilim ölçüsünü bulalım.



The Question Part:

The number of babies brought to the emergency service of a hospital in the last 11 days is below:
8,9,6,5,5,7,10,60,8,5,9

Find the most appropriate central tendency measure for the number of babies brought to the emergency service.

Figure 4.43 Example Task 2: Hospitalization (Bilen, 2017, p.234)

The teacher mentioned that this particular example would help the students better understand when to use a specific measure of central tendency compared to other examples from the book. In those examples, the data values led to similar results for the arithmetic mean, mode, and median, so the students did not have the opportunity to grasp the distinct characteristics of each measure for different contexts.

This segment was the only one that scored High for the mathematical and statistical explanation. During the segment, the students gave answers and explained their reasoning about which measures of central tendency represented the typical number of

children coming to the emergency hospital per day. After discussing it as a whole class, the teacher answered and explained why not use the arithmetic mean and mode. She also explained why to use the median.

.....If I talk according to the arithmetic mean, an average of 12 babies visit the emergency each day. I am looking at the numbers 8,9,6,5,5,7,10,60,8,5,9, and look; it is close to 10. If the numbers here were close to 10, 11, 12, 13, 14, okay, arithmetic mean could be meaningful... The data from here and the arithmetic mean are very different from each other... What is the reason behind this? This number is 60. The distribution here is stable, but here this day came 60 babies; this [number] 60 increased arithmetic mean. What happened? Sixty patients came all of a sudden. The “60” increased the [arithmetic] mean. Calculating the arithmetic mean was unhealthy because my [arithmetic] mean is high because of 60. Then we said that we should not look at mode. Our mode is 5. Mode is our minimum number here. At least five patients were coming to the hospital already... Why did we not use mode? Because the minimum value is 5, so we canceled it. Now we came to the median. The Median is 8. These numbers, except 60, are close to 8. The number that best represents the mean is 8; it is healthier to look at the median... there is 60, we cannot cancel 60, but the healthiest number representing these [shows the data] is 8, okay?

[Esra 17.04.2018-7D 22:30-30:00]

The mathematical and statistical explanation was the focus of instruction in the segment. The teacher heard students’ reasoning and responded accordingly, emphasizing why questions shaped the segment. Furthermore, while solving the previous tasks in the lesson, the teacher stated that she would provide a more precise task to differentiate which measures of central tendency to choose for a given context by referring to this task.

4.1.2.2.2.3. Mathematical and Statistical Sense-Making

Figure 4.44 shows the distribution of scores assigned to lesson segments for mathematical and statistical sense-making in Esra’s instruction. This section explores teaching practices in attending to the meaning-making process. The results show that four (7.02%) lesson segments did not include mathematical and statistical sense-making. In eight (14.04%) lesson segments, the focus on the meaning was minimal. Most lesson segments (n=33, 57.89%) contained meaning-making incidents more than briefly.

Only in 12 (21.05%) lesson segments the focus on mathematical and statistical sense-making was substantial and sustained during those segments. Esra focused on meaning during calculations in the construction of the graphs, interpretations of graphs, and discussing average concepts within context during segments with sense-making.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the mathematical and statistical sense-making code was 2.93 out of 4, with a standard deviation of 0.80. This section provides the descriptions and classroom examples of each score assigned to the Mathematical and Statistical Sense-Making code in Esra's statistics instruction.

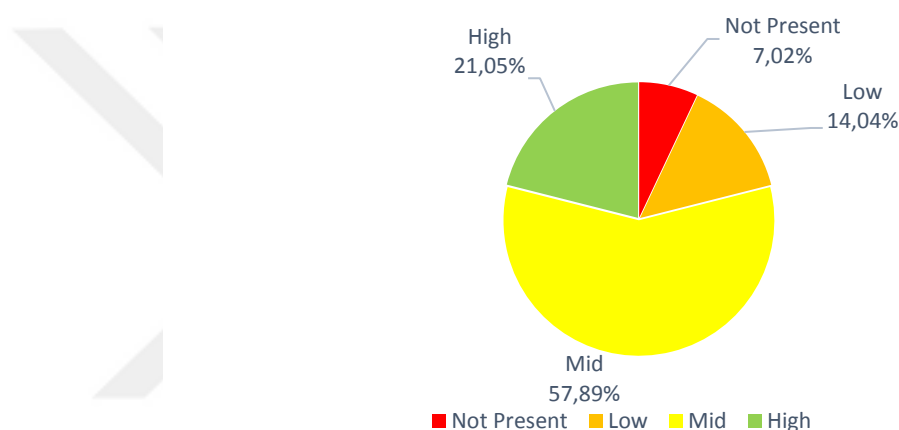


Figure 4.44 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Sense-making” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I scored “Not Present” in four (7.02%) lesson segments for mathematical and statistical sense-making of Esra’s teaching practices. One segment was the one with no mathematical and statistical work. In the segment, the teacher checked the homework. The other three segments consisted of partially correct and incorrect sense-making processes. As the MQI suggested (Hill et al., 2014), I present incorrect portions in the Errors and Imprecision section. I considered the correct portion of the segments when scoring for mathematical and statistical sense-making. However, the remaining parts included the teacher’s and students’ descriptions of procedures without attending to the meaning. Two of the three segments included teacher and student descriptions of

finding the mode and median of a data set. Some descriptions were not correct or complete, and others focused on conveying sole definitions or procedural aspects without meaning-making. The last one was about drawing a line graph. The context given for the line graph was incorrect; during drawing, the teacher only focused on the procedural accuracy of the line graph. Therefore, those segments were not examples of mathematical and statistical sense-making.

I scored Low to eight (14.04%) lesson segments for mathematical and statistical sense-making because Esra briefly focused on the meaning-making process. At the beginning of the lessons, the teacher wrote the task on the board in some segments. Then she asked students to read the task while she was writing. The segments consisted of mathematical and statistical work because students focused on the task individually in most 7.5-minute segments. When the teacher finished writing, she asked the questions in the task to students. Therefore, a small portion of the segments included sense-making. In some segments, the teacher and students focused on the procedures. The brief moments were present to give meaning to procedures. The teacher asked students to find the arithmetic mean of given data sets in a segment. The students added all values in the data set and divided them by the number of values. For example, when a student divided 116 by four, the teacher helped the student solve the division step by step. She asked how many fours were in 11 and how many fours were in 36. The teacher helped students to make sense of the division process. However, they did not focus on the meaning of arithmetic mean.

I assigned a Mid score to 33 (57.89%) lesson segments for mathematical and statistical sense-making. In some segments, Esra focused on the sense-making process more than briefly. For example, Esra helped students make sense of circle graphs, focusing on proportional reasoning while calculating the angle or percentage of the categories provided. However, the work on calculations did not include a sustained focus on sense-making. Some segments also included several instances of the relationship among numbers during calculations. Additionally, Esra and her students also focused on making sense of tables and graphs. For example, Esra helped students make sense of the increase and decrease in a line graph. She also discussed the raw data format of a frequency table for students to make sense of the frequency table more than briefly.

Those segments were scored as Mid because the instances with sense-making were not sustained or substantial, but the focus on meaning was present more than briefly or frequently.

In 12 (21.05%) lesson segments, the focus on mathematical and statistical sense-making was substantial and sustained. Therefore, those segments were scored as High for mathematical and statistical sense-making. Esra and her students focused on meaning-making on statistical concepts while covering each topic's introductory examples. For example, Esra started the statistics unit by defining “statistics.” She gave the definition and stated, “It does not mean anything [to you], right? when you say it as a definition”. Then she elaborated on the definition by describing the elements of the statistical problem-solving process. Then she provided an example for students to experience the statistical problem-solving process. With such descriptions and examples, they made sense of the definition of statistics and the statistical problem-solving process. In the same example, she also introduced the circle graph. The students engaged in constructing the circle graph for the first time. The teacher focused on proportional reasoning by working out the angles of the sectors. Then she demanded that students read the graph without guiding questions. She wanted students to make sense of the circle graph.

Furthermore, Esra also focused on the meaning of average concepts in a sustained way. For instance, she covered average tasks that required students to choose an appropriate measure of central tendency for a given context. She followed examples from a textbook. In the textbook, all measures of central tendency were very close to each other, while the textbook demanded that students choose one. Therefore, the teacher wanted students to calculate all measures and see how close they were. She chose the example presented at the only segment with a high explanation score to make sense of the different average types for different data sets. She wanted students to make sense of the distribution of the data set and listened to their reasoning for selecting a measure of central tendency for the example. In the end, she concluded that:

If we were talking about this arithmetic mean, what would we say? We were going to say that this hospital brought an average of 10 babies for 11 days, right? [finding the mode] If we were to look at mode, an average of 5 babies

were brought to this hospital for 11 days, for the emergency room If I were speaking accordingly, I would say what I would say, an average of 8 children came to the hospital. Look, they are all very different from each other. The previous examples were always close to each other; the arithmetic mean, mode, and median were close to each other in all previous examples; we could use all three. Now look, here is the question we need to understand which one to use. Which one to use in the questions we have already solved so far is very well distinguished in this question.

[Esra 17.04.18-7D 30:00-36:00]

Esra emphasized why she chose the example; and how the example was different from the other examples. She focused on the distribution of a data set. She did not follow what the textbook asked and explained why they did not calculate one measure for describing the data set in the given examples. With the example above, she emphasized what kinds of examples were meaningful for them to choose appropriate measures of central tendency.

4.1.2.2.2.4. Multiple Procedures or Solution Methods

Figure 4.45 displays the distribution of scores for Multiple Procedures or Solution Methods (MPSM) in Esra's statistics instruction, which consisted of discussions on using multiple procedures and solution methods. However, in 49 (85.96%) of the lesson segments, Esra did not use or discuss multiple procedures or solution methods. Five (8.77%) lesson segments included brief statements on a second procedure or method.

In three (5.26%) lesson segments, multiple procedures or solution methods were examined at length, but no detailed comparison or discussion was present. Esra and her students discussed a second method at various lengths, especially in constructing a circle graph and calculating the angle, percentages, or quantities corresponding to the categories of given data. Furthermore, none of the segments had multiple procedures and solution methods at a high level.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the MPSM code was 1.19 out of 4, with a standard deviation of 0.52. This section provides descriptions and classroom examples of each score

assigned to the Multiple Procedures or Solution Methods code in Esra's statistics instruction.

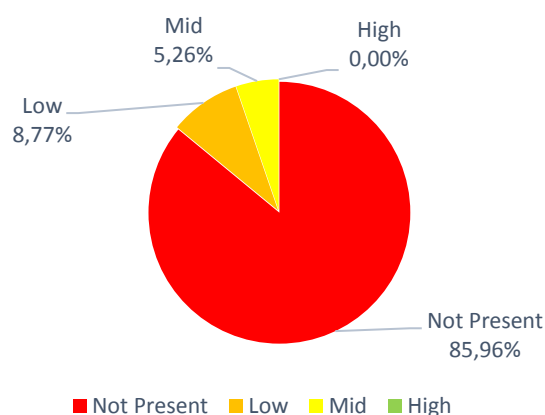


Figure 4.45 The Distribution of Scores Assigned to Lesson Segments for “Multiple Procedures or Solution Methods” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I assigned a score of Not Present to 49 (85.96%) lesson segments for MPSM because Esra used only one method or procedure during those segments. Furthermore, the segments with reading graphs did not generally require students to deal with multiple procedures and solution methods. She only used one method for some graph tasks, including data analysis questions. In addition, those segments did not include the work on average tasks. When the teacher asked if there was another way to find the arithmetic mean in the statistical content interview, the teacher stated:

There is only one formula, right? There is only one formula. We can calculate mentally in a practical way. We can find which number is close, or the difference [among numbers], but the formula we teach is one. We teach with one way [method].

[Esra- Statistical Content Interview 10.06.2018]

The teacher focused on one procedure for solving average tasks while different average interpretations were present. She also reported teaching only one method for finding the arithmetic mean.

I scored Low to 5 (8.77%) lesson segments for MPSM because Esra or her students mentioned a second method but did not discuss it in detail or use it explicitly. In a

circle graph task, the central angle corresponding to a category was asked by giving the percentage of the sector and the total number corresponding to the whole circle. A student solved the question by finding the number corresponding to the category, then attempted to use the result to find the angle. Then the teacher interrupted the student and explained that she did not have to find the number corresponding to the category. She stated that finding 25% of 360 would give the result asked. Then she wanted students to solve with her method. The first method the student attempted was incomplete and discussed briefly and scored as Low for MPSM.

In another example, the student was expected to find the arithmetic mean of a data set provided in a frequency table. The student added the values without considering the frequencies at first. Then another student explained that they needed to think about the frequencies to multiply the value with the frequencies. They worked on the method the student suggested. However, the teacher briefly mentioned the second method of adding all values instead of multiplying with the frequencies.

I scored three (5.26%) lesson segments Mid for MPSM because the teacher used more than one solution method or procedure. However, the discussion of the two methods remained distinct, and the comparison among solution methods was not explicit or at length. For example, while representing 7D's favorite colors in the circle graph, Esra used a direct proportion and fractional approach to find each category's angle in the data. Firstly, she found the angle of the red category, which consisted of three students. Then she stated that they could use proportions for other categories similarly. Then she mentioned a second method; she used the frequency table to organize data for the ratio table. After finding the angle of the red category, she used a multiplicative relationship among the frequencies corresponding to each category to find the angles of each category. The second method was presented at length, but the teacher only compared the results with the two methods. They did not compare both methods in terms of advantages and disadvantages. In another example, there were four categories. The teacher used proportions to find the angle of a category. Then a student suggested that they could find each person's angle through the knowledge of the angle and quantity of a category. They divided 96 by 8 to find the angle corresponding to one person. Then the teacher stated they could find the angle corresponding to one person

by dividing 360 by the total number of people (30). To calculate the angle corresponding to other categories, they multiplied the quantities by the angle corresponding to per person. The teacher and students did not compare the two methods using the proportion and angle of each person; both methods remained distinct. However, both methods were presented explicitly and at length.

4.1.2.2.2.5. Patterns and Generalizations

I did not observe any work on generalizations in Esra's statistics instruction. She did not generalize the procedures or statistical concepts for other contexts. For example, the teacher and students worked on contextualized examples related to choosing appropriate measures of central tendency for a given context; they calculated all measures. It was important to examine examples of when not using the arithmetic mean. However, most examples did not have an extreme value that might affect the arithmetic mean and representativeness of the data set. Only one or two examples consisted of extreme values, but the work did not consist of a mathematical and statistical generalization; the conclusion remained within that context; it was not based on at least two examples.

Furthermore, the instruction did not include the work on patterns; Esra did not help students read beyond the data by examining the patterns in the data through tables and graphs. In addition, I did not observe the formation of a statistical definition by examining different examples. Esra asked the students about the definitions of statistical concepts. The students tried to explain what they understood from the word (e.g., mode, median, mean, statistics). She provided the definitions when the teacher realized the students lacked knowledge about the concepts. Therefore, the definitions were not built individually or collectively; they were directly given.

4.1.2.2.2.6. Mathematical and Statistical Language

Figure 4.46 shows the distribution of scores assigned to lesson segments for the Mathematical and Statistical Language code in Esra's statistics instruction. This section aims to uncover Esra's mathematical and statistical language and her encouragement for fluent and correct mathematical and statistical language. Only one

segment did not contain any mathematical and statistical work. Therefore, the segment did not include mathematical and statistical terms because the teacher just did a homework check. Ten (17.54%) lesson segments comprised a few mathematical and statistical words. In 43 (75.44%) lesson segments, Esra used mathematical and statistical language fluently and explicitly with middling density. In addition, Esra supported students' use of correct mathematical and statistical language or dense mathematical and statistical language only in three (5.26%) lesson segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the Mathematical and Statistical Language code was 2.84 out of 4, with a standard deviation of 0.53. This section provides the descriptions and classroom examples of each score assigned to the Mathematical and Statistical Language code in Esra's statistics instruction.

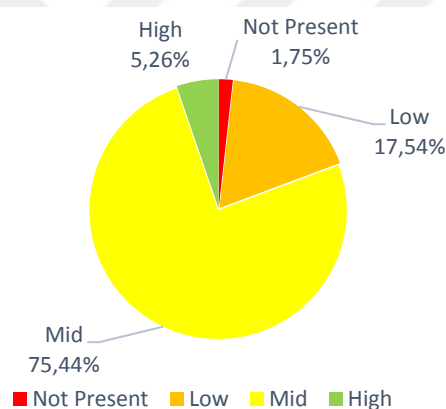


Figure 4.46 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Language” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I scored Low to ten (17.54%) lesson segments for mathematical and statistical language because the density of mathematical and statistical language was low. Esra solved most tasks from the textbook on the blackboard. Especially the segments with the tasks followed by the introductory tasks; she used similar mathematical and statistical terms throughout the segment. For example, a segment with a line graph task consisted of the increase, the decrease, the line, and the difference repeatedly. In addition, a Low segment with average concepts included add, divide, arithmetic mean, median, mode, frequent, and middle repeatedly.

I gave a Mid score to 43 (75.44%) lesson segments for the mathematical and statistical language. In those segments, the teacher used the language at a middling density. In the segments covered circle graph, she accurately used the terminology related to circles, percentages, quantities, proportions, and ratios, including angle, the sector, central angle, and cross-product method. In addition, she also used average-related terminologies. The segments with average concepts consisted of arithmetic mean, average, median, mode, middle point, sorting, and the difference among values, categories, numbers, tables, and frequencies. The segments did not rise to High scores for mathematical and statistical language because the teacher did not press students to use mathematical and statistical language accurately. Mostly the segments consisted of teacher-dominant talk.

I scored High in three (5.26%) lesson segments for the mathematical and statistical language. Those segments included mathematical and statistical language at middling density, encouraging students to use correct language. When the teacher covered a circle graph in a segment, she used circle and circle graph-related terms. In addition, she used terms related to ratios and proportions correctly. When the students read the graph, she corrected the students' incorrect use of language. For example, a student stated, "Green lovers were the fourth one of the circle sector." The teacher stated, "One-fourth of the whole class because the sector of the circle was one-fourth of [whole circle]. The student used "circle sector" in the wrong place, and the teacher reminded students to use correct terms for describing the circle. Later another student reached the same conclusion about the number of students with the favorite color of green by the correct use of language after the teacher corrected the previous student. In another segment, the teacher used technical language accurately and fluently. In addition, when a student asked, "The question says that to find the thing of the onion, should not we find about potatoes ..." and the teacher said that "In the fourth question? What does it ask about an onion? The central angle. It asked for the kilogram equivalent of the potato production". When the student stated what the question asked for incompletely, the teacher corrected the sloppiness in the student's technical language. The teacher emphasized the proper use of mathematical and statistical terminology to reach the correct conclusion in the following.

4.1.2.2.2.7. Overall Richness of the Mathematics and Statistics

Esra's instruction was analyzed to determine the overall richness of the mathematics and statistics code used in her lesson segments. Figure 4.47 presents the distribution of scores in this dimension. The results show that out of 57 lesson segments, six (10.53%) had incorrect or missing in terms of the richness of mathematics and statistics. In most of the lesson segments ($n=31$, 54.39%), a low level of mathematics and statistics was presented to the students.

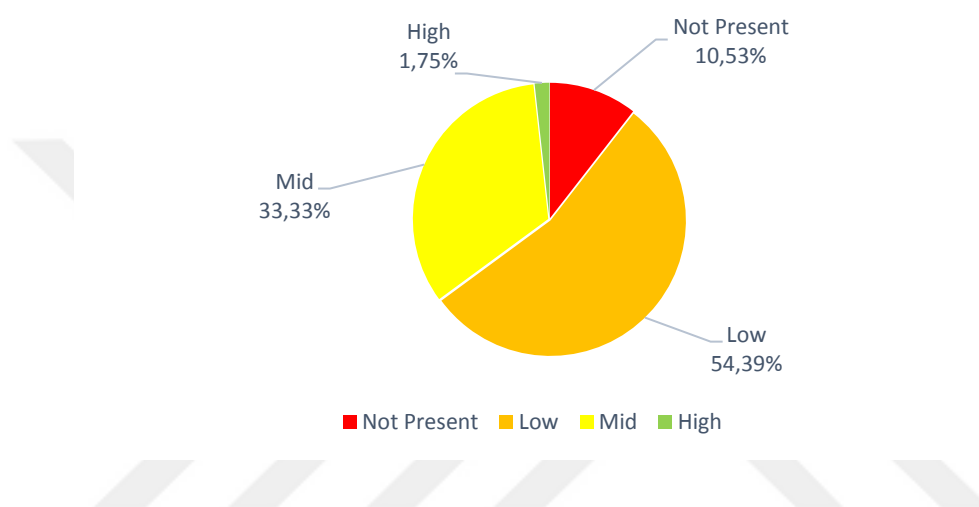


Figure 4.47 The Distribution of Scores Assigned to Lesson Segments for the “Overall Richness of Mathematics and Statistics” Code in Esra’s Instruction

Note. $N=57$ (the total number of lesson segments)

The codes in richness were present at a middling density at 19 (33.33%) lesson segments. There were Mid scores for mathematical and statistical sense-making, explanations, multiple procedures, and solution methods at some segments. In addition, there were some isolated High scores for mathematical and statistical sense-making, but it did not raise the segment to a High score of overall richness. Furthermore, only one segment (1.75%) consisted of high-quality richness codes. The segment was about choosing an appropriate measure of central tendency. It consisted of rich practices regarding explanations and mathematical and statistical sense-making related to the average concept. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on the overall richness of the mathematics and statistics code was 2.26 out of 4, with a standard deviation of 0.67.

4.1.2.2.2.8. Overview of “The Richness of Mathematics and Statistics” Dimension

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the frequencies of occurrences and co-occurrences of The Richness of Mathematics and Statistics codes with their scores as visualized on the MAXQDA map (see Figure 4.48)

As stated earlier, MAXQDA utilizes a similarity matrix that is transformed into a distance matrix to determine the spacing between codes on the map. According to the MAXQDA Manual, the proximity between codes is established by their occurrence together within the data. When two codes always appear together, their distance is zero, and when they frequently appear together, they are situated closer to each other on the map. On the other hand, when two codes never appear together, they are situated farthest apart on the map. This analysis is founded on the frequency of co-occurrences between codes within a particular segment. In addition, the size of code symbols on the map is determined by the frequency of their occurrence within a given segment. Codes with larger symbols appear more frequently are represented. Also, codes with smaller symbols appear less often.

The MAXQDA map in Figure 4.48 shows that the richness of mathematics and statistics in Esra’s statistics instruction could be characterized by low quality to mid quality. The frequencies of “Not Present” scores assigned to MPSM, Linking Representations, Explanations, Patterns and Generalization, were high and co-occurred. In addition, Mid scores on Mathematical and Statistical Sense-Making and Mathematical and Statistical Language co-occurred with Not Present scores in these codes. Those relationships led to low scores for the overall richness of mathematics and statistics, which were more frequent than other scores at the same code.

Esra instruction did not include linking between representations, multiple procedures or solution methods, and patterns and generalizations in the majority of the segments. First, she linked representations in a few segments. The nature of statistics instruction consisted of multiple representations from the same family of representations. Therefore, even if the different graphs were introduced, they did not link with other representations, such as distribution and frequency tables. In addition, the teacher

might have linked different representations for summarizing data; she would introduce average concepts with different types of graphs and different interpretations of average (e.g., leveling approach, the mathematical balance point approach).

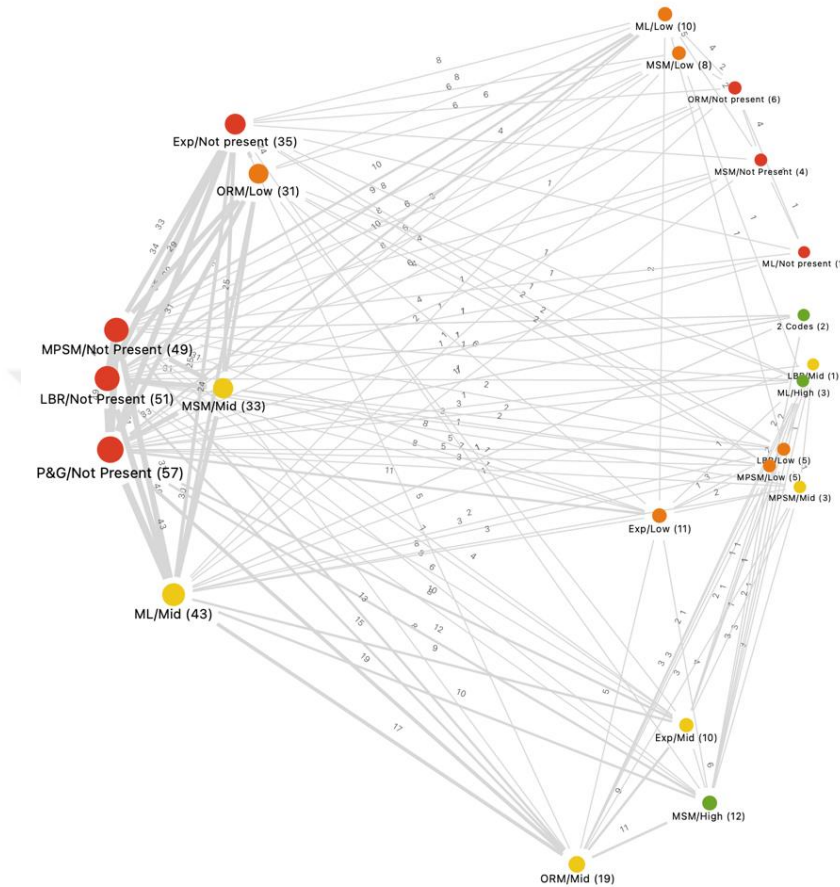


Figure 4.48 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “The Richness of Mathematics and Statistics” Dimension in Esra’s Instruction

Note: LBR- Linking between Representations, Exp- Explanations, MSM- Mathematical and Statistical Sense Making, MPSM- Multiple Procedures and Solution Methods, ML- Mathematical and Statistical Language, P&G- Patterns and Generalizations, ORM- Overall Richness of Mathematics and Statistics

The second issue was about MPSM, where the teacher provided multiple procedures and solution methods at segments with calculations in a circle graph. The circle graph tasks consisted of calculations considering ratio and proportion. However, the students were expected to calculate the arithmetic mean, median, and mode. The teacher did not introduce another way to find an average of a data set. In the statistical content

interview, she also stated that only one way was introduced to calculate the arithmetic mean.

In addition, none of the segments included Patterns and Generalizations, as explained in that section. The teacher accurately used mathematical and statistical language at middling density; the teacher focused on making sense of mathematics and statistics more than briefly. In addition, she explained provided explanations based on why in less frequent segments.

The results regarding the overall richness of the mathematics and statistics code show that most segments could also be characterized as Mid. As shown in Figure 4.48, the co-occurrence of Mid quality codes or the presence of High-quality isolated codes might lead to the characterization of the richness of Esra's instruction as mid-quality to some extent. In addition, Low and Mid scores at Explanation, MPSM, and Linking between representations co-occurred with Mid scores at the overall richness code. The results also show that Mid and High scores at richness codes were less frequent in Esra's statistics instruction. However, the co-occurrences of Mid and High scores at different codes led to a high quality of overall richness. Especially High quality at Mathematical and Statistical Sense-making and Explanation occurred together with High-quality overall richness. As a result, the richness map showed that richness codes interacted at similar levels.

4.1.2.2.3. Working with Students and Mathematics and Statistics

This section describes Esra's work with students on the mathematical and statistical content. For this reason, three codes in the dimension were scored; Remediation of Student Errors and Difficulties (RSE&D), Teacher Uses Student Mathematical and Statistical Contributions (TUSMC), and Overall Working with Students and Mathematics and statistics (OWSM). Table 4.12 indicates that Esra's statistics instruction could be characterized by low quality in understanding and responding to students' mathematical and statistical contributions.

According to Table 4.12, Esra did not remediate student errors and difficulties in 32 (56.14%) lesson segments; she did not use student mathematical and statistical contributions in eight (14.04%) segments. Overall Working with Students and

Mathematics and Statistics was scored as Not Present for five (8.77%) lesson segments. Most lesson segments for TUSMC and OWSM were scored Low (respectively, n=37, 64.91%; n=45, 78.95%). In addition, 21 (36.84%) lesson segments were scored Low for RSE&D.

Table 4.12

“Working with Students and Mathematics and Statistics” Dimension Scores in Esra’s Instruction

	1		2		3		4		M	SD
	Not Present		Low		Mid		High			
	n	%	n	%	n	%	n	%		
Remediation of Student Errors and Difficulties	32	56.14	21	36.84	4	7.02	0	0.00	1.51	0.63
Teacher Uses Student Mathematical and Statistical Contributions	8	14.04	37	64.91	11	19.30	1	1.75	2.09	0.63
Overall Working with Students and Mathematics and Statistics	5	8.77	45	78.95	6	10.53	1	1.75	2.05	0.51

Note. N=57 (the total number of lesson segments)

Furthermore, Esra remediated student errors and difficulties at mid-level (n=4, 7.02%). No lesson segments included high-level remediation. In 11 (19.30%) lesson segments, the teacher used student mathematical and statistical contributions at mid-level. Six (10.53%) lesson segments included OWSM at Mid-level. Only one (1.75%) lesson segment consisted of high-quality practices regarding TUSCM and OWSM.

As shown in Figure 4.49, the teacher uses of mathematical and statistical contributions of students and overall working with students and content codes were primarily present in Esra’s statistics instruction. However, most lesson segments did not include remediation. The segments with remediation were of low quality in general. Furthermore, a few lesson segments consisted of high-quality practices at the codes in this dimension. The details of the Working with Students and Mathematics and Statistics codes will be presented in the following sections.

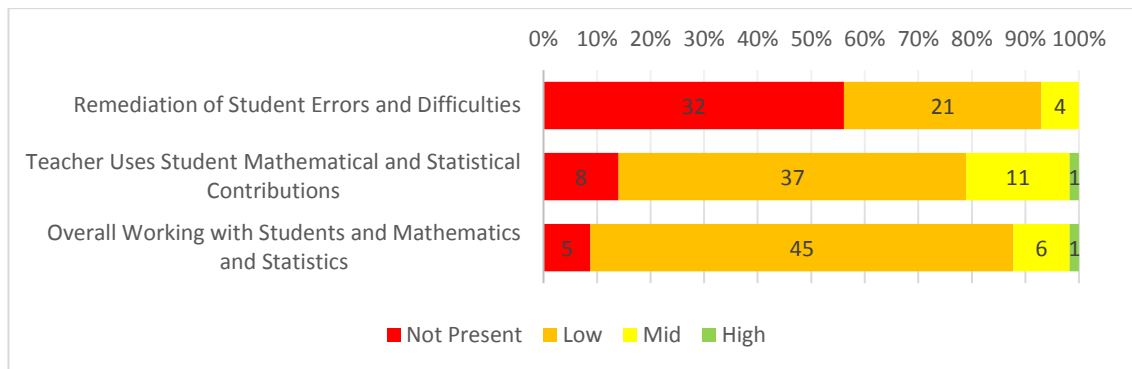


Figure 4.49 “Working with Students and Mathematics and Statistics” Dimension in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

4.1.2.2.3.1. Remediation of Student Errors and Difficulties

Figure 4.50 shows the distribution of scores assigned to lesson segments for Esra's statistics instruction in the Remediation of Student Errors and Difficulties code. Esra did not remediate student errors or imprecisions at 32 (56.14%) lesson segments. Esra did conceptual remediation briefly or procedural remediation briefly or moderately at 21 (36.84%) lesson segments. In four (7.02%) lesson segments, Esra considered and acted upon student errors and difficulties conceptually at a moderate level or procedurally in a substantial manner. None of the segments consisted of systematical remediation of student errors and difficulties. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.51 out of 4, with a standard deviation of 0.63. This section will provide the descriptions and classroom examples of each score assigned to the Remediation of Student Errors and Difficulties code in Esra’s statistics instruction.

I gave a score of Not Present to 32 (56.14%) lesson segments for remediation of student errors and difficulties. The teacher talk was dominant in specific segments, or students worked on tasks individually. Therefore, student misunderstandings or difficulties were not observed. As mentioned earlier, the students did not know the definition of statistics term or concept since they did not cover the topic in the previous year. When the teacher heard this information from the students during the segment, she did not determine their plausible misconceptions. She wanted students to

experience a statistical problem-solving process briefly before starting the first objective in the curriculum (e.g., circle graph). Therefore, she directly introduced the content with an example task. Furthermore, some segments included student errors, but the teacher corrected them without remediating them.

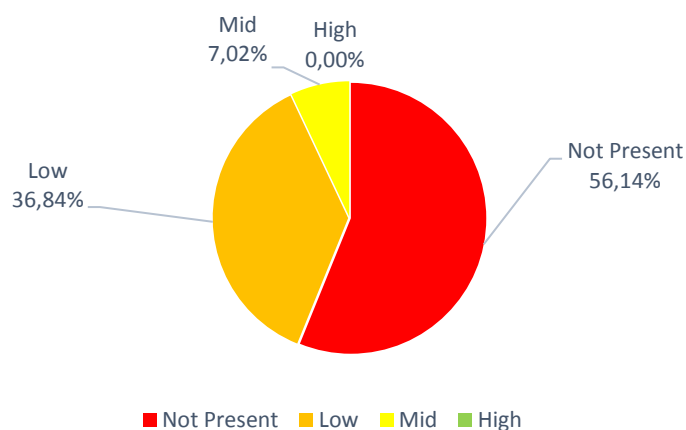


Figure 4.50 The Distribution of Scores Assigned to Lesson Segments for “Remediation of Student Errors and Difficulties” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I scored Low on 21 (36.84%) lesson segments to remedy student errors and difficulties. The segments included occasions with student mistakes or student difficulties while conducting procedures. When the students had difficulties with ratios and proportions for calculating angles or quantities corresponding to the categories on a circle graph, the teacher helped them to solve the equations by simplifying the numbers. The teacher procedurally remediated students’ difficulties briefly. Furthermore, a student added the values from a data set to find the arithmetic mean. However, the student made a mistake during the addition operation and struggled with the division. The teacher demonstrated the addition procedure by pointing out the regrouping; then, she helped students divide 116 by four step by step. This segment included moderate procedural remediation.

Furthermore, when reading a line graph, the teacher asked the students, “Which day did the price of gold increase?”. A student gave an incorrect answer. The teacher made brief conceptual remediation by showing the student the change in the day. The teacher

said, “Look, there is a decrease here [by showing the decrease].” She showed how the student’s answer was incorrect, then asked the student again which day it would increase. While the student gave the correct answer, the teacher showed the increase to the whole class on the line graph.

I scored Mid to four (7.02%) lesson segments for RSE&D. The segments included either extensive procedural or moderate conceptual remediation. A student had difficulty with calculations in a segment while finding the arithmetic mean of a data set shown in a frequency table. The teacher did not answer directly to the student. She helped the student on the board by adding all values and dividing by the number of values. In another average example, the students were expected to find the typical value of the data set with the context of the number of shoes (See Average Example: The Number of Shoes in Figure 4.27). The data was shown in a frequency table. This segment consisted of pre-remediation by reading the table to the students explicitly. In addition, when a student came to the board to find the central tendency measures, the student made an error by adding the frequencies to find the total values for the arithmetic mean. Another student reported a similar misunderstanding. The teacher conceptually remediated the students' misunderstanding by demonstrating the data in open form to tell what the frequency table means.

T: It is not like that, but you will not collect 41, 96, are you? You will visualize it.

Student: Yes, teacher, 96 and 23 equal 119, 41 more equals 160, and 172 from there.

T: But now, look, do you know what that means? 1 1 1 1, how many of them are there 41 there are, for example, those with two shoes, they said 2 2 2 2, how many people said two, there are 96 said 2. I go as 3 3 3 3 3, but how many times did it repeat? It repeated 23 times... So you will calculate accordingly.

[Esra 17.04.2018-7D 07:30-15:00]

The teacher addressed the source of student errors. Students added the numbers they encountered without considering the context and the data. Therefore, she described the data on the table; what the values and the frequencies meant. All in all, the segment involved brief pre-remediation and moderate conceptual remediation. If the conceptual

remediation was at length and consisted of discussing possible errors stemming from the misunderstanding, the segment could be scored High.

4.1.2.2.3.2. Teacher Uses Student Mathematical and Statistical Contributions

Figure 4.51 displays the distribution scores assigned to lesson segments for Esra's use of student mathematical and statistical contributions. The teacher did not utilize student ideas to enhance the teaching of mathematics and statistics in 14.04% of the eight lesson segments. In 37 (64.91%) of the lesson segments, the teacher employed pro-forma use of student contributions without fully leveraging them to develop the teaching of mathematics and statistics. However, she used student contributions to develop mathematics and statistics towards high-quality instruction in 11 (19.30%) lesson segments. Only one (1.75%) lesson segment consisted of responding to student ideas by giving credit to the student and moving the instruction forward based on the idea.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.09 out of 4, with a standard deviation of 0.63. This section will provide the descriptions and classroom examples of each score assigned to the Teacher Uses Student Mathematical and Statistical Contributions code in Esra's statistics instruction.

I scored Not Present for eight (14.04%) lesson segments to use student mathematical and statistical contributions to develop mathematics and statistics. Those segments were dominated by teacher talk. Some of them did not include student contributions. In other segments, there were a few student contributions. The students gave answers to questions while explaining the solution to a task. However, the teacher stopped the students, said, "Stop and wait," and described the solution. Therefore, the teacher ignored student contributions. In other segments, the students worked on a worksheet individually. The students struggled with a task in the worksheet; the teacher helped them solve it. However, the teacher solved the question without or with minimal student contributions. Those segments were also dominated by teacher talk.

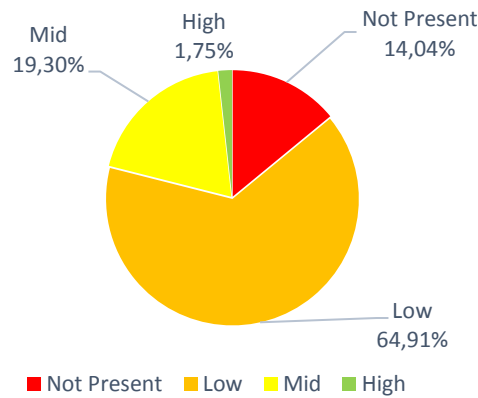


Figure 4.51 The Distribution of Scores Assigned to Lesson Segments for the “Teacher Uses Student Mathematical and Statistical Contributions” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

In low-scored lesson segments ($n=37$, 64.91%), student mathematical and statistical contributions were present. However, the students gave one or two short answers to questions, stated their mathematical and statistical ideas, and solved tasks on board. The teacher responded to the contributions in a way that she was the one who shaped the instruction. For example, while solving a circle graph task, the student said, “We need to use proportion.” The teacher described how they could use proportion, began to solve the task, and asked students to answer bound questions during the solution process, like “What is the percent of potatoes?” When she got the answer to her question, she continued the solution. In another example, the teacher asked students to define mode while solving a task in the context of shoe size. A student stated mode as “repeating number,” and the teacher corrected it as “the most frequent number in a data set.” She emphasized that the students needed to understand the table first. However, she was the one who read the table and solved the task with a few student utterances. She used student contribution in a proforma way; she was the one who developed the mathematics and statistics in the segment.

In Mid-scored lesson segments ($n=11$, 19.30%), students contributed to the lesson in the segments, and the teacher responded to students' mathematical and statistical contributions to shape the instruction. For instance, the teacher expected students to

contribute mathematics and statistics while reading a graph task. Then she moved the instruction based on student analysis and interpretations of the graph. For example, in the first lesson, the teacher collected data from the students and expected students to read the data from the table. In addition, students read the circle graph they constructed on the board. The student and teacher interacted over the content. The teacher moved the instruction forward with students' interpretations based on the circle graph. Furthermore, the students were given data in multiple segments. They were expected to calculate both the quantities and angles for each category of the data. In some parts, the teacher's talk was dominant, but the teacher asked students how they solved the question, and she followed a student's solution path. The common characteristics of these segments are that the teacher elicited students' ideas. Then she responded to those ideas to shape the instruction to some extent based on the student contributions.

One (1.75%) lesson segment received a High score because Esra used student contributions and moved the instruction forward based on those contributions. In the segment, the teacher constructed a circle graph with the context of the ingredients of strawberry jam. In the first step, she asked students how to set the proportion. A student gave an incorrect answer; the teacher described setting the proportion with certain student utterances. She used student contributions to some extent at the beginning of the segment. However, in the later parts, she asked students if there was another way to find the other ingredients. The students' answers were at length, and the teacher commented on the students' mathematical and statistical ideas. For example, she said, "Emircan says if eight kg of strawberries is 240 degrees, he says, I will focus on 1 kg because eight divided by eight is one, 240 divided by eight is 30 degrees." Then another student used the key ideas for finding the last ingredients. She acknowledged the students' statements and focused on the answers' key ideas (e.g., focusing on the unit).

4.1.2.2.3.3. Overall Working with Students and Mathematics and Statistics

Figure 4.52 presents the distribution of the scores assigned to lesson segments for the Overall Working with Students and Mathematics and Statistics code in Esra's statistics instruction. What stands out in this figure is that the instruction can be characterized

as Low quality in terms of OWSM (n=45, 78.95%). The lesson segments scored Low consisted of using student contribution in a proforma way or brief remediation. There were no student contributions in five (8.77%) lesson segments or the use of student ideas when it was present. In addition, there were no student errors and difficulties or no remediation even if there were errors and difficulties.

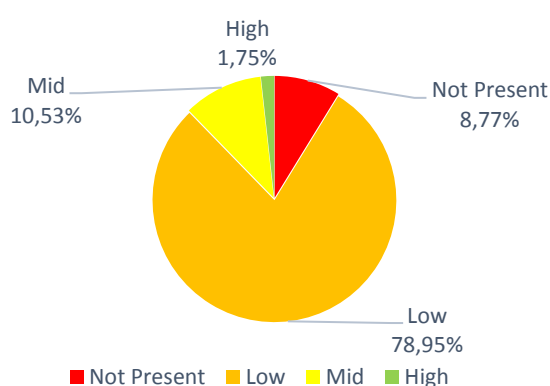


Figure 4.52 The Distribution of Scores Assigned to Lesson Segments for the “Overall Working with Students and Mathematics and Statistics” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

The overall evaluation of the dimension indicated that Esra worked with students over content at Mid-level in 6 (10.53%) lesson segments. Esra used student ideas to move instruction forward and moderately remediated student errors and difficulties. However, only one (1.75%) lesson segment was shaped by responding to student ideas and remediating errors and difficulties of the students during student-teacher interaction while constructing a circle graph. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.05 out of 4, with a standard deviation of 0.51.

4.1.2.2.3.4. Overview of “Working with Students and Mathematics and Statistics” Dimension

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the

frequencies of occurrences and co-occurrences of Working with Students and Mathematics and Statistics codes with their scores as visualized on the MAXQDA map (see Figure 4.53)

The MAXQDA map shows that Esra's work with students over mathematics and statistics instruction could be characterized by low quality. The frequencies of Low scores assigned to Remediation of Student Errors and Difficulties, Teacher Uses Student Mathematical and Statistical Contributions, and Overall Working with Students and Mathematics and Statistics codes were the highest in the dimension. In addition, the Low scores at all codes were intersected in the same segment in the highest frequency. In many segments, I observed that the teacher's talk was dominant. There were few student utterances. The teacher procedurally responded to student errors and difficulties briefly. In addition, the teacher used student contributions to move the instruction forward in a proforma way. She utilized student contribution, but the teacher developed the mathematics and statistics of the segment.

Furthermore, Not Present scores assigned to Remediation of Student Errors and Difficulties also co-occurred with high frequency with Low scores at remaining codes in the dimension. Those segments did not include teacher practices regarding conceptual or procedural remediation of student misunderstandings but consisted of student contributions. Thus, the teacher either gave correct answers directly to the students or observed any difficulty during these segments.

Figure 4.53 also indicates that the Mid scores assigned to all codes intersected in the segments less frequently than the Low scores. Moderate-level quality practices in working with students over content occurred together. In the segments, the teacher asked students' ideas about the content. The student gave answers to the questions, and they commented on a topic. The teacher used such contributions to shape the instruction to some extent. The segments consisted of teacher use of student contributions to move the instruction forward. Moreover, during student-teacher interaction over content, the students had difficulties calculating the circle graph and measures of central tendency. The teacher made extensive procedural remediation. In

specific segments, moderate conceptual remediation was present (e.g., selecting appropriate measures of central tendency).

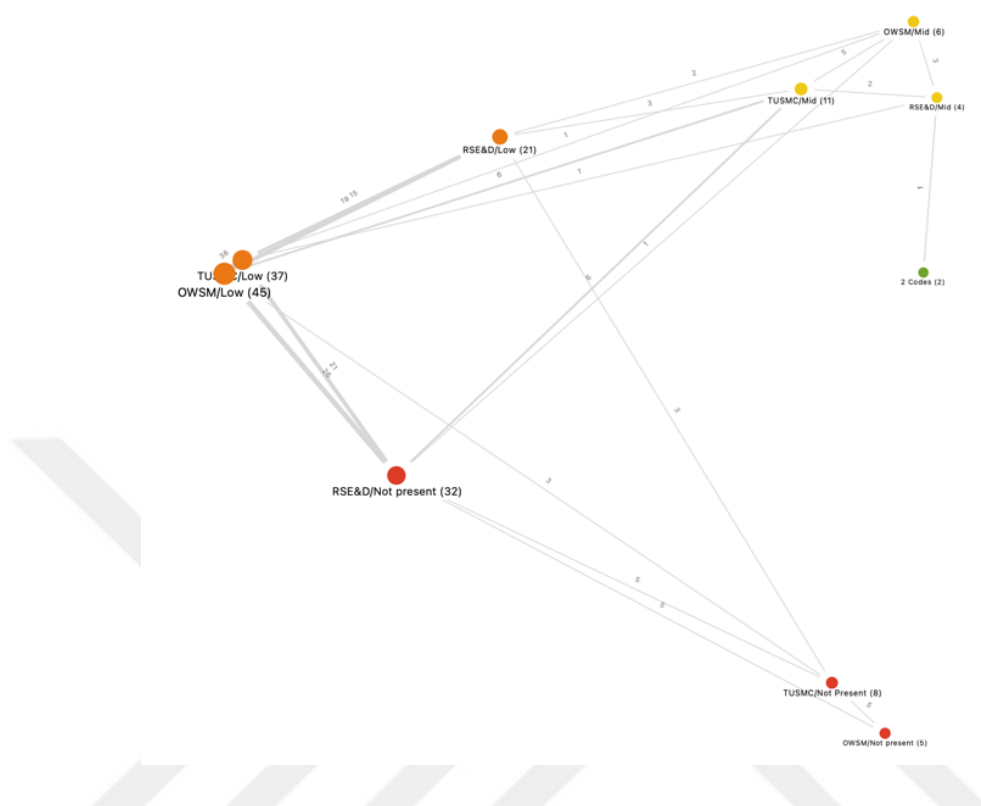


Figure 4.53 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Working with Students and Mathematics and Statistics” Dimension in Esra’s Instruction

Note: RSE&D- Remediation of Student Errors and Difficulties, TUSMC- Teacher Uses Student Mathematical and Statistical Contributions, OWM- Overall Working with Students and Mathematics and Statistics

In addition, some segments were assigned different scores for the codes. For example, one incident with high-quality practices regarding TUSMC interacted with remediation at the Mid-level, resulting in high-quality overall working with students and mathematics and statistics. Esra’s statistics instruction did not include systematic conceptual remediation. However, there were a few interactions at different levels except Not Present for RSE&D. As a result, the map in Figure 4.53 shows that working with students and mathematics and statistics codes interacted at similar levels.

4.1.1.2.4. Errors and Imprecision

In this section, I aim to explore teacher errors and imprecision in terms of content and language; and the lack of clarity in Esra's statistics instruction. Table 4.15 indicates the results regarding four codes in the dimension: Mathematical and Statistical Content Errors (MCE), Imprecision in Language or Notation (ILN), Lack of Clarity in Presentation of Mathematical and Statistical Content (LCPM), and Overall Errors and Imprecision (OE&I). The dimension has reverse scoring than other dimensions. I reversely assigned numerical points to each score when calculating the mean and standard deviation (4=Not Present, 3=Low, 2=Mid, 1=High). Therefore, it is necessary to emphasize that lower scores assigned to the codes mean fewer errors and imprecisions.

According to Table 4.13, Esra's statistics instruction demonstrated a low occurrence of mathematical and statistical content errors (n=48, 84.21%), imprecise language or notation (n=49, 85.96%), lack of clarity in presenting mathematical and statistical content (n=47, 82.46%), and overall errors and imprecision (n=38, 66.67%). Five (8.77%) lesson segments included content errors by distorting the mathematics and statistics of the segment. Content errors were present at one (1.75%) lesson segment for some parts of the segment and three (5.26%) lesson segments during the whole segment.

Errors in language were present at seven (12.28%) lesson segments, but the minor errors did not distort the mathematics and statistics of the lesson. Only one (1.75%) segment included High-level language error. Esra's statistics instruction had problematic aspects in terms of the clarity in the presentation of the content at seven (12.28%) lesson segments at the Low level, two (3.51%) lesson segments at Mid-Level, and one (1.75%) lesson segment at High Level. Lastly, overall errors and imprecision were present at 13 (22.81%) lesson segments with minor errors; the number of errors was higher for six (10.52%) lesson segments (n=3, 5.26% for Mid scores and n=3, 5.26% for High scores).

Table 4.13

“Errors and Imprecision” Dimension Scores in Esra’s Instruction

	4 Not Present		3 Low		2 Mid		1 High		M	SD
	n	%	n	%	n	%	n	%		
Mathematical and Statistical Content Errors	48	84.21	5	8.77	1	1.75	3	5.26	3.72	0.75
Imprecision in Language or Notation	49	85.96	7	12.28	0	0.00	1	1.75	3.82	0.50
Lack of Clarity in Presentation of Mathematical and Statistical Content	47	82.46	7	12.28	2	3.51	1	1.75	3.75	0.61
Overall Errors and Imprecision	38	66.67	13	22.81	3	5.26	3	5.26	3.51	0.83

Note. N=57 (the total number of lesson segments)

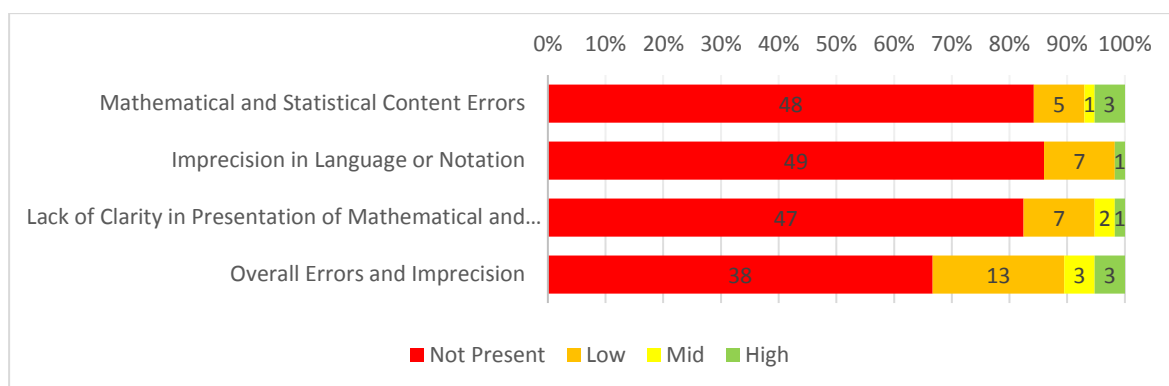


Figure 4.54 The “Errors and Imprecision” Dimension in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

Figure 4.54 shows that most lesson segments were error-free in content errors, language imprecision, and the presentation's lack of clarity. The errors obscuring the mathematics and statistics of the segment did not frequently occur. However, some minor errors were present; but they did not obscure the mathematics and statistics of the segment. The details of the Errors and Imprecision dimension codes will be presented in the following sections.

4.1.1.2.4.1. Mathematical and Statistical Content Errors

Figure 4.55 shows the distribution of scores assigned to lesson segments for the Mathematical and Statistical Content Errors code in Esra’s statistics instruction, describing the content errors that occurred and were not corrected during Esra’s statistics instruction. Figure 4.55 shows that 48 (84.21%) lesson segments did not include any mathematical and statistical content error. Five (8.8%) lesson segments included brief content errors. One (1.75%) lesson segment consisted of content errors to some extent. In three (5.26%) lesson segments, the content errors obscured the mathematics and statistics of the lesson. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.72 out of 4, with a standard deviation of 0.75. This section provides the descriptions and classroom examples of each score assigned to the Mathematical and Statistical Content Errors code in Esra’s statistics instruction.

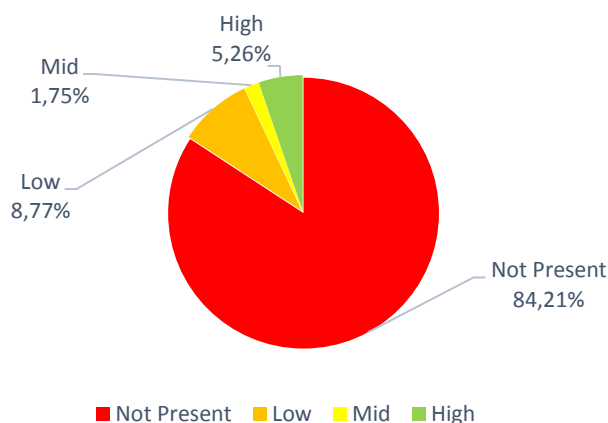


Figure 4.55 The Distribution of Scores Assigned to Lesson Segments for “Mathematical and Statistical Content Errors” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I scored Low for five (8.77%) lesson segments because Esra made brief content errors. Esra defined the mode as the most frequent “number,” not “value.” Therefore, she forgot the critical condition in the definition. The context given in the tasks was appropriate for the most frequent number. Therefore, the definitional error did not obscure the mathematics and statistics of the lesson. However, she provided the general definition, so it was necessary to define it as a value and emphasize that the number was also a value. Furthermore, a student made a calculation error on the board in another segment, and the teacher left it on it. The student calculated 35% of 720 kg as 232 instead of 252. The teacher said, “How many kg? 232 kg. Okay”. This incident happened at the end of the segment; all steps in the solution process were correct except for minor calculation errors.

One (1.75%) lesson segment was scored Mid for the Mathematical and Statistical Content Error code. The lesson segment included the teacher’s work on Average Example: The Number of Shoes (see Figure 4.27). They calculated the arithmetic mean of the given data set. Then they found the mean number of shoes in the classroom as 2.03 shoes per person. In the following parts of the segment, the statistical content error occurred. The teacher said to students that “If we rounded the [number], we could say this [arithmetic mean] as 2? The number is very close to 2. Could not we decide

this [the average] as two by using other things?”. The teacher added that the most frequent number was two and repeated 96 times in 172 data points.

The teacher emphasized that the arithmetic mean and the mode were very similar, and they could use the mode to find the data set's average. However, the emphasis on the closeness of the values is not statistically correct for interchangeably using both measures of central tendency. In this case, the teacher equated two non-identical statistical terms. This conclusion might lead to misconceptions regarding the purpose of using arithmetic mean and mode. Furthermore, when similar contexts were provided, the students would not prefer to use the arithmetic mean to describe the average because it would be easier to find the mode. This segment was not scored High because the content error obscured a part of the segment. The majority of the segment included calculations of the arithmetic mean of a given data set. In those parts, students made mathematical errors; but the teacher corrected the errors within the segment.

I assigned a score of High to three (5.26%) lesson segments for the Mathematical and Statistical Content Errors code. The teacher covered two segments of line graph tasks in a mock exam context. The first task was about a line graph about a student called Fatih's number of correct answers in the math section of a series of mock exams. It is necessary to note that the same task was covered in Ali's lesson and assigned a Mid score for mathematical and statistical content errors (see Figure 4.23- Graph Task 5).

Similarly, the teacher demanded that students construct a line graph with the given data in the table. There were five exam results. Although the teacher emphasized the change in time when introducing line graphs, we did not know exactly the time interval between mock exams. This situation might lead to misinterpretation of the graph; the conclusions about the increase or decrease in the number of correct answers would not be correct.

The second example was similar: the change in two students' number of correct answers in the consecutive mock exams with two lines in the graph. Similar content errors were present since the context was similar. Those segments consisted of a

critical error about forgetting a critical condition about the continuity of the time concept in the line graph.

The third segment was about the definition of the mode of a data set. In the segment, there were data sets without context. The students were expected to find the mode of the data sets on the board. As I stated earlier, there were two expressions for the mode in Turkish. One refers to the peak value [tepe değeri], and the second refers to the mode [mod]. Both expressions are used interchangeably. The expression “the peak value [tepe değeri]” was used in the task. After two students found the mode of two different data sets, Ayberk came to the board to find the mode of “4 4 4 4 4”. Then Ayberk was confused. The teacher discussed the mode of the data set:

T: What does peak value mean, the most repeated number? We call it the peak value, whichever is repeated more often than other numbers in a data set. Let us look at the numbers here—all 4. So? Is there any other number you can compare to other than 4? No. Then what do we do? What do we say in such situations? What will we say?

Multiple Students: 4

Multiple Students: No mode

T: Are we going to say that our peak value is 4? To talk about the peak value, for example, you just said that there are hills, the most repetitive peak, so I throw it like this. Let us consider a hill. There are even small hills. There is this and this.

[16.04.2018-2 Esra 07:30-15:00]



Figure 4.56 Esra's First Drawings for Describing the Mode

After her explanations about the mode, the students provided different answers. The teacher did not ask them why they gave such answers and continued to describe the mode for those whose answer was four as the mode of the given data set. She made some drawings to show the peaks (see Figure 4.56). The purpose of the drawings was not given in the beginning. She cleared the drawings and continued to discuss the mode with different examples and drawings:

T: Let us draw a nice picture, do not be misunderstood (deleting). Hmm. Look at [drawing]. The heights are all different. Which ones are equal, this and that? Now then, I showed it by analogy to the top. What is our peak? If that corresponds to 5, what is our peak? Five is okay? Look, there are other hills; how many peaks? There are 1 2 3 4 5 6 peaks, but two correspond to how many of these hills correspond to 5. Then can we say that our peak value is 5? We can say, why? Because there are other peaks, we can compare.

[16.04.2018-2 Esra 07:30-15:00]

Next, she drew a representation involving six numbers in the data set (Figure 4.57). The value of 5 was the largest number in the distribution and was repeated twice. She did not label other values, but they were smaller than 5 in magnitude. This example was not the same in the beginning. She aimed to show that if there were different values in the data set, they could mention the most frequent value in the data set.

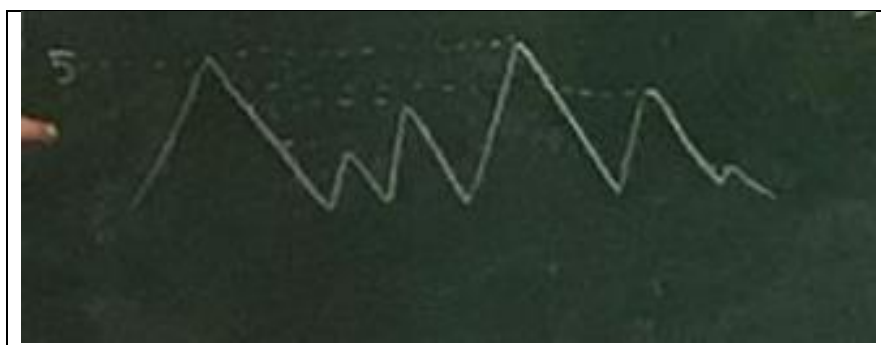


Figure 4.57 Esra's Second Drawing for Describing the Mode

Later she went back to the task of finding the mode of the “4 4 4 4 4” data set. She drew two hills to indicate that the peaks' lengths would be identical (see Figure 4.58). She concluded that there was no distribution mode because there were no other values to compare.

T: Now, here are all 4. If it is all 4 [let us do it], are there any other peaks you can compare? No, they are all the same anyway; you do not have a peak if you do not have other places to compare. Know what I mean? If the numbers in my dataset are all the same, we do not have a mode; that is, we do not have a peak value, and there is no other value to compare with, guys. So you will say there are no modes, no peaks.

[16.04.2018-2 Esra 07:30-15:00]



Figure 4.58 Esra's Second and Third Drawings for Describing the Mode

This segment included incorrect representations of the mode concepts. The teacher used the magnitude of a number while describing peaks. She was supposed to compare the frequencies of each value; the peak should correspond to the most frequent value. Although she gave somewhat correct definitions of mode (without the emphasis on “value” but “number”), she could not correctly represent what she defined. Therefore, there were definitional errors in the representations. The error was the focus of the entire segment. Hence the mathematics and statistics of the segment were obscured.

4.1.1.2.4.2. Imprecision in Language or Notation

Figure 4.59 indicates the distribution of scores assigned to lesson segments for the Imprecision in Language or Notation code, describing the problematic aspects in mathematical and statistical language or notation in Esra's statistics instruction. Out of the total lesson segments analyzed, there were no errors in language or notation in 49 (85.96%) lesson segments. Seven (12.28%) contained imprecision in language or notation. However, these imprecisions did not detract from the mathematical and statistical content of the segment. Only one (1.75%) segment had imprecisions that obscured the understanding of mathematics and statistics. The results of the descriptive

analysis indicate that the mean score assigned to the lesson segments on this code was 3.82 out of 4, with a standard deviation of 0.50. This section provides the descriptions and classroom examples of each score assigned to the Imprecision in Language or Notation code in Esra's statistics instruction.

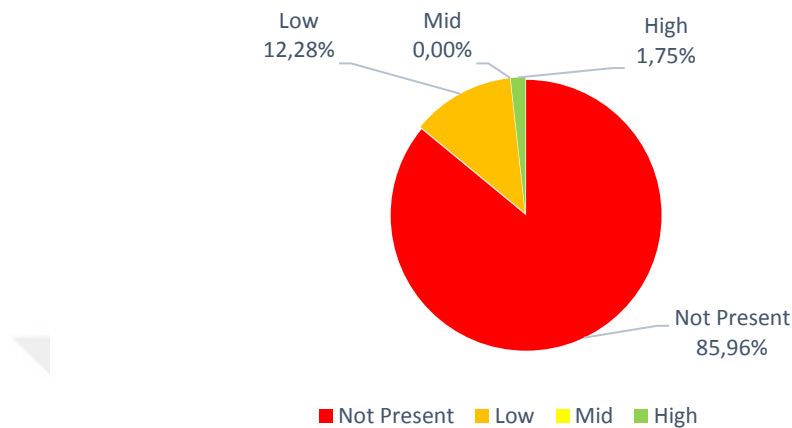


Figure 4.59 The Distribution of Scores Assigned to Lesson Segments for “Imprecision in Language or Notation” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

The lesson segments scored Low consisted of imprecise language while covering the measures of central tendency. The teacher described all measures of central tendency as average (i.e., ortalama). She described the purpose of all was to find the typical value in the distribution. Then she described the mean as arithmetic mean (i.e., aritmetik ortalama) with the algorithm. In seven (12.28%) lesson segments, the teacher used average (i.e., ortalama) instead of the arithmetic mean (i.e., aritmetik ortalama) briefly. The error might stem from the structure of the Turkish language. The Turkish correspondence of average was like “the mean,” and the teacher used a short version of the arithmetic mean as the mean. The imprecision in language did not obscure the mathematical and statistical point of the segment. The students understood what the teacher asked for even if she used “ortalama” as the shorter version of the arithmetic mean and although she defined “ortalama” as the measures of central tendency.

Only one (1.75%) lesson segment received a high score for the Imprecision in Language and Notation code, the only segment to score high for the Mathematical

Content Errors code. The teacher made an error with the general language. She used the concept of mode as “peak value” (i.e., tepe değeri) technically. It was described as the most frequent number (value) in the distribution. However, while describing the peak value, she drew peaks like the magnitude of the numbers (values) in the distribution. The meaning of peak value was different in everyday language and mathematical and statistical language. However, the teacher's descriptions and drawings did not convey correct mathematical and statistical concepts and were imprecision, obscuring the segment's mathematics and statistics.

4.1.1.2.4.3. Lack of Clarity in Presentation of Mathematical and Statistical Content

Figure 4.60 indicates the distribution of scores assigned to lesson segments for the Lack of Clarity in the Presentation of Mathematical and Statistical Content code. Esra's statistics instruction did not lack clarity in 47 (82.46%) lesson segments. In seven (12.28%) lesson segments, the lack of clarity was present, but the mathematics and statistics of the segment remained clear. Two (3.51%) lesson segments had unclear parts that affected the segments' mathematical and statistical aspects, but the teacher made them more transparent later. The mathematical and statistical point was not evident in one (1.75%) lesson segment. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.75 out of 4, with a standard deviation of 0.61. This section provides the descriptions and classroom examples of each score assigned to the Lack of Clarity in the Presentation of Mathematical and Statistical Content code in Esra's statistics instruction.

The segments with a Low score ($n=7$, 12.28%) consisted of a brief lack of clarity that did not distort the segment's mathematical and statistical points. While solving equations during calculations in circle graph tasks, the equality of both sides was not emphasized. Equations were not solved by adding/subtracting or multiplying/dividing into both sides of equations. Esra used a shorter version of solving equations without considering equality. She used inverse operations for leaving x alone. For example, in one case, the teacher said, “Let us throw 30 to the opposite side as over 30”. In the incidents, the teacher did not clearly describe solving equations. However, the main

mathematical and statistical point was not the equation solving but the components of a circle graph.

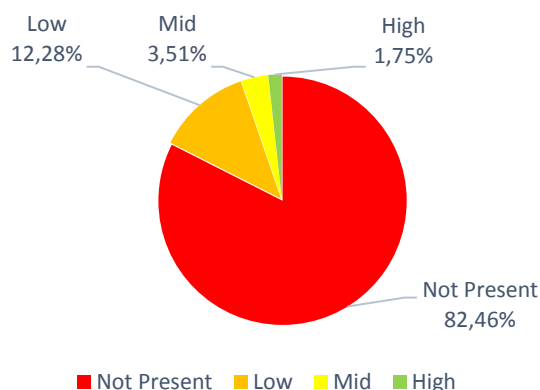


Figure 4.60 The Distribution of Scores Assigned to Lesson Segments for “Lack of Clarity in the Presentation of Mathematical and Statistical Content” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

Furthermore, while covering measures of central tendency in a task, the teacher briefly stated the purpose of the arithmetic means by looking in more “detail.” She did not provide other explanations. Thus, it was unclear what she meant by “detail.” However, I scored Low because it was a brief incident that did not affect mathematical and statistical points.

I assigned a Mid score to two (3.51%) lesson segments for the LCPM code due to some lack of clarity in certain parts of the segment, which only had a minor impact. The teacher asked the students what they understood from peak value by the word in a segment. She did not give a statistical definition of the peak value (mode). The students came up with ideas, but they considered possible Turkish meanings. They mentioned hills, the largest number, and the rounding. The teacher did not guide students' mathematical and statistical meaning of the word for a while. In this part, the lack of clarity was present in the presentation of the mode concept. The students seemed confused. The mathematical and statistical point in the part was obscured. In later parts, she defined the mode and made it more straightforward for students. Furthermore, the teacher explained why the arithmetic mean was not used for the given task in the second segment. However, she gave some explanations without clear

reasons. She stated that they should talk about distance and proximity. Then she mentioned that they should not use the arithmetic mean if there were [big] differences among numbers, the inequality in distribution, and very odd numbers, when the numbers were inconsistent and increased and decreased suddenly. What the teacher meant by inequality in distribution, odd numbers, and inconsistent numbers were unclear. Therefore, the mathematical and statistical point was distorted. However, I scored Mid because the explanation was only a part of the segment.

The only segment that scored high for LCPM was the one that received high scores for MCE and IL&N, with incorrect and obscured mathematical and statistical points throughout. The general language errors also contributed to a lack of clarity in all segments. The teacher's initial definition of mode was inconsistent with her representations and descriptions while drawing the figures, thus affecting the mathematical and statistical point of the segment.

4.1.1.2.4.4. Overall Errors and Imprecision

Figure 4.61 displays the distribution of the scores assigned to lesson segments for the Overall Errors and Imprecision in Esra's instruction. 38 (66.67%) lesson segments did not include errors and imprecision. Minor errors were present in 13 (22.81%) lesson segments, but those errors did not obscure the mathematics and statistics of the segments.

Three (5.26%) lesson segments had errors that distorted the mathematical and statistical points. Two segments with a Mid score in overall errors and imprecision consisted of the lack of clarity at the Mid-level. The third segment consisted of content errors at Mid-level. Therefore, more than one error was present that affected some parts of the segments. Moreover, three (5.26%) had significant content errors while constructing a line graph and describing the mode. Most lesson segments in Esra's instruction were error-free regarding content, language, and clarity. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 3.51 out of 4, with a standard deviation of 0.83.

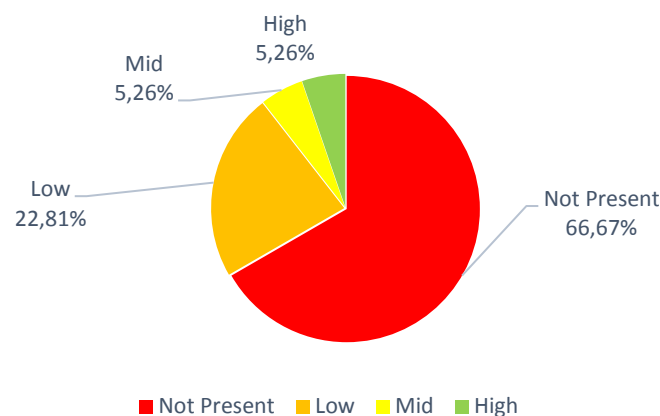


Figure 4.61 The Distribution of Scores Assigned to Lesson Segments for the “Overall Errors and Imprecision” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

4.1.1.2.4.5. Overview of “Errors and Imprecision” Dimension

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the frequencies of occurrences and co-occurrences of Errors and Imprecision codes with their scores as visualized on the MAXQDA map (see Figure 4.62). As MAXQDA Map shows, Esra’s teaching practices in statistics instruction were primarily error-free regarding content, language, notation, and imprecisions. The results show that Not Present scores for all dimensions' codes were more frequently present. Furthermore, the Not Present scores at all codes co-occurred in the same segment with the highest frequency. In addition, the overall code with a Not Present score only overlapped with other codes with Not Present scores. The findings confirmed that most lesson segments were error-free in Esra’s statistics instruction. Moreover, the overall errors and imprecision code was not present when other codes were absent in the segments.

Similarly, the codes with Mid scores were also less frequent and clustered. The errors in those segments were more than brief and distorted the mathematical and statistical points of a part of the segment. There were also incidents in the codes with Mid scores overlapped with ILN code with a low score and the codes with Not Present scores. Mainly middling density content error did not occur with the lack of precision in

content presentation or vice versa. Figure 4.62 also indicates that the codes with High scores co-occurred less frequently. Their distance also showed that they were present in similar segments. Those segments included errors affecting the instruction negatively throughout the segment. All in all, the map indicates that errors and imprecision codes co-occurred with each other at similar levels in most segments. In addition, it also shows that most lesson segments were error-free in Esra's instruction. The content errors at any level were less frequent and occurred in similar lesson segments.



Figure 4.62 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Errors and Imprecision” Dimension in Esra’s Instruction

Note: MCE- Mathematical and Statistical Content Errors, ILN- Imprecision in Language or Notion, LCPM- Lack of Clarity in Presentation of Mathematical and Statistical Content, OE&I-Overall Errors and Imprecision

Furthermore, Low, Mid, and High scores on each Errors and Imprecision dimension code were less frequent. Therefore, the errors in Esra’s instruction were less frequent.

As shown in Figure 4.62, the codes with low scores were gathered around the map. The errors present in those segments were brief and did not obscure the mathematics and statistics of the segments. However, the codes with Low scores also interacted with those with Not Present scores. However, the distance between the codes' scores was not close; their co-occurrence was not likely to occur as the scores gathered around. Therefore, for instance, the incidents with no errors in content did not overlap with brief errors in imprecision in language.

4.1.1.2.5. Student Participation in Meaning-Making and Reasoning

I explored how students interacted with the mathematical and statistical content in Esra's statistics instruction. Therefore, I examined this dimension with six codes in the MQI. Table 4.14 indicates the results regarding the Student Participation in Meaning Making and Reasoning dimension in the MQI. Student participation in meaning-making and reasoning in Esra's statistics instruction can generally be characterized as absent to low. On the other hand, Students Communicate about the Mathematics and Statistics of the Segment (SCM) code, and Students Work with Contextualized Problems (SWCP) were present as mid-quality above the quarter of all segments.

Table 4.14 presents the results for student participation in meaning-making and reasoning based on four different codes: Students Provide Explanation (SPE), Student Mathematical and Statistical Questioning and Reasoning (SMQR), Task Cognitive Demand (TCD), and overall code. The analysis revealed no high-quality lesson segments in terms of SMQR, SCM, TCD, SWCP, and overall code. The codes with relatively higher occurrence than others were SCM, with 31 segments (54.39%) scoring mid and 15 segments (26.32%) scoring low, and SWCP, with 16 segments (28.07%) scoring mid and 30 segments (52.63%) scoring low.

However, a vast majority of the segments were scored as absent or not present in terms of the Students Provide Explanation (SPE) code (n=44, 77.19%), Student Mathematical and Statistical Questioning and Reasoning (SMQR) code (n=50, 87.72%), and Task Cognitive Demand (TCD) code (n=42, 73.68%), as well as the overall code (n=39, 68.42%).

Table 4.14

“Student Participation in Meaning-Making and Reasoning” Dimension Scores in Esra’s Instruction

	1						2						3						4					
	Not Present			Low			Mid			High														
	n	%		n	%		n	%		n	%		n	%		n	%		n	%		M	SD	
Students Provide Explanations	44	77.19		10	17.54		1	1.75		2	3.51											1.32	0.69	
Student Mathematical and Statistical Questioning and Reasoning (SMQR)	50	87.72		7	12.28		0	0.00		0	0.00											1.12	0.33	
Students Communicate about the Mathematics and Statistics of the Segment	5	8.77		15	26.32		31	54.39		6	10.5											2.67	0.79	
Task Cognitive Demand	42	73.68		9	15.79		6	10.53		0	0.00											1.37	0.67	
Students Work with Contextualized Problems	11	19.30		30	52.63		16	28.07		0	0.00											2.09	0.69	
Overall Student Participation in Meaning Making and Reasoning	39	68.42		14	24.56		4	7.02		0	0.00											1.39	0.62	

Note. N=57 (the total number of lesson segments)

Figure 4.63 also shows that Esra's statistics instruction did not include the students' engagement with the mathematical and statistical content in SPE, SMQR, TCD, and overall code in most lesson segments. However, the students communicated about the content and worked with contextualized problems in many lesson segments. A few segments included Low and Mid quality indicators in terms of SPE, SMQR, TCD, and overall code. The following sections present the details of the codes of this dimension.

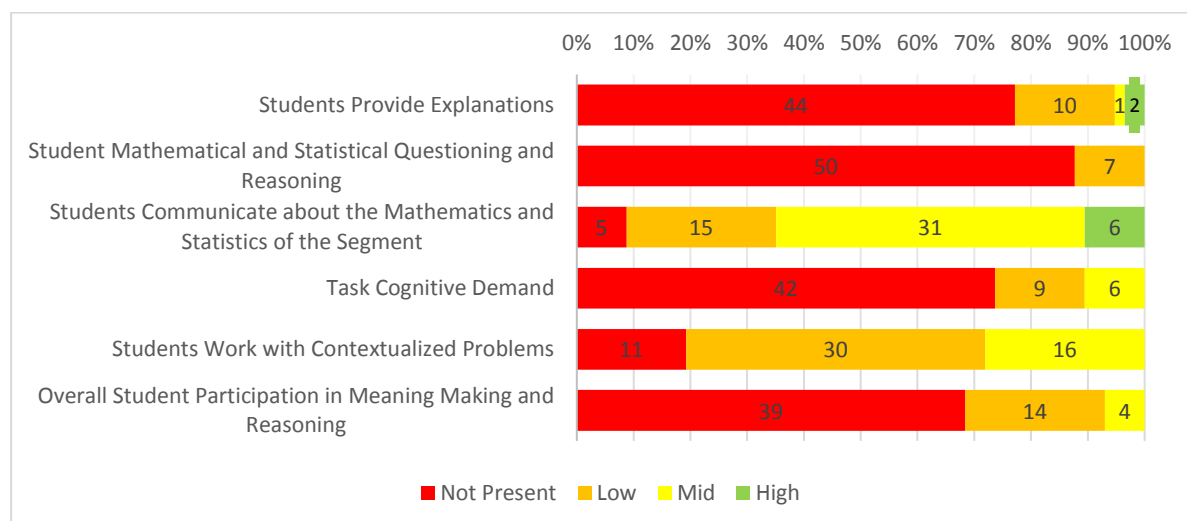


Figure 4.63 “Student Participation in Meaning Making and Reasoning” Dimension in Esra's Instruction

Note. N=57 (the total number of lesson segments)

4.1.1.2.5.1. Students Provide Explanations

Figure 4.64 shows the distribution of scores assigned to lesson segments for students' mathematical and statistical explanations in Esra's statistics instruction. In 44 (77.19%) lesson segments, the students did not provide explanations. There were brief student explanations at 10 (17.54%) lesson segments. Furthermore, the students provided explanations more frequently in one (1.75%) lesson segment. Student explanations characterized two (3.51%) lesson segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.32 out of 4, with a standard deviation of 0.69. This section provides the descriptions and classroom examples of each score assigned to the Students Provide Explanations code in Esra's statistics instruction.

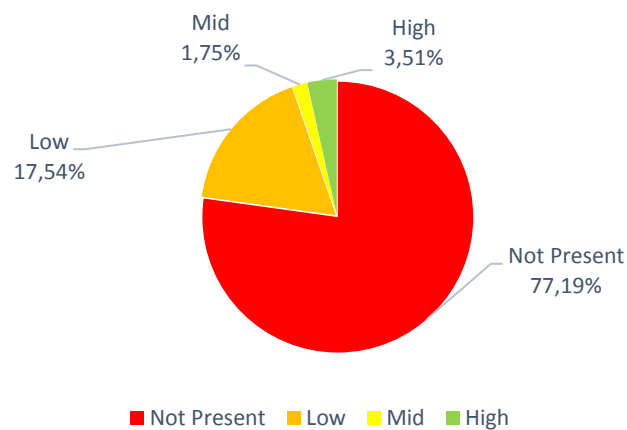


Figure 4.64 The Distribution of Scores Assigned to Lesson Segments for “Students Provide Explanations” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

The SPE code was scored as Not Present in 44 (77.19%) lesson segments. Throughout these segments, the teacher predominantly talked while students provided only one- or two-word answers to questions without participating in meaning-making or reasoning. Some segments consisted of reading a graph; the students just interpreted what they saw on the graph or analyzed its findings. However, the students did not read beyond the graph, which required making predictions based on justifications. Furthermore, the students worked individually; their communication with the teacher over the content was not hearable to the observer. Therefore, I scored Not Present for those segments. I scored Low to 10 (17.54%) lesson segments for student explanations because one or two brief student explanations were present. While doing calculations about a circle graph, the students explained the procedure by focusing on the meaning. For example, the students explained how they found a quantity by looking at multiplicative relationships while constructing a circle graph. Additionally, they explained how they found the percentage of a sector by adding other sectors up and subtracting from the whole, which was 100%. Other segments included examples of measures of central tendency. The students gave one or two brief explanations of why not to choose, for instance, the arithmetic mean for describing the data set. The student said they should not use the arithmetic mean because 1600 was in the data set. The student's explanation was not sustained.

One (1.75%) lesson segment consisted of more frequent student explanations. The students examined an average task in the segment and decided which central tendency measures to describe the data set. The data set consisted of money for a week “1000, 1250, 1500, 2500, 2750, 1500”. The students’ answers comprised the arithmetic mean, median, and mode. The students explained why they selected arithmetic mean or mode. The student explanations were incomplete; some of them were not correct. For example, a student explained why to find mode because there were large numbers. Another student explained why to calculate the arithmetic mean because the numbers were close. Some students gave correct answers by selecting the median, but the teacher did not ask why to find the median.

Student explanations characterized the segment in two (3.51%) lesson segments. In both segments, the average concepts were covered. The teacher expected students to decide which measures to choose based on the given context and why they chose the measure. The teacher challenged students’ explanations and pressured students to explain their reasoning. For example, one segment compared the grade averages of mathematics in two classes. The teacher and students calculated all measures of central tendency and then discussed which one to use for comparing both groups. The arithmetic means of both groups were close to each other. The teacher asked if they could use the arithmetic mean for comparing the groups. Based on the calculations, the teacher asked if 7A class was less successful than the 7B. The teacher challenged students’ thinking. The students explained why they should not use the arithmetic mean. One student stated that one group had an extreme value, so the arithmetic mean was affected by the extreme value. The group’s average became lower because of the extreme value. The student explanations were the focus of the segments.

4.1.1.2.5.2. Student Mathematical and Statistical Questioning and Reasoning

Figure 4.65 shows the distribution of scores assigned to lesson segments for Student Mathematical and statistical Questioning and Reasoning in Esra’s statistics instruction. There were no student mathematical and statistical questions and reasoning instances in 50 (87.72%) lesson segments. Seven (12.28%) lesson segments included brief one or two incidents about student mathematical and statistical questioning and reasoning.

Mid-quality and High-quality lesson segments were not present in SMQR. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.12 out of 4, with a standard deviation of 0.33. This section provides the descriptions and classroom examples of each score assigned to the Student Mathematical and Statistical Questioning and Reasoning code in Esra's statistics instruction.

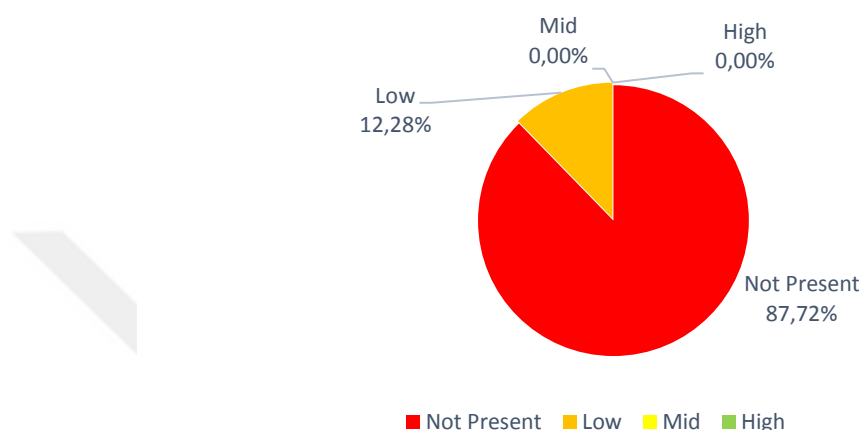


Figure 4.65 The Distribution of Scores Assigned to Lesson Segments for “Student Mathematical and Statistical Questioning and Reasoning” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

One or two brief student mathematical and statistical questioning or reasoning in lesson segments scored Low (n=7, 12.28%). For example, the students reasoned about the multiplicative relationship while calculating a circle graph. They also had some mathematically and statistically motivated questions. Their reasoning was incorrect; the students stated they could use what remained from the total number after using the part instead of the total number corresponding to 360 degrees. In another example, a student questioned the representation of a sector with 90 degrees in a graph.

There were some student questioning or reasoning examples in lesson segments with the average concept. For example, the teacher covered an average task where the median should be used because of extreme values. After the discussion, a student asked mathematically and statistically motivated questions by providing another example. To discuss using the arithmetic mean in the previous example, the student stated three

numbers, 98, 94, and 94. The incidents were not sustained, relatively brief ones. Therefore, I scored Low for the segments.

4.1.1.2.5.3. Students Communicate about the Mathematics and statistics of the Segment

Figure 4.66 shows the distribution of scores assigned to lesson segments for students' communication about the mathematics and statistics of the segment in Esra's instruction. Students did not contribute the mathematics and statistics of the segment while the teacher talk was dominant in five (8.77%) lesson segments. 15 (26.32%) lesson segments included brief student contributions; students gave one-or-two-word answers to the teacher's questions.

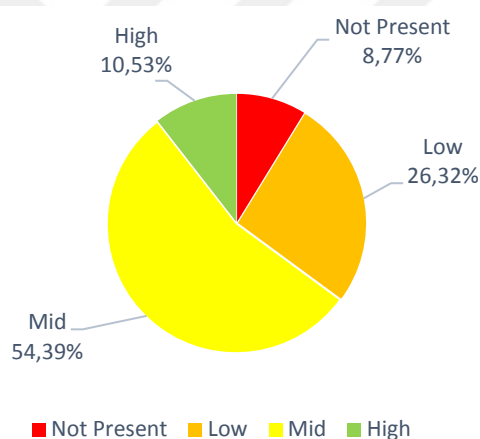


Figure 4.66 The Distribution of Scores Assigned to Lesson Segments for the “Students Communicate about the Mathematics and Statistics of the Segment” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

Substantive student contribution was present in 31 (54.39%) lesson segments. The student contributions characterized the segment at 6 (10.53%) lesson segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.67 out of 4, with a standard deviation of 0.79. This section provides the descriptions and classroom examples of each score assigned to the

Students Communicate about the Mathematics and Statistics of the Segment code in Esra's statistics instruction.

The students did not communicate about the mathematics and statistics of the segment in five (8.77%) lesson segments. The lesson segments included individual work time; students practiced the topic by solving questions. Some infrequent talk between students, and the teacher was present. Especially student contribution was not hearable during observations. Therefore, I scored Not Present for SCM. In addition, the student contribution was absent in some segments; the teacher's talk was dominant. Furthermore, one segment also did not consist of mathematical and statistical work.

The student contributions were present briefly in 15 (26.32%) lesson segments. The teacher dominated the lesson segments with their talk, while student contributions were limited to one or two-word answers to teacher questions, which were frequent but very short. During the first examples of the circle graph, the teacher provided the related calculations. The students gave answers to the calculations. The students briefly answered data analysis and interpretation questions in some segments with line and bar graph tasks. When they had errors and difficulties, the teacher corrected them by describing them in detail. Similarly, the lesson segments with measures of central tendency tasks included students' brief answers to the questions about the mode, the median, or the arithmetic mean of the data set. In the segments, the students' substantive contributions were not present.

The substantive student contributions were present in 31 (54.39%) lesson segments, but the contributions did not characterize the segments. In the lesson segments, the student solved the tasks on the board. Some students did not discuss their solutions, but the teacher described students' work on the board. Some students also described their solutions. The student contribution was considered substantive when the students worked on solution methods with or without words. In addition, while the teacher was solving the questions on board, the students offered explanations about finding the quantity, angle, or percentage corresponding to a circle sector. In other segments, the students described the meaning of a term.

In six (10.53%) lesson segments, the student contributions characterized the segments. The students' ideas were communicated dominantly in those segments. For example, in the first lesson of the statistics unit, the students analyzed and interpreted the collected data about the favorite color of their class. The student's interpretation of the graph was the focus of the segment. Furthermore, the students communicated about the measures of central tendency substantively; they examined the data set; and discussed which measures of central tendency should be used for describing the data set. Especially the lesson segments with a High score for SPE were scored as High for SCM because the contributions were the explanations on using the arithmetic mean, median, or mode for finding typical value for the given context characterized by the segment.

4.1.1.2.5.4. Task Cognitive Demand

Figure 4.67 shows the distribution of scores assigned to lesson segments for the Task Cognitive Demand code in Esra's instruction. In 42 (73.68%) lesson segments, students did not engage in cognitively demanding activities. Nine (15.79%) lesson segments comprised brief students' engagement in cognitively demanding activities. Six (10.53%) lesson segments included demanding and undemanding activities. Students did not engage in high cognitive demand in any segments. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.37 out of 4, with a standard deviation of 0.67. This section provides the descriptions and classroom examples of each score assigned to the Task Cognitive Demand (TCD) code in Esra's statistics instruction.

I gave a score of Not Present to 42 (73.68%) lesson segments for students' engagement in cognitively demanding activities. The lesson segments included dominant teacher talk with no student input or a few students' input. In addition, the students applied well-established procedures while constructing circle graphs or calculating measures of central tendency. One segment included unsystematic explorations; the students tried to guess the definition of mode based on the word meaning of peak value (i.e., Tepe değeri). The teacher took all the guesses from the students without directing or guiding them to conceptualize the meaning of the mode. Some students described the

mode as the largest number because of the word's meaning. Therefore, this section contained exploration without leading to the correct conceptualization of the mode.

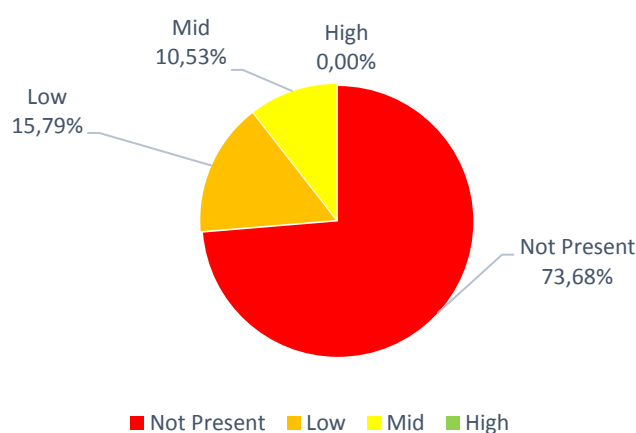


Figure 4.67 The Distribution of Scores Assigned to Lesson Segments for “Task Cognitive Demand” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

I scored Low to nine (15.79%) lesson segments for the Task Cognitive Demand code. Some segments included one or two student explanations or reasoning about the calculations during solving circle graph tasks. The initial cognitive demand of the tasks and the cognitive demand during the set-up phase were not cognitively demanding because the tasks were routine. In addition, the nature of the instruction was teacher-directed, not student-centered. However, I scored Low on the segments for student contributions on the solution process of circle graph tasks. Other segments included measures of central tendency tasks. The tasks were non-routine to some extent. However, the teacher heavily scaffolded students during the enactment phase; she gave hints about using the arithmetic mean, median, and mode.

Lesson segments scored Mid (n=6, 10.53%) included cognitively demanding and undemanding activities. The task enactment started as cognitively demanding; they included open-ended tasks. The student explanations and reasoning were more present than in the Low lesson segments. In the first lesson of the statistics unit, the teacher asked students what they inferred from the circle graph. Therefore, the students contributed to making sense of the circle graph. However, the teacher also contributed

to the development of mathematics and statistics in the segment. She summarized what the students stated and directed the students' contributions. Therefore, cognitively demanding and undemanding activities occurred during the segment.

Similarly, the lesson segments with average tasks also had cognitively demanding parts with some student directions. After the lesson segments scored Low, the teacher opened a room for students to discuss the average concepts for the following lesson segments. The students made significant contributions to the development of mathematics and statistics of the segments with their explanations. However, the teacher-directed students to reason about average concepts after hearing student ideas. The teacher's guidance lowered the cognitive demand at some points.

4.1.1.2.5.5. Students Work with Contextualized Problems

Figure 4.68 shows the distribution of scores assigned to lesson segments for students' work with contextualized problems in Esra's statistics instruction. In 11 (19.30%) lesson segments, students did not work with contextualized problems. In 30 (52.63%) lesson segments, students worked on contextualized problems as routine exercises, while in 16 (28.07%) segments, they reasoned about contextualized problems during the solution process with the teacher providing guidance. However, none of the segments included high-level student engagement with contextualized problems.

The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 2.09 out of 4, with a standard deviation of 0.69. This section provides the descriptions and classroom examples of each score assigned to the Students Work with Contextualized Problems code in Esra's statistics instruction.

The lesson segments scored Not Present ($n=11$, 19.30%) for SWCP did not consist of students' work on contextualized problems. In some segments, there were no tasks, but some content was discussed (e.g., a discussion on the definitions of measures of central tendency). In other segments, the teacher provided data sets without contexts. The students worked on the data sets for calculating the measures of central tendency. Two lesson segments were about a line graph: The first one introduced a line graph

with examples of related contexts but without a problem. The second one was about drawing a line graph. Even if there was a context while drawing, the students did not work on the context; but focused on the elements of a line graph (e.g., the intersection of x and y axes; the joining of two points by a line).

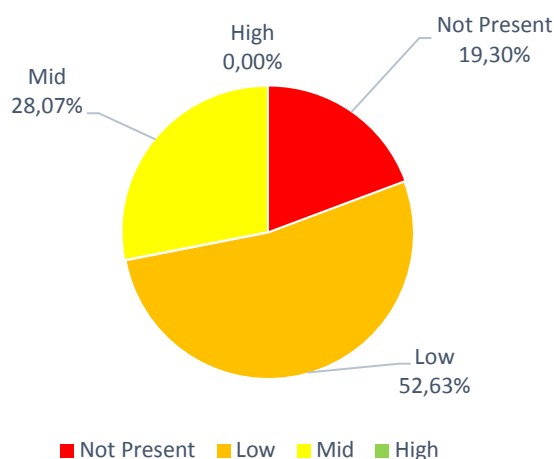


Figure 4.68 The Distribution of Scores Assigned to Lesson Segments for “Students Work with Contextualized Problems” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

The contextualized problems were routine exercises in 30 (52.63%) lesson segments. The teacher scaffolded the students during the segments. The teacher told students what to do during calculations about graphs and measures of central tendency. In some cases, the students were on board, writing what the teacher said. In other cases, students were working on routine exercises individually. The teacher heavily scaffolded the students by solving the questions the students had difficulty with the content. Especially in circle graph tasks, the teacher helped students solve the equations based on proportional reasoning. She stated which procedure to apply; how to solve proportions. She also helped students having difficulties with division. Furthermore, she directed students to choose which measures of central tendency to use for a given context. For example, one task included a categorical variable, and she asked students which measures of the central tendency they should use. The students said mode. The teacher asked if they could calculate the arithmetic mean of the data set. Then she said no and answered the question for the students. The contextualized problem was not routine, but the teacher heavily scaffolded the students’ reasoning.

Some student reasoning about contextualized problems was present at 16 (28.07%) lesson segments. However, teacher scaffolding was also present to some extent. The lesson segments included contextualized problems about circle graphs and measures of central tendency. The segments with circle graph tasks consisted of students' finding a quantity/a percentage/angle corresponding to a circle's sector. The students contributed to how to solve the tasks. The teacher helped students with calculations. However, the student's role was observable in the solution process. Furthermore, the segments also consisted of the interpretation of graphs or data analysis on graphs. The contextualized problems started as open-ended, but the teacher's directions were present after a while.

Similarly, the contextualized problems with measures of central tendency were also started as non-routine. In the segments scored as Mid, the students and the teacher discussed which measures of central tendency to use for a given context after calculating all the measures. The students contributed to the solution of the problem. During student contributions, the teacher gave hints about choosing a specific measure.

4.1.1.2.5.6. Overall Student Participation in Meaning-Making and Reasoning

Figure 4.69 shows the distribution of the scores assigned to lesson segments for the Overall Student Participation in Meaning-Making and Reasoning code in Esra's instruction. In 39 (68.42%) lesson segments, students did not participate in any cognitively demanding activities; they were primarily routine exercises or heavily scaffolded by the teacher. The type of instruction was mainly IRE pattern; teacher talk was dominant. Students contributed, explained, and reasoned about the content or cognitively demanding tasks briefly in 14 (24.56%) lesson segments. However, the activities the students engaged in were primarily procedural.

In four (7.02%) lesson segments, students participated in meaning-making and reasoned about the content at a mixed level. They engaged with cognitively demanding tasks at a middling level. They worked with contextualized problems and contributed to the content substantively.

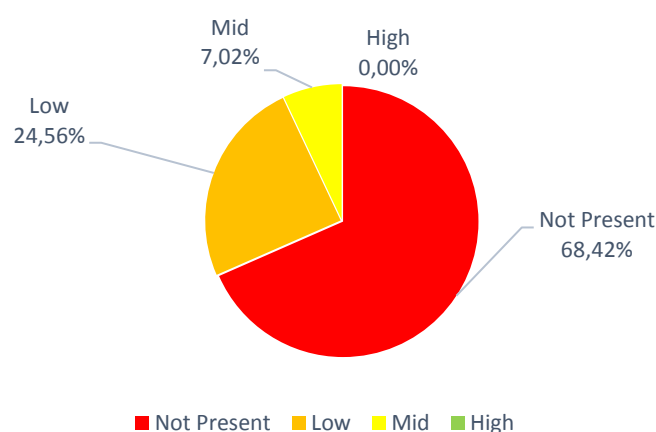


Figure 4.69 The Distribution of Scores Assigned to Lesson Segments for “Overall Student Participation in Meaning Making and Reasoning” Code in Esra’s Instruction

Note. N=57 (the total number of lesson segments)

No segments included a High score for overall code because no lesson segments were characterized by substantive student contributions with cognitively demanding activities. The results of the descriptive analysis indicate that the mean score assigned to the lesson segments on this code was 1.39 out of 4, with a standard deviation of 0.62.

4.1.1.2.5.7. Overview of “Student Participation in Meaning-Making and Reasoning” Dimension

The MAXQDA map generated for the current study consists of segment-level codes in the MQI with their assigned scores. This section discusses the findings based on the frequencies of occurrences and co-occurrences of Errors and Imprecision codes with their scores as visualized on the MAXQDA map (see Figure 4.70). As seen in the MAXQDA Map, student participation in meaning-making and reasoning was not present to a low quality in Esra’s statistics instruction. The occurrences of student explanations, questioning and reasoning, and task cognitive demand codes were most frequently absent in the overall code. The overall code with a Not Present score also co-occurred with the codes with Not Present scores frequently.

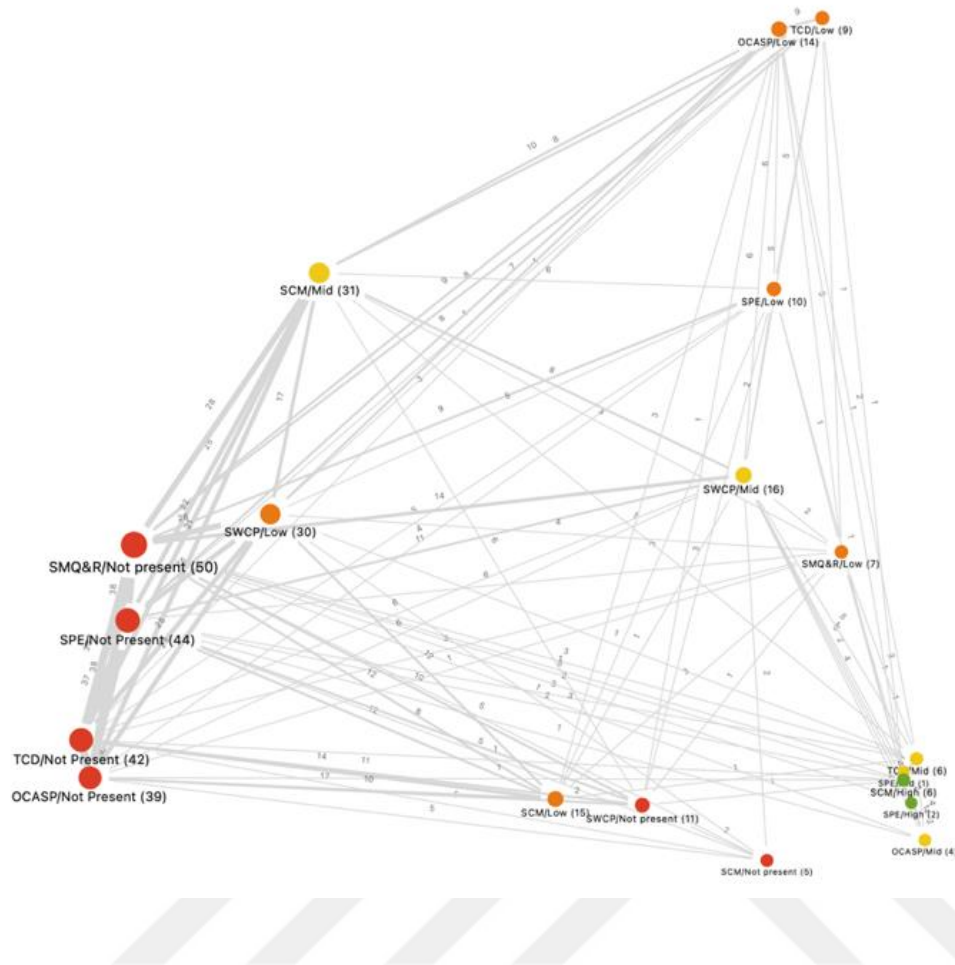


Figure 4.70 The Map of the Occurrences and Co-occurrences of the Scores Assigned to the Codes in “Student Participation in Meaning-Making and Reasoning”

Dimension in Esra’s Instruction

Note: SPE- Students Provide Explanation, SMQ&R- Student Mathematical and Statistical Questioning and Reasoning, SCM- Students Communicate about the Mathematics and Statistics of the Segment, TCD- Task Cognitive Demand, SWCP- Students Work with Contextualized Problems, OCASP- Overall Student Participation in Meaning-Making and Reasoning

Furthermore, the codes mentioned earlier with Not Present scores co-occurred with SWCP with Low and Mid scores and SCM with Low and Mid scores. The co-occurrences' frequencies were lower than those of Not Present scores of SPE, SQMR, and TCD. Accordingly, the students did not explain or reason the process. They did not engage in cognitively undemanding activities in many segments together. However, the segments generally consisted of brief student contributions, but some were at length. In addition, the segments also consisted of the students’ work with contextualized problems with varying levels of scaffolding from the teacher.

The frequency of Low, Mid, and High scores on many codes in the dimension was low, indicating that student participation in meaning-making and reasoning was infrequent in Esra's instruction. Figure 4.70 shows that Mid scores assigned to SPE, TCD, and overall code and High scores assigned to SPE and SCM were present in Esra's statistics instruction with the lowest frequencies. The distinct codes also co-occurred in similar segments (see the length of lines between each code and the width of the connected lines in Figure 4.70). The segments with High scores on SPE and SCM interacted with Mid scores on other codes because Esra's statistics instruction did not have the characteristics of High scores in other codes. Furthermore, SWCP with Mid scores and SMQR with Low scores co-occurred with the codes mentioned above with Mid and High scores, but the length of the lines among those codes was more prolonged than other codes. Thus, they were less applied to similar segments. The results indicate that substantive student explanations and contributions were present with middling task cognitive demand. When the instruction in the segments allowed students to engage in cognitively demanding activities, the students contributed to the development of statistical concepts. They explained their solution process and reasoning about using an average concept for a given context and the data set.

As a result, the map indicates that the codes in the Student Participation in Meaning Making and Reasoning dimension co-occurred at similar levels in most segments. In addition, the map also shows that most lesson segments did not include student participation in meaning-making and reasoning in Esra's instruction. Student participation was less frequent and occurred in similar lesson segments.

4.1.2.3. Whole Lesson Mathematical and Statistical Quality of Instruction (MQI): A Lesson Level Analysis in Esra's Statistics Instruction

This section aims to provide an overview of Esra's statistics instruction as a whole by evaluating the overall quality of instruction in terms of subject-specific teaching practices in mathematics and statistics. The analysis was conducted on twelve lessons of Esra's statistics instruction, evaluated at both the segment and lesson levels. To assess the overall quality of instruction, I used a specific code on a 5-point scale and

assigned scores to all lessons. Additionally, a descriptive analysis was conducted on all scores assigned to Esra's statistics instruction lessons.

The results show that Esra's instruction's overall mathematical and statistical quality can be characterized as Mid quality ($n=11$, 91.67%), with a mean of 3.08 and a standard deviation of 0.29. The Mid quality lessons did not exhibit Low and High-quality characteristics and were considered typical lessons. However, they were generally free from errors and imprecisions. The Mid/High-quality lessons were scored higher because they had richer mathematics and statistics, more substantive student contribution and participation, and responsive teacher-student interaction. However, they were not scored as High quality because not all high-quality codes were present and combined in the same lesson.

In summary, Esra's instruction demonstrated an overall Mid quality, with most lessons exhibiting typical but error-free and well-organized instruction. In the next section, I present a cross-case analysis and compare Ali's and Esra's teaching practices in terms of statistical content, problem-solving process, and the mathematical and statistical quality of instruction.

4.2. Cross-Case Analysis

This section consists of the findings of the cross-case analysis comparing data from Ali's and Esra's statistics teaching practices to identify patterns, themes, and relationships (Creswell, 2013; Yin, 1981). Specifically, the analysis focuses on three aspects: the alignment of statistical content with curricular expectations, the statistical problem-solving process, and the mathematical and statistical quality of instruction.

4.2.1. Statistical Content

The curriculum expected teachers to dedicate 15 lesson hours to cover four objectives in the statistics unit, which included constructing a circle graph, constructing a line graph, calculating central tendency measures, and transforming graphs. In this study, Ali spent 14 lesson hours teaching statistics and another three hours collecting student data to understand the classroom context. Similarly, Esra spent 12 lesson hours

teaching statistics and allowed three hours for student data collection. Additionally, Esra spent one extra hour for students to take an individual test on statistics to gauge their understanding of the topic. The total time dedicated to teaching statistics and conducting this study was consistent with what the curriculum expected from the teachers.

For the statistical content, the unit of analysis was the episode, which refers to a unique set of data that the class worked on (Litke & Hill, 2020). Ali's statistics instruction comprised 47 episodes, while Esra's contained 31. Figures 4.71 and 4.72 show the episodes allocated for the statistical content, mainly graphs and measures of central tendency.

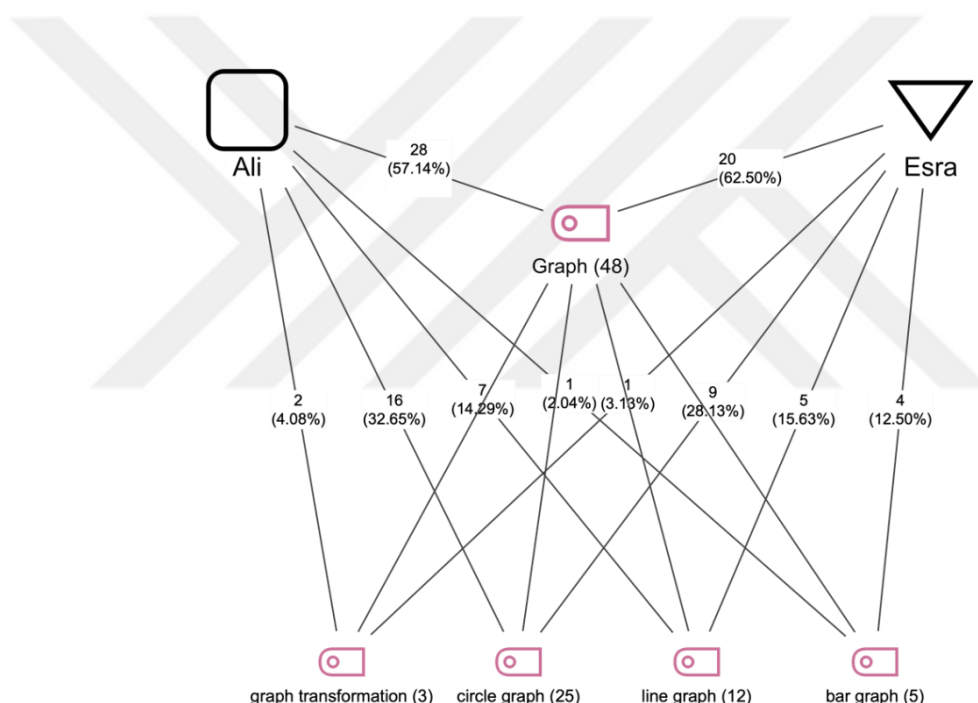


Figure 4.71 Graphs as the Statistical Content in Both Teachers' Statistics Instruction

When comparing the number of episodes allocated to each topic, both Ali and Esra spent similar amounts of episodes for graphs (57.14% and 62.50%, respectively). When the graphs were categorized, both teachers spent a similar number of episodes on the line graphs (14.29% for Ali and 15.63% for Esra) and graph transformation (4.08% for Ali and 3.13% for Esra) objectives. However, Esra's instruction (12.50%) included more episodes related to bar graphs than Ali's instruction (2.04%). In

addition, a slight difference was present in the number of episodes in the circle graph (32.65% for Ali and 28.13% for Esra).

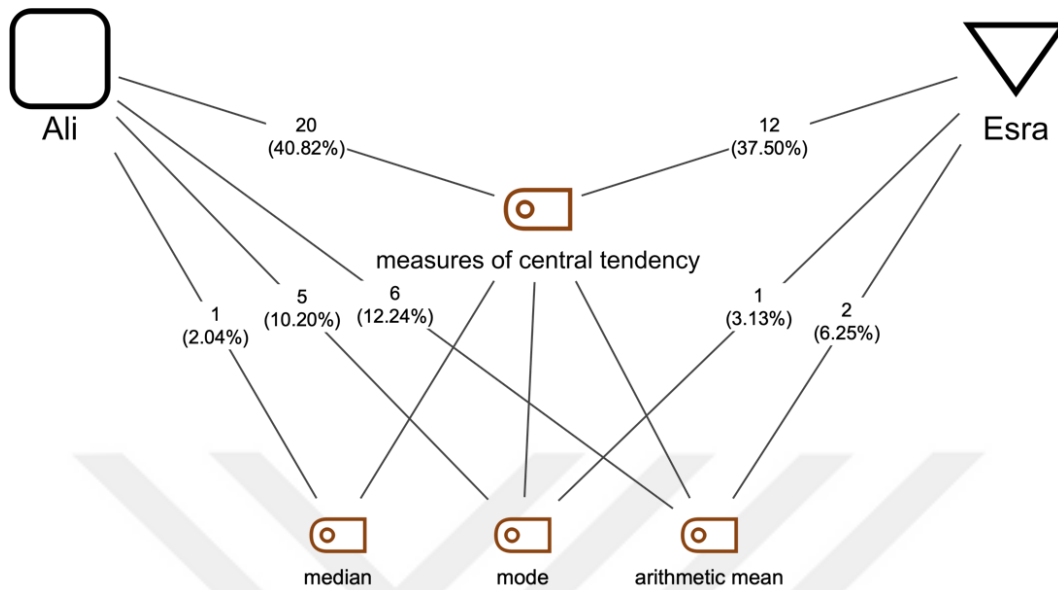


Figure 4.72 Measures of Central Tendency as the Statistical Content in Both Teachers' Statistics Instruction

Considering the measures of central tendency, Ali and Esra spent similar amounts of episodes (40.82% and 37.50%, respectively). Both teachers covered the same measures, including mean, mode, and median. However, Ali had one episode that focused only on the median, while Esra did not have a similar episode. As shown in Figure 4.72, both teachers often covered more than one measure in a single episode. Esra's instruction covered 9 out of 12 episodes with more than one measure, while Ali's instruction had 8 out of 20 episodes with more than one measure.

4.2.2. Statistical Problem-Solving Process

In this section, I compared Esra and Ali's statistics instruction regarding the presence of each element in the statistical problem-solving process, which consists of formulating questions, collecting data, analyzing data, and interpreting results. The Transcript Coding Protocol Using GAISE Framework had three categories: not observed (Level 1), observed to some extent without depth specified in GAISE (Level

2), and more fully aligned with GAISE expectations and implemented with a more in-depth focus on statistical ideas or concepts (Level 3).

4.2.2.1. Formulation of Questions

Figure 4.73 shows that Ali and Esra did not formulate questions in most episodes (93.88% and 96.87% at Level 1, respectively). The tasks were not based on a statistical investigative question in those episodes.

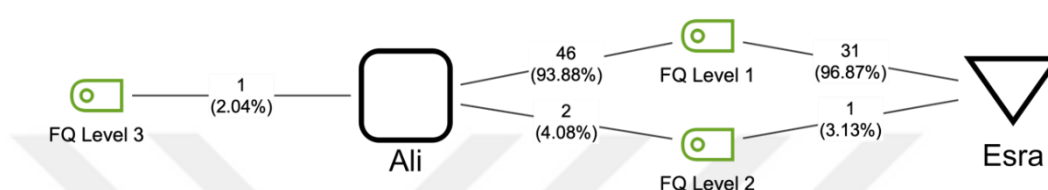


Figure 4.73 “Formulation of Questions” Element in Both Teachers’ Statistics

Instruction

Note: FQ- Formulation of Questions

Only 2.04% of the episodes in Ali’s instruction included research questions formulated by students (at Level 3), whereas, in Esra's instruction, students did not formulate any questions. However, even though students generated statistical questions in Ali’s instruction, it did not lead to any investigation because he did not intend to conduct a statistical inquiry. Furthermore, the questions were survey items and were not corrected or guided by the teacher. Both Ali and Esra posed statistical questions in at least one episode (4.08% and 3.13% at level 2, respectively), but Esra's question led to a statistical investigation, whereas Ali's questions did not. A data collection process followed Esra's question, whereas Ali's questions were based on pre-existing data provided by the teacher.

4.2.2.2. Data Collection

Figure 4.74 indicates that data collection was not usually present in Ali’s and Esra’s statistics instruction (97.96% and 96.87% at Level 1, respectively). Both teachers implemented the data collection in only one episode at Level 2.

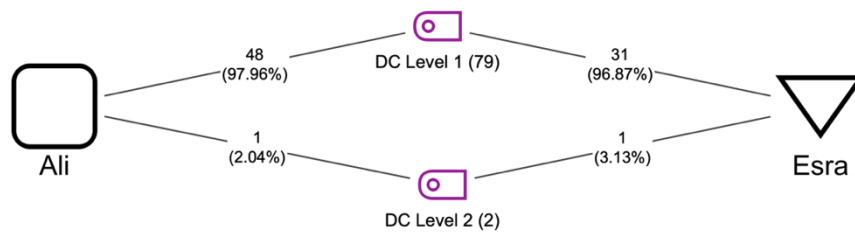


Figure 4.74 “Data Collection” Element in Both Teachers’ Statistics Instruction

Note: DC- Data Collection

The nature of data collection was different for both teachers. Ali collected data without meaningful context, whereas Esra collected data based on student interests. Ali distributed color strips to the students, recorded the data, and created tables and graphs. Still, the purpose of this activity was unrealistic as it only aimed to introduce a circle graph. In contrast, Esra collected data on students' favorite colors, which was more relevant and realistic for them. Both teachers created tables and graphs based on their collected data, but Esra's approach was more contextualized and relevant for the students.

4.2.2.3. Data Analysis

Figure 4.75 shows that both teachers used data analysis in most episodes. Data analysis was not carried out in 2.04% of episodes of Ali’s instruction, while all episodes in Esra’s instruction included the data analysis elements. 79.87% of Esra’s episodes included data analysis at Level 2 by constructing and reading a graph and calculating central tendency measures.

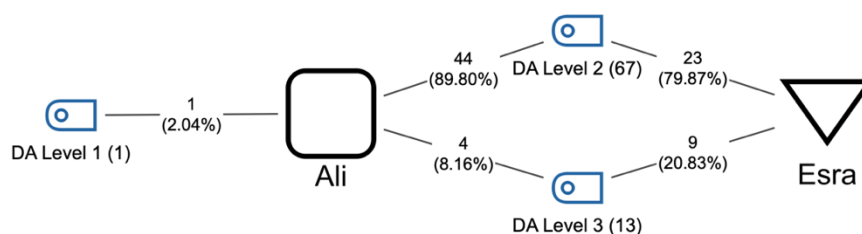


Figure 4.75 “Data Analysis” Element in Both Teachers’ Statistics Instruction

Note: DA- Data Analysis

Similarly, Ali's instruction included 89.80% of episodes for data analysis at Level 2. At Level 2, data analysis consisted of answering primarily descriptive questions, and they were procedural. Only 8.5% of Ali's episodes and 20.83% of Esra's episodes had the data analysis element at level 3. Esra provided more episodes describing the distribution, comparing groups, or observing relationships among variables than Ali. For example, both teachers used the same examples for finding appropriate central tendency measures for a given context. Although both teachers had different approaches to describing a distribution in those questions, they had opportunities to connect the context and the data by understanding the distribution with the average concept. Ali's and Esra's instructions included data analysis for comparing groups. Unlike Ali, Esra used graphs to describe a distribution or compare groups in the graphs.

4.2.2.4. Interpretation

According to Figure 4.76, a higher percentage of Ali's episodes (81.63%) did not involve any interpretation compared to Esra's episodes (59.37%). The result suggests that Ali did not encourage students to interpret the results obtained from data analysis in most episodes. However, in 12.8% of episodes, Ali's instruction did address “within data” questions from graphs. Furthermore, there was only one episode where a within-data question about central tendency measures was answered.

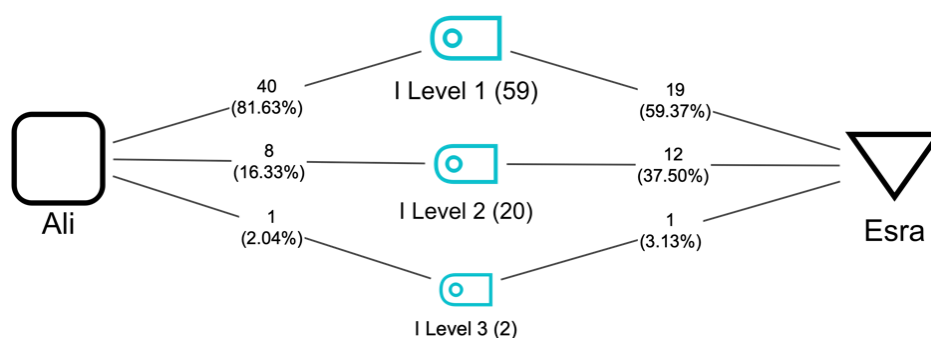


Figure 4.76 “Interpretation” Element in Both Teachers’ Statistics Instruction

Note: I- Interpretation

In contrast, Esra's instruction had a higher percentage (37.50%) of episodes at Level 2, indicating that she emphasized interpretation more than Ali (16.33% at Level 2). Specifically, Esra consistently asked students to interpret the information presented on a graph. Moreover, some of the average tasks given by Esra required students' interpretation of results from the calculations.

Both Ali and Esra's instruction had one episode with the interpretation of findings by discussing the variability, making predictions, and comparing groups (2.04% for Ali, 3.13% for Esra). In Ali's episode, they discussed the outliers' effects on interpreting the results and then reaching incomplete or incorrect interpretations while comparing groups. The context was different, but the purpose was similar to Esra's episode. Both teachers also concluded the necessity of using the median in data analysis to make accurate interpretations.

4.2.2.5. Overview of Statistical Problem-Solving Process

After examining the statistical problem-solving process concerning each element's occurrences, the episodes were analyzed in terms of co-occurrences of the categories assigned to each element, aimed to look at the statistical investigation as a whole process. The results show that neither teacher conducted statistical investigations except for Esra's one episode. In this episode, Esra's instruction consisted of the teacher's statistical question, collecting data from the class, analyzing the data to describe the distribution, comparing different parts of the collected data, and interpreting the results within the data. None of the formulated questions led to a statistical investigation in Ali's instruction, and Esra and her students did not pose a statistical question except in one episode.

All in all, both teachers did not carry out the statistics instruction as a statistical problem-solving process. They followed the objectives in 7th grade in general. In addition, the examples in the textbooks generally consisted of data analysis questions. Therefore, when the occurrences and co-occurrences of the elements were examined, the focus on the data analysis was pronounced. Neither teacher considered the spiral structure of the curriculum during the instruction.

When asked what would be better to do in the instruction in the last interview, Esra stated that it would be better to start with an investigation related to students' daily lives. The students could collect data from their surroundings and bring them to the class. Based on the data, tables, and graphs could be constructed and interpreted. Ali did not focus on the statistical investigative process in his interviews as in his teaching practices. However, he also stated that he understood the importance of interpreting analyzed data.

4.2.3. Mathematical and Statistical Quality of Instruction: Segment Level Analysis

I utilized the Mathematical Quality of Instruction (MQI) instrument to examine the statistics instruction of Ali and Esra. In this section, I compare their practices in terms of the quality codes from the segment-level analysis. Specifically, I compare the instruction of both teachers across 21 codes within four dimensions. The segment-level analysis involved using 7.5-minute lesson segments as the unit of analysis. I divided the lessons into 7.5-minute segments and evaluated each segment using 21 segment-level codes as quality indicators. Each code, except for "Classroom Work is Connected to Mathematics and Statistics," was scored on a scale of Not Present (1), Low (2), Mid (3), or High (4). Ali's statistics instruction comprised 61 segments, whereas Esra's consisted of 57.

4.2.3.1. Classroom Work is Connected to Mathematics and Statistics

All of Ali's lesson segments were connected to mathematics and statistics, but Esra's only one was not connected to mathematics and statistics (See Figure 4.77). In his first interview, he described arranging his teaching so that no students could engage in nonmathematical work for a while. The observation data supported this statement. Even if there were some distractions, the teacher warned students to engage in mathematics and statistics. While checking homework, he wanted students to continue practicing the content.

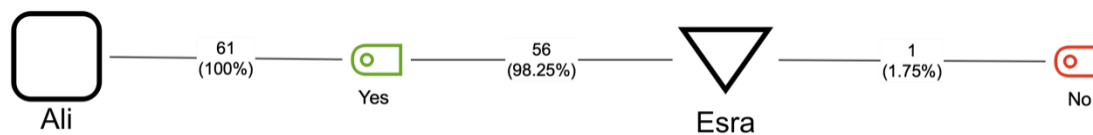


Figure 4.77 “The Classroom Work is Connected to Mathematics and Statistics” Code in Both Teachers’ Instruction

The only segment in Esra was related to checking homework. It took one segment and some other parts of the following segment. During checking homework, other students were waiting for the teacher to check theirs. Furthermore, in Ali’s instruction, even if Ali was teaching to the whole class and some students already understood and solved the task. Ali allowed those students to move on; he called on them when an important point was mentioned. Ali’s instruction focused on staying with mathematics and statistics in the class. Esra’s instruction also focused on mathematical and statistical content during most segments except for one. However, she did not have a similar teaching style to Ali. She wanted all students to work on and listen to what was told. In Esra’s segments, some parts were spent writing the task (e.g., drawing tables and graphs in the tasks) on the blackboard. However, the students were expected to read or work on the tasks in those parts. Therefore, the focus stayed on the content in most segments.

4.2.3.2. The Richness of the Mathematics and Statistics

This section compares Ali’s and Esra’s statistics instruction regarding “the depth of the mathematics and statistics offered to the students” (Hill et al., 2014). This dimension consists of seven codes scoring from Not Present to High (1 to 4). One code was the overall richness code. The other six codes belonged to two categories. One category was about the meaning of facts and procedures in the instruction. Linking Between Representations (LBR), Explanations, and Mathematical and Statistical Sense-Making were the codes for this category. The second category was about critical mathematical and statistical practices in the instruction. Multiple Procedures or Solution Methods (MPSM), Patterns and Generalizations (P&G), and Mathematical and Statistical Language (ML) were the codes for this category. I will compare Ali’s

and Esra's statistics instructions by following the codes in the categories. In the end, I will compare both in terms of overall richness.

4.2.3.2.1. Linking Between Representations

Figure 4.78 compares the scores for the "Linking between representation" code assigned to lesson segments in the statistics instruction of both teachers. First, both teachers used symbolic representations, tables, or stories as representational families. However, they were not linked or used within the same segments. Some segments included the construction of graphs or finding graphs' elements without reference to the other types of representations. For example, for the transforming graph objective, both teachers used different graphs but were from the same representational families. Therefore, this code was scored as Not Present in 80.33% of Ali's lesson segments and 89.47% of Esra's.

Furthermore, both teachers had similar incidents when they linked the representations at a low level (13.11% for Ali, 8.77% for Esra). While constructing circle graphs, they linked the angle with the shape of a sector corresponding to the given data. However, their links were brief and not explicit. They both linked the representations during a segment while teaching average. Esra's link was low; she linked the frequency table with an open version of the data to find the arithmetic mean. However, the link was not explicit. Besides, Ali used the open form of the data from bar graphs to show the mode of the data set. His link was explicit, but it was not the focus of the instruction, so it was scored as Mid (out of 4.92% Mid scores).

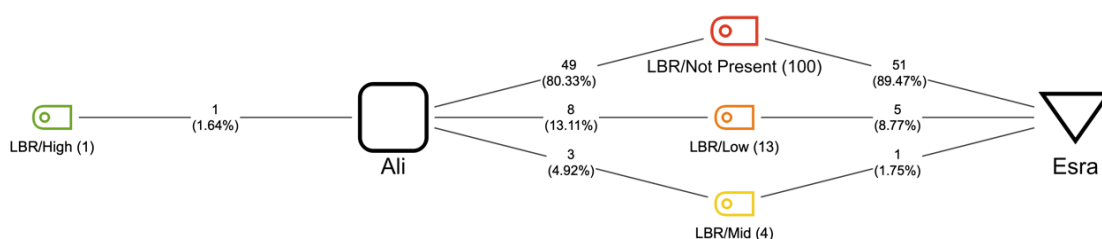


Figure 4.78 "Linking Between Representations" Code in Both Teachers' Statistics Instruction

Note: LBR- Linking between Representations

Only one segment was scored Mid for LBR in Esra's instruction (1.75%). The segment was about constructing a circle graph. The teacher and students reasoned with procedures related to the circle components but did not elaborate on the link among the representations. Similarly, Ali had one segment with explicit links while constructing a circle graph with a high score (1.64%). Unlike Esra's mid-scored segment on LBR, Ali elaborated on the connections among representations. He purposefully selected the numbers in the example; he distributed the colors to the students and wanted them to be a circle with similar colors side by side. He connected students' concrete representation, the frequency table, and the circle graph explicitly and in detail. These practices were the focus of the instruction in the segment.

Overall, Ali and Esra's statistics instruction was similar in LBR; they mostly lacked the link among representations. When the link was present, they were at a low level and about graphs in general. Both teachers had one segment with good quality in LBR, especially while constructing circle graphs in their first examples.

4.2.3.2.2. Explanations

Figure 4.79 presents a comparison of scores for the "Explanations" code assigned to lesson segments in the statistics instruction of both teachers. Ali and Esra did not provide mathematical and statistical explanations in most segments (55.74%, and 61.40%, respectively). Both teachers described how to find the elements of a circle graph, draw a line graph, or calculate central tendency measures, but they did not answer why to use the procedures or why the solution method worked.

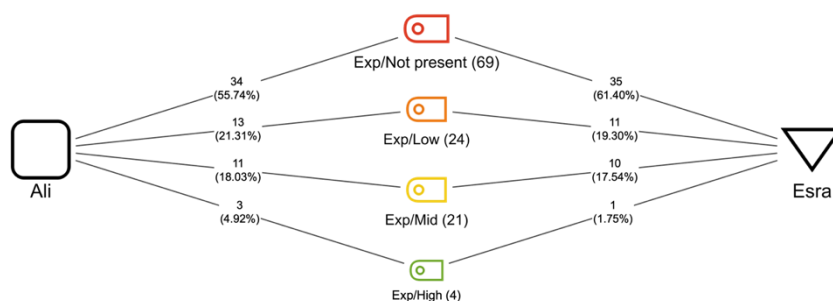


Figure 4.79 "Explanations" Code in Both Teachers' Statistics Instruction

Note: Exp -Explanations

Furthermore, both teachers' instruction consisted of explanations at a low level (21.31% for Ali, 19.30% for Esra). The explanations in those segments were isolated and brief in general. While solving a task, they briefly focused on why the solution was correct or incorrect, why the proportion was used in a specific way, or why a particular graph was used were subject to brief incidents of explanations in both teachers' instruction.

Similarly, they used one or two brief explanations or more than brief explanations (18.03% for Ali and 17.54% for Esra). The segments included why a procedure works while constructing a circle graph or using a specific central tendency measure for a given context. However, those explanations were not the focus of the instruction. Ali's and Esra's instruction included mathematical and statistical explanations, which were the focus of the segment while especially discussing context-dependency of average concepts (4.92% for Ali and 1.75% for Esra). Those segments focused on using median over arithmetic mean when extreme scores were in a data set in general.

4.2.3.2.3. Mathematical and Statistical Sense-Making

Figure 4.80 presents a comparison of scores for the “Mathematical and Statistical Sense-Making” code assigned to lesson segments in the statistics instruction of both teachers. While all segments in Ali's instruction included mathematical and statistical sense-making practices, 7.02% of Esra's lesson segments lacked these processes entirely and instead contained incorrect sense-making practices with no mathematical or statistical work.

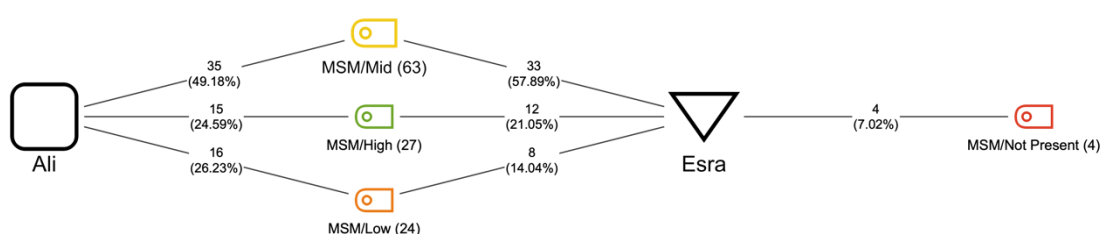


Figure 4.80 “Mathematical and Statistical Sense-Making” Code in Both Teachers’
Statistics Instruction

Note: MSM- Mathematical and Statistical Sense-Making

Furthermore, while Ali's instruction consisted of mathematical and statistical sense-making at a low level in 26.23% of lesson segments, Esra briefly focused on meaning in 14.04% of lesson segments. Both teachers focused on sense-making more than briefly for most segments (49.18% for Ali and 57.89% for Esra). In addition, sense-making was the focus of the segment in 24.59% of Ali's lesson segments and 21.05% of Esra's lesson segments. During segments scored Mid and High, Ali and Esra made sense in constructing, analyzing, and interpreting graphs. They used proportional reasoning during circle graph tasks. In addition, they focused on the context during some average tasks to make sense of different central tendency measures.

Overall, both teachers made sense of statistical concepts and mathematical calculations except for four segments in Esra's instruction. In both teachers' instruction, the nature of the sense-making process was changed when introducing a concept or doing exercises in later parts of teaching the concept.

4.2.3.2.4. Multiple Procedures or Solution Methods

Figure 4.81 presents a comparison of scores for the "Multiple Procedures or Solution Methods" code assigned to lesson segments in the statistics instruction of both teachers. The MPSM code was missing in a significant portion of both teachers' instruction, with a percentage of 77.05% for Ali and 85.96% for Esra. In most lesson segments, the teachers did not utilize multiple procedures or solution methods, instead focusing on a single method or procedure, particularly in tasks involving reading graphs and calculating averages.

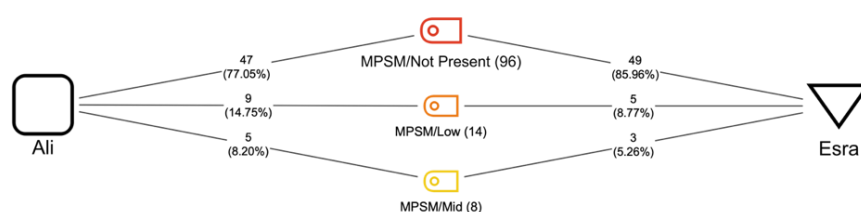


Figure 4.81 "Multiple Procedures or Solution Methods" Code in Both Teachers' Statistics Instruction

Note: MPSM- Multiple Procedures or Solution Methods

Only a few segments that involved average tasks included MPSM practices. Also, teachers taught students to construct a circle graph using multiple procedures or solution methods, emphasizing proportionality from different perspectives, such as within or between ratios. Therefore, both Ali and Esra used multiple procedures or solution methods in their lessons, but at a low level (14.75% for Ali and 8.77% for Esra). Although they briefly mentioned a second method or procedure, they did not discuss them at length. Some lesson segments included the description of multiple procedures or solution methods at length. Still, comparing methods or procedures was not explicit enough for students to see the connections among each method or procedure (8.20% for Ali, 5.26% for Esra). However, none of the lesson segments included high-quality practices on MPSM for both teachers. In summary, the nature of both teachers' practices regarding MPSM was similar regarding the presence of multiple procedures or solution methods.

4.2.3.2.5. Patterns and Generalization

Figure 4.82 presents a comparison of scores for the “Patterns and Generalization” code assigned to lesson segments in the statistics instruction of both teachers. None of Esra’s instruction segments included practices related to patterns and generalizations, while only 3.28% of lesson segments of Ali’s instruction consisted of working on a generalization or examining patterns.

Ali’s work on patterns and generalization was not the focus of the segments and was low. All in all, this code was also not present in both teachers’ instruction in general. The teachers did not work on patterns in the data or read beyond the data by examining patterns in the data. They did not work on forming definitions based on generalizations.

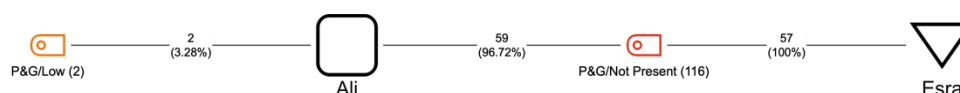


Figure 4.82 “Patterns and Generalization” Code in Both Teachers’ Statistics

Instruction

Note: P&G- Patterns and Generalization

4.2.3.2.6. Mathematical and Statistical Language

Figure 4.83 presents a comparison of scores for the “Mathematical and Statistical Language” code assigned to lesson segments in the statistics instruction of both teachers. In Ali’s instruction, all segments included mathematical and statistical language. Esra’s instruction did not include mathematical and statistical language in one segment with no mathematical and statistical work.

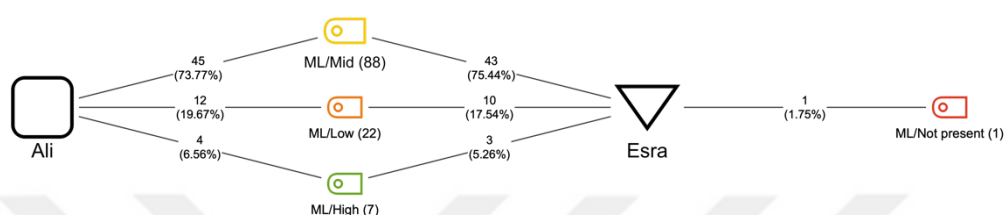


Figure 4.83 “Mathematical and Statistical Language” Code in Both Teachers’
Statistics Instruction

Note: ML- Mathematical and Statistical Language

Ali and Esra used low mathematical and statistical language density in similar coverage (19.67% and 17.54%, respectively). Similarly, both teachers used middling density mathematical and statistical language while covering graphs and average concepts (73.77% for Ali, 75.44% for Esra). Other teachers used some mathematical language when discussing circle graph tasks that involve proportionality. They generally used statistical terms like data, analysis, average, arithmetic mean, median, mode, middle point, tables, and graphs.

In contrast, Mid-scored segments in both teachers’ instruction lacked guidance on students’ appropriate mathematical and statistical language use. High-scored segments (6.56% for Ali and 5.26% for Esra), on the other hand, featured students using mathematical and statistical terminology or being encouraged to do so by the teachers. Therefore, it can be concluded that the density of language was higher in segments assigned High scores, indicating that the teachers provided more opportunities for students to engage with the mathematical and statistical language.

Overall, both teachers were similar in their use of mathematical and statistical language. They used mathematical and statistical language in their instruction except for one segment for Esra. However, they did not have high density. In most lesson segments with mid-density mathematical language, the teachers needed to engage with students regarding terminology.

4.2.3.2.7. Overall Richness of the Mathematics and Statistics

Figure 4.84 presents a comparison of scores for the “Overall Richness of the Mathematics and Statistics” code assigned to lesson segments in the statistics instruction of both teachers.

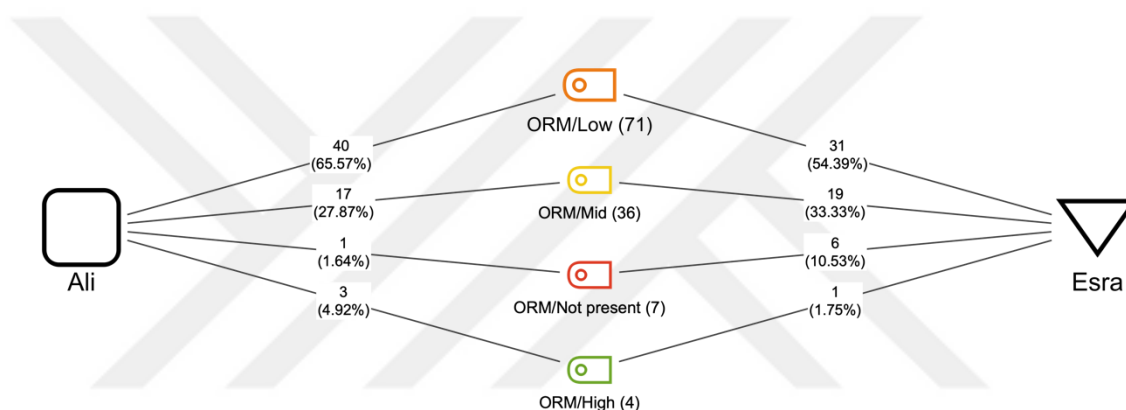


Figure 4.84 “Overall Richness of the Mathematics and Statistics” Code in Both Teachers’ Statistics Instruction

Note: ORM- Overall Richness of the Mathematics and Statistics

While Ali's instruction had an overall richness of only 1.64% of lesson segments, Esra's instruction had no overall richness in 10.53% of lesson segments. In 65.57% of lesson segments, Ali's instruction had richness codes present at a minimum level, while Esra's instruction had a low level of richness in 54.39% of lesson segments. However, Esra's instruction (33.33%) had slightly more Mid scores in overall richness than Ali's instruction (27.87%). Both teachers used richness codes for more than a brief period in those segments. In addition, 4.92% of lesson segments in Ali's instruction and 1.75% of Esra's instruction had high-quality practices in terms of overall richness code.

4.2.3.3. Working with Students and Mathematics and Statistics

In this section, I compare the teaching practices of Ali and Esra in terms of how they respond to students' contributions or errors using three codes ranging from Not Present to High (1 to 4). One code includes overall interactions between teachers and students around the content. The other codes are called Remediation of Student Errors and Difficulties (RSE&D), and Teacher Uses Student Mathematical and Statistical Contributions (TUSMC).

4.2.3.3.1. Remediation of Student Errors and Difficulties

Figure 4.85 presents a comparison of scores for the “Remediation of Student Errors and Difficulties” code assigned to lesson segments in the statistics instruction of both teachers. Both teachers had a dominant talk in some segments, so student errors and difficulties were not observed in some segments (32.79% for Ali and 56.14% for Esra).

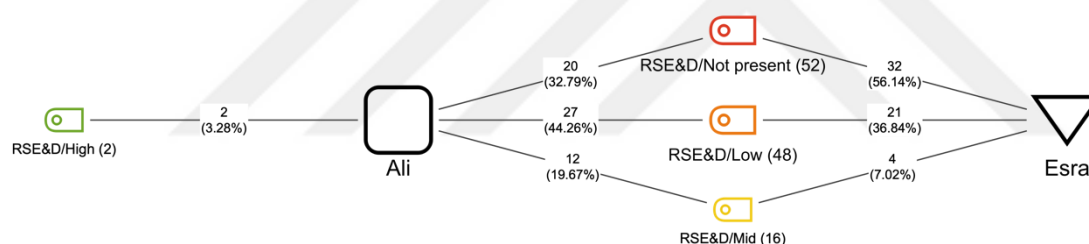


Figure 4.85 “Remediation of Student Errors and Difficulties” Code in Both Teachers’ Statistics Instruction

Note: RSE&D- Remediation of Student Errors and Difficulties

They remediated students’ errors or difficulties during doing calculations in circle graph tasks or calculating measures of central tendency at a brief or moderate level (44.26% for Ali and 36.84% for Esra) procedurally. In other segments, teachers conducted extensive procedural remediation during calculations of similar concepts or moderate conceptual remediation on the concept of measures of central tendency (19.67% for Ali and 7.02% for Esra). None of the segments included systematic and at-length remediation in Esra’s instruction. However, Ali’s instruction included such segments (3.28%), which covered the teacher’s actions for foreseeing possible

students' misunderstandings or difficulties in a line graph and asking students to reveal the misunderstandings for extensive conceptual remediation.

In summary, Ali's instruction had more practice in remediating students' difficulties than Esra's. Their instruction mainly differed in terms of the amount of extensive procedural remediation or moderate conceptual remediation. However, the amount of brief conceptual or procedural remediation was similar in both instructions.

4.2.3.3.2. Teacher Uses Student Mathematical and Statistical Contributions

Figure 4.86 presents a comparison of scores for the “Teacher Uses Student Mathematical and Statistical Contributions” code assigned to lesson segments in the statistics instruction of both teachers.

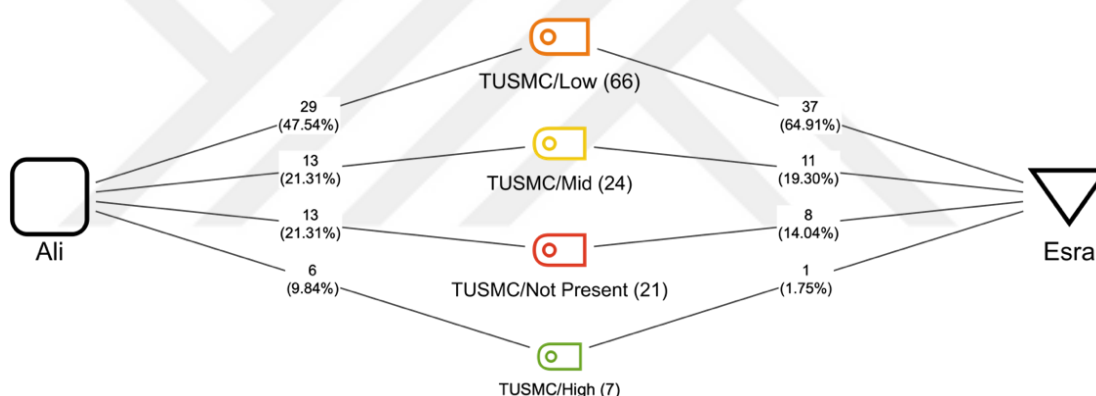


Figure 4.86 “Teacher Uses Student Mathematical and Statistical Contributions”
Code in Both Teachers’ Statistics Instruction

Note: TUSMC- Teacher Uses Student Mathematical and Statistical Contributions

In some lesson segments, the students worked individually, so no contribution was present. In other segments, both teachers did not allow students to contribute to the mathematics and statistics of the segment. Even if the segments had student contributions, teacher talk generally dominated the segments (21.31% for Ali, 14.04% for Esra). There were some segments with students’ contributions present, but teachers’ take-up of those contributions was limited, like one or two-word answers.

Those incidents were more present in Esra's instruction (64.91%) than in Ali's (47.54%) instruction.

Both teachers engaged in students' contribution to developing mathematics and statistics of the segment (21.31% for Ali, 19.30% for Esra). However, those practices did not include a specific call on the student's ideas; the teachers did not acknowledge the students' ideas at length, and at some point, they took ownership of the ideas to develop mathematics and statistics of the segment as the authority.

In one lesson segment (1.75%), Esra used student contributions and moved the instruction forward according to those contributions. In Ali's instruction, 9.84% of lesson segments included high-quality practices regarding TUSMC. Ali shaped the content with the help of students' contributions and gave credit to those students' ideas. Esra's instruction used students' contributions while reading graphs in general, and in some cases, Esra used students' errors to move instruction forward. In addition, Ali used students' contributions while introducing a concept; he shaped the instruction with students' ideas or prior knowledge. Overall, even if the nature of TUSCM was different for both teachers' instruction, the amount of their use of student contributions was similar.

4.2.3.3.3. Overall Working with Students and Mathematics and Statistics

Figure 4.87 presents a comparison of scores for the "Overall Working with Students and Mathematics and Statistics" code assigned to lesson segments in the statistics instruction of both teachers. While Ali's instruction did not include his use of student contribution or remediation at 1.64% of lesson segments, the practices were not present in 8.77% of lesson segments in Esra's instruction. While 70.49% of lesson segments in Ali's instruction included brief remediation and limited use of student contributions, Esra's instruction was also low in 78.95% of lesson segments.

However, Ali's instruction (26.23%) consisted of more Mid scores using student contributions or substantive remediation than Esra's instruction (10.53%). Both teachers used all codes more than briefly in those segments. Both teachers used student

contributions to shape the instruction in only one lesson segment and conceptually remediated the students' misconceptions (1.64% for Ali, 1.75% for Esra).

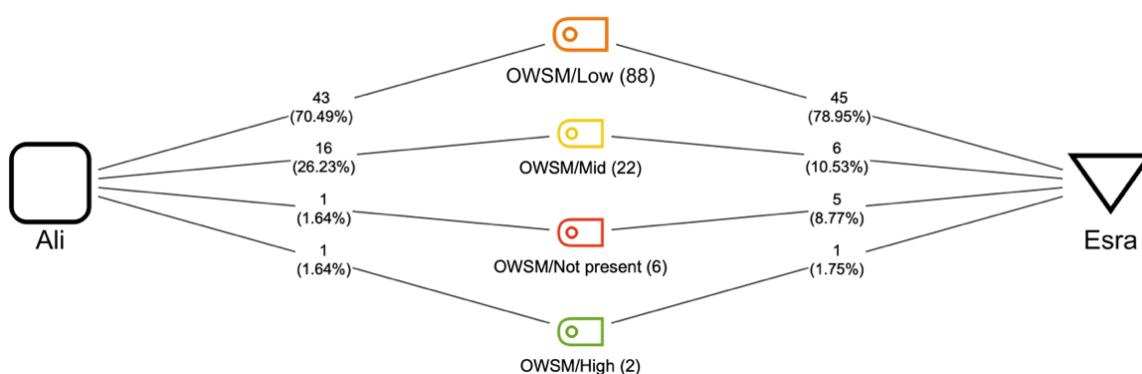


Figure 4.87 “Overall Working with Students and Mathematics and Statistics” Code in Both Teachers’ Statistics Instruction

Note: OWSM- Overall Working with Students and Mathematics and Statistics

4.2.3.4. Errors and Imprecision

This section compares both teachers regarding errors or imprecision in content, language, and lack of clarity. Due to reverse scoring, three codes were scored Not Present to High scores (4 to 1). One code is the overall number of errors called Overall Errors and Imprecision (OE&I). The other codes are called Mathematical and Statistical Content Errors (MCE), Imprecision in Language or Notation (ILN), and Lack of Clarity in the Presentation of Mathematical and Statistical Content (LCPM).

4.2.3.4.1. Mathematical and Statistical Content Errors

Figure 4.88 presents a comparison of scores for the “Mathematical and Statistical Content Errors” code assigned to lesson segments in the statistics instruction of both teachers. Both teachers made no errors over 80% of their lesson segments (See Figure 4.88).

They made brief content errors while forgetting a critical condition of the definition of the mode. They both mentioned mode as the most frequent number instead of value (11.48% for Ali and 8.77% for Esra). Both teachers’ instruction included mid-scored

lesson segments in terms of MCE (4.92% for Ali and 1.75% for Esra). Content errors were present in some portions of the segment. When the errors were present during the whole segment, a high score was assigned (3.28% for Ali and 5.26% for Esra).

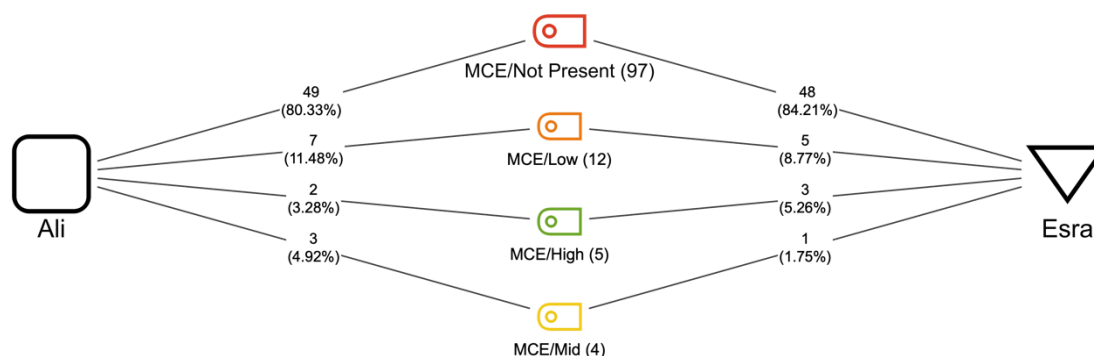


Figure 4.88 “Mathematical and Statistical Content Errors” Code in Both Teachers’ Statistics Instruction

Note: MCE- Mathematical and Statistical Content Errors

The content errors stemmed from the inconsistencies between the purpose of graphs and the context of constructing a graph. For example, both teachers directly used line graph tasks from the textbook. The context was about the number of correct answers during mock exams over time. However, the mock exams were considered time, but they were ordinal. Furthermore, both teachers had content errors while representing a data set to show the mode. For example, both made major errors while representing uniform data with four 4s. All in all, their instructions were similar regarding the amount and the nature of content errors.

4.2.3.4.2. Imprecision in Language or Notation

Figure 4.89 presents a comparison of scores for the “Imprecision in Language or Notation” code assigned to lesson segments in the statistics instruction of both teachers. Both teachers did not have imprecision in language or notation in over 85% of the lesson segments (88.52% for Ali and 85.96% for Esra).

Ali’s statistics instruction (9.84%) and Esra’s statistics instruction (12.28%) had brief instances of imprecision while talking about “arithmetic mean” and “average.” In

Turkish, average and mean correspond to the same word, “ortalama.” Even if both teachers defined average as three central tendency measures, they sometimes said average instead of the arithmetic mean. Those segments did not distort the mathematics and statistics of the segments.

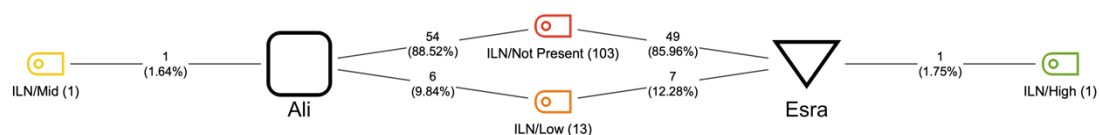


Figure 4.89 “Imprecision in Language or Notation” Code in Both Teachers’ Statistics Instruction

Note: ILN- Imprecision in Language or Notation

In one segment (1.64%), Ali used the term “average” instead of “the arithmetic mean”; the segment was confusing for students because it came just after defining average as three measures of central tendency, and the imprecision was present in some portions of the segment. None of the lesson segments were scored Mid for ILN in Esra’s instruction. However, one segment (1.75%) consisted of an imprecision that obscured the mathematics and statistics of the segment. This segment was related to representing the mode, the teacher.

4.2.3.4.3. Lack of Clarity in Presentation of Mathematical and Statistical Content

Figure 4.90 presents a comparison of scores for the “Lack of Clarity in Presentation of Mathematical and Statistical Content” code assigned to lesson segments in the statistics instruction of both teachers. Both teachers had explicit instruction in over 80% of the lesson segments (80.33% for Ali and 82.64% for Esra). 13.11% of Ali’s lesson segments and 12.28% of Esra’s lesson segments lacked clarity briefly without affecting mathematical and statistical points. In addition, a lack of clarity obscured the mathematics and statistics of some parts of the lesson segments (3.28% for Ali and 3.51% for Esra) when the teachers discussed choosing appropriate measures of central

tendency. The teachers could not give clear reasons to students for choosing the mode for the given context instead of the other two measures.

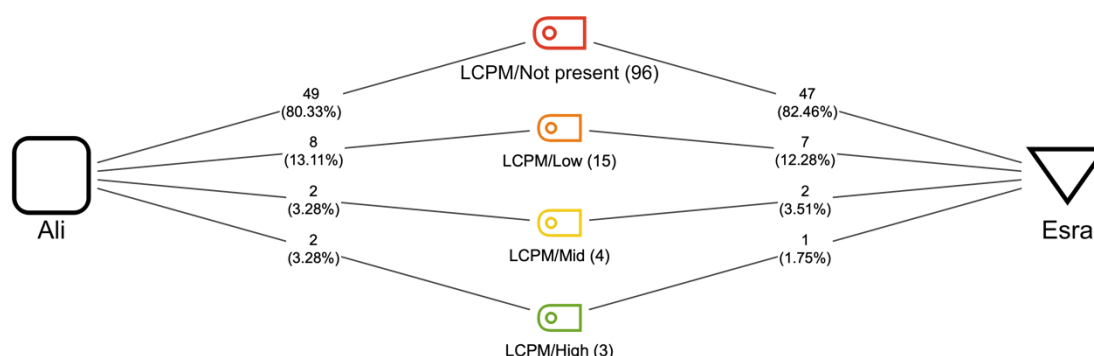


Figure 4.90 “Lack of Clarity in Presentation of Mathematical and Statistical Content” Code in Both Teachers’ Statistics Instruction

Note: LCPM- Lack of Clarity in Presentation of Mathematical and Statistical Content

Lastly, one segment (175%) in Esra’s instruction lacked clarity during all segments. Representing the mode of a data set example was confusing for students. In Ali’s instruction, the “transforming the circle graph into a line graph” example took two segments (3.51%) and was confusing for students due to insufficient information for drawing a line graph. Overall, both teachers’ instructions were similar in LCPM, and their instruction was explicit in general. Furthermore, segments that lacked clarity were generally related to teachers’ incorrect mathematical and statistical practices (e.g., the concept of mode and line graph).

4.2.3.4.4. Overall Errors and Imprecision

Figure 4.91 presents a comparison of scores for the “Overall Errors and Imprecision” code assigned to lesson segments in the statistics instruction of both teachers. The results are consistent with the other codes in the dimension. Most lesson segments in Ali’s instruction (59.02%) and Esra’s (66.67%) did not include errors or imprecision.

While 27.87% of lesson segments in Ali’s instruction included temporary and minor errors, Esra made such errors in 22.81%. In addition, some lesson segments at Ali’s (8.20%) and Esra’s instruction (5.26%) consisted of one or more errors. Lastly, both

teachers had many or one significant errors in some segments obscuring the mathematics and statistics (4.91% for Ali, 5.26% for Esra).

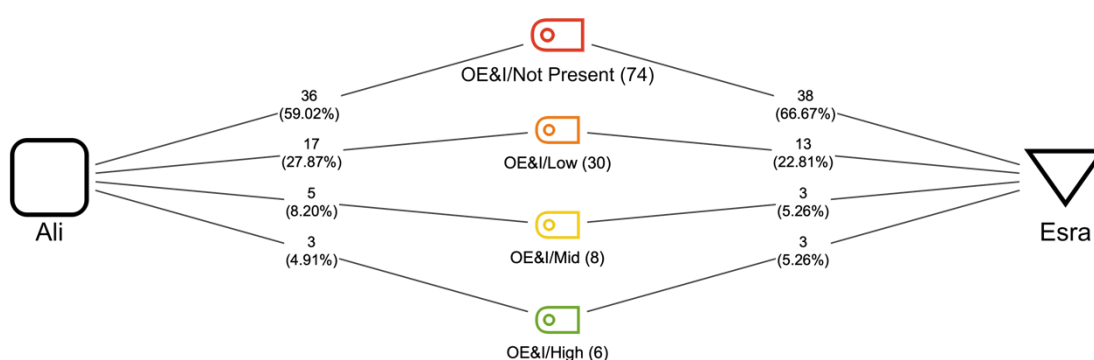


Figure 4.91 “Overall Errors and Imprecision” Code in Both Teachers’ Statistics

Instruction

Note: OE&I- Overall Errors and Imprecision

4.2.3.5. Student Participation in Meaning-Making and Reasoning

This section compares teachers’ instruction regarding students’ participation in meaning-making and reasoning processes (Hill et al., 2014). There are three codes with the Not Present to High (1 to 4). One code is about overall student participation. The other codes are called Students Provide Explanations (SPE), Student Mathematical and Statistical Questioning and Reasoning (SMQR), Students Communicate about the Mathematics and Statistics of the Segment (SCM), Task Cognitive Demand (TCD), and Students Work with Contextualized Problems (SWCP).

4.2.3.5.1. Students Provide Explanation

Figure 4.92 presents a comparison of scores for the “Students Provide Explanation” code assigned to lesson segments in the statistics instruction of both teachers. Both instructions did not include students’ explanations over half of the lesson segments (60.66% for Ali, 77.19% for Esra). Although those segments included student contributions while students were reading graphs or doing calculations, those contributions were generally one-or-two-word answers without justifications.

Furthermore, both teachers' instructions included one or two brief student explanations (18.03% for Ali and 17.54% for Esra), while more frequent student explanations were present in 18.03% of Ali's lesson segments and 1.75% of Esra's lesson segments. Therefore, Ali's instruction consisted of more Mid-scored lesson segments than Esra's. Lastly, student explanations have characterized the segments in 3.3% of Ali's and 3.51% of Esra's lesson segments.

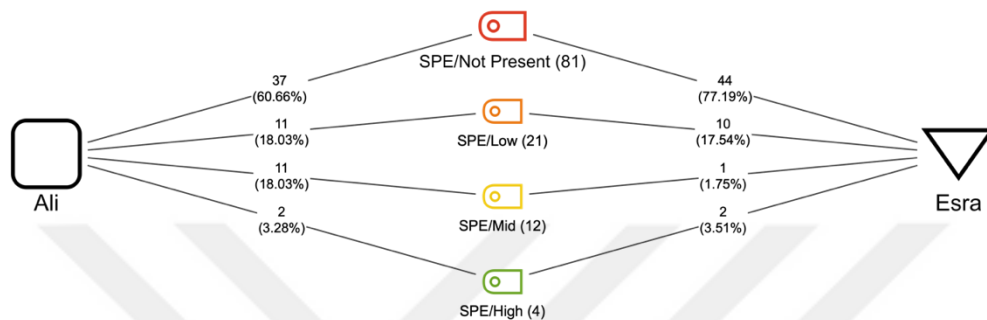


Figure 4.92 “Students Provide Explanation” Code in Both Teachers’ Statistics

Instruction

Note: SPE- Students Provide Explanation

Overall, Ali's instruction consisted of more student explanations than Esra's. However, both instructions were similar regarding the nature of students' explanations in most cases. When students briefly provided explanations, the explanations were generally about procedures conducted during circle graph tasks. When the explanations were more than briefly present, they were mostly about why to choose a specific graph and central tendency measures.

4.2.3.5.2. Student Mathematical and Statistical Questioning and Reasoning

Figure 4.93 presents a comparison of scores for the “Student Mathematical and Statistical Questioning and Reasoning” code assigned to lesson segments in the statistics instruction of both teachers. Most lesson segments in Ali's (81.97%) and Esra's (87.72%) lacked SMQR. 12.28% of lesson segments in Esra's instruction were low-quality in terms of SMQR because they consisted of brief questioning or reasoning while doing calculations about a circle graph, drawing a circle graph, or finding the median of a data set. 16.39% of Ali's lesson segments were also low

regarding SMQR because there were some brief questioning and reasoning on the context in the graphs or finding the median or mode of a data set.

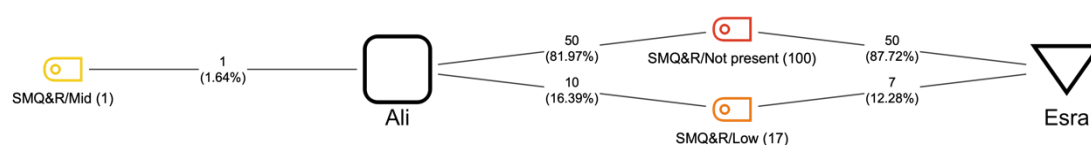


Figure 4.93 “Student Mathematical and Statistical Questioning and Reasoning” Code in Both Teachers’ Statistics Instruction

Note: SMQ&R- Student Mathematical and Statistical Questioning and Reasoning

In addition, Esra’s instruction had no lesson segments with a Mid score, and only one segment (1.64%) in Ali’s instruction had more frequent student mathematical and statistical questioning and reasoning while finding the mode of a data set that contained the same values. None of the lesson segments of both teachers’ instruction consisted of student mathematical and statistical questioning or reasoning as the segment’s focus. Therefore, both instructions were similar regarding SMQR; both lacked SMQR. When SMQR was present, their quality was low except for one segment in Ali’s instruction.

4.2.3.5.3. Students Communicate about the Mathematics and Statistics of the Segment

Figure 4.94 presents a comparison of scores for the “Students Communicate about the Mathematics and Statistics of the Segment” code assigned to lesson segments in the statistics instruction of both teachers. Student contributions were not observed in Ali’s instruction in 21.31% of lesson segments due to the time spent on the individual work and the amount of teacher talk, while only 8.77% of lesson segments in Esra’s instruction time did not consist of SCM.

Brief student contributions like one or two-word answers to the questions were present in 19.67% of Ali’s instruction and 26.32% of Esra’s instruction. These contributions were students’ brief answers to teachers’ answer-bound questions mostly. The students contributed more than briefly, but those contributions did not characterize the segment

because of teachers' involvement in those contributions (44.26% for Ali and 54.39% for Esra).

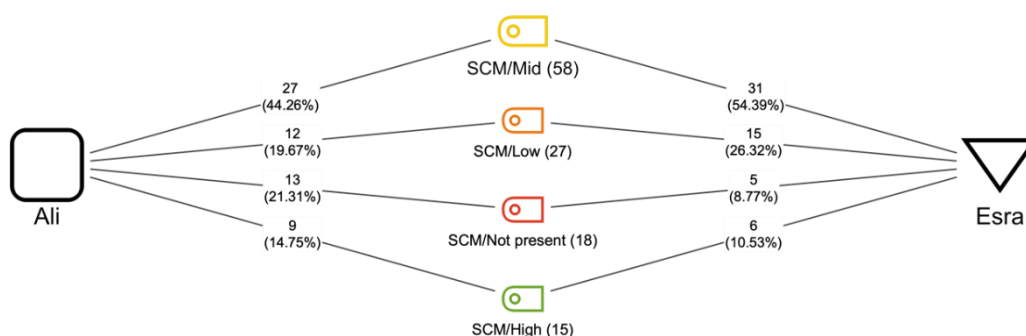


Figure 4.94 “Students Communicate about the Mathematics and Statistics of the Segment” Code in Both Teachers’ Statistics Instruction

Note: SCM- Students Communicate about Mathematics and Statistics of the Segment

Student contributions were dominant in 14.5% of Ali’s lesson segments and 10.5% of Esra’s. The segments were also assigned a High score for SPE or SMQR. When students worked on tasks on board with or without descriptions or explanations, the students’ contributions were at least Mid in both instructions. Overall, both instructions were similar in terms of SCM.

4.2.3.5.4. Task Cognitive Demand

Figure 4.95 presents a comparison of scores for the “Task Cognitive Demand” code assigned to lesson segments in the statistics instruction of both teachers. Ali’s instruction did not include TCD in 44.27% of lesson segments, while Esra’s instruction lacked TCD in 73.68%. Primarily Esra’s instruction consisted of less student input in many segments and conducting more well-established procedures related to circle graphs and central tendency measures.

Some student reasoning and questioning were present even if the tasks were routine or teacher-directed in 39.34% of Ali’s and 15.79% of Esra’s lesson segments. Furthermore, Ali’s (16.39%) and Esra’s (10.53%) instructions consisted of cognitively

demanding tasks, encouraging students to reason or question the mathematical and statistical idea behind the task.

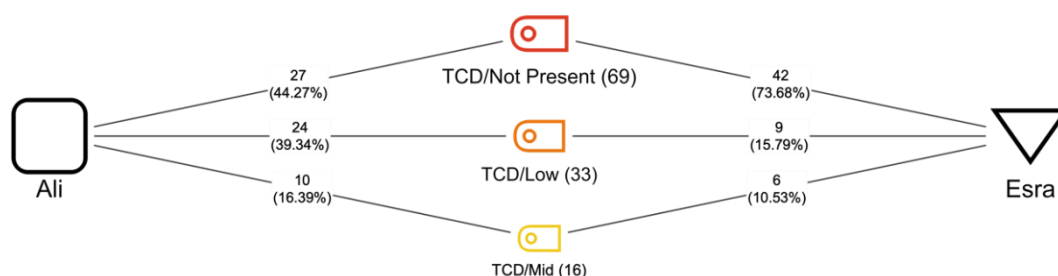


Figure 4.95 “Task Cognitive Demand” Code in Both Teachers’ Statistics Instruction

Note: TCD- Task Cognitive Demand

These tasks were mainly present while both teachers presented new statistical concepts to the students. In addition, average tasks with contexts also opened a room for discussing students’ learning on three measures of central tendency in both instructions. However, the teachers’ intervention was also present to lower the demanding aspect of the task. None of the lesson segments included a high-level TCD in both teachers’ instruction. As a result, Ali’s instruction had more practices related to TCD, but their practices' nature was similar when TCD was present in the segments.

4.2.3.5.5. Students Work with Contextualized Problems

Figure 4.96 presents a comparison of scores for the “Students Work with Contextualized Problems” code assigned to lesson segments in the statistics instruction of both teachers. Some lesson segments did not include contextualized problems in both teachers’ instruction (14.75% for Ali and 19.30% for Esra). When contextualized problems were present, they were worked on as routine exercises in 47.54% of Ali’s and 52.63% of Esra’s lesson segments. In addition, students’ reasonings were present in contextualized problems with some teacher directions (37.70% for Ali, 28.1% for Esra). However, none of the segments included contextualized problems reasoned by students with full involvement without teachers’ directions or input.

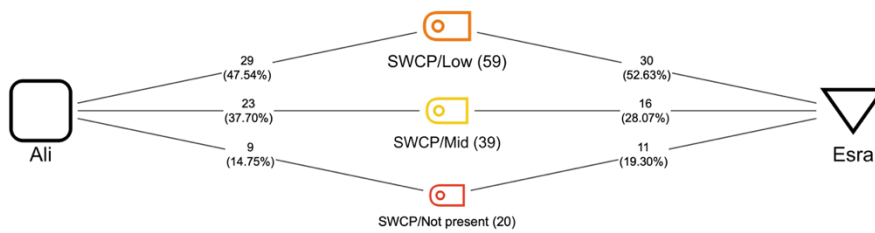


Figure 4.96 “Students Work with Contextualized Problems” Code in Both Teachers’ Statistics Instruction

Note: SWCP- Students Work with Contextualized Problems

All in all, both instructions were similar in terms of SWCP because both teachers used students’ reasoning at some level while working on contextualized problems (e.g., graph tasks). Also, both instructions consisted of average tasks without context, just calculations.

4.2.3.5.6. Overall Student Participation in Meaning-Making and Reasoning

Figure 4.97 presents a comparison of scores for the “Overall Student Participation in Meaning-Making and Reasoning” code assigned to lesson segments in the statistics instruction of both teachers.

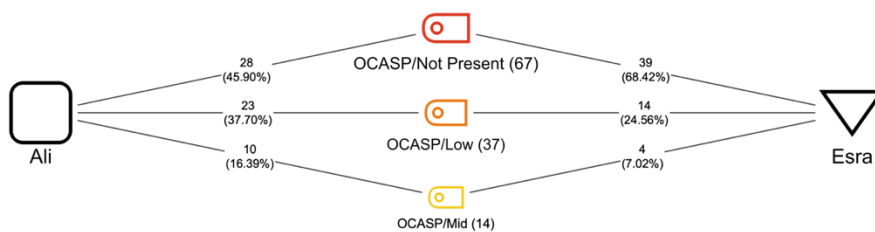


Figure 4.97 “Overall Student Participation in Meaning-Making and Reasoning” Code in Both Teachers’ Statistics Instruction

Note: OCASP- Overall Student Participation in Meaning-Making and Reasoning

Students' participation in cognitively demanding activities was limited, or the activities they participated in were heavily guided by the teacher or procedural in nature in a significant number of lesson segments for both Ali (45.90%) and Esra (68.42%). In

terms of student engagement in meaning-making and reasoning, this was brief in 37.70% of Ali's lesson segments and 24.56% of Esra's.

Furthermore, student engagement in mathematical and statistical activities included both high and low-quality indicators regarding their overall participation in meaning-making and reasoning (16.39% for Ali and 7.02% for Esra). Overall, Ali's instruction had more instances of student participation practices than Esra's, but neither teacher had any lesson segments with high-quality overall student participation.

4.2.4. Mathematical and Statistical Quality of Instruction: Lesson Level Analysis

This section compares the quality of Ali's and Esra's instructions based on a whole-lesson code that describes overall instructional quality. The unit of analysis for this section is the lesson, with all lessons assigned scores on a 5-point scale in terms of whole-lesson quality indicators. Ali's instruction included 14 lessons, while Esra's consisted of 12. The whole-lesson mathematical and statistical quality of Ali's instruction can be characterized by Mid quality (50%) and Mid/High quality (50%), while Esra's instruction can be characterized by Mid quality (91.66%). Mid-quality lessons were generally error-free, but they had some problematic aspects in mathematical and statistical richness. Those lessons were typical without special features. Mid/High-quality lessons consisted of better instructional practices than Mid quality lessons with richer mathematics and statistics, more substantive student contribution and participation, and responsive teacher-student interaction. Ali's instruction consisted of more such practices than Esra's instruction.

CHAPTER 5

DISCUSSION AND CONCLUSION

This section aims to provide a comprehensive discussion of the findings of the study regarding statistics teaching practices in mathematics classrooms. This study focused on statistical teaching practices, including the statistical content covered in mathematics instruction, how teachers approached the four-step model of statistical problem-solving processes, and the quality of statistics instruction within their broader mathematics instruction.

First, I begin by reflecting on the findings related to the first research question, which focused on the statistical practices employed by mathematics teachers in the classroom. I analyzed the statistical content covered in their instruction and described how teachers approached the four-step model of statistical problem-solving processes.

Next, I discuss the findings related to the second research question, which explored the quality of mathematics teachers' statistics instruction within their broader mathematics instruction. I examined how classroom instruction was connected to mathematics and statistics, explored the richness of mathematics and statistics, described how teachers worked with students to promote their learning of mathematics and statistics and described student participation in meaning-making and reasoning.

In the end, I will discuss the implications of the findings for mathematics and statistics education research, the limitations of the study, and suggest recommendations for further research.

5.1. Statistical Teaching Practices

In this section, I will discuss the findings related to discipline-related practices. While considering the discipline-related practices, first, I examined the content teachers covered and assessed if the content aligned with the curricular objectives or requirements. I also explored the connection between the covered content and the big ideas in statistics.

5.1.1. Statistical Content

Curriculum and curricular materials are crucial in shaping what teachers do in the classrooms (Batur et al., 2021; Litke & Hill, 2020). In Turkish public schools, teachers are required to cover all the objectives outlined in the curriculum. However, there is a need to investigate teachers' understanding of these curricular objectives and the extent to which they cover the specific objectives. According to Batur et al. (2021), the statistics objectives covered in the Turkish curriculum are comparatively limited compared to other countries like New Zealand and the United States. The researchers found that the depth of the curricular objectives in statistics was shallower in Turkey. Ader's (2015, 2018) analysis of the Turkish national curricula demonstrated an improvement in the inclusion of statistics objectives over the years, with an increase in the number of objectives and curricular approaches. However, as Batur et al. (2021) noted, there is still room for further development and refinement of these objectives.

This study's findings indicate that teachers generally followed the objectives and covered all of them, although the weight given to each objective varied (e.g., more time spent on circle graphs than line graphs). One possible reason for this discrepancy is the weight assigned to specific concepts in the additional textbooks teachers used. Additionally, circle graphs involved more mathematical calculations and provided a context for students to learn about circles and proportionality. Many tasks brought into the classroom focused on calculations, such as finding a sector of a circle.

Furthermore, the study revealed that teachers allocated less time to the objective of graph transformation. One possible reason for this is that it was the last objective in

the curriculum, covered after the average objectives. As a result, the connection between individual objectives related to graphs and graph transformation diminished due to shifts in focus. Another reason identified by the teachers was their difficulty in understanding the purpose of the objective. They struggled to comprehend the strengths and weaknesses of different graphs within specific contexts. Despite these challenges, the results show that teachers covered all the objectives. As Watson et al. (2018) previously stated, primary and middle school students learned measures of centers as part of descriptive statistics. They were able to construct and interpret limited types of graphs.

As mentioned earlier, both teachers followed the same additional textbook, which differed from the textbook provided by the Ministry of Education to each student free of charge. The additional textbook aligned its structure with the objectives in the curriculum, except for the final objective concerning graph transformation. For that objective, both teachers used examples from official textbooks or other resources to address the topic. In addition, the literature emphasizes that teachers' interpretation and implementation of objectives are strongly influenced by the content presented in textbooks (Balci, 2023; Remillard & Heck, 2014; Weiland, 2019; Yanik et al., 2017). Weiland (2019) asserted that “in the case of statistics, where teachers often do not have much prior experience with the subject, textbooks may be highly influential in shaping the enacted curriculum of the classroom” (p. 18). In the current study, both teachers acknowledged that they did not receive any specific training on how to teach statistics during their pre-service education or any professional development regarding statistics during their in-service education. Consequently, they relied on textbooks and their own experiences to guide their teaching of statistics.

Despite teachers covering all the objectives outlined in the curriculum, there was a limited focus on the big ideas in statistics, leading to a lack of conceptual understanding of statistics (Garfield & Ben-Zvi, 2004). For instance, most teaching practices primarily revolved around the center concept, precisely the three measures of central tendency. Center, which is considered a significant big idea in statistics according to the literature (Shaughnessy, 2007, 2019a; Watson et al., 2007), was mainly approached through calculations and procedures to find an average (Jacobbe

& Carvalho, 2011). This emphasis on procedural rather than conceptual understanding has been discussed in the literature (Shaughnessy, 2007).

The literature presents various interpretations of the average concept (Konold & Pollatsek, 2002; Mokros & Russell, 1995). The findings of this study indicate that the teaching practices predominantly viewed average as an algorithm (Mokros & Russell, 1995) or a typical value (Konold & Pollatsek, 2002). Both teachers conceptualized the average as an algorithm and a typical value reflected in their instructional practices. These findings align with research on teachers' conceptualization of mean and median as average concepts (e.g., Jacobbe, 2012). These findings are also consistent with Bütüner's (2020) research on the concept of average as covered in Turkish mathematics textbooks. The textbook example questions primarily focused on procedural aspects, emphasizing the add-and-divide algorithm. Moreover, the concept of average in these textbooks emphasized finding a single value rather than representativeness. Consequently, the teachers' conceptualization of average and the content of the textbooks are potential factors shaping the teaching practices observed in this study.

Furthermore, the seventh-grade objectives in the curriculum do not explicitly include measures of spread. However, in the sixth-grade curriculum, students are introduced to the arithmetic mean as an expectation big idea and the concept of range as a variability big idea. Esra's class did not cover the statistics objectives in sixth grade based on students' reports. Instead, Esra introduced statistics by defining what statistics is and engaging students in statistical investigations. However, Esra did not introduce the concept of range, which was one of the missing objectives from sixth grade. On the other hand, Ali's students reported that they had covered statistics in previous years. However, Ali did not incorporate the concept of range, for instance, when comparing two groups. As Garfield and Ben-Zvi (2004) pointed out, "After one topic has been studied, there is little mention of it again, and students fail to see how the big ideas are actually providing a foundation for course content" (p.401).

Research highlights the importance of discussing the average along with the variability around the center (Garfield & Ben-Zvi, 2008). Konold and Pollatsek (2002) describe

the “signal in noise” interpretation, which involves considering both concepts together to comprehend the stable feature within a variable system. Variation is a distinguishing feature of statistics compared to mathematics (Watson, 2006). The developmental progression model of variation proposed by Watson et al. (2007) emphasizes establishing the connection between variation and expectation, particularly in comparative contexts. However, in this study, although some teaching practices involved comparing two samples or classes in statistical tasks, the discussion mainly revolved around determining which average to use; specifically, the arithmetic mean or median, in the presence of an outlier value. The link between expectation and variation was not explicitly addressed, nor was the concept of the range covered in previous years utilized.

The curriculum in this study does not explicitly include big ideas other than expectation and variation. However, Ali briefly discussed the concept of “random” when addressing the concept of a sample and its representativeness. Nevertheless, this concept was not thoroughly introduced or emphasized as a big idea. In certain instances, teaching practices implicitly focused on the concept of distribution when comparing two sets of data, using circle graphs to describe data distribution, or discussing the choice of center in an outlier's presence. Including different types of graphs in the curriculum might serve as an opportunity to introduce the big idea of distribution.

Frischemeier et al. (2022) highlighted that the Turkish primary mathematics curriculum does not explicitly incorporate the distribution concept, despite its focus on data representations. They also emphasized that the German and Irish curricula include data representations and big ideas of variation and distribution. However, in the middle school mathematics curriculum, there is an objective in sixth grade that expects students to use the arithmetic mean and range to compare two groups. Another objective expects students to form statistical investigative questions to compare two groups and create tables and bar graphs for comparing two data sets (MoNE, 2018). Nevertheless, Batur et al. (2021) also noted that the Turkish K-12 mathematics curriculum lacks explicit inclusion of the distribution concept, despite a limited range of graphs. Reflecting on this finding, the presentation of big ideas in the classroom

may be influenced by a curriculum that does not explicitly address these concepts. In addition, the choice of curricular materials used by the teachers also impacts their approach to covering statistical content. Therefore, the textbooks employed in their teaching practices may have limitations in fostering a conceptual understanding of statistical content, often focusing primarily on procedural aspects (e.g., Bütüner, 2020; Yanık et al., 2017).

5.1.2. Statistical Problem-Solving Process

The statistical problem-solving process is considered a fundamental aspect of teaching and learning statistics (Wild & Pfannkuch, 1999), and it is a key objective of statistics curricula worldwide (Arnold & Franklin, 2021; Gal & Garfield, 1997; Leavy, 2015). Recent research has emphasized various elements of statistical investigations and the cyclical nature of the process (Burgess, 2011). One purpose of this study is to describe how middle school teachers approached the four-step model of the statistical problem-solving process. To achieve this aim, the teaching practices were analyzed to identify which specific elements the teachers focused on within the process when they were not addressing the process as a whole (e.g., Litke & Hill, 2020). In addition, the teaching practices were examined to determine whether the teachers covered the entire cycle of the statistical problem-solving process.

The results of the study indicate that statistics teaching practices generally did not include the posing of statistically investigative questions, except for limited instances observed in each teacher's practices. Research on this particular element is relatively new, leading to limited studies on posing statistically investigative questions (Frischemier & Leavy, 2020). As mentioned earlier, the seventh-grade curriculum does not explicitly address this element. However, the fifth and sixth-grade curricula include objectives related to forming investigative questions. In Esra's class, since the sixth-grade statistics unit was not covered in previous years, she initiated the statistics unit by asking the class about their favorite color. This was followed by data collection from the class, creating a table and a circle graph, and interpreting the graphs. As observed, starting with an investigative question in Esra's class facilitated the completion of the entire statistical investigation cycle.

Similarly, in Ali's class, two episodes involved the teacher posing somewhat statistically investigative questions, but they did not lead to data collection. Instead, the teacher created imaginary data rather than utilizing primary or secondary real data. As Watson et al. (2018) noted, one possible reason why teachers did not emphasize posing statistically investigative questions in their practices is the perception that it may be challenging for students or that it may require additional time, which could be allocated to calculations in data analysis, aligning with the objectives in the seventh-grade curriculum.

In Ali's class, one episode involved student-generated statistical questions. However, these questions did not lead to investigations because the students' questions were more focused on data collection rather than being investigative. Arnold and Franklin (2021) noted that statistical questions should encompass interrogative, investigative, data collection, and analysis components throughout the entire process. However, the results of this study align with Arnold's (2008) study, where pre-test results show that the questions formed by students did not lead to statistical investigations but instead focused on data analysis prior to the teaching experiment. Allmond and Katar (2010) found that students not exposed to statistical questions or inquiry-based teaching struggled to formulate meaningful statistical questions that led to investigations. However, these studies also demonstrate that specific interventions designed to enhance the formation of statistically investigative questions can improve students' abilities in this regard. The results of this study suggest that similar interventions could be implemented for middle school students to enhance their ability to pose such questions. Furthermore, teaching practices related to this element could be improved by designing interventions for teachers or pre-service teachers, similar to the study conducted by Leavy and Frishemeier (2022), which involved feedback from peers and experts.

Regarding the data collection element, the results indicate that neither teacher extensively employed this element in classroom practices, except for one episode each. Bargagliotti et al. (2020) emphasized that statistical investigative questions should be answered through data collection, either from primary data or by considering data from secondary sources. This element also encompasses the planning phase of data

collection. However, data collection is often perceived as time-consuming and unpredictable (Ben-Zvi et al., 2018). One possible reason for the limited inclusion of data collection in teaching practices is that the learning outcomes in the seventh-grade curriculum are primarily focused on data analysis and interpretation. Consequently, due to time constraints, the time allocated for data analysis and interpretation may be prioritized over data collection (Ben-Zvi et al., 2018). Hedges et al. (2022) also noted that the data collection element includes an active data collection period and discussions on study design and related concepts such as sampling methods. Thus, teachers require sufficient time to adequately cover the data collection element (Hedges et al., 2022). In the current study, teachers expressed in their final interviews that they would have devoted more time to statistical investigations, particularly the data collection phase if they had more time dedicated to the subject matter.

The findings of the study indicate that the teachers primarily focused on the data analysis element of the statistical problem-solving process without giving equal attention to the entire cycle. As noted by Watson et al. (2018), the instruction mainly revolved around constructing graphs and calculating measures of central tendency. Furthermore, research in statistics education has often emphasized data analysis with “well-defined problems” (Delmas, 2004; Makar & Fielding-Wells, 2011), which aligns with the findings of this study. However, according to NCTM (2009), even data analysis, when approached as calculations and procedures, encompasses all four components of the statistical investigations because analyzing data involves considering the data required to answer statistical questions and lead to interpretations. Nevertheless, well-defined problems, which resemble mathematical problem-solving rather than statistical problem-solving, can distort the perspective on data analysis as encompassing all components of the statistical problem-solving process. As Shaughnessy (2007) noted, problem-solving in statistics differs from problem-solving in mathematics because statistical problems arise from real-life contexts, resulting in poorly defined problems. Additionally, Batur et al. (2019) mentioned that Turkish mathematics textbooks did not incorporate actual data adequately in the statistics unit, despite including some example questions with real-life contexts.

The data analysis element was predominantly observed at Level 2 in the current study. This level of engagement involved students in statistical tasks that included constructing graphs, calculating measures of center, and conducting descriptive analysis based on the graphs. Also, there were some instances at Level 3 where discussions on distribution occurred, mainly when an outlier was present. These discussions involved deciding which measure of center to use or comparing groups based on a measure of center or within graph distributions. These findings align with the expectations outlined in Level A of the GAISE document (Bargagliotti et al., 2020). It is worth noting that the protocol used in the study was developed for the primary years in the context of the United States, primarily referencing Level A learning outcomes. Batur et al. (2021) mentioned that the Turkish K-12 mathematics curriculum predominantly consists of Level A learning outcomes. As a result, the teaching practices observed in relation to the data analysis element were aligned with the curricular documents.

The final element in the statistical problem-solving process is the interpretation of the results, which involves making conclusions based on the analyzed data. However, the findings of this study revealed that most teaching practices of both teachers did not include the interpretation of results; the focus was primarily on the data analysis section, including calculations, procedures, and graph construction. Nevertheless, in episodes where the interpretation element was present, students primarily engaged in reading the graphs to answer questions related to the data itself. A review conducted by Friel et al. (2001) on the taxonomy of skills required to answer graph-related questions indicates that the episodes in this study generally aligned with the elementary level, where information is directly read from the graphs. In a few limited instances, the teaching practices involved interpreting the results of relationships within the data and comparing two data sets while considering the variability in the data, particularly by looking at outliers. Esra's teaching practices included more episodes with an interpretation element than Ali's practices. These differences might stem from Esra's perspectives on teaching and learning statistics. She emphasized the importance of making sense of data through interpretations and highlighted that data, once collected and analyzed, lacks meaning without proper interpretation.

The findings of this study indicate that the teaching practices observed did not encompass the entire statistical problem-solving process, from formulating investigative questions to collecting, analyzing, and interpreting data, except for one episode. These results are consistent with the findings of Litke and Hill (2020), who found that statistics instruction in real classrooms primarily focuses on calculations and graphs rather than the complete cycle of statistical investigation despite the increased emphasis on this issue in research and curricular documents. One possible reason for this discrepancy is that while the curriculum emphasizes statistical investigations, it explicitly states which elements of statistical investigation should be emphasized in specific grades separately (MoNE, 2018). Furthermore, the Turkish K-8 mathematics curriculum includes a values education section that addresses scientificity and underscores the importance of the statistical problem-solving process, highlighting relevant objectives related to statistics (MoNE, 2018). However, despite the presence of a focus on the complete cycle of statistical investigation throughout the K-8 curriculum, it is limited or underdeveloped within each year. One issue in statistics education is that teachers often fail to consider the cyclical nature of the curriculum, neglecting to incorporate elements of statistical investigation and concepts from previous years, such as forming investigative questions or the range concept from sixth grade. Although there has been an increase in the focus on statistical content and statistical investigations in the Turkish curriculum (Ader, 2016, 2018), as noted by Makar and Fielding-Wells (2011), teaching practices in school mathematics still predominantly revolve around calculations and graph constructions related to average concepts. While the adoption of statistical investigations is present in curricula worldwide, including those of Australia, New Zealand, US based on curricular documents (e.g., GAISE), there is limited evidence on how these approaches are implemented in real classrooms (Litke & Hill, 2020).

This study focused on seventh-grade classrooms because students at that grade level have already been exposed to all the K-8 statistics concepts and elements of the statistics problem-solving process. It was expected that the teaching practices would consider all the concepts and elements covered until the seventh grade when teaching statistics. However, the study's findings demonstrate, in line with the claims of Makar and Fielding-Wells (2011), that teaching practices in statistics were still predominantly

traditional. Özmen and Baki (2019) discussed including statistical problem-solving processes in the Turkish K-8 mathematics curriculum. However, they noted the need for a systematic and holistic approach to considering the statistical problem-solving process. The current curriculum version is seen as insufficient to promote statistical investigations effectively. As a result, teaching practices in statistics have remained mainly traditional.

The literature demonstrates that students can develop a deeper understanding of statistics when provided with a learning environment that allows them to pose investigative questions within a specific context and utilize the various elements of statistical investigations to answer those questions. Previous studies have shown that engaging students in statistical investigations fosters their statistical understanding (e.g., Silva et al., 2022 with the context of Environmental Studies; Watson and English, 2015 with the context of being environmentally friendly). Similar findings have also been observed in research conducted on pre-service mathematics teachers. When pre-service teachers engage in a more connected cycle of statistical investigations through lesson study, they can develop a firmer grasp of statistical concepts and improve their statistical understanding (Yılmaz & Yetkin-Özdemir, 2021). However, as discussed, the emphasis on statistical investigations and the incorporation of statistical concepts in real classroom practices often fall short compared to what is outlined in the literature and the curriculum.

The curriculum has certain limitations regarding the presence and interconnectedness of big ideas, elements of statistical investigations, and the overall cycle of statistical problem-solving (Özmen & Baki, 2019; Watson & Callingham, 2015). This suggests that the curriculum is one possible factor influencing teaching practices. Alongside the written curriculum, the Turkish Ministry of Education has published the “General Competencies of the Teaching Profession,” which encompasses three areas: professional knowledge, skills, and attitudes/values (MoNE, 2017b). These competencies include indicators related to curriculum-related knowledge and skills. According to the document, teachers are expected to possess content knowledge, pedagogical content knowledge, and regulations-related knowledge. One specific

indicator within the pedagogical content knowledge area is that teachers should be able to “explain the program of the field, including all its elements” (MoNE, 2017b, p.13).

Additionally, teachers are expected to have professional skills in areas such as planning teaching and learning, creating conducive learning environments, managing the teaching and learning process, and assessment and evaluation. The document also includes specific indicators related to the written curriculum, such as preparing teaching materials aligned with the objectives, organizing learning environments, and planning in accordance with the field's objectives. Based on these indicators, it is evident that teachers are expected to align their actions with the written curriculum.

When considering the role of the curriculum in shaping teaching practices, it is vital to acknowledge the transformative role of teachers. The curriculum encompasses various components, including “the national curriculum document, curriculum materials designed for use by teachers, or the curriculum enacted in the classroom” (Pfannkuch, 2018, p. 389). Thus, the role of teachers in transforming the curriculum needs to be highlighted.

The transformation process of the written curriculum begins with the intended curriculum, which refers to teachers' interpretation of the written curriculum (Eichler & Erens, 2014). While the competencies document emphasizes the written curriculum, the expectations placed on teachers to have knowledge related to the program and its elements are related to the teacher's interpretation and transformation of the written curriculum. In addition, the document includes other indicators related to teacher knowledge and skills.

The enacted curriculum encompasses classroom processes and can be influenced by teacher knowledge and beliefs, as they play a significant role in shaping the intended curriculum (Blömeke et al., 2020; Eichler, 2011; Hiebert & Grouws, 2007). Eichler and Erens (2014) discuss how the enacted curriculum emerges through the interaction among the teacher, students, and content. Ultimately, the enacted curriculum is transformed into student learning as students develop their content knowledge and beliefs.

The literature highlights the significance of teacher knowledge, particularly in the field of statistics, as a potential factor influencing discipline-related practices such as conducting statistical investigations (Burgess, 2011; Litke & Hill, 2020). Groth (2013) emphasizes the importance of teachers' content knowledge and pedagogical content knowledge in developing effective teaching strategies specific to statistics. Burgess (2011) argues that teacher content knowledge and pedagogical content knowledge are essential for effective statistical problem-solving and for enhancing student learning in statistics.

The findings of this study regarding statistical content indicate that the teachers possessed a certain level of content knowledge in statistics, as evidenced by their consideration of outliers when selecting appropriate measures of center within a given context. This is further supported by the statistical content interview, where the teachers could define contexts for using specific measures of center. However, as mentioned earlier, the teaching practices primarily focused on calculations, procedures, and graphing. In contrast, Litke and Hill (2020) found that the teachers' observations predominantly focused on calculations without delving into the reasons for selecting specific averages. These differences may stem from differences in curricular objectives, as the curriculum in the current study explicitly emphasized the understanding of which measure of center is more appropriate for a given dataset (MoNE, 2017a). Another possible explanation for the disparity between the findings of this study and the literature is the availability of well-defined worked-out examples in the curricular materials (e.g., hospital examples), which the teachers could make use of those examples to be reflected on their content and pedagogical content knowledge.

Still, the findings of the study indicate that the teaching practices did not adequately support statistical investigations across each element, with the exception of the data analysis element and the overall cycle. Burgess (2011) suggested that this may be attributed to teachers' knowledge and ability to teach statistical investigations. Pre-service teacher education plays a crucial role in enabling teachers to develop their statistical knowledge for teaching. Franklin et al. (2015) emphasized that statistics is a significant component of teacher education programs, given the increased emphasis on statistics in the PreK-12 curriculum. Pre-service teachers need to have a strong

foundation in statistical content and processes (Franklin et al., 2015; Lovett & Lee, 2018).

In the current study, both teachers acknowledged having some prior experiences with statistics during their undergraduate studies, but these experiences were primarily focused on mathematical statistics, including calculations and procedures. They also expressed a lack of specific training on how to teach statistics to middle school students. Their learning process primarily relied on their years of teaching experience and using official and additional textbooks. These findings underscore the significance of providing in-service teacher education that explicitly addresses teaching statistics. This would enable teachers to develop their statistical knowledge for teaching (SKT) and effectively teach statistical content within the context of statistical investigations (Lee et al., 2022).

The literature suggests teachers' beliefs about statistics can influence their teaching practices (Eichler, 2011). In the current study, the participating teachers perceived statistics as a part of mathematics rather than a distinct discipline. They viewed mathematics as a necessary foundation for engaging with statistical concepts. One teacher expressed a dislike for statistics as a topic but felt confident in handling statistical content involving calculations and graph construction. The other teacher mentioned a personal interest in statistics in real-life contexts, such as following basketball statistics. These beliefs might affect discipline-related practices, influencing whether teachers focus on conceptual understanding rather than solely procedural aspects or engage in the entire statistical investigation cycle.

Groth (2008) highlights the importance of understanding teachers' perceptions of the intended curriculum to support the implementation of recommendations like those outlined in the GAISE. In Eichler and Erens' study (2015), secondary school teachers considered stochastic (statistics and probability) as "an applied domain of mathematics" due to its connections to real-world problems (p. 192). This perspective aligns with viewing statistics as a part of mathematics rather than a separate discipline. This viewpoint resonates with the participating teachers' beliefs in the current study and might be reflected in their intended and enacted curricula.

5.2. The Quality of Statistics Instruction within Broader Mathematics Instruction

Instructional quality is crucial in determining effective teaching practices and promoting student learning (Schoenfeld, 2014). Researchers have utilized various frameworks and observational instruments to assess instructional quality to examine teaching practices in different subject areas. This study employed the Mathematical Quality of Instruction (MQI) framework to assess the quality of statistics instruction within broader mathematics instruction. The MQI framework, developed by LMT (2011), focuses on the practices observed while delivering mathematical content to students. It utilizes the instructional triangle to examine the teacher, students, and content interactions. The framework encompasses subject-specific practices, including domains such as arithmetic, algebra, geometry, and statistics. Although statistics is increasingly recognized as a distinct discipline, it is covered within the broader context of school mathematics in this study. In addition to discipline-specific practices, it is essential to consider the quality of instructional practices in mathematics education more broadly. The MQI framework provides a comprehensive perspective on the quality of mathematics-specific instructional practices across different domains, as emphasized by Charalambous and Praetorius (2018).

The findings of the study indicate that both teachers' teaching practices predominantly focused on mathematics and statistics activities rather than non-mathematical and statistical ones. Only one segment in Esra's instruction involved a homework check, where no mathematical or statistical activities were observed. Ali's statistics lessons were rated as Mid to Mid/High quality when examining the overall quality of mathematics instruction. In contrast, Esra's statistics lessons were rated as Mid quality within mathematics lessons. These slight differences in overall lesson quality highlight the need to delve into the segment-level analysis, focusing on four dimensions of the MQI framework. These dimensions include the richness of mathematics and statistics presented to students, the teachers' ways of working with students and mathematics and statistics, the mathematical and statistical errors and imprecision present in the instruction, and student participation in meaning making and reasoning process. The results of the codes within each dimension will be discussed, followed by an overview of the overall MQI findings.

5.2.1. The Richness of Mathematics and Statistics

The dimension of “The Richness of Mathematics and Statistics” focuses on the depth of the mathematical and statistical content being taught in the lesson. It examines the interaction between the teacher and the content within the instructional triangle, as described by Hill et al. (2014). To analyze this dimension, the statistical teaching practices of both teachers were assessed in terms of their ability to link representations, provide explanations, and engage in mathematical and statistical sense-making, which were categorized as meaning-oriented codes. In addition, the analysis considered the use of multiple procedures and solution methods, patterns and generalizations, and mathematical and statistical language, which were classified as practice-oriented codes.

Upon analyzing the teaching practices of both teachers, it was observed that the overall richness of mathematics and statistics in most lesson segments ranged from low to mid-quality. The following sections will discuss the results of each code within this dimension.

5.2.1.1. Linking between Representations

First, the code for linking between representations refers to the practices that establish explicit connections among mathematical or statistical concepts. According to NCTM (2014), the use and connections among different representational families in mathematics, such as contextual, visual, verbal, physical, and symbolic representations, are crucial for students' mathematical sense-making (adapted from Lesh et al., 1987). It is emphasized that using multiple representations and establishing links among them can help students make sense of the abstract nature of mathematical concepts. Additionally, linking different mathematical representations is essential for describing relationships within data (Burrill & Biehler, 2011).

In this study, I analyzed the teaching practices in terms of the linking between representations code, with a focus on conveying the same relationships through different representations as suggested by the code. The results of the study revealed

that in most segments, both teachers did not effectively establish links between representations, despite their use of different representations in the context of statistics. The teachers taught various types of graphs to the students in line with the curriculum objectives. On certain occasions, they did make connections between visuals and symbolic forms by discussing angles, fractions, and sectors of a circle while constructing circle graphs. These instances of linking between representations aimed to facilitate students' understanding of graphs during the construction process.

As discussed earlier, in specific lesson segments in Ali's lessons, there were discussions about which graphs to use in specific contexts. However, visual representations were not visually present during these discussions. Moreover, both teachers covered the transformation of graphs. Friel et al. (2001) emphasized that different graphical representations in statistics can lead to different interpretations of the data. Lee et al. (2014) supported the idea that using multiple graphical representations can convey different aspects of the data and provide a holistic view. However, it is important to note that these different graphical representations belong to the same representational family. In order to have high-quality practices regarding the linking between representations, explicitly connecting the chosen graphs with the context could be beneficial.

When discussing the transformation of graphs, explicitly focusing on the context can also contribute to effective practices in linking between representations. In the MQI framework (Hill et al., 2014), the emphasis is on the explicitness and details of the connections among different representations rather than solely the presence of individual representations. Gerstenschlager (2019) suggested that while statistics as a discipline utilizes mathematical representations, choosing appropriate graphs to represent a dataset is influenced by the nature of the discipline and is connected to the context, which is an integral part of statistics. Therefore, using different types of graphs and selecting the most appropriate ones are key practices in teaching and learning statistics. However, from the perspective of the MQI framework, it is important to have an explicit link between two representational families. Thus, it is crucial to explicitly connect the graphical representation with the contextual

representation to discuss whether the focus is on the meaning-making process during the content delivered to the students in this code.

Burrill and Biehler (2011) explained that graphs are used to represent data. According to Wild and Pfannkuch (1999), transnumeration, a form of statistical thinking, involves using various representations of data to comprehend real-world data. It entails transforming real-world data by organizing it into tables and specific graphs to make sense of data (Chick, 2004). However, in the observed teaching practices, the teachers did not utilize real-world data, as discussed in the section on the statistical problem-solving process. In one specific segment of Esra's instruction, where a statistical investigation was presented, there was no presentation or connection between the table where the collected data was organized and the circle graph used for analysis and interpretation. Consequently, the transnumeration process did not involve linking between the table and the circle graph.

On the other hand, Ali incorporated the linking between visual and symbolic representations when teaching average concepts. In a reverse transnumeration process using given data, the teacher wrote down all the numbers in an open form, allowing students to connect the data frequencies in an open form with the heights of the bars in the bar graph to determine the mode.

However, both teachers did not utilize different representations when teaching the concepts of arithmetic mean and median. The literature discusses various conceptualizations of averages held by teachers and students, including the standard algorithm, fair-share, and mathematical balance points perspectives (Jacobbe, 2012; Mokros & Russell, 1995; Santos & Ponte, 2013). These different conceptualizations can be represented using various visualizations to enhance understanding of the average concept. In particular, the balance point conceptualization is crucial for comprehending the average within a variable system, as it relates to the “signal in the noise” process (Gfeller et al., 1999; Konold & Pollatsek, 2002). However, apart from the add-and-divide algorithm, the teachers did not create a learning environment where different representations were present and connected to conceptualize the average.

The lack of emphasis on connecting representations in the teaching practices may be linked to the teachers' mathematical knowledge for teaching (MKT). During the statistical content interview, when asked if there were other ways to find the average, the teachers' responses were limited, with only Ali somewhat mentioning the balance point approach but not using it in the classrooms. These findings align with existing literature highlighting how in-service and pre-service teachers often possess add-and-divide algorithm definitions of arithmetic mean and do not utilize different representations (Gfeller et al., 1999; Groth & Bergner, 2006; Jacobbe, 2012).

Furthermore, the absence of different representations in the textbooks could be another contributing factor. Bütüner (2020) analyzed the use of average concepts in the official textbooks provided by the Ministry of Education. The findings revealed that the examples in the textbooks predominantly relied on the add-and-divide algorithm, with limited inclusion of different types of representations. Consequently, the lack of representation links may stem from the absence of such examples in textbook materials.

5.2.1.2. Explanations

Explanations in the teaching practices refer to how teachers focus on why questions and engage in classroom discourse that promotes understanding (Hill et al., 2014). In the analyzed segments, it was observed that a significant portion of the teaching practices did not involve explanations that addressed why a procedure or solution method works or why a particular answer is true or false. These segments mainly consisted of descriptions rather than providing in-depth explanations to answer why questions.

However, there were other segments where mathematical explanations were present, particularly when discussing proportions while constructing or solving circle graph tasks. These segments generally required mathematical reasoning involving calculations of percentages, ratios, and proportions. While these practices did not directly address the topic of comparing two groups based on proportions, Franklin et al. (2015) highlighted the importance of proportional reasoning in statistics for

“understanding magnitudes of differences within a context” (p.68). Although not explicitly focused on comparing groups, using proportional reasoning, such as considering the sector of a circle in relation to the whole circle, could lay the foundation for making statistical comparisons based on proportions. The teachers did not utilize such statistical understanding because the purpose was not to compare two groups based on proportions. Still, the sector of a circle compared to the whole circle with the help of proportional reasoning could create a basis for comparing two groups based on proportions to make sense of the data statistically. All in all, teachers need to provide explanations regarding the rationale behind procedures and solution methods in statistics. In a broader sense, the extent to which teachers utilize proportional reasoning might contribute to students' understanding of the role of mathematics within the context of statistics.

In the teaching practices, there were limited segments where the teachers provided explanations regarding why specific types of graphs or measures of central tendency were selected. However, these explanations were generally of higher quality than those regarding procedures or solution methods. This finding aligns with the MQI framework, which includes specific descriptions of providing explanations for data analysis topics (Hill et al., 2014). However, it should be noted that most segments did not include such mathematical and statistical explanations.

One possible reason for this is related to teacher knowledge, including content and pedagogical content knowledge (Inoue, 2009). Lee and Santagata (2020) highlighted the conceptualization of teacher knowledge proposed by Ball et al. (2008), which involves knowing how to explain concepts and related procedures or algorithms to students. In the statistical content interview, both teachers demonstrated an understanding of why specific types of graphs and averages are used to some extent. They applied their understanding of statistical concepts to some extent during the teaching process. However, the number of explanations focused on “why” questions was limited.

The literature highlights the nature of instructional explanations through reasoning to the why questions as the center of the instruction, navigating the relationship among the teacher, students, and the content (Leinhardt & Steele, 2005; Charalambous et al.,

2011). The quality of teacher explanations has been linked to student learning outcomes (Leinhardt & Steele, 2005). Providing instructional explanations facilitates student learning. Thus, there is a need for teachers to possess content knowledge and pedagogical content knowledge to provide effective explanations. Therefore, it is crucial for teachers to receive adequate training and support in developing their instructional explanation skills.

Research suggests that pre-service teacher education programs can be designed to improve instructional explanations in mathematics (Charalambous et al., 2011; Inoue, 2009; Kinach, 2002). However, there is a lack of research on how improvements in instructional explanations during pre-service teacher education translate into actual classroom practices (Charalambous et al., 2011). This study might inform how pre-service teacher education with a limited locus of statistics reflected on teaching practices in terms of providing explanations in real classrooms. As a result, in the current study, the limited explanations provided by the teachers could be attributed to a lack of specific focus on instructional explanations in both in-service and pre-service teacher education.

5.2.1.3. Mathematical and Statistical Sense Making

Mathematical and statistical sense-making involves going beyond the mere description of procedures and focusing on the meaning and understanding of mathematical and statistical concepts (Hill et al., 2014; NCTE, n.d.). Good instructional explanations provide meaning to procedures by connecting them to relevant contexts and engaging in proportional reasoning rather than simply describing the steps in a procedure (Charalambous et al., 2011).

The findings of the study indicate that the teachers in this study generally engaged in mathematical and statistical sense-making to a considerable extent. They connected mathematical operations to the context, emphasized proportional reasoning, and made sense of graphs while interpreting results. In most segments, the teachers demonstrated a sense-making process beyond a mere description of procedure steps. However, there were a few instances in which the teachers made sense of averages or graphs during

whole segments. Only some segments in Esra's instruction did not consist of the meaning-making process because the focus was more on describing the procedures' steps rather than the meaning behind them. The teaching practices observed in terms of mathematical and statistical sense-making align with the NCTM (2009) definition, which emphasizes developing an understanding of an issue through connections with prior knowledge and considering the context or concept in mathematics and statistics.

As discussed previously, the teaching practices observed in this study did not encompass the entire statistical problem-solving process, with a focus mainly on the data analysis element alone and occasionally combining data analysis with the interpretation element. However, the literature emphasizes that making sense of statistics requires students to engage in the complete statistical investigative process (NCTM, 2009; MacGillivray & Pereira-Mendoza, 2011). Even if only one step is covered, it should be discussed in relation to the other elements to facilitate understanding (MacGillivray & Pereira-Mendoza, 2011).

A sense-making process in statistics and probability includes four key elements: data analysis, modeling variability, connecting statistics and probability, and interpreting designed statistical studies (NCTM, 2009). The elements of data analysis and interpreting designed statistical studies are particularly relevant to the learning outcomes of the Turkish middle school mathematics curriculum. According to NCTM (2009), students can develop a sense of statistics by engaging in the process of answering statistically investigative questions through data collection and analysis. Understanding data through tables and graphs is an essential aspect of this process. Friel et al. (2001) defined graph sense as a critical component for making sense of collected data. In the current study, the teachers somewhat demonstrated graph sense by connecting graphs to the context and data, as well as to tables (Friel et al., 2001).

NCTM (2009) also emphasized the importance of interpreting designed statistical studies as a key element for sense-making in statistics and probability. They highlighted the significance of statistical inference, including informal inference, as the curriculum does not explicitly cover formal measures of inference, such as probability distributions (Makar et al., 2011). In the current study, the teachers utilized

graph sense to interpret findings derived from graphs, drawing conclusions based on the data represented in the graphs (Friel et al., 2001; Makar & Rubin, 2009; Shaughnessy, 2007).

5.2.1.4. Multiple Procedures or Solution Methods

Multiple procedures or solution methods, the first practice-oriented code, examine how teachers provide students different approaches to solve problems or use multiple methods within a given problem type (Hill et al., 2014; NCTE, n.d.). In the context of statistics, the discipline uses mathematical skills such as “arithmetic procedures, measurement skills, and proportional reasoning involving fractions, decimals, ratios, and percentages” throughout the statistical problem-solving process (Watson, 2016, p.133). “Modeling with mathematics” is identified as one of the core statistical practices (Franklin et al., 2015, p.11). This particular code considers the quality of instructional practices regarding multiple procedures or solution methods to explore the opportunities offered to students to learn statistics using mathematical skills and concepts for an effective statistical problem-solving process.

The results of the study indicated that most of the lesson segments in both teachers lacked teaching practices regarding multiple procedures or solution methods. Teachers used a single method or procedure in those segments while solving statistical tasks. Specifically, the segments that included arithmetic mean were scored as not present since the teachers did not offer students different ways of finding the arithmetic mean. Specifically, when finding the arithmetic mean, the teachers relied solely on the standard algorithm, neglecting other possible approaches, such as the leveling or balance point method (Van de Walle et al., 2010). During the statistical content interview, Esra questioned the existence of alternative methods for finding the arithmetic mean. At the same time, Ali mentioned the leveling approach but admitted that he did not use it in class because students were more familiar with the standard algorithm.

These findings suggest a potential lack of knowledge from the teachers regarding multiple solution methods for finding the arithmetic mean. This aligns with the findings of Sleep and Eskelson (2012), who found a similar relationship between

teachers' mathematical knowledge for teaching and their use of multiple solution methods when teaching fractions. Moreover, Bütüner (2020) demonstrated that the official textbooks used in Turkish classrooms did not provide multiple strategies in the worked-out examples. This lack of exposure to multiple solution methods in curricular resources may have influenced the teaching practices observed, which lacked the use of alternative approaches when teaching the arithmetic mean.

Furthermore, some segments involving circle graphs had low to mid-quality practices regarding multiple procedures or solution methods. Both teachers utilized proportional reasoning within and between ratios to find angles or percentages of sectors in the circle graph. However, in the segments with low-quality practices, alternative solution methods were briefly mentioned but not explicitly explored in the classroom. These examples were mainly related to constructing circle graphs or extracting information from them. The content was similar in the segments with mid-quality practices.

Diezmann and Lowrie (2009) emphasized the importance of students developing various fluencies, including mathematical, conceptual, perceptual, and graphical fluencies. In this case, the teachers primarily focused on mathematical fluency by using mathematical operations that contributed to graphical fluency, specifically in interpreting circle graphs through proportional reasoning. There were also low to mid-quality practices in segments related to average concepts. For example, when finding the median of a dataset with an even number of data points, Ali offered two solution methods in the classroom. Esra employed multiple solution methods while calculating the arithmetic mean from data presented in a frequency table, which helped students make sense of the procedures involved in the standard algorithm.

In this study, there were only a limited number of segments where multiple procedures and solution methods were utilized for graphs and averages. Improving teaching practices in this aspect is crucial because focusing on multiple procedures or solution methods enables learners to engage with different perspectives and use mathematical and statistical ideas more flexibly (Grosse & Renkl, 2006). Moreover, emphasizing multiple procedures and solution methods is linked to student motivation as it allows them to approach problems in different ways (Grosse, 2014; Schukajlow & Krug,

2014). Overall, increasing student engagement in understanding mathematics and statistics through various solution approaches contributes to student learning (Grosse, 2014; Richland et al., 2017). As Diezmann and Lowrie (2009) pointed out, teachers need to develop their content knowledge and understanding of students' needs in order to create a learning environment where classroom discussions can revolve around multiple solutions to specific problems, such as graphs or averages, thus promoting students' mathematical and statistical proficiencies.

5.2.1.5. Patterns and Generalization

Patterns and generalization, as a quality indicator, involve examining patterns and developing generalizations based on those patterns (Charalambous & Litke, 2018; Hill et al., 2014; NCTE, n.d.). Cobb and Moore (1997) discussed the role of patterns and generalizations in mathematics and statistics. They highlighted that mathematics focuses on exploring patterns and developing generalizations to achieve abstraction, while statistics incorporate patterns and generalizations within a contextual framework. However, in the current study, both teachers demonstrated limited teaching practices related to this code, except for two segments in Ali's instruction, where examples were briefly explored to draw conclusions about similar cases. This lack of emphasis on patterns and generalization may be attributed to the perception that these practices are primarily associated with algebra (MoNE, 2018; Van de Walle et al., 2010).

Nonetheless, patterns and generalization can also be utilized in statistics by reading beyond graphs in which there was a need for examining patterns in the data to make data-driven assumptions (Delmas, 2004; Friel et al., 2001). Delmas (2004) noted that using graphs to “identify patterns may require a level of abstraction that students find just as difficult as the abstract patterns encountered in mathematics” (p.86). In addition, Franklin et al. (2015) discussed the core statistical practice of “look for and express regularity in repeated reasoning,” emphasizing the importance of “identifying patterns in long-run variability” (p.12).

The absence of big ideas, such as statistical inference, in the curriculum and curricular materials was evident in the teaching practices observed in this study. Statistical inference involves drawing conclusions and generalizing based on data (Makar & Rubin, 2009; Zieffler et al., 2008). While formal concepts of statistical inference, which rely on probability concepts like p-values, confidence intervals, and statistical significance, are typically introduced at a later stage, informal inference can be integrated into the curriculum earlier (Harradine et al., 2011; Makar et al., 2011). The Turkish curriculum introduces basic probability concepts in eighth grade, preceding formal statistical inference (MoNE, 2018). The reasonings in informal inference could be used before the introduction of the probability concepts, described as “the reliance on the concept of aggregate through the use of generalizations about the group; acknowledgment of a mechanism or tendency that extended beyond the data at hand” (Makar & Rubin, 2009, p.85). Therefore, the study's results indicate the need for improving instructional practices regarding patterns and generalizations by considering the link between the quality of instructional practices and curriculum (Charalambous et al., 2012; Hill & Charalambous, 2012a).

5.2.1.6. Mathematical and Statistical Language

Mathematical and statistical language consists of teaching practices with the use of technical language fluently and explicitly (Hill et al., 2014; NCTE, n.d.). The density of terminologies and encouraging students to use those terminologies made this practice high-quality (Hill et al., 2014). The literature discusses the importance of spoken and written language in learning mathematics (Kaplan et al., 2009). NCTM (2000) pointed out that students need to “use the language of mathematics to express mathematical ideas precisely”(p.60). In the current study, most of the lesson segments in both teachers and students used mathematical and statistical language at middling density; when necessary, they used statistical language, including the terms related to averages and graphs, and a statistical problem-solving process. One aim of teaching and learning statistics is to make people statistically literate to make sense of the world around them (Bargagliotti et al., 2020). To do so, it is necessary to have “the understanding and using basic language and tools of statistics” (Garfield & Franklin, 2011, p.135). Therefore, the findings of this study highlight the importance of effective

communication through mathematical and statistical language in statistics instruction within mathematics classrooms.

Promoting the use of mathematical and statistical terms by students is an essential aspect of teaching practices in statistics instruction. When teachers actively encourage students to use such language, it can enhance the quality of instruction. However, in this study, there were only a limited number of segments where teachers explicitly encouraged students to use statistical and mathematical language. Makar and Confrey (2005) discuss the use of language in statistics to improve this practice, highlighting both nonstandard and standard forms. They emphasize the role of nonstandard language in developing a relational understanding of statistical concepts. Thus, encouraging students to use their own language to make sense of statistical concepts might foster a deeper understanding. Teachers need to understand and recognize students' use of nonstandard mathematical and statistical language and leverage this understanding to help students grasp statistical concepts. As students develop their understanding, teachers can gradually introduce and encourage the use of standard language to align with accepted mathematical and statistical conventions. The current study suggests incorporating a balance between nonstandard and standard language for students to develop their statistical reasoning and communication skills.

5.2.2. Working with Students and Mathematics and Statistics

Working with Students and Mathematics and Statistics dimension involves an essential aspect of teaching practices on how teachers engage in students' contributions, including their mathematical and statistical ideas, and addressing their errors and difficulties (Hill et al., 2014; NCTE, n.d.). In this dimension, I analyzed the teaching practices of two teachers in statistics instruction regarding remediation of student errors and difficulties, and the teacher uses student mathematical and statistical contributions codes. In this study, the overall quality of teaching practices in this dimension was found to be low in the majority of lesson segments for both teachers. The specific results of each code within this dimension will be discussed in the following sections to provide a more detailed examination of the teaching practices.

5.2.2.1. Remediation of Student Errors and Difficulties

Remediation of student errors and difficulties is crucial to teaching practices in statistics instruction. It involves teachers actively addressing and responding to student errors, misconceptions, and difficulties. Hill et al. (2014) distinguished between procedural and conceptual remediations as two types of remediation strategies. National Research Council (NRC) (2001) states that the differentiation between these types of remediation stems from students' learning in mathematics.

Student understanding is an outcome of student learning, and it is “a measure of the quality and quantity of connections that new ideas have with existing ideas” (Van de Walle et al., 2010, p.24). Conceptual understanding occurs when students can make connections between new concepts and their existing knowledge, while procedural understanding pertains to skill efficiency and the ability to apply algorithms. Both forms of understanding are essential for student learning. However, it is important to note that algorithms can be applied without a deep understanding, leading to errors or misconceptions. Developing conceptual understanding and procedural fluency are intertwined mathematical proficiencies that need to be addressed in mathematics classrooms (NRC, 2001; Hiebert & Grouws, 2007; Van de Walle et al., 2010).

Student errors and difficulties provide opportunities for students to engage in problem-solving and develop more sophisticated solutions (Kazemi, 1998). Therefore, when errors and difficulties arise in the classroom, teachers need to identify and remediate them effectively. This helps to improve students' conceptual and procedural understandings of the content (Hill et al., 2008; Larrain & Kaiser, 2022). The results of the study indicate that the majority of the lesson segments lacked remediation of student errors or only included brief procedural remediation. Based on the literature, this suggests that there may be missed opportunities to address students' errors and misconceptions and to deepen their conceptual understanding. In segments with no remediation, it is possible that either no student errors were observed or the teacher's dominant talk and direct corrections did not allow for remediation to take place.

However, in the segments where teachers attended to and responded to student errors and difficulties, the remediation provided was either extensive procedural or moderate

conceptual. Extensive procedural remediation occurred when students were engaged in mathematical calculations, such as calculating measures of center or constructing and interpreting circle graphs. Teachers extensively remediated student errors related to proportion calculations, particularly in circle graph tasks. In addition, teachers conceptually remediated student errors in tasks involving reading frequency tables, for example, when finding the arithmetic mean and the mode. Particular example segments on this issue also included pre-remediation, where the teacher anticipated those errors and difficulties and acted on them before the errors and imprecision appeared. These findings are consistent with Batanero et al. (1994), who highlighted that frequency tables and graphs can be sources of errors and difficulties for students, particularly in reading, interpreting, and comparing data. In the current study, students had difficulties in reading the data presented in frequency tables when finding the arithmetic mean and the mode.

The current study highlights the importance of encouraging teachers to integrate teaching practices that address student errors, difficulties, and misconceptions, as these are inherent aspects of the learning process. This aligns with the principles outlined by NCTM (2000) and NCTM (2009) that emphasize the role of teachers in responding to and addressing student misconceptions. To effectively respond to student errors and misconceptions, teachers need to enhance their pedagogical content knowledge, which involves understanding the content and knowing how to transform student misconceptions in their instruction (Shulman, 1998).

Improving pedagogical content knowledge in statistics enables teachers to possess knowledge about the content and their students, including understanding common student errors and misconceptions (Groth, 2013). Additionally, enhancing teachers' diagnostic competence in error situations, as discussed by Larrian and Kaiser (2022), empowers teachers to identify and respond effectively to student errors and difficulties. By utilizing teaching practices such as procedural or conceptual remediation, teachers can help students understand where they have errors and difficulties within mathematics and statistics.

The literature suggests that teacher competence, which includes teacher knowledge, beliefs, and diagnostic skills, is closely connected to high-quality teaching practices (Blömeke et al., 2020). The findings of this study underscore the challenging nature of attending and responding to student errors and difficulties, emphasizing the need for teacher education programs to focus on developing teachers' knowledge and diagnostic skills to remediate student errors and difficulties effectively.

5.2.2.2. Teacher Uses Student Mathematical Contributions

In mathematics classrooms, effective instruction involves correcting student errors and difficulties and actively incorporating student contributions to enhance the learning experience. Student contributions, such as answers, comments, explanations, and representations, can enrich classroom discourse and foster a student-centered learning environment (Hill et al., 2014). To achieve this, teachers need to anticipate, monitor, select, sequence, and connect student responses, while also attending to their thinking and demanding mathematical explanations and justifications (Stein & Kaufman, 2010).

Traditional classroom discourse practices often rely on close-ended questions that limit students' opportunities to explain their thinking or engage in collaborative discussions about mathematical ideas (Brendefur & Frykholm, 2000). In the current study, the majority of lesson segments exhibited low-quality uptake of student ideas, utilizing student contributions in an answer-bounded way or merely explaining the whole idea. This pattern of discourse, known as the IRE pattern (Initiation-Response-Evaluation), tends to revolve around closed-ended questions with a single correct answer, focusing primarily on evaluating the correctness of ideas. These findings suggest that traditional teacher-centered discourse practices may hinder the development of high-quality mathematics learning (Franke et al., 2007). In line with the recommendations of NCTM (2009), it is crucial to create a learning environment where students are active participants in a mathematical community, empowered to communicate their ideas and contribute to instruction. The role of the teacher is pivotal in fostering a mathematical community by orchestrating effective classroom discussions that revolve around student ideas (Smith & Stein, 2011).

The results of the study not only highlight the need for understanding how teachers utilize student contributions and underscore the necessity to provide support for teachers to enhance their teaching practices. As introduced by Lineback (2015), responsive teaching in mathematics emphasizes the utilization and response to student contributions to foster meaningful classroom discourse. Redirecting, a key aspect of responsive teaching, involves connecting student ideas and promoting high-quality teaching practices for effectively incorporating student contributions (Lineback, 2015).

The current study revealed a lack of responsive teacher uptake of student ideas in most lesson segments. Teachers often failed to establish meaningful connections with student contributions to improve instruction or displayed delayed responses. For instance, in the case of Ali, while he would address student struggles during individual work time by solving questions on the board, there was limited focus on addressing the specific struggles or fully incorporating student contributions. Hence, providing support to teachers in implementing responsive teaching practices and effectively utilizing student contributions is crucial for improving the quality of mathematics instruction.

Stockero et al. (2020) argue that understanding how teachers use student contributions is important, but equally crucial is understanding why they use them. They propose that teachers' orientations towards student mathematical thinking as a resource significantly influence their teaching practices. Their research reveals a continuum of orientations, with high-potential orientations valuing student thinking as a resource and promoting direct student interaction through explanations, questioning, critique, and discussion of ideas. These high-potential orientations align with high-quality teaching practices that prioritize the utilization of student contributions.

However, in the current case study, there were limited instances where teachers exhibited high-potential orientations. Instead, the majority of the lesson segments were characterized by low-quality practices, indicating that teachers might be influenced by hindering orientations in their teaching. Stockero et al. (2020) describe these orientations as heavily scaffolded learning experiences, where the teacher's role

primarily involves explaining and demonstrating mathematical ideas. Therefore, the results of the current study suggest that teachers reflect on their orientations to transform their teaching practices on this issue.

5.2.3. Errors and Imprecision

This dimension focuses on teaching practices related to teacher errors or imprecision (Hill et al., 2014; NCTE, n.d.). In this dimension, I examined the teaching practices of two teachers in statistics instruction, specifically regarding mathematical and statistical content errors, imprecision in language and notation, and the lack of clarity in presenting mathematical and statistical content. Overall, the results of this dimension indicate that the majority of segments in both teachers' statistics instruction were free from errors and imprecision. However, there were specific segments where errors and imprecision were observed, which will be discussed in this section.

Errors and imprecision in teaching mathematics are a part of real classrooms and might have a negative impact on student learning (Ball & McDiarmid, 1990; LMT, 2011). While there is existing research on possible errors, misconceptions, and difficulties in learning statistics, there is a lack of research on teacher errors and imprecisions, specifically in teaching statistics. To illustrate, specific studies have focused on student errors and difficulties in interpreting graphs (e.g., Aydın-Güç et al., 2022 for scatterplots; Capraro et al., 2005 for bar, line, and circle graphs; Edwards et al., 2017 for boxplots). Other studies also examined pre-service teachers' errors and imprecision in interpreting graphs (e.g., Işık et al., 2012 for line graphs; Ulusoy & Çakıroğlu, 2013 for histogram). However, there is a lack of research on what kinds of teacher errors and imprecision are present in teaching statistics. Therefore, it is essential to explore teacher errors and imprecision to learn from them for not to transfer inaccurate information teachers possess to the students (Ball & McDiarmid, 1990). By exploring teacher errors and imprecisions in teaching practices, it is possible to identify areas of improvement, leading to more effective and engaging instruction and better student outcomes. Therefore, studying teacher errors and imprecision is essential for promoting high-quality mathematics education (LMT, 2011).

The results of the study revealed that both teachers made mathematical and statistical errors in more than 15% of the lesson segments. One example of an error was their definition of the mode as the most frequent number instead of the most frequent value. They did not focus on the data type in their lessons and mostly worked on examples with quantitative data while teaching average. Therefore, their focus on examples with quantitative data while teaching averages did not obscure the statistics in those examples. However, errors emerged when students encountered examples with categorical data, as they mistakenly reported frequency numbers as the mode of the data set. Teacher definitional errors in defining the mode could potentially lead to student errors.

The findings align with Santos and Ponte's (2013) study, which showed that pre-service teachers often defined the mode as “something that happens 'more times'” (p.5). This definition is problematic because the highest value in a dataset may not necessarily represent the mode in a quantitative context. The mode is determined by the frequency of occurrence, not the magnitude of the values. It is essential to understand this distinction, especially when working with quantitative variables, as the highest value could be an outlier or an extreme value rather than the mode. The current study highlights the need to properly understand the mode, especially when working with quantitative variables.

Leavy (2015) points out the misconception that the mode is simply the number of occurrences without considering the center concept. It is often perceived as a relatively straightforward concept that only requires the ability to count the frequency of a particular data value. In their lesson study, Leavy (2015) uncovered the misconceptions of pre-service teachers who defined the mode as “the most frequently occurring number in a data set” (p.303). This misconception arises from a lack of common content knowledge and the failure to distinguish between categorical and quantitative variables. Similarly, in the current study, although most incidents involved examples with quantitative data, the definition of the mode provided by the teachers could lead to student errors and difficulties when confronted with examples featuring categorical data in a frequency table. The classroom examples examined in

this study revealed instances where students had difficulties finding the averages from data organized in a frequency table.

In line with the results of Santos and Ponte's (2013) and Leavy's (2015) studies, the teaching practices in the current study lacked an emphasis on the distinction between quantitative and categorical variables or data. Franklin et al. (2015) pointed out that not having such distinction might create issues in understanding, first, the nature of data, leading to incorrect ways of dealing with summary statistics, such as adding all frequencies for finding the arithmetic mean of the data set in a categorical variable. The importance of understanding the data types is also discussed in GAISE documents. GAISE I document described data types as quantitative and categorical, currently considered traditional. The students are expected to "become comfortable with analyzing these traditional variable types" (p. 9). GAISE II expands on the traditional variable by presenting different data types, including pictures, sounds, or videos. After engaging in traditional ones, the students could experience other types of data. The mathematical and statistical content errors observed in the teaching average in the current study may stem from the teachers' lack of content knowledge in statistics.

Leavy's (2015) study shows that pre-service teachers can develop their common content knowledge and the knowledge of students and content through engaging in lesson study. Furthermore, the missing elements in the curriculum and curricular materials might be the other reason the issues arose from teaching the concept of mode. First, official textbooks distributed by the Ministry of Education defined the mode as "the most repeated number in a data set" (Bilen, 2017, p.237), which aligns with the errors identified in this study. Furthermore, the curriculum does not explicitly address any learning outcomes related to data types, so the curricular materials do not explicitly cover different types of data. The textbooks might be one of the sources of the content error that appeared. Textbooks are crucial in teacher learning, and their content can influence teaching practices (Stein & Kim, 2009).

In both teachers' teaching practices, a significant error and imprecision arose in the same example of data set "4 4 4 4 4". The example was about finding the mode of the data set from the additional textbook the teachers used in the classrooms. Before covering the example, both teachers sought clarification from the researcher on how

to present such examples to the class. I reminded my role as the researcher as the nonparticipant observer. However, I observed that both struggled with understanding the mode within such kinds of distributions even before bringing up this example to the class. Their struggles before the class were reflected in the classroom observations; they made significant errors, and their use of language and presentation of the content were affected by this struggle.

Esra, in particular, experienced confusion between the magnitude of the value and the frequency of occurrences, leading to incorrect representations. Her imprecise language, such as relating “peaks” in her drawings to the meaning of peak value, further contributed to these errors. On the other hand, Ali contextualized the example by using a bar graph with days on the x-axis and the numbers in the data set on the y-axis. However, the data set in the question consisted of only six individual data points. Ali's contextualized example included 24 data points, introducing inaccuracies into the discussion. Moreover, both teachers failed to address student errors and difficulties that arose during the discussion of the example, leaving them uncorrected.

The main issue in this example and other teaching practices is that they do not consider “variability” as a core idea in statistics. Variability is what differentiates statistics from mathematics and makes it a distinct discipline. Cobb and Moore (1997) pointed out that “the need for such a discipline arises from the omnipresence of variability” (p. 801). As a measure of central tendency, the mode falls under the expectation big idea. Watson et al. (2013) described the purpose of expectation as “to harness variation and summarise data” (Callingham et al., 2021, p.332). Therefore, it is necessary for teachers to have SKT for choosing what to bring to the classroom while teaching statistics based on the primary purpose of statistics as a discipline.

Additionally, the textbooks need to cover cognitively demanding statistical tasks rather than examples that present unnecessary challenges, leading to confusion and struggle for teachers and students. Furthermore, the results of the study also emphasize the importance of context in statistics as a discipline. Providing such examples without context might have led to more teacher errors and imprecision and subsequently resulted in student errors and difficulties. Context is an inseparable part of statistics as

a discipline (Watson et al., 2018). Similar to the current study's results, Litke and Hill (2020) pointed out that teachers mostly practice average concepts as a series of calculations without the context provided, which leads to no meaning attached to the value found based on the calculations. The current study shows the need for integrating context into the numbers while teaching statistics to make sense of summary statistics. Groth (2013) indicates the need for developing the knowledge of content and teaching in SKT to provide students with meaningful and critical error-free statistics instruction.

Some segments included high content errors due to the inconsistencies between the purpose of the graph and the context provided in constructing a graph. For example, both teachers used ordinal data on the x-axis in a line graph task. Patahuddin and Lowrie (2019) emphasized that the purpose of line graphs is to represent the data in terms of the change over time. Alacacı et al. (2011) pointed out that line graphs are used for trend analysis, analyzing the patterns over time. Işık et al. (2012) also added that continuity is one critical aspect of line graphs. Therefore, it is not always appropriate to transform all bar graphs into line graphs, as there are specific conditions and considerations to be taken into account (Alacacı et al., 2012).

The results of the current study are consistent with Işık et al.'s (2012) study, which revealed erroneous teaching practices in line graphs in the current study due to forgetting key conditions in the use of line graphs; the variable in the x-axes was time-dependent but not continuous; it would be appropriate to use bar graph instead of line graphs to represent the data. Işık et al. (2012) also found that the majority of pre-service teachers created a context for line graphs that included noncontinuous data, including the changes in the number of correct questions over quizzes. In addition to the teachers' incorrect choice of graphs according to the context given, Li and Shen's (1992) study also showed that while representing qualitative variables, secondary school students constructed line graphs (as cited in Batareno et al., 1994). In our context, one example was about the change in the number of correct questions over mock exams. Those examples were present in the additional textbooks used by the teachers, and the teachers brought the examples to the class to introduce line graphs.

In the study, it was expected that the teachers would transform bar, line, and circle graphs appropriately based on the types of data and the purposes of the graphs. Different types of graphs are present due to their purposes and the types of data (Alacacı et al., 2011). For example, Alacacı et al. (2011) highlighted that line graphs are used for observing patterns over time using continuous data, while bar and circle graphs predominantly use categorical data. Bar graphs are used to compare the frequency counts, while circle graphs are used to compare different categories within a whole (Alacacı et al., 2011; Friel et al., 2001).

While transforming the graphs, it is necessary to have knowledge regarding the data types and the purposes of the graphs. In Ali's teaching practices, the transformation process lacked clarity while constructing a circle graph based on a given context and asking students which graphs could be transformed. First, the context the teacher provided was incomplete and inappropriate for a circle graph in the beginning. Then the focus moved to understand the context with some missing information rather than transformation. The results regarding errors and imprecision appeared while teaching graphs conveyed a similar message in teaching average. In addition to using numbers in context while teaching averages as a part of the statistics discipline, it is important to use appropriate graphs while representing the data in context.

As the last issue, imprecision in language was observed in specific segments, particularly when discussing average concepts. In previous sections, we discussed the importance of using mathematical and statistical language. In this dimension, the results of the study show that both teachers used imprecise language in the arithmetic mean. Yılmaz and Yetkin-Özdemir (2017) indicated similar results regarding Turkish pre-service teachers' average conceptions that were equated with the arithmetic mean concept. When they encounter a context asking for an average, they tend to discuss only arithmetic mean, overlooking the median and mode.

The imprecision in statistical language could be attributed to general language use. To be more precise, in the Turkish language, arithmetic mean means "aritmetik ortalama". In introducing average concepts, the teachers introduced average as "ortalama" including arithmetic mean, median, and mode. During the process, they used ortalama

both for average and arithmetic mean. In some cases, this creates confusion for students. The sources for the imprecision might come from the curriculum stating “Bir veri grubuna ait ortalama, ortanca ve tepe değeri bulur ve yorumlar.” As translated in English, “Finds and interprets the mean, median and peak value of a data group.” In the official textbook, it is stated that “Aritmetik ortalama, ortanca ve tepe deer istatistikte kullanılan ortalama çeşitleridir.” As translated in English, “Arithmetic mean, median, and mode are the types of means used in statistics.” The teachers used this definition in introducing average concepts. To be more precise, the teachers should have used “arithmetic mean” instead of “ortalama,” or alternatively, they could have defined all three concepts as measures of central tendency to avoid imprecision in language.

Furthermore, the use of “ortalama” in general language creates imprecision and distorts its specific meaning in statistics as in the use of formal language. In the current study, the word “ortalama” was perceived as “approximately” due to the general language rather than a value representing a data set or a typical value similar to Yılmaz and Yetkin Özdemir (2017) where the pre-service teachers in Turkish context considered the word “approximately” as “arithmetic mean” in a given context. Similarly, the mode concept is used in two versions in the lessons: mode and peak value. In a segment, the use of mode as a peak value referring to the daily use of “peak” as a hill and representing hills to teach the mode concept created imprecision. It created confusion on the magnitude of the data or the occurrences of the data.

Biehler (1997) stated that the difficulties and errors in understanding the data might stem from the lack of formal language (Biehler, 1997, as cited in Makar & Confrey, 2005). Franklin et al. (2015) emphasized attending to precision in statistics while communicating through the use of precise statistical language. Hill et al. (2014) indicated that the use of correct terminology needs special attention to “distinguishing everyday meanings different from their mathematical meanings” (p.20). The study’s results, along with the previous discussions on errors and imprecision, indicate that the teachers’ SKT needs to be improved in order to decrease possible imprecision in language. Also, the use of formal language might be another source of such imprecision. Therefore, by improving SKT, it is possible to bring fewer errors and

imprecision stemming from other sources. All in all, detecting and learning from teacher errors and imprecision might prevent possible misconceptions in student learning (Ball & McDiarmid, 1990).

5.2.4. Student Participation in Meaning-Making and Reasoning

This dimension focuses on teaching practices related to understanding students' interaction with the content, including their participation in mathematics and statistics (Hill et al., 2014; NCTE, n.d.). While the previous dimension focuses on how the teacher provides a learning environment for students' work with the content, this dimension examines students' engagement in the content itself in the presence of the teacher (NCTE, n.d.).

In this dimension, the teaching practices of two teachers in statistics instruction were analyzed based on five codes categorized into two groups. The first group includes codes related to mathematical and statistical explanations, questioning, reasoning, and communication, which consider how students talk about mathematics and statistics. The second group includes codes related to students' work with contextualized problems and engagement in cognitively demanding tasks, which refer to “the types of mathematical tasks students are involved in” (NCTE, n.d.). The overall result of this dimension indicates that the quality of teaching practices in terms of students' interaction with the content is generally missing or of low quality. This dimension will be further discussed under its two categories.

5.2.4.1. Students Talk about Mathematics and Statistics

Vygotsky highlighted the constructive nature of knowledge, which emerges through exchanging and exploring ideas, ultimately leading to informed judgments (Hiebert & Grouws, 2007). Within this context, communication of mathematical ideas plays a crucial role in mathematics learning, involving students articulating their own understanding through spoken or written language and engaging in analyzing, evaluating, and explaining others' arguments (NCTM, 2000). As I delve into the discussion on student talk about mathematics and statistics in this study, I will

specifically focus on student explanations, questioning, reasoning, and the overall communication of mathematical content.

The results of the study show that there were limited incidents of students actively in mathematical and statistical discourse through talking about mathematics and statistics, particularly when providing explanations, questioning, and reasoning. Most lesson segments in both codes were scored absent in both teachers' teaching practices due to teacher-dominant talk or lack of observation during individual working time. This suggests that both teachers mostly had teacher-centered classroom discourse, where the students did not find many opportunities for sharing their ideas and solution methods (Franke et al., 2007; Sherin, 2002).

Students in a student-centered learning environment become “doers of mathematics” to make mathematical arguments through questioning and reasoning processes and answer why questions (Schoenfeld, 2013). As stated in Eight Standards of Mathematical Practices, students need to “S1-make sense of problems and persevere in solving them” and “S3-construct viable arguments and critique the reasoning of others” (CCSSI, 2010, p. 6-8; NCTE, n.d.). This process not only supports students in developing their agency (i.e., perceiving themselves as capable mathematical learners who can deal with challenging aspects of mathematics) and authority (i.e., being aware of what they know in mathematics, using that knowledge and using mathematical reasoning to decide what is valid or invalid.) (Schoenfeld et al., 2016; Schoenfeld, 2014).

According to NCTM (2014), one of the important roles of teachers is to facilitate meaningful mathematical discourse as one of the teaching practices needs to be attended through “discourse among students to build a shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments” (p.10). However, in our case, the teachers used the IRE discourse pattern; this pattern often involved the teachers posing a question, expecting a correct answer from the students, and then evaluating the correctness of the response. Instead of fostering interactions among students and promoting student engagement with the content and their peers' thinking, the authority in the classroom discourse remained primarily with

the teacher, who determined the correctness of the answers (Cazden, 1988; Franke et al., 2007; Stein & Kaufman, 2010).

What is discussed regarding student talk in mathematics also applies to statistics as a discipline. For example, in the adaptation of eight standards, Watson and English (2017c) stated that students need to engage in “make sense of statistical questions posed and preserve in answering them” and “construct viable arguments based on evidence from data and critique the reasoning of others.” While the first standard was not observed in the statistical teaching practices, there were scarce examples that the students created arguments based on data while dealing with statistical measures and questioned the context to make sense of, for example, the existence of bimodal or multimodal distributions based on the given definition of mode, as the most frequent number. Moreover, Gerstenschlager (2019) stated that mathematical thinking could be used to calculate measures in statistics. Using multiplicative reasoning is an example of student mathematical and statistical questioning and reasoning. However, statistical thinking is necessary for making sound inferences and constructing arguments based on data.

Furthermore, Ben-Zvi and Garfield (2004) described statistical reasoning as making interpretations based on data, representations, and statistical measures. In the current study, student explanations, questioning, and reasoning of at least low quality were based on their statistical reasoning on interpreting graphs beyond just reading the elements of the graphs and selecting appropriate graphs to the context. Also, in a few instances, students demonstrated statistical reasoning while questioning and explaining why specific statistical measures should be used, although some of the reasoning was incorrect and incomplete.

The results of the study indicate that student communication of the content was more prevalent during the lesson segments, as opposed to providing explanations, questioning, and reasoning. It is important to note that the majority of these lesson segments exhibited mid-quality student communication about the content. In other words, in both teachers' statistics instruction, the students made substantive contributions to the discussion, demonstrating a reasonable level of understanding.

However, these contributions did not fully characterize the entire segment. These findings highlight the observed IRE pattern of classroom discourse, where the teacher-initiated interactions, students responded, and the teacher evaluated or provided feedback (Cazden, 1988; Franke et al., 2017). However, student talk primarily focused on communicating the content rather than engaging in higher-order cognitive processes such as providing explanations, questioning deeply, or demonstrating reasoning skills. Also, the teachers evaluate the correctness of the responses by not mathematical and statistical questioning and reasoning (Cazden, 1988). Based on their talk about mathematics and statistics, these findings underscore the need to foster and encourage student engagement in meaning-making and reasoning.

5.2.4.2. The Types of Mathematical and Statistical Tasks That Students Work On

Mathematical tasks are the units of instruction in mathematics as they provide a medium for students to solve problems, reason, and understand mathematical concepts (Remillard, 2000). Teachers engage in the cycle of planning and implementing these tasks as part of their daily teaching practices (Georgius, 2014; Stein & Kaufman, 2010). It is through instructional tasks that teachers use to help students for learning mathematics. Tasks that demand high-level thinking processes from students encourage students to engage in mathematical and statistical processes. Therefore, one of the responsibilities of mathematics teachers is to plan and implement cognitively demanding mathematical tasks (Stein & Kaufman, 2010). In the MQI, the task cognitive demand code is a component of the student participation in meaning making and reasoning dimension; It focuses on how students engage in cognitively demanding activities while carrying out mathematical and statistical tasks rather than solely on how teachers plan, intend, and implement these tasks (Hill et al., 2014).

The study's findings reveal that in most lesson segments, students were engaged in cognitively undemanding tasks. This can be attributed to the nature of the tasks brought to the class (e.g., tasks at memorization and procedures without connection levels). In addition, the classroom discourse was predominantly teacher-centered, consisting of teacher-dominated talk with little student input. In some instances, there was evidence of unsystematic exploration, such as the graph transformation task in

Ali's instruction, engaging students cognitively in undemanding tasks. Ponte and Quaresma (2016) argue that in traditional classrooms where teachers hold the aces of the discussion through close-ended questions, undemanding low-level tasks that focus on recalling specific information or explaining the procedures without connecting to the meaning tend to dominate (Ponte & Quaresma, 2016).

Furthermore, the teaching practices observed in the study also included limited instances of low or mid-level task cognitive demand segments, where students briefly engaged in cognitively demanding activities, or their engagement was at length by starting at a high level and declining due to specific reasons such as teachers' interventions and the use of IRE discourse patterns in these segments. This decline in high cognitive demand can be attributed to various reasons, as Stein et al. (1996) found possible reasons for the decline in high cognitive demand, such as adapting challenging parts of tasks as routine because of students' requests or teachers' handling issues related to maintaining high levels of cognitive demand. Similarly, Boston and Wilhelm's (2015) study also showed that even though most tasks were found to be cognitively demanding in the planning phase, task potentials were not used in cognitively challenging ways in the implementation due to similar reasons.

As described in the MQI (2014), task cognitive demand with high scores included student engagement through "determining the meaning of mathematical concepts, processes, or relationships, drawing connections among different representations or concepts, making and test conjectures, looking for patterns, examine constraints, and explain and justify" (p.28). These examples of cognitively demanding activities align closely with the codes related to students talking about the content. As discussed earlier, the minority of lesson segments included students' statistical and mathematical reasoning, questioning, and their justifications for why questions. Especially, there were not any high-quality practices related to student mathematical and statistical questioning and reasoning in both teachers' lesson segments.

This study did not consider the task's quality as planned because it aims to uncover teaching practices as it is enacted, specifically the enacted curriculum. However, some lesson segments consisted of undemanding tasks in the enactment phase due to the

nature of tasks brought to the class. Therefore, it is important to understand task quality as planned and intended to improve teaching practices. A high-quality mathematical task considers various factors such as students' prior knowledge and experiences, grade level, and age (Stein & Smith, 1998a). It consists of significant content that is challenging and meaningful to the students and requires justifications (Boaler & Staples, 2008; Hiebert et al., 1997).

Moreover, a high-quality task offers multiple entry points, allowing students to engage in tasks from different perspectives (Lotan, 2003; Peressini et al., 2004). A good mathematical task is open-ended and supports multiple representations (Lotan, 2003). If a high-quality task is not challenging, the students reveal limited high-level thinking and reasoning (Stein & Lane, 1996). Furthermore, a high-quality task encourages learners to use multiple strategies (Stein & Lane, 1996). Therefore, including challenging aspects and encouraging multiple strategies can enhance students' engagement and mathematical thinking.

As Doyle (1988) stated, teachers need to have a goal while selecting a task, define what the task requires, and explain the process to students to reach high-quality engagement of students in cognitively demanding activities. In addition, as Hsu (2013) found out, high cognitive demand could not be maintained through close dialogue or answer-bound dialogue between teachers. Therefore, the roles of teachers in maintaining high-level tasks are essential. What affects the high-quality lesson or student work is the teachers' responsibility (Peressini et al., 2004).

Teacher-related factors have been identified as influential in the enactment of high-level cognitively demanding tasks. These factors include teacher knowledge (Charalambous, 2010; Wilhelm, 2014; Boston, 2013), teacher skills (Choppin, 2011; Foley et al., 2012), and teaching practices such as curriculum use (Boston & Wilhelm, 2017; Stein & Kaufman, 2010) and classroom discourse (Otten & Soria, 2014). The current study's results align with previous research, suggesting that teacher knowledge in mathematics and statistics, their diagnostic skills, the use of curriculum and curricular materials, and classroom discourse might be possible reasons for the low-quality teaching practices in task cognitive demand.

By integrating discipline-related practices in statistics instruction, such as tasks related to the statistical problem-solving process, productive student participation could be met. In the example of Şeker's (2018) study, the teacher presented cognitively demanding tasks that were situated into the context of a statistical investigation. The statistical question and the data set were given; organizing, analyzing, and interpreting were expected from the students. The questions in the task demanded students to use a representation to make comparisons, analyze raw data, interpret and evaluate an incorrect argument, and communicate the results. In the enactment phase, their results show that the teacher implemented what the task demanded and pressed students to make explanations and reason-why questions. Students engaged in cognitively demanding activities. In the end, Şeker's (2018) study provided some evidence that bringing and implementing cognitively demanding statistical tasks might be related to an increase in student learning. This highlights the importance of the curriculum and curricular materials for facilitating the students' engagement in challenging mathematical and statistical activities (Stein & Kaufman, 2010).

The teachers in the study used additional textbooks but benefitted from the official textbooks provided by the Ministry of Education. Yilmaz's (2022) study investigated the cognitive levels of the statistics objectives presented in the Turkish curriculum for grades 1-8, concluding that they were at the level of understanding and applying in general. When analyzing the cognitive demand of tasks in the textbooks aligned with the curriculum, it was found that a significant portion of the tasks focused on procedures, with or without connections. Specifically, most tasks in 7th-grade textbooks were found as low cognitive demands, but 36% were found to be cognitively demanding.

Yilmaz (2022) shed light on the current situation of the curriculum and curricular materials in terms of the opportunities provided to students to engage in cognitively demanding tasks by analyzing the written curriculum. The results of this study align with the findings from the literature, suggesting the need for a link between written, intended, and enacted curricula. In addition, these results highlight the importance of considering the factors affecting students' engagement in cognitively demanding activities.

In addition to task cognitive demand, this category addresses students' work with contextualized problems, including story problems and real-life applications. The activities students work on need to be contextualized and require students to use their statistical and mathematical thinking and reasoning (Hill et al., 2014). As discussed, the role of context is different in mathematics and statistics. The context in mathematics is used as a vehicle to reach an abstraction by applying mathematical thinking and reasoning to the context. For example, the teachers in the current study spared some segments for practicing the algorithms in the average concepts, as Watson et al. (2018) stated. They did not need the context while carrying out algorithms. In the current study, most lesson segments included student work with contextualized problems at a low level because they were mostly routine or heavily guided by the teacher. For example, the circle graph-related tasks consisted of heavily guided instruction at the beginning for supporting proportional reasoning, followed by routine exercises.

Context is inevitable in statistics, and real data is essential for statistical learning (Cobb & Moore, 1997). However, discipline-related practices were not the main focus of the instruction in the current study. Therefore, the students could not be engaged in contextualized problems because they did not engage in the real data from the need to answer statistical investigative questions. Weiland (2019) examined high school textbooks to determine how contextualized situations were present. Their study highlighted that the contextualized situations in the textbooks were often fictional as they emphasized not touching on societal issues as statistics are used in the real world. In the current study, the tasks brought to the class from textbooks were not primarily statistical in nature. They were not required to think about the whole process of statistical investigations, even though one element was more prevalent. Therefore, they might be related to students' engagement in the context. The limited presence of discipline-related practices and contextualized problems in the instruction might have contributed to the limited engagement of students with the context. In order to promote students' meaningful engagement with statistical concepts and their application in real-world contexts, it is essential to incorporate contextualized problems into instruction, allowing students to experience the entire process of statistical investigations.

5.2.5. Overall discussion on MQI

Overall, when considering all findings regarding the MQI, it could be concluded that the teaching practices in statistics generally lacked high-quality indicators. Instead, they could be defined as typical, including weak and robust features. At the lesson level, the overall quality of the lessons in both teachers was mainly typical, but one teacher had a higher proportion of mid to high-quality lessons. Mathematics and statistics richness were considered low to mid-quality at the segment level in both teachers' teaching practices. In contrast, the teaching practices in working with students and the content dimension were low quality. Also, student participation in meaning-making and reasoning was lacking or at low quality in general in both teachers' teaching practices. The lessons lacked mathematical and statistical errors and imprecision in most segments. These results were aligned with the studies conducted on the mathematical quality of instruction in the US in general, highlighting the typical nature of instruction (e.g., Hill & Lovison, 2022; Hill et al., 2018; Kane & Stigler, 2012).

In their study, Hill et al. (2018) described teachers at the 50% percentile in their sample could provide direct instruction in general and described into two profiles. First, their lessons included few richness elements with no mathematical errors, no challenging students to work on mathematics, and no promotion of student participation in meaning-making and reasoning. The richness elements present were generally there as routine practices. The second profile consists of both strong and weak elements in their instruction regarding the quality indicators in the MQI. They consisted of richness elements with student participation in the meaning-making process, but minor mathematical errors were present. The current results are consistent with both profiles of teachers at 50% in Hill et al.'s (2018) study. Similarly, Kane and Stigler (2012) also found similar results in the MET study, stating that their observations included teaching practices regarding the richness of mathematics, working with students, and student participation at low quality in most lessons, although most lessons included mathematical work. The results from Hill and Lovison (2022) supported that most lessons were typical in terms of the practices in the MQI dimensions.

In the Dutch context, Thurlings et al. (2019) analyzed fraction lessons based on the Quality of Instruction for Fraction (QIFI) Framework, combining both MQI and general pedagogical strategies. Their results indicate six clusters of teachers based on the QIFI: connectivist teaching, experimenting teaching, thorough teaching, autopilot teaching, teaching to the book, and ineffective teaching. The current study results show that both participating teachers could be in the groups of autopilot teaching or teaching by book clusters due to their “transmissionist nature” (p.181) and their dependence on curricular materials. Also, in both clusters, the teaching practices in the MQI were described as low quality, mainly with some weak and strong features together, especially in the Autopilot teaching cluster.

The current study indicates the need to improve teaching practices in the MQI. In a small-scale study by Park and Kim (2023), the teaching practices of two beginning teachers were analyzed over three years. It was found that the teaching practices were not improved throughout the years, except one teacher made fewer errors and imprecision in the last observation. Therefore, experience over the years does not enhance mathematical teaching practices. Instead, the results support that the root causes of low MQI needed to be addressed with joint contributions of curricular materials, MKT, teacher beliefs, as well as student-related factors (Blömeke et al., 2020; Borko et al., 2014; Charalambous & Hill, 2012; Koehler & Grouws, 1992).

Furthermore, professional development (PD) programs could contribute to improvements in instructional quality. Borko et al. (2014) showed that a mathematics PD workshop on the Problem-Solving Cycle, including a video-based discussion, improved instructional quality, especially working with students and mathematics. In addition, Gallagher et al. (2019) highlighted that pre-service mathematics teachers could develop their noticing of the high-quality teaching practices related to the richness dimension through reflective practice. Similarly, quality of instructional quality in statistics could be developed through a PD program or by reflective practice designed with the MQI elements considering the outcomes of the current study.

5.3. Implications for the Practice and Recommendations

The importance of statistics education has been emphasized in mathematics education research during the last years around the world. Being statistically literate is necessary for active citizenship (Engel, 2017). It is primarily understood that being statistically literate is curial during the COVID-19 crisis because the world is faced with “real problems involving high levels of uncertainty” (Watson & Smith, 2022, p.176). Therefore, teaching statistics becomes more essential to cope with such crises as a community (Watson & Smith, 2022).

Throughout the years, the literature has uncovered the curricular demands, students' and teachers' understanding of statistical concepts, their statistical literacy, and reasoning. Despite the critical role of teaching statistics to raise statistically literate people, as Litke and Hill (2020) stated, there is a limited study on what is happening in the classrooms during statistics instruction. The current study substantially fills the gap in the literature by investigating the statistical teaching practices of mathematics teachers in mathematics classrooms. Thus, this study extends our understanding of this issue by using two robust frameworks to evaluate teachers' statistical teaching practices and the quality of their instruction.

The findings from this study have several implications for practice. First, as discussed, the teachers covered the objective from the statistics unit as expected in the Turkish curriculum. Findings also highlighted a need for an integrated approach to mathematics and statistics instruction; through an increasing focus on big statistical ideas to provide learners more conceptual understanding of statistics within mathematics classrooms. Statistics is a discipline that can provide tools for solving problems in various disciplines, not just mathematics (Usiskin & Hall, 2015). By integrating statistics instruction into other disciplines, students can see how statistics can be used to make sense of the world around them.

Batur et al. (2021) stated that curriculum has an important place in shaping statistics instruction. Still, their study showed that the Turkish curriculum has narrower scope compared to other countries. Yılmaz (2022) also shows that with the narrow scope, most objectives did not require a high cognitive level, including evaluating and

creating. As per the findings of Batur et al. (2021) and Yılmaz (2022), a reassessment and improvement of the current mathematics curriculum could help to achieve this. It is, therefore, recommended that policymakers revisit the current curriculum and embed more statistics content to align with the goals of data and statistics education (Litke & Hill, 2020).

Second, the current research investigates how they approach the four-step model of statistical problem-solving processes as discipline-related practices. The results indicate that despite the emphasis on situating statistics education in the context of statistical problem-solving processes to solve real problems, the teachers did not use statistical investigations as a whole cycle in their practices. This detailed examination of discipline-related practices provides insights into what is going on in real classrooms despite the call for integrating statistical investigations into the practice for providing statistical learning opportunities for students.

There are possible reasons why teachers might not use statistical investigations in their practices. In addition to all possible reasons discussed, including limited SKT and lack of discipline-related beliefs, the teachers in the current study reported that they did not have time or resources to cover the whole cycle of statistical investigations. Whatever the reason, it is essential to address this issue to ensure equal opportunity for students to develop their statistical understanding. Some specific implications could be provided based on the results of this issue.

First, the findings suggest that there is a need for PD opportunities for teachers on how to integrate statistical investigations into their teachings. The PD might benefit from “a model of learning to teach statistical inquiry” proposed by Makar and Fielding-Wells (2011, p.354). With a PD, the teacher might have developed SKT and MKT specifically for teaching statistics in mathematical classrooms. For example, although there are time-related issues, as discussed, one element of the statistical problem-solving process (e.g., data analysis) could be contextualized within the process to give a sense of the experience of statistical investigations with the knowledge developed. The findings also suggest that there is a need for more resources to support teachers in using statistical investigations in their classrooms. Finally, the teachers need to be

communicated the importance and the need to address statistical investigations into their classrooms for their students to be statistically literate, active citizens.

Last, the current study evaluates the quality of statistics instruction within broader mathematics instruction using the Mathematical Quality of Instruction (MQI) instrument. This enables a comprehensive understanding of the richness of mathematics and statistics content covered in lessons, teachers' engagement with students to support their understanding, the mathematical and statistical errors made by teachers, and how students participate in meaning-making and reasoning during instruction. The use of the MQI instrument, particularly in the context of statistics instruction within broader mathematics teaching, is a novel contribution to the literature.

As Charalambous and Litke (2018) noted, the MQI provides the fine-grain analysis of instruction. Thus, the results of the study identified the areas where the teachers could improve their teaching, and it can also help teacher educators and policymakers to identify areas where they can provide support to teachers while teaching statistics. Teacher educators and policymakers can utilize the results of this study based on the specificity of the MQI codes and dimensions to prioritizing aspects of instruction that they see crucial for student learning in their context. For example, the results regarding errors and imprecision dimension inform teacher educators on where to start addressing errors and imprecision in teaching practices to create a teacher education (pre-service and in-service) to resolve them before they are reflected in the classroom practices. However, it is worth noting that MQI is not a tool for evaluating teachers; instead, it aims to understand the characteristics of high-quality instruction (Charalambous & Litke, 2018). The quality of classroom practices should not be judged solely based on MQI. Factors affecting classroom processes, like teacher MKT and SKT, teacher beliefs, and students' characteristics, should also be considered (Koehler & Grouws, 1992).

In sum, the current study fills a critical gap in the literature by providing a comprehensive and practical examination of statistical practices in the mathematics

classroom, contributing to the ongoing efforts to improve statistics instruction within broader mathematics education.

5.4. Limitations of the Study

This section provides the limitations of the study and the steps taken to mitigate the limitations. In the introduction chapter, I provided the delimitation of the studies based on the number of participants, time and place, and the data collection method. In this section, I will consider these delimitations within the perspective of limitations and discuss other study limitations.

The first possible limitation of the study stems from the number of participants, as the case study consists of two teachers' teaching practices. Recent studies on the quality of instruction within the context of the U.S. primarily relied on a higher number of participant observations (e.g., Hill et al., 2018; Kane & Stigler, 2012). For example, Hill et al.'s (2018) study consisted of 329 teachers with above 1.5 thousand videos. As a large-size study, the MET study consisted of over a thousand teachers who participated (Kane & Stigler, 2012). However, in the recent literature, small-size studies could be explicitly found in different contexts, such as South Korea or the Netherlands (Park & Kim, 2023; Thurlings et al., 2017).

One might point out the limited generalizability of the findings, but this study aims in-depth examination of statistics teaching practices within broader mathematics rather than generalizability. Therefore, this qualitative case study aims to provide rich descriptions of the teaching practices within the context of the participants involved. Still, it is important to discuss the transferability of the findings to similar contexts. As discussed in the method chapter, rich and thick descriptions of the participants and their context, the data collection, and analysis were provided to ensure transferability.

The second limitation of this study might result from "observation" as the primary data collection method. With the observation, the focus was on the enacted curriculum. Therefore, the evidence regarding the reasons behind low-quality teaching practices was not obtained directly through observations. Written and intended curriculum and

teacher-related factors, including teacher knowledge, beliefs, and skills, might have been explicitly linked to the enacted curriculum to obtain a fuller picture. In the current study, I provided some evidence related to the written and intended curriculum by collecting data through curricular sources and interviews with the teachers to mitigate this limitation. Providing in-depth contextual information related to the written and intended curricula, including factors affecting teacher knowledge and beliefs related to statistics, provides a broader picture to make sense of observational data.

Furthermore, observational data did not consider a direct measure of student-related factors, including student learning. Again, this study incorporated student-related background information, including their attitudes toward mathematics and statistics, achievement scores, and statistical literacy scores. Furthermore, classroom artifacts were collected to make sense of observational data with more detailed field notes. Through the in-depth description of student-related contextual information, I aim to mitigate possible limitations from the observational data.

Observer bias is another concern related to observation as a primary data collection method. I incorporated a reflexivity stance by taking reflective notes during the data collection process. In other parts of the research process, I defined my role as a researcher in depth. In addition, I analyzed video data through structured existing frameworks in two different time stamps. Using a high inference instrument, like the MQI, is another limitation that arose in the data analysis of observational data due to the presence of a high level of subjective judgment. I mitigated the possible limitations of observer bias and a high-inference observation instrument by incorporating peer examination. Furthermore, the presence of an observer might have influenced the behaviors of the teachers and their students, which might distort the observational data. Such possible limitation was mitigated by spending prolonged time in the field.

In the following section, I will discuss the implications of the study's findings for further research by considering implications for the practice, recommendations, and the study's limitations.

5.5. Implications for Future Research

To start with, this study investigated the enacted curriculum, which has been overlooked in the literature. Most studies focus on the written curriculum, intended curriculum, and student outcomes separately, suggesting looking back to the teaching practices and teacher-related factors (e.g., Akyüz et al., 2013; Batur et al., 2021; Jones & Jacobbe, 2014; Toluk Uçar & Akdoğan, 2009; Yanık et al., 2017). The current study looked at what is happening in real classrooms and the quality of teaching practices, which shows to what extent the possible outcomes in the written and intended curriculum and student outcomes are realized in practice. In particular, there is evidence of whether the written objectives are reflected in practice and whether the intended practices are carried out. Especially with the in-depth examination of these two frameworks, the pieces of evidence from the study which areas need improvement in the written and intended curriculum. However, it is still recommended that future research should be conducted in the field of statistics education that deals with the relationships between these written, intended, and enacted curricula holistically with considering student outcomes. Specifically, further research might utilize the fine-grained results of this study to investigate the factors contributing to the disconnect between the written, intended, and enacted curriculum and the ways to improve the alignment of the written, intended, and enacted curriculum to improve student learning of statistics.

While research exists on students' understanding of statistical concepts and processes in the literature, not enough studies investigate what and how statistical concepts are taught to students (e.g., Litke & Hill, 2020). However, Charalambous & Litke (2018) noted that observational instruments, like MQI, are often ill-suited to measure student learning directly. Therefore, future research should consider using additional measures to measure students' understanding of mathematics and statistics. This way, a comprehensive understanding of the relationship between statistical teaching practices and statistical student understanding in mathematics classrooms could be obtained.

As discussed, the results related to MQI codes informed the practice in detail about what is needed to improve to increase the quality of statistics instruction in

mathematics classrooms. Future research can benefit from each detailed result from the real classrooms to develop a PD specific to the area of development. In addition, Furtak et al. (2012) discussed the importance of learning progressions for both in-service and pre-service teacher education to support high-quality teaching practices because they included “knowledge of student ideas and how students learn, and in some cases suggest strategies or actions help students learn” (p.405). The results regarding student participation in meaning-making and reasoning dimensions provide insight into different aspects of student-content interaction in the presence of a teacher while developing a learning progression for the use of PD.

More specifically, the current study suggests addressing the link between SKT and the MQI. The MQI is constructed based on MKT; many studies showed the link between MKT and MQI. As from the discussion of statistics as a separate discipline, it is important to determine to what extent the MQI meets the needs of statistics as a discipline. The MQI provides a broader view of the quality of instruction within a mathematics class, allowing for an analysis that extends beyond just statistical instruction, thereby acknowledging that statistical instruction does not happen in isolation but is embedded within the broader teaching of mathematics. However, the current study suggests the need for a study in which teachers with high and low SKT and practicing discipline-related practices to see if MQI indicates high-quality statistical teaching practices and address the improvements in terms of the needs of statistics as a discipline.

Finally, by connecting the observational research with the current literature on mathematics and statistics education, the current study addresses the call for more research on teacher practices related to statistics instruction within mathematics classrooms. This bridging of the gap can inform teacher professional development, curriculum design, and future research aiming to enhance statistical practices in the mathematics classroom.

REFERENCES

- Ader, E. (2018). Programlardaki veri ve olasılık öğrenme alanı içeriklerine karşılaştırmalı bir bakış (A comparative look at data and probability content area in middle school mathematics programs). In M. F. Özmantar, H. Akkoç, B. Kuşdemir Kayıran & M. Özyurt (Eds.), *Ortaokul matematik öğretim programları: Tarihsel bir inceleme (A historical analysis of middle school mathematics programs)* (pp. 275-306). Pegem Yayıncılık
- Ader, E., (2016). Programlardaki veri öğrenme alanı içeriklerine bakış: Program verilerinin karşılaştırmalı incelemesi (A close look of data content area in mathematics programs: A comparative study of program data). In M. F. Özmantar, A. Öztürk & E. Bay (Eds.), *Reform ve değişim bağlamında ilkökul matematik öğretim programları (Primary school mathematics programs in a context of reform and change)* (pp. 267-292). Pegem Yayıncılık.
- Akkoç, H., & Selçuk, A. S. (2017). Dokuzuncu sınıf öğrencilerinin merkezi eğilim ve yayılım ölçüleri konusundaki informel istatistiksel çıkarımları üzerine bir inceleme. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 45(45), 1-21. <https://doi.org/10.15285/maruaebd.308618>
- Akyuz, D., Dixon, J. K., & Stephan, M. (2013). Improving the quality of mathematics teaching with effective planning practices. *Teacher Development*, 17(1), 92-106. <https://doi.org/10.1080/13664530.2012.753939>
- Alacaci, C., Lewis, S., O'Brien, G. E., & Jiang, Z. (2011). Pre-service elementary teachers' understandings of graphs. *Eurasia: Journal of Mathematics, Science and Technology Education*, 7(1), 3-14. <https://doi.org/10.12973/ejmste/75171>
- Allmond, S., & Makar, K. (2010, July). *Developing primary students' ability to pose questions in statistical investigations*. In Proceedings of the 8th international conference on teaching statistics. Voorburg, The Netherlands: International Statistical Institute.
- Arnold, P. (2008). What about the P in the PPDAC Cycle? An initial look at posing questions for statistical investigation. In M. Santillan (Ed.), *Proceedings of the 11th International Congress on Mathematical Education (ICME-11)*, (pp. 1-8). Monterrey, Mexico.
- Arnold, P. (2009, July 10-16). *Context and its role in posing investigative questions* [Paper presentation]. The Sixth International Forum for Research on Statistical Reasoning, Thinking and Literacy, Brisbane, Australia.

- Arnold, P. (2013). *Statistical investigative questions. An enquiry into posing and answering investigative questions from existing data* [Unpublished doctoral dissertation]. The University of Auckland.
- Arnold, P., & Franklin, C. (2021). Statistical investigative questions. An enquiry into posing and answering investigative questions from existing data?. *Journal of Statistics and Data Science Education*, 29(1), 122-130. <https://doi.org/10.1080/26939169.2021.1877582>
- Arnold, P., & Pfannkuch, M. (2018, July). Critiquing investigative questions. In M. A. Sorto, A. White, and L. Guyot (Eds.), *Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)*, Kyoto, Japan.
- Aydın-Güç, F., Özmen, Z. M., & Güven, B. (2022). Difficulties scatter plots pose for 11th grade students. *The Journal of Educational Research*, 115(5), 298-314. <https://doi.org/10.1080/00220671.2022.2128018>
- Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147– 168). Springer. https://doi.org/10.1007/1-4020-2278-6_7
- Balcı, M. (2023). *Ortaokul matematiğinde istatistik öğretimi: öğretim programı ve ders materyalleri açısından bir inceleme* [Unpublished master's thesis]. Hacettepe University. <https://www.openaccess.hacettepe.edu.tr/xmlui/handle/11655/29399>
- Ball, D. L., & McDiarmid, G. W. (1990). The subject matter preparation of teachers. In W. R. Houston, M. Haberman, & J. Sikula (Eds.), *Handbook of research on teacher education* (pp. 437-449). Macmillan.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389-407. <https://doi.org/10.1177/0022487108324554>
- Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *PreK-12 guidelines for assessment and instruction in statistics education II (GAISE II). A framework for statistics and data science education*. National Council of Teachers of Mathematics. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf
- Batanero, C. (2015). Understanding randomness: Challenges for research and teaching. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (pp. 34-49). Prague: European Society for Research in Mathematics Education.
- Batanero, C., Burrill, G., & Reading, C. (2011). Overview: Challenges for teaching statistics in school mathematics and preparing mathematics teachers. In C.

- Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics - Challenges for teaching and teacher education: A joint ICMI/IASE study* (pp. 407- 418). Springer. <https://doi.org/10.1007/978-94-007-1131-0>
- Batanero, C., Estepa, A., & Godino, J. (1997). Evolution of students' understanding of statistical association in a computer-based teaching environment. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 198–212). International Statistical Institute. <https://doi.org/10.52041/SRAP.96304>
- Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R., & Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematical Education in Science and Technology*, 25(4), 527-547. <https://doi.org/10.1080/0020739940250406>
- Batur, A., Elmas-Baydar, H., & Güven, B. (2019). Ortaokul matematik ders kitaplarının GAISE raporu açısından incelenmesi. In A. Baki, B. Güven, & M. Güler (Eds.), *4. Uluslararası Türk Bilgisayar ve Matematik Eğitimi Sempozyumu bildiriler e-kitabı* (pp. 10-18). TÜRKBİLMAT Eğitim Hizmetleri.
- Batur, A., Özmen, Z. M., Topan, B., Akoğlu, K., & Güven, B. (2021). A cross-national comparison of statistics curricula. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(1), 290-319. <https://doi.org/10.17762/turcomat.v12i1.278>
- Begg, A. (1997). Some emerging influences underpinning assessment in statistics. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 17-25). IOS Press.
- Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, 30(1), 3-29. <https://doi.org/10.1080/09243453.2018.1539014>
- Ben-Zvi, D., & Garfield, J. B. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-16). Kluwer academic publishers. <https://doi.org/10.1007/1-4020-2278-6>
- Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473–502). Springer. https://doi.org/10.1007/978-3-319-66195-7_16
- Berlin, R., & Cohen, J. (2018). Understanding instructional quality through a relational lens. *ZDM Mathematics Education*, 50, 367-379. <https://doi.org/10.1007/s11858-018-0940-6>

- Biehler, R. (1994). Probabilistic thinking, statistical reasoning, and the search for causes—Do we need a probabilistic revolution after we have taught data analysis? In J. Garfield (Ed.), *Research papers from the Fourth International Conference on Teaching Statistics (ICOTS 4)* (pp. 20–37). Minneapolis, MN: University of Minnesota.
- Biehler, R. (1997). Students' difficulties in practicing computer supported data analysis—Some hypothetical generalizations from results of two exploratory studies. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 169–190). ISI. <https://doi.org/10.52041/SRAP.96303>
- Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben-Zvi, K. Makar & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 139–192). Springer. https://doi.org/10.1007/978-3-319-66195-7_5
- Bilen, O. (2017). *Secondary school mathematics textbook 7*. Ankara: Gizem Publishing, ISBN: 978-975-7000-79-2.
- Blazar, D. (2015). Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review*, 48, 16-29. <https://doi.org/10.1016/j.econedurev.2015.05.005>
- Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis*, 39(1), 146–170. <https://doi.org/10.3102/0162373716670260>
- Blazar, D., & Pollard, C. (2017). Does test preparation mean low-quality instruction? *Educational Researcher*, 46(8), 420-433. <https://doi.org/10.3102/0013189X17732753>
- Blazar, D., Braslow, D., Charalambous, C. Y., & Hill, H. C. (2017). Attending to general and mathematics-specific dimensions of teaching: Exploring factors across two observation instruments. *Educational Assessment*, 22(2), 71-94. <https://doi.org/10.1080/10627197.2017.1309274>
- Blazar, D., Litke, E., & Barmore, J. (2016). What does it mean to be ranked a “high” or “low” value-added teacher? Observing differences in instructional quality across districts. *American Educational Research Journal*, 53(2), 324-359. <https://doi.org/10.3102/0002831216630407>
- Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, 79 (101600), 1-11. <https://doi.org/10.1016/j.learninstruc.2022.101600>
- Blömeke, S., Kaiser, G., König, J., & Jentsch, A. (2020). Profiles of mathematics teachers' competence and their relation to instructional quality. *ZDM*

Mathematics Education, 52, 329-342. <https://doi.org/10.1007/s11858-020-01128-y>

- Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record*, 110(3), 608–645. <https://doi.org/10.1177/016146810811000302>
- Bogdan, R.C., & Biklen, S.K. (2007). *Qualitative research for education: An introduction to theories and methods* (5th ed.). Pearson Education.
- Borko, H. (2016). Methodological contributions to video-based studies of classroom teaching and learning: A commentary. *ZDM Mathematics Education*, 48, 213-218. <https://doi.org/10.1007/s11858-016-0776-x>
- Borko, H., Eisenhart, M., Brown, C. A., Underhill, R. G., Jones, D., & Agard, P. C. (1992). Learning to teach hard mathematics: Do novice teachers and their instructors give up too easily? *Journal for Research in Mathematics Education*, 23(3), 194-222. <https://doi.org/10.5951/jresmetheduc.23.3.0194>
- Bostic, J., Lesseig, K., Sherman, M., & Boston, M. (2021). Classroom observation and mathematics education research. *Journal of Mathematics Teacher Education*, 24, 5-31. <https://doi.org/10.1007/s10857-019-09445-0>
- Boston, M. D. (2013). Connecting changes in secondary mathematics teachers' knowledge to their experiences in a professional development workshop. *Journal of Mathematics Teacher Education*, 16, 7-31. <https://doi.org/10.1007/s10857-012-9211-6>
- Boston, M. D., & Wilhelm, A. G. (2017). Middle school mathematics instruction in instructionally focused urban districts. *Urban Education*, 52(7), 829-861. <https://doi.org/10.1177/0042085915574528>
- Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics classroom observation protocols. *Mathematics Teacher Educator*, 3(2), 154-175. <https://doi.org/10.5951/mathteaceduc.3.2.0154>
- Boston, M.D., Candela, A.G. (2018). The Instructional Quality Assessment as a tool for reflecting on instructional practice. *ZDM Mathematics Education*, 50, 427–444. <https://doi.org/10.1007/s11858-018-0916-6>
- Braham, H. M., & Ben-Zvi, D. (2019). Design for reasoning with uncertainty. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research: International perspectives* (pp. 97–121). Springer International Publishing.
- Brendefur, J., & Frykholm, J. (2000). Promoting mathematical communication in the classroom: Two preservice teachers' conceptions and practices. *Journal of Mathematics Teacher Education*, 3(2), 125-153. <https://doi.org/10.1023/A:1009947032694>

- Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687-699. <https://doi.org/10.1177/001316448104100307>
- Bulut, S. (1994). *The effects of different teaching methods gender on probability achievement and attitudes toward probability*. [Unpublished doctoral dissertation]. Middle East Technical University.
- Burgess, T. (2002, July). Investigating the “data sense” of preservice teachers. In B. Phillips (Ed.), *Proceedings of the sixth international conference on teaching statistics* (pp. 1-6). International Statistical Institute and International Association for Statistics Education.
- Burgess, T. (2008). Teacher knowledge for teaching statistics through investigations. In Batanero, C., Burrill, G., Reading, C. & Rossman, A. (Eds.), *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI study 18 and 2008 IASE Round Table Conference* (pp. 1-6). ICMI and IASE. https://www.stat.auckland.ac.nz/~iase/publications/rt08/T2P12_Burgess.pdf
- Burgess, T. (2009). Statistical knowledge for teaching: Exploring it in the classroom. *For the Learning of Mathematics*, 29(3), 18-21.
- Burgess, T. A. (2011). Teacher knowledge of and for statistical investigations. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics—Challenges for teaching and teacher education* (pp. 259-270). Springer. https://doi.org/10.1007/978-94-007-1131-0_26
- Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education: A Joint ICMI/IASE Study: The 18th ICMI study* (pp. 57-69). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_10
- Bütüner, S. Ö. (2020). Türkiye’de okutulan ortaokul matematik ders kitaplarının aritmetik ortalama kavramına ilişkin öğrencilere sunduğu öğrenme fırsatları. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 11(1), 157-187. <https://doi.org/10.16949/turkbilmat.627826>
- Callingham, R., Watson, J., & Oates, G. (2019). Statistics and probability: From research to the classroom. In D. Siemon, T. Barkatsas, & R. Seah (Eds.), *Researching and using progressions (trajectories) in mathematics education* (pp. 181-204). Brill Sense.
- Callingham, R., Watson, J., & Oates, G. (2021). Learning progressions and the Australian curriculum mathematics: The case of statistics and probability. *Australian Journal of Education*, 65(3), 329-342. <https://doi.org/10.1177/000494412111036521>

- Campos, T. M. M., Cazorla, I. M., & Kataoka, V. Y. (2011). Statistics school curricula in Brazil. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education: A Joint ICMI/IASE Study: The 18th ICMI study* (pp. 5–8). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_1
- Canada, D. (2006). Elementary pre-service teachers' conceptions of variation in a probability context. *Statistics Education Research Journal*, 5(1), 36-63. <https://doi.org/10.52041/serj.v5i1.508>
- Capraro, R. M., Capraro, M. M., Parker, D., Kulm, G., & Raulerson, T. (2005). The mathematics content knowledge role in developing preservice teachers' pedagogical content knowledge. *Journal of Research in Childhood Education*, 20(2), 102-118. <https://doi.org/10.1080/02568540509594555>
- Çatman-Aksoy, E., & Işıksal-Bostan, M. (2021). Seventh graders' statistical literacy: An investigation on bar and line graphs. *International Journal of Science and Mathematics Education*, 19, 397-418. <https://doi.org/10.1007/s10763-020-10052-2>
- Cazden, C. B. (1988). *Classroom discourse: The language of teaching and learning*. Heinemann.
- Charalambous, C. Y. (2010). Mathematical knowledge for teaching and task unfolding: An exploratory study. *The Elementary School Journal*, 110(3), 247–278. <https://doi.org/10.1086/648978>
- Charalambous, C. Y., & Hill, H. C. (2012). Teacher knowledge, curriculum materials, and quality of instruction: Unpacking a complex relationship. *Journal of Curriculum Studies*, 44(4), 443-466. <https://doi.org/10.1080/00220272.2011.650215>
- Charalambous, C. Y., & Kyriakides, E. (2017). Working at the nexus of generic and content-specific teaching practices: An exploratory study based on TIMSS secondary analyses. *The Elementary School Journal*, 117(3), 423-454. <https://doi.org/10.1086/690221>
- Charalambous, C. Y., & Litke, E. (2018). Studying instructional quality by using a content-specific lens: The case of the mathematical quality of instruction framework. *ZDM Mathematics Education*, 50, 445-460. <https://doi.org/10.1007/s11858-018-0913-9>
- Charalambous, C. Y., & Praetorius, A. K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM Mathematics Education*, 50, 355-366. <https://doi.org/10.1007/s11858-018-0914-8>
- Charalambous, C. Y., Hill, H. C., & Ball, D. L. (2011). Prospective teachers' learning to provide instructional explanations: How does it look and what might it

take? *Journal of Mathematics Teacher Education*, 14, 441-463.
<https://doi.org/10.1007/s10857-011-9182-z>

- Charalambous, C. Y., Hill, H. C., & Mitchell, R. N. (2012). Two negatives don't always make a positive: Exploring how limitations in teacher knowledge and the curriculum contribute to instructional quality. *Journal of Curriculum Studies*, 44(4), 489-513. <https://doi.org/10.1080/00220272.2012.716974>
- Charles, R. I. (2005). Big ideas and understandings as the foundation for elementary and middle school mathematics. *Journal of Mathematics Education Leadership*, 7(3), 9-24.
- Chick, H. (2004). Tools for transnumeration: Early stages in the art of data representation. In Putt, I., Faragher, R., & McLean, M. (Eds.), *Mathematics Education for the Third Millennium: Towards 2010. Proceedings of the Twenty-seventh Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 167–174). Sydney, Australia: MERGA.
- Chick, H. (2018). Understanding randomness: From gimmicks to effective learning. In M. A. Sorto, A. White & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (pp. 1-8). The Netherlands: International Statistical Institute.
- Chick, H. L., & Watson, J. M. (2002). Collaborative influences on emergent statistical thinking—a case study. *The Journal of Mathematical Behavior*, 21(3), 371-400. [https://doi.org/10.1016/S0732-3123\(02\)00135-9](https://doi.org/10.1016/S0732-3123(02)00135-9)
- Chin, C., & Kayalvizhi, G. (2002). Posing questions for open investigations: What questions do pupils ask? *Research in Science & Technological Education*, 20(2), 269-287. <https://doi.org/10.1080/0263514022000030499>
- Choppin, J. (2011). The impact of professional noticing on teachers' adaptations of challenging tasks. *Mathematical Thinking and Learning*, 13(3), 175–197. <https://doi.org/10.1080/10986065.2010.495049>
- Cobb, G., & Moore, D. (1997). Mathematics, statistics and teaching. *American Mathematical Monthly*, 104(9), 801-823. <https://doi.org/10.1080/00029890.1997.11990723>
- Cohen, D. K. (2011). *Teaching and its predicaments*. Harvard University Press.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119-142. <https://doi.org/10.3102/01623737025002119>
- Cohen, J., Hutt, E., Berlin, R., & Wiseman, E. (2022). The change we cannot see: Instructional quality and classroom observation in the era of common core. *Educational Policy*, 36(6), 1261-1287. <https://doi.org/10.1177/0895904820951114>

- Coker, D. C. (2022). A thematic analysis of the structure of delimitations in the dissertation. *Interational Journal of Doctoral Studies*, 17, 141-159. <https://doi.org/10.28945/4939>
- Common Core State Standards Initiative (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices and the Council of Chief State School Officers. https://www.nctm.org/uploadedFiles/Standards_and_Positions/Common_Core_State_Standards/Math_Standards.pdf
- Conway IV, B., Gary Martin, W., Strutchens, M., Kraska, M., & Huang, H. (2019). The statistical reasoning learning environment: A comparison of students' statistical reasoning ability. *Journal of Statistics Education*, 27(3), 171-187. <https://doi.org/10.1080/10691898.2019.1647008>
- Copur-Gencturk, Y. (2012). *Teachers' mathematical knowledge for teaching, instructional practices, and student outcomes* [Unpublished doctoral dissertation]. University of Illinois at Urbana-Champaign.
- Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. <https://doi.org/10.4324/9780203939185>
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Sage Publications.
- Creswell, J. W. (2009). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson Education.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage Publications.
- Curcio, F. R. (1989). *Developing graph comprehension: Elementary and middle school activities*. National Council of Teachers of Mathematics
- Danielson, C. (2013). *The framework for teaching evaluation instrument*. Danielson Group.
- delMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79–95). Springer. https://doi.org/10.1007/1-4020-2278-6_4
- Denzin, N. K., & Lincoln, Y. S. (2017). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (pp. 1-26). Sage Publications.

- Diezmann, C. M., & Lowrie, T. (2009). The role of fluency in a mathematics item with an Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, 23(2), 167–180. <https://doi.org/10.1007/s11858-009-0212-6>
- Drageset, O. G. (2014). Redirecting, progressing, and focusing actions-a framework for describing how teachers use students' comments to work with mathematics. *Educational Studies in Mathematics*, 85(2), 281–304. <https://doi.org/10.1007/s10649-013-9515-1>
- Edwards, T. G., Özgün-Koca, A., & Barr, J. (2017). Interpretations of boxplots: Helping middle school students to think outside the box. *Journal of Statistics Education*, 25(1), 21-28. <https://doi.org/10.1007/s10649-013-9515-1>
- Eichler, A. (2011). Statistics teachers and classroom practices. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 175–186). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_19
- Eichler, A., & Erens, R. (2014). Teachers' beliefs towards teaching calculus. *ZDM Mathematics Education*, 46(4), 647–659. <https://doi.org/10.1007/s11858-014-0606-y>
- Eichler, A., & Erens, R. (2015). Domain-specific belief systems of secondary mathematics teachers. In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education* (pp. 179–200). Springer International Publishing.
- Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44–49. <https://doi.org/10.52041/serj.v16i1.213>
- Engel, J., & Sedlmeier, P. (2005). On middle-school students' comprehension of randomness and chance variability in data. *ZDM Mathematics Education*, 37, 168-177. <https://doi.org/10.1007/s11858-005-0006-4>
- English, L., & Watson, J. (2016). Making decisions with data: Are we environmentally friendly?. *Australian Primary Mathematics Classroom*, 21(2), 3-7.
- English, L., Watson, J., & Fitzallen, N. (2017). Fourth-graders' meta-questioning in statistical investigations. In A. Downton, S. Livy, & J. Hall (Eds.), *Proceedings of the 40th annual conference of the Mathematics Education Research Group of Australasia* (pp. 229–236). Melbourne, Australia: MERGA.
- Erol, E. (1989). *Prevalance and correlates of math anxiety in Turkish high school students*. [Unpublished master's thesis]. Boğaziçi University.

- Fielding-Wells, J. (2010). Linking problems, conclusions and evidence: Primary students' early experiences of planning statistical investigations. In C. Reading (Ed.), *Proceedings of the 8th international conference on teaching statistics*, Ljubljana, Slovenia. International Statistical Institute.
- Fitzallen, N., & Watson, J. (2011). Graph creation and interpretation: Putting skills and context together. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thorton (Eds.), *Proceedings of the joint conference of the Australian association of mathematics teachers and mathematics education research group of Australasia* (pp. 253–260). MERGA.
- Flick, U. (2014). Mapping the Field. In U. Flick (Eds.), *The SAGE handbook of qualitative data analysis* (pp. 3-18), SAGE Publications.
<https://doi.org/10.4135/9781446282243>
- Foley, G. D., Khoshaim, H. B., Alsaed, M., & Nihan Er, S. (2012). Professional development in statistics, technology, and cognitively demanding tasks: Classroom implementation and obstacles. *International Journal of Mathematical Education in Science and Technology*, 43(2), 177–196. <https://doi.org/10.1080/0020739X.2011.592616>
- Fraenkel, J. R., & Wallen, N. E. (2014). *How to design and evaluate research in education*. (8th ed.). New York, NY: McGraw-Hill.
- Franke, M. L., Kazemi, E., & Battey, D. (2007). Understanding teaching and classroom practice in mathematics. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256). Information Age Publishing.
- Franklin C., Bargagliotti A. E., Case C. A., Kader G. D., Schaeffer R. L., & Spangler D. A. (2015). *The statistical education of teachers*. American Statistical Association.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K–12 curriculum framework*. American Statistical Association. https://www.amstat.org/asa/files/pdfs/gaise/gaiseprek-12_full.pdf
- Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124–158. <https://doi.org/10.2307/749671>
- Friel, S., O'Connor, W., & Mamer, J. (2006). More than “meanmedianmode” and a bar graph: What’s needed to have a statistical conversation? In G. Burrill and P. Elliott (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth Yearbook* (pp. 117–137). National Council of Teachers of Mathematics

- Frischemeier, D., & Biehler, R. (2018). Preservice teachers comparing groups with tinkerplots—An exploratory laboratory study. *Statistics Education Research Journal*, 17(1), 35-60. <https://doi.org/10.52041/serj.v17i1.175>
- Frischemeier, D., & Leavy, A. (2020). Improving the quality of statistical questions posed for group comparison situations. *Teaching Statistics*, 42(2), 58-65. <https://doi.org/10.1111/test.12222>
- Frischemeier, D., Kazak, S., Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2022). International perspectives on early statistical thinking: Comparison of primary school curricula in different countries. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the gap: Empowering & educating today's learners in statistics. proceedings of the 11th international conference on teaching statistics (ICOTS11 2022)*, Rosario, Argentina. International Association for Statistical Education.
- Furtak, M. E., Thompson, J., Braaten, M., & Windschitl, M. (2012). Learning progressions to support ambitious teaching Practices. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 405-443). Brill. <https://brill.com/view/book/edcoll/9789460918247/BP000018.xml>
- Gal, I., & Garfield, J. (1997). Curricular Goals and Assessment Challenges in Statistics Education. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 1-13). AIOS Press.
- Gallagher, M. A., King, L. A., Suh, J. M., & Hargrove, D. L. (2019). The richness of mathematics noticed by teacher candidates in a professional development school model. *School-University Partnerships*, 12(1), 36-44.
- Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3), 1-12. <https://doi.org/10.1080/10691898.2002.11910676>
- Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer.
- Garfield, J., & Ben-Zvi, D. (2004). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 397-409). Kluwer academic publishers.
- Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, 4(1), 92-99. <https://doi.org/10.52041/serj.v4i1.527>
- Garfield, J., & Franklin, C. (2011). Assessment of learning, for learning, and as learning. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics—challenges for teaching and teacher education: A*

joint ICMI/IASE study: *The 18th ICMI study* (pp. 133–145). Springer Netherlands.

Gattuso, L. (2006). Statistics and mathematics. Is it possible to create fruitful links? In A. Rossman, & B. Chance (Eds.), *Proceedings of the seventh international conference on teaching statistics*, Salvador (Bahia), Brazil. International Association for Statistical Education and International Statistical Institute, http://www.stat.auckland.ac.nz/wiase/publications/17/1C2_GATT.pdf

Gattuso, L., & Ottaviani, M. G. (2011). Complementing mathematical thinking and statistical thinking in school mathematics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study* (pp. 121–132). Springer Netherlands.

Georgius, K. (2014). *Planning and enacting mathematical tasks of high cognitive demand in the primary classroom*. [Doctoral thesis, University of Nebraska-Lincoln]. Proquest Dissertations and Theses Global.

Gerstenschlager, N. E. (2019, January 2). Re-Envisioning the Mathematics Teaching Practices as the Statistics Teaching Practices. *Statistics Teacher*. <https://www.statisticteacher.org/2019/01/02/statistics-teaching-practices/>

Gfeller, M. K., Niess, M. L., & Lederman, N. G. (1999). Preservice teachers' use of multiple representations in solving arithmetic mean problems. *School Science and Mathematics*, 99(5), 250-257. <https://doi.org/10.1111/j.1949-8594.1999.tb17483.x>

Gil, E., & Ben-Zvi, D. (2011). Explanations and context in the emergence of students' informal inferential reasoning. *Mathematical Thinking and Learning*, 13(1-2), 87-108. <https://doi.org/10.1080/10986065.2011.538295>

Gómez-Blancarte, A. L. & Chávez Aguilar, R. D. (2022). The intersection of statistical literacy, reasoning, and thinking. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the gap: Empowering & educating today's learners in statistics. proceedings of the 11th international conference on teaching statistics (ICOTS11 2022)*, Rosario, Argentina. International Association for Statistical Education. DOI: 10.52041/iase.icots11.T8G2

González, M. T., & Pinto, J. (2008). Conceptions of four pre-service teachers on graphical representation. In Batanero, C., Burrill, G., Reading, C. & Rossman, A. (Eds.), *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI study 18 and 2008 IASE Round Table Conference* (pp. 1–6). ICMI and IASE. http://www.ugr.es/~icmi/iase_study

González, M. T., Espinel, M. C., & Ainley, J. (2011). Teachers' graphical competence. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school*

mathematics-Challenges for teaching and teacher Education: A joint ICMI/IASE study: The 18th ICMI study (pp. 187–197). Springer Netherlands.
https://doi.org/10.1007/978-94-007-1131-0_20

Goodchild, S. (1988). School pupils' understanding of average. *Teaching Statistics*, 10, 77-81. <https://doi.org/10.1111/j.1467-9639.1988.tb00017.x>

Gould, R., Machado, S., Ong, C., Johnson, T., Molyneux, J., Nolen, S., Tangmunarunkit, H., Trusela, L., & Zanonian, L. (2016). Teaching data science to secondary students – the mobilize introduction to data science curriculum. In J. Engel (Ed.), *Promoting understanding of statistics about society. Proceedings of the IASE Roundtable Conference* (pp. 1-12), Berlin, Germany. <https://iase-web.org/documents/papers/rt2016/Gould.pdf>

Graeber, A., & Tirosh, D. (2008). Pedagogical content knowledge: Useful concept or elusive notion. In P. Sullivan & T. Wood (Eds.), *The international handbook of mathematics teacher education* (pp. 117-132). Sense Publishers.

Graham, A. T. (1987). *Statistical investigations in the secondary school*. Cambridge University Press.

Grossman, P. L. (1995). A psychological view of teachers - teachers' knowledge. In L. W. Anderson (Ed.), *The international encyclopedia of teaching and teacher education* (Vol. 2, pp. 20-24). Pergamon Press.

Große, C. S. (2014). Mathematics learning with multiple solution methods: Effects of types of solutions and learners' activity. *Instructional Science*, 42(5), 715–745. <https://doi.org/10.1007/s11251-014-9312-y>

Große, C. S., & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction*, 16(2), 122-138. <https://doi.org/10.1016/j.learninstruc.2006.02.001>

Groth R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, 15(2), 121-145. <https://doi.org/10.1080/10986065.2013.770718>

Groth, R. (2003). High school students' levels of thinking in regard to statistical study design. *Mathematics Education Research Journal*, 15(3), 252-268. <https://doi.org/10.1007/BF03217382>

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, 38(5), 427–437. <https://doi.org/10.2307/30034960>

Groth, R. E., & Bargagliotti, A. E. (2012). GAISEing into the common core of statistics. *Mathematics Teaching in the Middle School*, 18(1), 38-45. <https://doi.org/10.5951/mathteacmiddscho.18.1.0038>

- Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning*, 8(1), 37-63. https://doi.org/10.1207/s15327833mtl0801_3
- Groth, R., & Meletiou-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. In D. Ben-Zvlim (Ed.), *International handbook of research in statistics education* (pp. 327-355). Springer. https://doi.org/10.1007/978-3-319-66195-7_10.
- Guerreiro, A., Ponte, J. P., & Serrazina, L. (2013). Mathematical communication: Teachers' recognition of the singularity of students' knowledge. In B. Ubuz, C. Haser, M.A. Mariotti (Eds.), *Proceedings of the eight congress of the European society for research in mathematics education* (pp. 3075-3084). Middle East Technical University.
- Hall, J. (2011). Engaging teachers and students with real data: Benefits and challenges. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 335-346). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_32
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S., Brown, J. L., Cappella, E., Atkins, M., Rivers, S. E., Brackett, M. A., & Hakigami, A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, 113(4), 461-487. <https://doi.org/10.1086/669616>
- Hannigan, A., Gill, O., & Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers' conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teacher Education*, 16, 427-449. <https://doi.org/10.1007/s10857-013-9246-3>
- Harradine, A., Batanero, C., & Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 235-246). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_24
- Harrison, T. R. (2020). *Decision-Making of Secondary Statistics Teachers*. [Unpublished doctoral dissertation]. North Carolina State University.
- Hedges, S., & Given, K. (2023). Addressing confirmation bias in middle school data science education. *Foundations of Data Science*, 5(2), 223-243. <https://doi.org/10.3934/fods.2021035>
- Hedges, S., Harkness, S. S., & Given, K. (2022). Attending to middle school students' understanding of study design: The often overlooked or contrived step in the statistical problem-solving process. *Middle Grades Research Journal*, 13(1), 53-65.

- Hiebert, J., & Grouws, D. A. (2007). The effect of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Wearne, D., Murray, H., & Human, P. (1997). *Making sense: Teaching and learning mathematics with understanding*. Heinemann.
- Hill, H. & Learning Mathematics for Teaching (2014). Mathematical quality of instruction (MQI): 4-Point version. <https://hu.sharepoint.com/sites/GSE-CEPR/MQI-Training/Shared%20Documents/MQI%204-Point.pdf>
- Hill, H. C., & Charalambous, C. Y. (2012a). Teacher knowledge, curriculum materials, and quality of instruction: Lessons learned and open issues. *Journal of Curriculum Studies*, 44(4), 559-576. <https://doi.org/10.1080/00220272.2012.716978>
- Hill, H. C., & Charalambous, Y. C. (2012b). Teaching (un)connected mathematics: Two teachers' enactment of the pizza problem. *Journal of Curriculum Studies*, 44(4), 467-487. <https://doi.org/10.1080/00220272.2012.716972>
- Hill, H. C., & Lovison, V. (2022). US Middle School Mathematics Instruction, 2016. <https://edworkingpapers.org/sites/default/files/ai22-618.pdf>
- Hill, H. C., Blazar, D., & Lynch, K. (2015). Resources for teaching: Examining personal and institutional predictors of high-quality instruction. *AERA Open*, 1(4), 1–23. <https://doi.org/10.1177/2332858415617703>
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511. <https://doi.org/10.1080/07370000802177235>
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012a). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, 41(2), 56-64. <https://doi.org/10.3102/0013189X12437203>
- Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., Humez, A., Litke, E. & Lynch, K. (2012b). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment*, 17(2-3), 88-106. <https://doi.org/10.1080/10627197.2012.715019>
- Hill, H. C., Litke, E., & Lynch, K. (2018). Learning lessons from instruction: Descriptive results from an observational study of urban elementary classrooms. *Teachers College Record*, 120(12), 1-46. <https://doi.org/10.1177/016146811812001207>

- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American educational research journal*, 42(2), 371-406. <https://doi.org/10.3102/00028312042002371>
- Hill, H. C., Umland, K., Litke, E., & Kapitula, L. R. (2012c). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education*, 118(4), 489-519. <https://doi.org/10.1086/666380>
- Hill, H., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard educational review*, 83(2), 371-384. <https://doi.org/10.17763/haer.83.2.d11511403715u376>
- Hsu, W. M. (2013). Examining the types of mathematical tasks used to explore the mathematics instruction by elementary school teachers. *Creative Education*, 4 (6), 396–404. <https://doi.org/10.4236/ce.2013.46056>
- Inoue, N. (2009). Rehearsing to teach: Content-specific deconstruction of instructional explanations in pre-service teacher training. *Journal of Education for Teaching*, 35(1), 47-60. <https://doi.org/10.1080/02607470802587137>
- International Data Science in Schools Project Curriculum Team. (2019). *Curriculum frameworks for Introductory Data Science*. http://idssp.org/files/IDSSP_Frameworks_1.0.pdf.
- Işık, C., Kar, T., İpek, A. S., & Işık, A. (2012). Difficulties encountered by pre-service classroom teachers in constructing stories about line graphs. *International Online Journal of Educational Sciences*, 4(3), 644-658.
- Jacobbe, T. (2007). *Elementary school teachers' understanding of essential topics in statistics and the influence of assessment instruments and a reform curriculum upon their understanding*. [Unpublished doctoral dissertation]. Clemson University.
- Jacobbe, T. (2012). Elementary school teachers' understanding of the mean and median. *International Journal of Science and Mathematics Education*, 10, 1143-1161. <https://doi.org/10.1007/s10763-011-9321-0>
- Jacobbe, T., & Carvalho, C. (2011). Teachers' understanding of averages. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 199–209). Springer Netherlands.
- Jacobbe, T., Foti, S., Case, C., & Whitaker, D. (2014, July). High school (ages 14-18) students' understanding of statistics. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth*

- International Conference on Teaching Statistics (ICOTS9)* (pp.1-6), Flagstaff, Arizona USA. International Statistical Institute.
- Jacobbe, T., & Horton, R. M. (2010). Elementary school teachers' comprehension of data displays. *Statistics Education Research Journal*, 9(1), 27-45. <https://doi.org/10.52041/serj.v9i1.386>
- Janik, T., Seidel, T., & Najvar, P. (2009). Introduction: On the power of video studies. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 7-23). Waxmann Publishing.
- Jones, D. L., & Jacobbe, T. (2014). An analysis of the statistical content in textbooks for prospective elementary teachers. *Journal of Statistics Education*, 22(3), 1-18. <https://doi.org/10.1080/10691898.2014.11889713>
- Kader, G., Jacobbe, T. & Wilson, P. D. (2013). *Developing essential understanding of statistics grades 6-8*. National Council of Teachers of Mathematics.
- Kahan, J. A., Cooper, D. A., & Bethea, K. A. (2003). The role of mathematics teachers' content knowledge in their teaching: A framework for research applied to a study of student teachers. *Journal of Mathematics Teacher Education*, 6(3), 223-252. <https://doi.org/10.1023/A:1025175812582>
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Measures of Effective Teaching Project. Bill & Melinda Gates Foundation. <https://eric.ed.gov/?id=ED540960>
- Kaplan, J. J., Fisher, D. G., & Rogness, N. T. (2009). Lexical ambiguity in statistics: What do students know about the words association, average, confidence, random and spread? *Journal of Statistics Education*, 17(3), 1-19. <https://doi.org/10.1080/10691898.2009.11889535>
- Kazak, S., & Confrey, J. (2007). Elementary school students' intuitive conceptions of random distribution. *International Electronic Journal of Mathematics Education*, 2(3), 227-244. <https://doi.org/10.29333/iejme/211>
- Kazak, S., Fujita, T., & Turmo, M. P. (2023). Students' informal statistical inferences through data modeling with a large multivariate dataset. *Mathematical Thinking and Learning*, 25(1), 23-43. <https://doi.org/10.1080/10986065.2021.1922857>
- Kazak, S., Fujita, T., & Wegerif, R. (2016). Students' informal inference about the binomial distribution of "bunny hops": A dialogic perspective. *Statistics Education Research Journal*, 15(2), 46-61. <https://doi.org/10.52041/serj.v15i2.240>
- Kazemi, E. (1998). Discourse that promotes conceptual understanding. *Teaching Children Mathematics*, 4(7), 410-415. <https://doi.org/10.5951/TCM.4.7.0410>

- Kinach, B. M. (2002). Understanding and learning-to-explain by representing mathematics: Epistemological dilemmas facing teacher educators in the secondary mathematics “methods” course. *Journal of Mathematics Teacher Education*, 5(2), 153-186. <https://doi.org/10.1023/A:1015822104536>
- Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: Investigating effects if teaching and learning in Swiss and German mathematics classrooms. In T. Janik, & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.
- Knoblauch, H., Tuma, R., & Schnettler, B. (2014). Video analysis and videography. In U. Flick (Eds.), *The SAGE handbook of qualitative data analysis* (pp.435-449). Sage Publications.
- Koehler, M.S., & Grouws, D.A., (1992). Mathematics teaching practices and their effects. In: Grouws, D.A. (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp 115–126). Macmillan
- Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals and noisy processes. *Journal for Research in Mathematics Education*, 33(4), 259-289. <https://doi.org/10.2307/749741>
- Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers’ pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM Mathematics Education*, 40, 873-892. <https://doi.org/10.1007/s11858-008-0141-9>
- Kuckartz, U., & Rädiker, S. (2019). *Analyzing Qualitative Data with MAXQDA: Text, Audio, and Video*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-15671-8>
- Kurt, G. (2015). An investigation of understanding of preservice elementary mathematics teachers (PEMT) about data displays. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the ninth congress of the European society for research in mathematics education (CERME9)* (pp. 651-657). Charles University in Prague, Faculty of Education and ERME.
- Kyriakides, E., Tsangaridou, N., Charalambous, C., & Kyriakides, L. (2018). Integrating generic and content-specific teaching practices in exploring teaching quality in primary physical education. *European Physical Education Review*, 24(4), 418-448. <https://doi.org/10.1177/1356336X16685009>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174. <https://doi.org/10.2307/2529310>
- Landtblom, K., & Sumpter, L. (2021). Teachers and prospective teachers’ conceptions about averages. *Journal of Adult Learning, Knowledge and Innovation*, 4(1), 1-8. <https://doi.org/10.1556/2059.03.2019.02>

- Larrain, M., & Kaiser, G. (2022). Interpretation of students' errors as part of the diagnostic competence of pre-service primary school teachers. *Journal für Mathematik-Didaktik*, 43(1), 39-66. <https://doi.org/10.1007/s13138-022-00198-7>
- Lavigne, N. C., & Lajoie, S. P. (2007). Statistical reasoning of middle school children engaged in survey inquiry. *Contemporary Educational Psychology*, 32, 630–666. <https://doi.org/10.1016/j.cedpsych.2006.09.001>
- Learning Mathematics for Teaching Project (LMT). (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, 14(1), 25-47. <https://doi.org/10.1007/s10857-010-9140-1>
- Leavy, A. (2015). Looking at practice: revealing the knowledge demands of teaching data handling in the primary classroom. *Mathematics Education Research Journal*, 27, 283-309. <https://doi.org/10.1007/s13394-014-0138-3>
- Leavy, A., & Frischemeier, D. (2022). Developing the statistical problem-posing and problem-refining skills of prospective teachers. *Statistics Education Research Journal*, 21(1), 1-27. <https://doi.org/10.52041/serj.v21i1.226>
- Leavy, A., & O'Loughlin, N. (2006). Preservice teachers understanding of the mean: Moving beyond the arithmetic average. *Journal of Mathematics Teacher Education*, 9, 53-90. <https://doi.org/10.1007/s10857-006-9003-y>
- Lee, H. S., Kersaint, G., Harper, S. R., Driskell, S. O., Jones, D. L., Leatham, K. R., Angotti, R. L., & Adu-Gyamfi, K. (2014). Teachers' use of transnumerations in solving statistical tasks with dynamic statistical software. *Statistics Education Research Journal*, 13(1), 25–52. <https://doi.org/10.52041/serj.v13i1.297>
- Lee, H., Mojica, G., Thrasher, E., & Baumgartner, P. (2022). Investigating data like a data scientist: key practices and processes. *Statistics Education Research Journal*, 21(2), 3. <https://doi.org/10.52041/serj.v21i2.41>
- Lee, J., & Santagata, R. (2020). A longitudinal study of novice primary school teachers' knowledge and quality of mathematics instruction. *ZDM Mathematics Education*, 52, 295-309. <https://doi.org/10.1007/s11858-019-01123-y>
- Lehrer, R., Kim, M. J., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in measuring and modeling variability. *International Journal of Computers for Mathematical Learning*, 12, 195-216. <https://doi.org/10.1007/s10758-007-9122-2>
- Leinhardt, G., & Steele, M. D. (2005). Seeing the complexity of standing to the side: Instructional dialogues. *Cognition and Instruction*, 23(1), 87-163. https://doi.org/10.1207/s1532690xci2301_4

- Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Eds.), *Problems of representations in the teaching and learning of mathematics* (pp. 33–40). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Li, K. Y., & Shen, S. M. (1992). Students' weaknesses in statistical projects. *Teaching Statistics*, 14(1), 2-8. <https://doi.org/10.1111/j.1467-9639.1992.tb00195.x>
- Lineback, J. E. (2015). The redirection: An indicator of how teachers respond to student thinking. *Journal of the Learning Sciences*, 24(3), 419-460. <https://doi.org/10.1080/10508406.2014.930707>
- Litke, E. (2020). The nature and quality of algebra instruction: Using a content-focused observation tool as a lens for understanding and improving instructional practice. *Cognition and Instruction*, 38(1), 57-86. <https://doi.org/10.1080/07370008.2019.1616740>
- Litke, E. G. (2015). *The state of the gate: A description of instructional practice in algebra in five urban districts*. [Unpublished doctoral dissertation]. Harvard University.
- Litke, E., & Hill, H. C. (2020). Exploring central tendencies: The teaching of data and statistics in elementary mathematics classrooms. *The Elementary School Journal*, 120(3), 480–506. <https://doi.org/10.1086/707007>
- Litke, E., Boston, M., & Walkowiak, T. (2021). Affordances and constraints of mathematics-specific observation frameworks and general elements of teaching quality. *Studies in Educational Evaluation*, 68, 1-17. <https://doi.org/10.1016/j.stueduc.2020.100956> Get rights and content
- Lotan, R. A. (2003). Group-worthy tasks. *Educational Leadership*, 60(6), 72-75.
- Lovett, J. N., & Lee, H. S. (2018). Preservice secondary mathematics teachers' statistical knowledge: A snapshot of strengths and weaknesses. *Journal of Statistics Education*, 26(3), 214-222. <https://doi.org/10.1080/10691898.2018.1496806>
- Lynch, K., Chin, M., & Blazar, D. (2017). Relationships between observations of elementary mathematics instruction and student achievement: Exploring variability across districts. *American Journal of Education*, 123(4), 615-646. <https://doi.org/10.1086/692662>
- MacGillivray, H., & Pereira-Mendoza, L. (2011). Teaching statistical thinking through investigative projects. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 109–120). Springer Netherlands.

- Makar, K. (2013). Predict! Teaching statistics using informal statistical inference. *Australian Mathematics Teacher*, 69(4), 34-40.
- Makar, K. (2014). Young children's explorations of average through informal inferential reasoning. *Educational Studies in Mathematics*, 86(1), 61–78. <https://doi.org/10.1007/s10649-013-9526-y>
- Makar, K. M., & Confrey, J. (2004). Secondary teachers' reasoning about comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), *The challenges of developing statistical literacy, reasoning, and thinking* (pp. 327–352). Kluwer.
- Makar, K., & Confrey, J. (2005). "Variation-talk": Articulating meaning in statistics. *Statistics Education Research Journal*, 4(1), 27-54. <https://doi.org/10.52041/serj.v4i1.524>
- Makar, K., & Fielding-Wells, J. (2011). Teaching teachers to teach statistical investigations. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 347–358). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_33
- Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1), 82-105. <https://doi.org/10.52041/serj.v8i1.457>
- Makar, K., & Rubin, A. (2018). Learning about statistical Inference. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 261–294). Springer International Publishing. https://doi.org/10.1007/978-3-319-66195-7_8
- Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, 13(1-2), 152-173. <https://doi.org/10.1080/10986065.2011.538301>
- Mantzicopoulos, P., French, B. F., & Patrick, H. (2019). The quality of mathematics instruction in kindergarten: Associations with students' achievement and motivation. *The Elementary School Journal*, 119(4), 651-676. <https://doi.org/10.1086/703176>
- MAXQDA Manual (n.d.). MAXQDA 2022 Manual. <https://www.maxqda.com/download/manuals/MAX2022-Online-Manual-Complete-EN.pdf>
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). John Wiley & Sons.

- Mesa, V., Duranczyk, I., Watkins, L., & the AI@CC Research group. (2019). The structure of EQIPM, a video coding protocol to assess the quality of community college algebra instruction. In U.T. Jankvist, M. Van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceeding of the eleventh congress of the European society for research in mathematics education* (pp. 2596-2603). Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University.
- Miller, J. (2000). *The quest for the constructivist statistics classroom: viewing practice through constructivist theory*. [Unpublished doctoral dissertation]. Ohio State University.
- Ministry of National Education [MoNE]. (2017a). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)* [Mathematics curriculum (primary and middle school grades 1–8)].
- Ministry of National Education [MoNE]. (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)* [Mathematics curriculum (primary and middle school grades 1–8)]. <http://mufredat.meb.gov.tr/Dosyalar/201813017165445-MATEMAT%C4%B0K%20%C3%96%C4%9ERET%C4%B0M%20PROGRAMI%202018v.pdf> National Council for Curriculum and Assessment.
- Ministry of National Education. [MoNE]. (2017b). General Competencies for Teaching Profession. Directorate General for Teacher Training and Development. https://oygm.meb.gov.tr/meb_iys_dosyalar/2018_06/29111119_TeachersGeneralCompetencies.pdf
- Mohr, M. J. (2006). *An assessment of middle grades preservice teachers' mathematics knowledge for teaching*. [Unpublished doctoral dissertation]. Texas A&M University.
- Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, 26, 20–39. <https://doi.org/10.2307/749226>
- Moore, D. S. (1990). Uncertainty. In L. A. Steen (Ed.). *On the shoulders of giants: New approaches to numeracy* (pp.95-137). National Academy Press
- Moore, D. S., & Cobb, G. W. (2000). Statistics and mathematics: Tension and cooperation. *The American Mathematical Monthly*, 107(7), 615-630. <https://doi.org/10.1080/00029890.2000.12005247>
- Moore, D. S., & McCabe, G. P. (1993). *Introduction to the practice of statistics* (2nd ed.). W. H. Freeman.
- Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B. P., & Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: how useful is the International System for Teacher Observation

and Feedback (ISTOF)? *ZDM Mathematics Education*, 50, 395-406.
<https://doi.org/10.1007/s11858-018-0921-9>

National Center for Teacher Effectiveness (NCTE) (n.d.). MQI training.
<https://hu.sharepoint.com/sites/GSE-CEPR/MQI-Training/Shared%20Documents/Forms/AllItems.aspx>

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2009). *Focus in high school mathematics: reasoning and sense-making*. National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematics success for all*. National Council of Teachers of Mathematics.

National Research Council [NRC]. (2001). *Adding it up: Helping children learn mathematics*. National Academies Press.

Nazlıççek, N., & Erkin, E. (2002, September 16-18). *İlköğretim matematik öğretmenleri için kısaltılmış matematik tutum ölçeği* [Paper Presentation]. V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Ankara, Turkey.
<https://toad.halileksi.net/olcek/matematik-tutum-olcegi-matt-kisaltilmis-formu-ilkogretim-icin/>

Newton, J., Dietiker, L., & Horvath, A. (2011). Statistics education in the United States: Statistical Reasoning and the Statistical Process. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 9–13). Springer Netherlands.
https://doi.org/10.1007/978-94-007-1131-0_2

Opolot-Okurut, C., & Eluk, P. O. (2011). Statistics school curricula for Uganda. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 15–19). Springer Netherlands.
https://doi.org/10.1007/978-94-007-1131-0_3

Otten, S., & Soria, V. M. (2014). Relationships between students' learning and their participation during enactment of middle school algebra tasks. *ZDM Mathematics Education*, 46, 815-827. <https://doi.org/10.1007/s11858-014-0572-4>

Öz, Ö., & Işıksal, M. (2021). Yedinci sınıf öğrencilerinin istatistiksel süreçlerinin incelenmesi. *İnönü Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, 8(15), 90-11. <https://doi.org/10.29129/inujse.836528>

- Özaltun-Celik, A. (2022). A preservice middle school mathematics teacher's knowledge of student thinking about line graphs. *Journal of Pedagogical Research*, 6(1), 231-244. <https://doi.org/10.33902/JPR.2022.164>
- Özmen Z., Baki A. (2019). 5-8. sınıf matematik öğretim programının istatistik okuryazarlığı bağlamında incelenmesi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 13(2), 1063-1082. <https://doi.org/10.17522/balikesirnef.603569>
- Park, M., & Kim, Y. (2023). Instructional quality of two beginning mathematics teachers for three years: what professional competency makes a difference? *Educational Studies in Mathematics*, 112, 505-529. <https://doi.org/10.1007/s10649-022-10195-9>
- Patahuddin, S. M., & Lowrie, T. (2019). Examining teachers' knowledge of line graph task: A case of travel task. *International Journal of Science and Mathematics Education*, 17, 781-800. <https://doi.org/10.1007/s10763-018-9893-z>
- Patton, M.Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Sage Publications.
- Peck, R., Gould, R., & Miller, S. (2013). *Developing essential understanding of statistics for teaching mathematics in grades 9–12*. National Council of Teachers of Mathematics.
- Peressini, D., Borko, H., Romagnano, L., Knuth, E., & Willis, C. (2004). A conceptual framework for learning to teach secondary mathematics: A situative perspective. *Educational Studies in Mathematics*, 56(1), 67-96. <https://doi.org/10.1023/B:EDUC.0000028398.80108.87>
- Peters, S. A. (2011). Robust understanding of statistical variation. *Statistics Education Research Journal*, 10(1), 52-88. <https://doi.org/10.52041/serj.v10i1.367>
- Pfannkuch, M. (2018). Reimagining curriculum approaches. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds), *International handbook of research in statistics education* (pp. 387–413). Springer Nature.
- Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, thinking and reasoning* (pp. 17–46). Kluwer Academic Press.
- Ponte, J. P., & Quaresma, M. (2016). Teachers' professional practice conducting mathematical discussions. *Educational Studies in Mathematics*, 93, 51-66. <https://doi.org/10.1007/s10649-016-9681-z>
- Praetorius, A. K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM Mathematics Education*, 50, 535-553. <https://doi.org/10.1007/s11858-018-0946-0>

- Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM Mathematics Education*, 50, 407-426. <https://doi.org/10.1007/s11858-018-0918-4>
- Reading, C. (2004). Student description of variation while working with weather data. *Statistics Education Research Journal*, 3(2), 84–105. <https://doi.org/10.52041/serj.v3i2.548>
- Reading, C., & Canada, D. (2011). Teachers' knowledge of distribution. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education: A joint ICM/IASE study: The 18th ICMI Study* (pp. 223–234). Springer Dordrecht.
- Reading, C., & Reid, J. (2006). An emerging hierarchy of reasoning about distribution: from a variation perspective. *Statistics Education Research Journal*, 5(2), 46-68. <https://doi.org/10.52041/serj.v5i2.500>
- Reading, C., & Reid, J. (2010). Reasoning about variation: Rethinking theoretical frameworks to inform practice. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the eighth international conference on teaching statistics (ICOTS-8), Ljubljana, Slovenia*. International Statistics Institute.
- Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201–226). Kluwer.
- Remillard, J. T. (2000). Can curriculum materials support teachers' learning? Two fourth-grade teachers' use of a new mathematics text. *The Elementary School Journal*, 100(4), 331-350. <https://doi.org/10.1086/499645>
- Remillard, J. T., & Heck, D. J. (2014). Conceptualizing the curriculum enactment process in mathematics education. *ZDM Mathematics Education*, 46, 705-718. <https://doi.org/10.1007/s11858-014-0600-4>
- Richland, L. E., Begolli, K. N., Simms, N., Frausel, R. R., & Lyons, E. A. (2017). Supporting mathematical discussions: The roles of comparison and cognitive load. *Educational Psychology Review*, 29, 41-53. <https://doi.org/10.1007/s10648-016-9382-2>
- Rodrigues, B. M. B., & da Ponte, J. P. (2022). Teacher education and didactics knowledge to teach statistics: A case study. *European Journal of Science and Mathematics Education*, 10(2), 225-243. <https://doi.org/10.30935/scimath/11717>
- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed). Sage Publications.

- Salinas-Herrera, J., & Salinas-Hernández, U. (2022). Teaching and learning the notion of normal distribution using a digital resource. *Canadian Journal of Science, Mathematics and Technology Education*, 22, 576–590. <https://doi.org/10.1007/s42330-022-00226-1>
- Sanchez, E., Borim da Silva, C., & Coutinho, C. (2011). Teachers' understanding of variation. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 211–221). Springer Netherlands.
- Santos, R., & Ponte, J. P. (2013). Prospective elementary school teachers' interpretation of central tendency measures during a statistical investigation. In B. Ubuz, C. Haser, M.A. Mariotti (Eds.), *Proceedings of the Eight Congress of the European Society for Research in Mathematics Education* (pp. 870-879). Ankara: Turkey. : Middle East Technical University
- Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics*, 102(6), 245-253. <https://doi.org/10.1111/j.1949-8594.2002.tb17883.x>
- Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM Mathematics Education*, 48, 29-40. <https://doi.org/10.1007/s11858-016-0765-0>
- Schlesinger, L., Jentsch, A., Kaiser, G., König, J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM Mathematics Education*, 50, 475-490. <https://doi.org/10.1007/s11858-018-0917-5>
- Schoenfeld, A. H. (2013). Classroom observations in theory and practice. *ZDM Mathematics Education*, 45(4), 607-621. <https://doi.org/10.1007/s11858-012-0483-1>
- Schoenfeld, A. H. (2014). What makes for powerful classrooms, and how can we support teachers in creating them? A story of research and practice, productively intertwined. *Educational Researcher*, 43(8), 404-412. <https://doi.org/10.3102/0013189X14554450>
- Schoenfeld, A. H. (2018). Video analyses for research and professional development: the teaching for robust understanding (TRU) framework. *ZDM Mathematics Education*, 50, 491-506. <https://doi.org/10.1007/s11858-017-0908-y>
- Schoenfeld, A. H., & the Teaching for Robust Understanding Project. (2016). An Introduction to the Teaching for Robust Understanding (TRU) Framework. Graduate School of Education. <http://truframework.org> or <http://map.mathshell.org/trumath.php>.

- Schukajlow, S., & Krug, A. (2014). Do multiple solutions matter? Prompting multiple solutions, interest, competence, and autonomy. *Journal for Research in Mathematics Education*, 45(4), 497-533.
- Şeker, V. (2018). Teaching average through implementing cognitively demanding statistical tasks. In M. A. Sorto, A. White & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (pp. 1-5). The Netherlands: International Statistical Institute.
- Senden, B., Nilsen, T., & Blömeke, S. (2021). Instructional quality: A review of conceptualizations, measurement approaches, and research findings. In K. Klette, M. Blikstad-Balas, & M. Tengberg (Eds.), *Ways of analyzing teaching quality. Potentials and pitfalls* (pp. 140–172). Universitetsforlaget. <https://doi.org/10.18261/9788215045054-2021-05>
- Shaughnessy, J. M. (2019a, May 5-10). *The Big Ideas in the statistics education of our students: Which ones are the biggest?* [Paper Presentation]. XV Conferencia Interamericana de Educación Matemática, Medellín, Colombia.
- Shaughnessy, J. M. (2019b). Recommendations about the big ideas in statistics education: A retrospective from curriculum and research. *Cuadernos de Investigación y Formación en Educación Matemática*, 14 (18), 44-58.
- Shaughnessy, M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1010). Information Age Publishing.
- Sherin, M. G. (2002). A balancing act: Developing a discourse community in a mathematics classroom. *Journal of Mathematics Teacher Education*, 5(3), 205–233. <https://doi.org/10.1023/A:1020134209073>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. <https://doi.org/10.3102/0013189X015002004>
- Shulman, L. S. (1998). Theory, practice, and the education of professionals. *The Elementary School journal*, 98(5), 511-526. <https://doi.org/10.1086/461912>
- Silva, A. C., Viseu, F., Aires, A. P., & Neto, T. B. (2022). Investigative activities for statistics learning with 1st grade Portuguese students. *Education Sciences*, 12(10), 665. <https://doi.org/10.3390/educsci12100665>
- Silver, E. A. , & Smith, M. S. (1996). Building discourse communities in mathematics classrooms: A worthwhile but challenging journey. In P.C: Elliott (Eds.), *Communication in mathematics, K-12 and beyond: 1996 yearbook* (pp.20-28). NCTM.

- Sleep, L., & Eskelson, S. L. (2012). MKT and curriculum materials are only part of the story: Insights from a lesson on fractions. *Journal of Curriculum Studies*, 44(4), 537-558. <https://doi.org/10.1080/00220272.2012.716977>
- Smith, M. S., & Stein, M. K. (2011). *5 practices for orchestrating productive mathematical discussions*. National Council of Teachers of Mathematics.
- Stake, R. E. (2000). Case studies. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (2nd ed.) (pp.435-454). Sage Publications.
- Stein, M. K., & Kaufman, J. H. (2010). Selecting and supporting the use of mathematics curricula at scale. *American Educational Research Journal*, 47(3), 663-693. <https://doi.org/10.3102/0002831209361210>
- Stein, M. K., & Kim, G. (2009). The role of mathematics curriculum materials in large-scale urban reform: An analysis of demands and opportunities for teacher learning. In J. T. Remillard, B. A. Herbel-Eisenmann & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 37-55). Routledge.
- 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. <https://doi.org/10.1080/1380361960020103>
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10(4), 313-340. <https://doi.org/10.1080/10986060802229675>
- Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, 33(2), 455-488. <https://doi.org/10.3102/00028312033002455>
- Stein, M. K., Remillard, J. T., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (Eds.), *Second handbook of research on mathematics teaching and learning* (pp. 319-369). Information Age Publishing.
- Stockero, S. L., Leatham, K. R., Ochieng, M. A., Van Zoest, L. R., & Peterson, B. E. (2020). Teachers' orientations toward using student mathematical thinking as a resource during whole-class discussion. *Journal of Mathematics Teacher Education*, 23, 237-267. <https://doi.org/10.1007/s10857-018-09421-0>
- Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 19(1), 64-80. <https://doi.org/10.5951/jresmetheduc.19.1.0064>

- Tarr, J. E., Chávez, Ó., 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. <https://doi.org/10.1111/j.1949-8594.2006.tb18075.x>
- Thurlings, M., Koopman, M., Den Brok, P., & Pepin, B. (2019). Portraying primary fraction teaching: A variety of mathematical richness, pedagogic strategies, and use of curriculum materials. *International Journal of Education in Mathematics, Science and Technology*, 7(2), 170-185. <https://doi.org/10.18404/ijemst.552452>
- Toluk Uçar, Z., & Akdoğan, E. N. (2009). 6-8. Sınıf öğrencilerinin ortalama kavramına yüklediği anlamlar. *İlköğretim Online*, 8(2), 391-400. <https://dergipark.org.tr/tr/pub/ilkonline/issue/8598/107023>
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
- Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, 34(1), 23-25. <https://doi.org/10.1080/00031305.1980.10482706>
- Ulusoy, F., & Çakıroğlu, E. (2013). İlköğretim matematik öğretmenlerinin histogram kavramına ilişkin kavrayışları ve bu kavramın öğretim sürecinde karşılaştıkları sorunlar. *İlköğretim Online*, 12(4), 1141-1156. <https://dergipark.org.tr/tr/pub/ilkonline/issue/8583/106614>
- Usiskin, Z., & Hall, K. (2015). The relationships between statistics and other subjects in the K-12 curriculum. *Chance*, 28(3), 4-18. <https://doi.org/10.1080/09332480.2015.1099361>
- Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2010). *Elementary and middle school mathematics: Teaching developmentally* (7th ed.). Allyn & Bacon.
- Varışlı, M. A. (2017). *Matemito: Akıllı matematik atölyem*. Arı Yayıncılık
- Visnovska, J., & Cobb, P. (2019). Supporting shifts in teachers' views of a classroom statistical activity: Problem context in teaching statistics. *Mathematical Thinking and Learning*, 21(4), 285-305. <https://doi.org/10.1080/10986065.2019.1576003>
- Walker, D. A. (2010). All possible kappa coefficient values and cell distribution combinations in a 2 x 2 matrix: The case of the small sample. *Multiple Linear Regression Viewpoints*, 36(2), 40-44.
- Walkington, C., & Marder, M. (2018). Using the UTeach Observation Protocol (UTOP) to understand the quality of mathematics instruction. *ZDM Mathematics Education*, 50, 507-519. <https://doi.org/10.1007/s11858-018-0923-7>

- Walkowiak, T. A., Berry, R. Q., Pinter, H. H., & Jacobson, E. D. (2018). Utilizing the M-Scan to measure standards-based mathematics teaching practices: Affordances and limitations. *ZDM Mathematics Education*, 50, 461-474. <https://doi.org/10.1007/s11858-018-0931-7>
- Watson, J. (2015, July 6-8). One son, five big ideas, and Tinkerplots. In N. Davis, K. Manuel, & T. Spencer, *Mathematics: Learn, lead, link. Proceedings of the 25th biennial conference of the Australian Association of Mathematics Teachers* (pp. 145-153). Adelaide, South Australia. The Australian Association of Mathematics Teachers (AAMT) Inc.
- Watson, J. (2006). *Statistical literacy at school: Growth and goals*. Lawrence Erlbaum Associates.
- Watson, J. (2014). Curriculum expectations for teaching science and statistics. . In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)* (pp.1-8), Flagstaff, Arizona USA. International Statistical Institute. http://icots.info/9/proceedings/pdfs/ICOTS9_1A1_WATSON.pdf
- Watson, J. M. (1997). Assessing statistical thinking using the media. *The Assessment Challenge in Statistics Education*, 12, 107-121.
- Watson, J. M. (2009). The influence of variation and expectation on the developing awareness of distribution. *Statistics Education Research Journal*, 8(1), 32–61. <https://doi.org/10.52041/serj.v8i1.456>
- Watson, J. M. (2016). Scaffolding statistics understanding in the middle school. In W. Widjaja, E. Yook-Kin Loong & L. A. Braag (Eds), *Maths explosion 2016: Fifty third annual conference of the mathematical association of Victoria* (pp. 132-139). The Mathematical Association Of Victoria For The Fifty Third Annual Conference. https://www.mav.vic.edu.au/Tenant/C0000019/00000001/downloads/Resources/annual-conferences/2016/2016_MAV_conference_proceedings.pdf#page=132
- Watson, J. M., & English, L. (2017a). Reaction time in grade 5: Data collection within the practice of statistics. *Statistics Education Research Journal*, 16(1), 262–293. <https://doi.org/10.52041/serj.v16i1.231>
- Watson, J. M., & English, L. D. (2017b). Statistical problem posing, problem refining, and further reflection in grade 6. *Canadian Journal of Science, Mathematics and Technology Education*, 17, 347-365. <https://doi.org/10.1080/14926156.2017.1380867>
- Watson, J. M., & Kelly, B. A. (2002). Can grade 3 students learn about variation? In B. Phillips (Ed.), *Proceedings of the sixth international conference on teaching statistics* (pp. 1-6), Cape Town. International Statistical Institute and International Association for Statistics Education. https://iase-web.org/documents/papers/icots6/2a1_wats.pdf?1402524960

- Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, 2, 11–50. https://doi.org/10.1207/S15327833MTL0202_2
- Watson, J. M., Callingham, R. A., & Kelly, B. A. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning*, 9(3), 83–130. <https://doi.org/10.1080/10986060709336812>
- Watson, J. M., Fitzallen, N., & Carter, P. (2013). Top Drawer Teachers: Statistics. Education Services Australia. <http://topdrawer.aamt.edu.au>
- Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3–46. <https://doi.org/10.52041/serj.v2i2.553>
- Watson, J., & Callingham, R. (2015). Lung disease, indigestion, and two-way tables. *Investigations in Mathematics Learning*, 8(2), 1–16. <https://doi.org/10.1080/24727466.2015.11790348>
- Watson, J., & English, L. (2015). Expectation and variation with a virtual die. *Australian Mathematics Teacher*, 71(3), 3–9.
- Watson, J., & English, L. (2015). Introducing the practice of statistics: Are we environmentally friendly?. *Mathematics Education Research Journal*, 27, 585–613. <https://doi.org/10.1007/s13394-015-0153-z>
- Watson, J., & English, L. (2017c). Eye color and the practice of statistics in grade 6: Comparing two groups. *Journal of Mathematical Behavior*, 49, 35–60. <https://doi.org/10.1016/j.jmathb.2017.06.006>
- Watson, J., & Fitzallen, N. (2019, September 19). Building understanding of randomness from ideas about variation and expectation. *Statistics Teacher*. <https://www.statisticsteacher.org/2019/09/19/building-understanding-of-randomness-from-ideas-about-variation-and-expectation/>
- Watson, J., & Fitzallen, N. (2022). Exploring the potential for student development of the big ideas of statistics with random trials: The case of the mystery spinner. In N. Fitzallen, C. Murphy, V. Hatisaru, & N. Maher (Eds.), *Proceedings of the 44th Annual Conference of the Mathematics Education Research Group of Australasia* (p. 530–537). MERGA.
- Watson, J., & Smith, C. (2022). Statistics education at a time of global disruption and crises: a growing challenge for the curriculum, classroom and beyond. *Curriculum Perspectives*, 42(2), 171–179. <https://doi.org/10.1007/s41297-022-00167-7>
- Watson, J., Fitzallen, N., English, L., & Wright, S. (2020). Introducing statistical variation in year 3 in a STEM context: Manufacturing licorice. *International*

Journal of Mathematical Education in Science and Technology, 51(3), 354–387. <https://doi.org/10.1080/0020739X.2018.1562117>

- Watson, J., Fitzallen, N., Fielding-Wells, J., & Madden, S. (2018). The practice of statistics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 105–137). Springer. https://doi.org/10.1007/978-3-319-66195-7_4
- Watson, J., Fitzallen, N., Wright, S., & Kelly, B. (2022). Characterizing student experience of variation within a STEM context: Improving catapults. *Statistics Education Research Journal*, 21(1), 1–25. <https://doi.org/10.52041/serj.v21i1.7>
- Weiland, T. (2019). The contextualized situations constructed for the use of statistics by school mathematics textbooks. *Statistics Education Research Journal*, 18(2), 18–38. <https://doi.org/10.52041/serj.v18i2.138>
- Weiss, I. R., & Pasley, J. P. (2004). What is high-quality instruction? *Educational Leadership*, 61, 24–28.
- Wessels, H. (2011). Statistics in the South African school curriculum. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: The 18th ICMI study* (pp. 21–25). Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_4
- Wild, C. J., Pfannkuch, M., Regan, M., & Horton, N. J. (2011). Towards more accessible conceptions of statistical inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(2), 247–295. <https://doi.org/10.1111/j.1467-985X.2010.00678.x>
- Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 5–36). Springer. <https://doi.org/10.1007/978-3-319-66195-7>
- Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–265. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>
- Wilhelm, A. G. (2014). Mathematics teachers' enactment of cognitively demanding tasks: Investigating links to teachers' knowledge and conceptions. *Journal for Research in Mathematics Education*, 45(5), 636–674. <https://doi.org/10.5951/jresmetheduc.45.5.0636>
- Wu, Y. (2004, July). *Singapore secondary school students' understanding of statistical graphs*. [Paper presentation]. The Tenth International Congress on Mathematics Education (ICME-10), Copenhagen, Denmark.

- Yanık, H. B., Özdemir, G., & Çevirgen, A. E. (2017). Investigating data processing related tasks in middle school mathematics textbooks. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 18(2), 45-61. <https://doi.org/10.17679/inuefd.323407>
- Yeo, J. B. W., & Toh, T. L. (2019). Big ideas in mathematics. In T. L. Toh & J. B. W. Yeo (Eds.), *Big ideas in school mathematics: Yearbook 2019 Association of Mathematics Educators* (pp. 71–94). World Scientific. https://doi.org/10.1142/9789811205385_0001
- Yin, R. K. (1981). The Case Study as a Serious Research Strategy. *Knowledge: Creation, Diffusion, Utilization*, 3(1), 97-114. <https://doi.org/10.1177/107554708100300106>
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Sage Publications.
- Yin, R. K. (2009). *Case study research: Design and methods* (5th ed.). Sage Publications.
- Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation*, 19(3), 321-332. <https://doi.org/10.1177/1356389013497081>
- Yin, R. K. (2016). *Qualitative research from start to finish* (Second edition). The Guilford Press.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.
- Yılmaz, N. (2022). Exploring pre-service mathematics teachers' knowledge of content and students through case reading and discussion. *Journal of Pedagogical Research*, 6(1), 171-195. <https://doi.org/10.33902/JPR.2022175841>
- Yılmaz, N., & Yetkin Özdemir, İ. E. (2017, August 21-25). *Pre-service middle school mathematics teachers understanding of average* [Paper presentation]. European Conference on Educational Research, Copenhagen, Denmark.
- Yılmaz, N., & Yetkin-Özdemir, İ. E. (2021). An Investigation of pre-service middle school mathematics teachers' discussion skills in the context of lesson study. *International Journal for Mathematics Teaching and Learning*, 22(2), 75-95. <https://doi.org/10.4256/ijmtl.v22i2.390>
- Yolcu, A. (2012). *An investigation of eighth grade students statistical literacy, attitudes towards statistics and their relationship* [Unpublished master's thesis]. Middle East Technical University.
- Zieffler, A., Garfield, J., & Fry, E. (2018). What is statistics education? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 37–70). Springer. https://doi.org/10.1007/978-3-319-66195-7_2

Zieffler, A., Garfield, J., DelMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40–58. <https://doi.org/10.52041/serj.v7i2.469>



APPENDICES

A. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS COMMITTEE

UYGULAMALI ETİK ARASTIRMA MERKEZİ APPLIED ETHICS RESEARCH CENTER	 ORTA DOĞU TEKNİK ÜNİVERSİTESİ MIDDLE EAST TECHNICAL UNIVERSITY
DUMLUPINAR BULVARI 06800 ÇANKAYA ANKARA/TÜRKİY T: +90 312 210 22 91 F: +90 312 210 79 59 Sayı: 28620816 / 229 www.ueam.metu.edu.tr	05 NİSAN 2018
Konu: Değerlendirme Sonucu	
Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)	
İlgi: İnsan Araştırmaları Etik Kurulu Başvurusu	
Sayın Prof.Dr. Erdiñ ÇAKIROĞLU	
Danışmanlığını yaptığınız doktora öğrencisi Vuslat ŞEKER'in " <i>Ortaokul Matematik Öğretmenlerinin Öğretim Pratikleri: İstatistiksel Görevlerin Uygulanması</i> " başlıklı araştırması İnsan Araştırmaları Etik Kurulu tarafından uygun görülerek gerekli onay 2017-EGT-012 protokol numarası ile 06.04.2018 - 30.09.2018 tarihleri arasında geçerli olmak üzere verilmiştir.	
Bilgilerinize saygılarımla sunarım.	
	
Prof. Dr. Ş. Halil TURAN Başkan V	
Prof. Dr. Ayhan SOL Üye	Prof. Dr. Ayhan Gürbüz DEMİR Üye
Doç. Dr. N. Yaşar KÖNDAKCI Üye	Doç. Dr. Zana ÇITAK Üye
Doç. Dr. Emre SELÇUK Üye	Dr. Öğr. Üyesi Pınar KAYGAN Üye

B. DATA COLLECTION TOOLS

B1. Field Notes (Adapted from Jacobbe, 2007, p.122)

Okul:		
Öğretmen:	Tarih:	Sınıf:
Tanımlayıcı Notlar:	Yansıtıcı notlar:	

B2. Interviews

B2.1. Statistical Content Interview (Adapted from Jacobbe, 2007, p.143-144)

Şimdi sana soracağım soruları kendi cümlelerin ile açıklamı istiyorum. Örneklerle açıklama yapabilirsin.

- 1- İstatistik nedir?
- 2- Veri ne demek? Veri toplama sürecinden bahseder misin biraz?
- 3- Verileri nasıl düzenleriz? Hangi tür grafikleri kullanıyoruz? Kullandığımız grafiklerin kullanım amaçlarını açıkla mısın?
- 4- Veri analizi ne demek?
 - a. Veri analizinin öneminden bahseder misin?
 - b. Veri analizi yaparken dikkate alınması gereken şeylerden bahseder misin?
 - c. Veri analizini öğretirken öğretim yaklaşımından bahseder misin?

- 5- Ortalama ne demek?
- Ortalamayı bulmak için birden fazla yolumuz var mı? Ortalama çeşitleri nelerdir?
 - Ortalama neyi ifade ediyor? Ortalamayı neden kullanıyoruz?
- 6- Aritmetik ortalama ne demek?
- Aritmetik ortalamayı nasıl buluruz? Birden fazla yolumuz var mı?
 - Aritmetik ortalama ne ifade ediyor? Aritmetik ortalamayı hesaplayarak neyi buluyoruz?
- 7- Medyan/ortanca ne demek?
- Ortancayı nasıl buluruz? Birden fazla yolumuz var mı?
 - Ortanca ne ifade ediyor? Ortancayı hesaplayarak neyi buluyoruz?
 - Medyan ve aritmetik ortalama arasındaki fark nedir?
- 8- Mod/tepe değeri ne demek?
- Modu nasıl buluruz? Birden fazla yolumuz var mı?
 - Mod ne ifade ediyor? Modu hesaplayarak neyi buluyoruz?
- 9- Neden üç çeşit ortalama kullanıyoruz?
- Hangi durumlarda A.O kullanıyoruz?
 - Hangi durumlarda ortanca/medyan kullanıyoruz?
 - Hangi durumlarda mod/tepe değeri kullanıyoruz?
- 10- Açıklık ne demek?
- Açıklığı nasıl buluruz? Birden fazla yolumuz var mı?
 - Açıklık ne ifade ediyor? Açıklığı hesaplayarak ne buluyoruz?
- 11- Uç değer (outlier) ne demek?
- Uç değeri nasıl buluruz? Birden fazla yolumuz var mı?
 - Uç değer ne ifade ediyor? Uç değer bulmak bize ne kazandırıyor?

B2.2. First Interview

Öğretmen Mülakat Soruları (İlk görüşme)

- 1- Öğretmen Geçmişi
- Eğitim geçmişiniz nedir? Hangi üniversiteden, kaç yılında mezun oldunuz?

- Kaç yıllık öğretmenlik deneyiminiz var? Kaç yıldır bu okulda çalışıyorsunuz?
- Çalıştığınız okulda kaç sınıfınız var? Haftalık kaç saat derse giriyorsunuz? Ortalama bir sınıfta kaç öğrenciniz var? Sosyoekonomik düzeyleri nasıl?
- Özel sektör deneyiminiz var mı? Dershane veya özel okulu gibi?
- Matematik öğretimine özel hizmetiçi eğitim seminerleri aldınız mı? Aldıysanız neler?
- Matematik öğretimine dair aldığınız dersler nelerdi?

2- Matematik

- Matematik öğretirken nasıl bir yol izliyorsunuz?
 - a. Herhangi bir araç ve yöntemi kullanıyor musunuz? Neler kullanıyorsunuz? Hangi araç ve ortamı daha çok kullanıyorsunuz?
- Matematiği öğretirken gerçek hayat durumlarından yararlanıyor musunuz? Yararlanıyorsanız nasıl?

3- Sınıf yönetimi

- Sınıf tartışmasını öğretiminizde kullanıyor musunuz? Kullanıyorsanız, açıklayınız.
- Sınıfta koyduğunuz kurallar var mı? Varsa nelerdir? (örneğin el kaldır, ödül veya ceza gibi sosyal kurallar veya matematiksel açıklama bekleme gibi sosyo matematiksel kurallar)

4- Öğrenciler

- Öğrencileriniz grup tartışmasına aşina mı? Grup çalışması yapıyor musunuz? Nasıl yapıyorsunuz?
- Öğrencilerinizin ders sırasında soru sormalarını nasıl karşılıyorsunuz?
 - a. Öğrencilerinizin derste problem çözümü ile ilgili fikir söylemeleri konusunda tavrınız nedir?

- b. Anlamadıkları zaman fikirlerini söylemeleri için veya soru sormaları için nasıl teşvik edersiniz?
- Çeşitli matematiksel kavramları öğrenirken öğrencilerinizin derse katılımlarını nasıl buluyorsunuz?
 - a. Matematiksel açıklama yapabiliyorlar mı? Yapabiliyorlarsa nasıl?
 - b. Yeterince aktifler mi? Sınıfta katılımı arttırmak için neler yapıyorsunuz?
 - c. Sınıfta çok konuşkan veya çok sessiz öğrencileriniz var mı? Nasıl başa çıkıyorsunuz bu durumla?
- Sınıftaki öğrencilerinizin başarı durumu nedir? Farklı becerilerde olan öğrencilerle nasıl başa çıkıyorsunuz?

B2.3. Statistics related Background Interview

Öğrenim Deneyimleri

1. Üniversitede istatistik dersi aldınız mı?
 - a- Aldığınız istatistik dersinin içeriğinde neler vardı? İstatistik ile ilgili neler öğrendiniz?
 - b- İstatistik öğretimine dair neler öğrendiniz?
 - c- Ders ile ilgili deneyimleriniz neler?
2. İstatistik öğrenirken;
 - a- Kendinizi en yeterli hissettiğiniz istatistik konusu hangisiydi?
 - b- Kendinizi en yetersiz hissettiğiniz istatistik konusu hangisiydi?
3. İstatistik öğretimine dair hizmetiçi eğitim aldınız mı? Aldıysanız içeriginden bahseder misiniz?

Öğretim Deneyimleri

4. İstatistik dersi veya istatistigin öğretimine dair ders almadıysanız bu konunun öğretimini şekillendiren unsurlar nelerdir?
 - a- Hangi kaynakları kullandınız? (Kitap, teknoloji vs.)
 - b- Bu kaynaklar öğretiminize nasıl katkı sağladı?
5. İstatistik öğretirken keyif alıyor musunuz?

- a- İstatistik öğretirken en keyif aldığınız konu veya konular hangileridir?
- b- İstatistik öğretirken en keyif almadığınız konu veya konular hangileridir?
- 6. İstatistik öğretirken kendinizi en yeterli hissettiğiniz istatistik konusu hangisi?
- 7. İstatistik öğretirken kendinizi hiç yetersiz hissettiğiniz oluyor mu?
 - a- Kendinizi en yetersiz hissettiğiniz istatistik konusu hangisi?
 - b- Bu durumla nasıl başa çıkıyorsunuz?
- 8. İstatistik öğretirken nasıl bir yol izlersiniz?
 - a- Konulara hazırlanır mısınız? Planlama için vakit ayırır mısınız?
 - b- Hangi kaynakları kullanırsınız?
 - c- İstatistik öğretirken gerçek hayat durumları kullanıyor musunuz? Nasıl? Örnek verebilir mısınız?
 - d- İstatistik öğretirken teknoloji kullanıyor musunuz?
- 9. İstatistik öğretirken en çok dikkate aldığınız şeyler neler olabilir?
- 10. İstatistik öğretirken en çok hangi konuyu öğretirken zorlanıyorsunuz?
- 11. Daha önce 7. Sınıf seviyesinde istatistik öğrettiniz mi?

Öğrenci

- 12. Öğrencilerinizin istatistik öğrenirken tepkileri neler?
 - a- İstatistik öğrenmekten keyif alıyorlar mı?
 - b- Matematik dersi ile ilişkisi hakkında ne düşünüyorlar?
 - c- Diğer derslerde gördükleri örneklerden bahsediyorlar mı?
- 13. Sizce öğrencilerin en keyif aldığı istatistik konusu hangisidir?
- 14. Öğrencilerin en çok zorlandıkları istatistik konuları hangisidir?
 - a- Öğrenciler zorlandığında siz ne yapıyorsunuz?

İstatistik ve Matematik

- 15. Sizce matematik ve istatistik birbirinden farklı mıdır? Açıklayınız.
- 16. İstatistik öğrenmek sizce önemli mi? Neden önemli?
- 17. Sizce neden istatistik öğretiyoruz?
 - a. Grafikleri neden öğrettiğimizi düşünüyoruz?
 - b. Veri analizini neden öğretiyoruz?
- 18. Öğrencilere istatistiği neden öğrendiklerini açıklıyor musunuz? Nasıl açıklıyorsunuz?

B2.4. Pre and Post Observation of Graphs and Average Content Interview (Adapted from Miller, 2000)

Sizinle öğreteceğiniz grafik/ortalama konusunu nasıl planladığınız hakkında görüşmek istiyorum. Şimdi size planlamanız ile ilgili sorular soracağım.

- 1- Grafikleri/ortalama kavramlarını nasıl öğretmeyi planlıyorsunuz?
Grafikleri/Ortalamayı öğretirken nasıl yol izleyeceksiniz?
 - a. Hangi kazanımları ele alacaksınız?
 - b. Hangi kaynakları kullanacaksınız?
 - c. Öğrenciler nasıl çalışacak? Bireysel mi? Grup çalışması mı?
- 2- Öğrencilerin grafik/ortalama konusunda olası hataları ve/veya kavram yanlışları neler olabilir?
 - a. Daire grafiğinde/çizgi grafiği /aritmetik ortalama/medyan/mod olası hatalar neler olabilir?
 - b. Bu hata ve kavram yanlışları ile nasıl başa çıkmayı düşünüyorsunuz?

Grafikler/Ortalama konusunu işlediniz. Şimdi grafikler işlediğiniz dersler üzerine görüşmek istiyorum.

- 1- Bu konunun öğretimi planladığınız gibi ilerledi mi?
- 2- Tahmin ettiğiniz hatalar ve kavram yanlışları ortaya çıktı mı? Tahmin etmediğiniz hatalar ve kavram yanlışları ortaya çıktı mı? Çıktıysa neler?
- 3- Öğrencilerin grafikleri öğrendiklerini düşünüyor musunuz?
 - a. Öğrencilerin öğrendiklerini nasıl anladınız? Birkaç örnek verebilir misiniz?
 - b. Öğrenemediğini düşündüğünüz öğrenciniz varsa sizce neden öğrenemediler?

B2.5. Last Interview

- 1- Dönemi genel olarak değerlendirebilir miyiz? Sizce nasıl geçti 7. Sınıflarla?
- 2- İstatistik öğretimi süreci nasıldı? Planladığınız gibi gitti mi?

- 3- Öğretim sürecinin sonucundabu tip örnekler vermek çok iyi oldu veya şu konuyu öğretirken şunu daha iyi yapabilirdim dediğiniz durumlar var mı?
- 4- İstatistik öğretiminizi pozitif ve/veya negatif etkileyen faktörler var mı?
- 5- Öğretmen olarak kendinizi yetersiz hissettiğiniz anlar oldu mu?
- 6- Öğrencilerin istatistik öğrendiklerini düşünüyor musun? İstatistikte en benimsedikleri ve en zorlandıkları konu hangisi/hangileri olmuş olabilir?
- 7- Sizce istatistik öğrencilerin ilgisini çekti mi? Çektiyse nasıl anladınız? Çekmediyse geriye dönük neler yapılabilirdi ilgi çekmesi için?
- 8- İstatistik kazanımlarından sonra konuya geri dönme fırsatınız oldu mu? Çeşitli kaynaklardan ödev verdiniz mi? Verdiyseniz ne tip örnekler içeriyordu?
- 9- Öğrencilerden istatistik kazanımlarına dair sorular geldi mi? Geldiyse içerikleri nasıldı?
- 10- İkinci yazılı sınavda ne tip sorular sordunuz? Öğrencilerin bu sorulardaki başarı durumunu değerlendirir misiniz?
- 11- İstatistikte var olan konuların öğretimine yönelik geri bildirimde bulunur musunuz? Bu konunun öğretiminin geliştirilmesi amaçlı neler yapılabilir?
- 12- İstatistik öğretimine yönelik bir hizmetiçi öğretim programı açılrsa katılır mısınız? İçeriğinin nasıl olmasını istersiniz?

C. TRANSCRIPT OF AN EPISODE

T: Now yes, let's open 149 and let's start the lesson. We skipped the topic, we go back there, ... we go back. 149. We see it in the area of the circle, sorry, right after the area of the circle. We start with the circle graph. Yes, we'll start with the circle graph. Let's get into some graphs now, okay? Do you remember any graphs from last year? The one you drew? The one you drew with teacher Ayşegül, the one you drew in 5th grade? [Teacher gives the floor to the student who raises his/her finger]

S: Teacher column graph

T: There was a bar graph, what other graphs did you know? Did you know a circle graph?

S: Yes [one or two people at the same time]

T: You knew the bar chart. Do you know that it has another name, circle graph?

S: Pie

T: It is called a pie chart in some places, isn't it, you saw it in primary school, you called it a pie chart. So what does this mean to us, pie chart. What is the difference from the other graphs, now we will think about these for a while [student ...] wait, we will think about them more. Now I will choose 12 people

S: How many people?

T: 12 people. ... quickly draw the color you like. Now let's draw something together, let's draw a pie chart [he had 12 students sitting by the window draw papers of different colors]... [unintelligible] someone has to take ... pink. [a student from a different region says I'll take it too]. Okay, Ali, I'll give it to you... How many people are there now? 2 4 6 8 9. 10 [Middle part to 3 people from the back] 11 take one more person, one more of you, okay 12.

S: Should we get a pen, teacher?

T: No, no, just take the pencil. We will create a circle graph together with you, we will make a very beautiful circle, only a little bit... come here, come this way a little bit, come this way a little bit. Now, those of you who have the same color should come side by side.

[students on the board looking at their colors]

T: Friends, come here a little more, don't block the board. Come, come, come, come. I want you to form a circle, a circle. Come on... Umut, stand straight, son, don't stand crooked, stand up straight. Now hold up the cards in your hand, do you think it looks like a circle graph?

S: No,

T: They don't look alike, do they? Let's make them look alike. Lift these... lift them up. You can even crouch down and lift them up so that they can see them with the board. You collapsed

S: As if praying, as if praying

T: Okay, now friends, let's think of these friends as a circle, let's see that I draw the circle with them here. I drew it. I drew it here, good.

S4: How beautiful you drew it.

T: I draw very straight. I drew it the way you are now. So how many yellows are there?

S: 12

S: 2

T: There are 2 yellows,

T: 3 green

T: There are 3 greens

S: 4

T: What is blue?

T: 4 blue

T: There are 4 blues

S: 3 pink, 3

T: 3 pinks, that's nice friends. Now friends, imagine that these friends of yours are sitting like here, but how many people are there in total?

S: 12

T: 12, right? There are 12 people. 12 cards, 12 people in total? So I will divide these 12 people, right, in the circle graph, according to the numbers here? So where do you think we should start dividing them?

S: Let's start from the middle

T: No, we will start from the center, but how will we divide it?

S: From the largest

T: Shall we start with the biggest?

S: We divide by 4

You tell me.

Let's start from T:3

T: Why from the threes [gives the student the floor by extending a finger to the student]

S6: He does half of it, ... 3 3 6

T: Why did you say 3? Hamza said, look, friends, I divided it into 2 halves, right?

T: Yes.

T: That's a quarter, right? 3 here, 3 pink ones here.

S: 3 in green.

T: 3 green ones here, are these done? [crosses out the data written on the board]

S: Done

Q: This is what's left, and I'll divide it between blue and yellow, right guys? Between blue and yellow. Well, if I divide it equally, let's do it like this, blue will come here and 2 what will come here?

S: Yellow

T: Anybody here know the angles? Raise your finger. Do I know this or that, do I know this or that, do I know that?

S: Yes, teacher

T: Do you have at least one or all of them that I know the angle of? Which one is

S7: My teacher has 2 teachers,

Q: Which one?

S7: Teacher 3 green and 3 blue

T: How many degrees?

S7: 90 degrees

T: 90 degrees, right? Because they are quarters. You could see it directly, without calculations. We know these things, don't we?

S: Yes.

T: Good. If he says let's go on with your seat, keep the papers or put them here. You were good... weren't you sitting? Now let's see, now what have we seen here? What does it mean to me that this is 90 degrees? What does it mean to me that these 3 people come to such a response or that these 3 people come to such a response?

S8: A quarter of it?

S9: The graph?

T: Tell me son

S10: Teacher, 1 fourth of the circle graph

T: It shows that it is one fourth of the circle graph, right. Well, it's good, it shows that it's a quarter, Mustafa also said that. Well, good, if I drew it like this, you remember the column graph, right? If I drew it as a column graph, I'm drawing something fast, I'm drawing it without any care. Here [Draws], let's make it like this, let's make it 3 here, let's make it 4 here, let's make it 2 here. Which ones were the 3 ones?

S: Green pink

T: Green and pink. Blue and yellow. This is a graph and this is a graph. Is there anything you can see in this one that you can't see in this one? What do you think is the difference? Why did I need such a thing? I mean, are they the same? Do you think they found the circle graph for nothing, I mean, when you look at it, does the circle graph tell you anything extra or does the bar graph tell you anything extra? Raise your fingers, let me see them.

S9: Teacher, there are angles in the circle graph.

T: There is an angle, but what does the angle give you? There is an angle

S9: Teacher field

T: You mean the field, not exactly the field. What do you think?

S10: A quarter S11: The area it covers

T: The quarter is nice, but what does it mean? If it wasn't a quarter, if it was a little bit different here?

(7.5 minute ended here)

S:11: Hocam, there, I mean, I mean. How can I explain

T: Are you talking about the area? The area it covers?

T: Yes, teacher

T: Exact... input.

S12: The next one where the exact thing causes the integer number of the column graph.

T: Do you mean the whole number?

S12: The average

T: We see the whole, not the average, right? You tell me the last

S13: The ratio of one color to another color

T: The ratio?

S13: I mean the number. Specific number

T: Comparison? I'll tell you something now. Can I easily compare this here?

S14: No

T: What?

Ö15: What are we comparing?

T: Now here, for example, I can compare which of green and pink is more or less, right? I look at the size... What this gives me is something like what Mustafa said, I can actually see how much share this has in the whole, but here I can't see the whole, I can't see all of them separately, I can't see how much share this has in the whole. I can't see it. This is the advantage of the circle graph for you, because you can see the whole in the circle graph because it can't be bigger than this, right friends, because you can see it, you can see how much share this has. You say that it is more than a quarter, okay friends, you can make comments like that. Okay, friends.

S: You see the big picture

T: You see the big picture, you see the big picture. Okay? Agreed?

S13: Understood

T: Good, the bar graph is already from last year. So what are we going to learn now?

Now we will learn that the angles here

S14: I found it, teacher

T: How did you find it?

S15: 110 and 70

T: Are you going to lift your head? 110 and 70. By eye? ... how?

S16: 90 a 90

T: Does anyone have any ideas, we'll find them together in a moment.

S: There are already 12 people, 1 circle is 360 degrees, we divide 360 by 12.

T: Do you think that's right?

S: No.

S: That's right, teacher [unintelligible]

T: He says, look, how many degrees is it here?

S: 360

T: There are 12 people, there are 12 people here, I divide them into slices, like pizza slices. I color 2 of the yellow ones. What does this look like, what subject does this look like?

S: Ratio, proportion, inverse proportion.

No, something easier.... The subject you don't like.

S: Fractions?

T: Heh, it looks like fractions, doesn't it. I divide it into 12 slices, I color 2 of them, I divide it into 12 slices, 4 of them. Is this okay, because there are 12, isn't the total number of people. That's okay. But let me ask you this Bahadır, let's say there are not 12 people but there are 180 people in total, can you divide the circle, excuse me, the circle by 180? Then you say no.

S: I have another idea. Are we going to find a proportion, teacher?

T: [warns with body language to raise your finger]. Say, say, he is looking blank right now, come on

S: I will remove it from 360, hodja

T: What

S: 180i

T: Why?

S: Since we cannot divide the result.

Q: What if it was 112?

T: We would add 360

T: Say Ayub.

S: Well, there are 12 people, hodja, are 3 of them equal to 90, hodja? Since there are 2-3-4, we add 20, hodja.

T: We add 20, does that make 110?

S: Yes, teacher

T: Why 20? Why not 25? Or can you add 15? I'll give the floor to one more person and I'll continue. Who is there that I haven't given the floor to?

S: Me.

T: Hüseyin and Hamza, Hamza tell me.

S: Shall I tell it on the blackboard?

T: Tell me.

S: 60, two of these are quarters, so half of them is 180, and when we add 2 and 4, we get 6. If we divide 180 by 6

T: You think we'll find someone

S: We find 1, and if we multiply by 4

T: You say yes. Okay, apart from that, ... what would have happened if I hadn't given it 90, if I had started the question from 0? I'm going to tell you now.

S: If my teacher is 90 out of 3

T: 90sa in 3. Which subject is this?

S: Proportion

T: Proportion, right? We are going to use proportion, let's start with the question in this book, let's see how we will set it up.

S: Teacher, with a finger ... I understand.

D. CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name:

Nationality:

Date and Place of Birth:

email:

EDUCATION

Degree	Institution	Year of Graduation
MA	Boğaziçi University	2014
	Primary Education	
BS	Boğaziçi University	2012
	Undergraduate Program in Primary Mathematics Education	

WORK EXPERIENCE

Year	Place	Enrollment
2022- 2022	Boğaziçi University	Part-time instructor
	Yeditepe University	
2018-2022	Boğaziçi University	Research Assistant
2017-2020	METU	Project Assistant
2013-2018	Boğaziçi University	Project Assistant

PUBLICATIONS

Journal Papers

Arıkan, S., Özer, F., Şeker, V., & Ertaş, G. (2020). The importance of sample weights and plausible values in large-scale assessments. *Journal of Measurement and Evaluation in Education and Psychology*, 11(1), 43-60.

Öztekin, E , Özer, F , Belin, M , Büyüksolak, A , Ertaş, G , Mutlu Gülbak, G , Pesen, M , Şeker, V , Türkmen, C . (2020). Öğretim üyeleri ve araştırma görevlileri perspektifinden fakültenin araştırma potansiyeli: Boğaziçi Üniversitesi Eğitim Fakültesi durumu [Research Potential of Faculty from the Perspectives of Faculty Members and Research Assistants: The Case of Bogaziçi University Faculty of Education]. *Boğaziçi Üniversitesi Eğitim Dergisi* , 37 (2) , 107-122

Şeker, V. & Ader, E. (2018). Profiling mathematics teachers regarding factors affecting the promotion of students' metacognition. *Boğaziçi Üniversitesi Eğitim Dergisi*, 35 (1), 51-70

Publications Published in the International Proceedings Book

Şeker, V. Teaching average through implementing cognitively demanding statistical tasks. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking Back, Looking Forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018)*, Kyoto, Japan. Voorburg, The Netherlands: International Statistical Institute

Vuslat Şeker, Engin Ader. A case study of a first-grade teacher's quality of implementation of mathematical tasks. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the 10th Congress of the European Society for Research in Mathematics Education* (pp. 3161-3168). Dublin: DCU Institute of Education and ERME.

Vuslat Şeker, Engin Ader. A study for profiling mathematics teachers regarding factors affecting the promotion of students' metacognition. Konrad Krainer; Nada Vondrova. CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Feb 2015, Prague, Czech Republic. pp.3129-3135, Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education.

Other Publications

Şeker, V. (2022, May). For A Better World Workshops Evaluation Study. The Educational Volunteers Foundation of Turkey [Türk Eğitim Gönüllüleri Vakfı (TEGV)].

Göl-Güven, M., Şeker, V., Erbil, F., & Özgünlü, M. (2021, February). Covid-19 pandemisinin aile yaşantısına yansımaları (Covid-19 Aile) rapor-3 [The reflections of Covid-19 pandemic to the family life in Turkey report 3] . https://www.researchgate.net/publication/349654953_Covid-19_Pandemisinin_Aile_Yasantisina_Yansimalari_Rapor-3_Subat_2021

Göl-Güven, M., Şeker, V., Erbil, F., Özgünlü, M., Alvan, G., & Uzunkök, B. (2020, July). Covid-19 pandemisinin aile yaşantısına yansımaları (Covid-19 Aile) rapor-2 [The reflections of Covid-19 pandemic to the family life in Turkey report 2]. <https://img1.wsimg.com/blobby/go/7cc36312-26f1-4303-a5b4-0d1598c91028/Covid-19%20Aile%20Rapor%202%20-0001.pdf>

Göl-Güven, M., Alvan, G., Erbil, F., Özgünlü, M., Şeker, V., & Uzunkök, B. (2020, July). Covid-19 pandemisinin aile yaşantısına yansımaları (Covid-19 Aile) rapor-1 [The reflections of Covid-19 pandemic to the family life in Turkey report 1]. <https://img1.wsimg.com/blobby/go/7cc36312-26f1-4303-a5b4-0d1598c91028/Covid-19%20Aile%20Rapor1%20.pdf>

Göl-Güven, M. & Şeker, V. (2020, July). For A Better World Workshops Evaluation Study. The Educational Volunteers Foundation of Turkey [Türk Eğitim Gönüllüleri Vakfı (TEGV)].

Ader, E. & Şeker, V. (2017). BÜMED Meç Okulları Matematik Öğretimi Projesi: Dostum Matematik [BÜMED MEÇ Schools Mathematics Instruction Project Handbook]. Merak Eden Çocuk Eğitim Kurumları A.Ş.

Conference Presentations

International

Şeker, V. (2021, September). Understanding basic statistical concepts: How mathematics teachers conceptualize what they teach in data handling. Paper presented at ECER 2021-Online, Geneva, Switzerland

Şeker, V. & Çakıroğlu, E.. (2019, September). Mathematics teachers' conceptualizations of average and instructional practices. Paper presented at Applied Statistics 2019, Bled, Slovenia.

Karaca, E. T., Çakıroğlu, E., & Şeker, V. (2019, September). First-grade students' problem-posing experiences. Paper presented at ECER 2019, Hamburg, Germany.

Şeker, V. (2018, September). Pre-service mathematics teachers' metacognitive activities during a problem-solving process. Paper presented at ECER 2019, Hamburg, Germany.

Şeker, V., Karaca, E. T., Sevinç, Ş., Çakıroğlu, E., & Işıksal Bostan, M. (2018, September). Developing first graders' number sense through subitizing activities. Paper presented at ECER2018, Bolzano, Italy

Şeker, V. (2018, September). A comparative study on teacher noticing in the context of algebra lesson case: Pre-service and in-service mathematics teachers. Paper presented at ECER 2018, Bolzano, Italy.

Şeker, V. (2018, July). Teaching average through implementing cognitively demanding statistical tasks. Paper presented at ICOTS10, Kyoto, Japan

Şeker V. (2017, September). Classroom teachers' motivational instructional practices in mathematics classrooms. Paper presented at ECER 2017, Copenhagen, Denmark.

Şeker, V. & Ader, E. (2017, February). A case study of a first-grade teacher's quality of implementation of mathematical tasks. Paper presented at 10th Congress of European Research in Mathematics Education, Dublin, Ireland

Şeker, V. & Çapa-Aydın, Y. (2015, September). Development and validation of a scale about teachers' practices on promotion of metacognition (TPPoM). Paper presented at ECER 2015, Budapest, Hungary

Şeker, V. & Ader, E. (2015, February). A study for profiling mathematics teachers regarding factors affecting promotion of students' metacognition. Paper presented at 9th Congress of European Research in Mathematics Education, Prague, Czech Republic

Şeker, V. & Ader, E. (2014, September). The development and validation of a scale measuring factors affecting mathematics teachers' promotion of metacognition. Paper presented at The 6th Biennial Meeting of the EARLI Special Interest Group 16 - Metacognition, İstanbul, Turkey

Şeker, V. & Ader, E. (2014, September). Self-regulation in research courses: A case study. Poster presented at The 6th Biennial Meeting of the EARLI Special Interest Group 16 -Metacognition, İstanbul, Turkey.

National

Çakıroğlu, E., Karaca, E. T., Işıksal-Bostan, M., Sevinç, Ş., & Şeker, V. (2018, April). Birinci sınıf öğrencilerinin problem kurma deneyimleri[First grade students' problem posing experiences]. Paper presented at 17. Uluslararası Sınıf Öğretmenliği Sempozyumu, Ankara, Turkey.

Acar, F., Ader, E., Pesen, M., Yörük, G., Şeker, V., & Boyacı, H. S. (2016, September). Matematik öğretmen adaylarının matematik görevlerini uygulama kaliteleri ve öğretmen bilgisi ile ilişkisi: Mikro öğretim durumu [The relationship between pre-service mathematics teachers' quality of implementation of mathematical tasks and teacher knowledge: The case of micro-teaching]. Paper presented at 12. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Trabzon, Turkey

Ader, E. & Şeker, V. (2015, May). Sınıf öğretmenlerinin matematik etkinliklerini uygulama kalitelerinin gelişiminin bir mesleki gelişim programı çerçevesinde incelenmesi [Investigation of elementary teachers' quality of implementation of

mathematical tasks: A professional development programme]. Paper Presented at 1. Türk Bilgisayar ve Matematik Eğitimi Sempozyumu, Adıyaman, Turkey.

Şeker, V., & Özel, S. (2013, June).Matematikte tutumun başarıyla ilişkisi [The relationship between attitude and success in mathematics]. Poster presented at 1. Türk Bilgisayar ve Matematik Eğitimi Sempozyumu, Trabzon, Turkey.

Başol, B., Erdik, E., & Şeker, V. (2012, April). Öğretmen adaylarının üstün yetenekli çocukların eğitimi hakkındaki bilgi ve farkındalıklarının incelenmesi [An examination of pre-service teachers' knowledge and awareness of gifted education]. Paper presented at I. Özel Eğitim Öğretmenleri Zirvesi, İstanbul, Turkey.



E. TURKISH SUMMARY / TRKE ZET

MATEMATİK RETMENLERİ İSTATİSTİĞİ NASIL RETİYOR? İSTATİSTİKSEL PROBLEM ZME SRECİ VE RETİM KALİTESİ ZERİNE KAPSAMLI BİR İNCELEME

1.Giriş

İstatistik ğretimi ve ğrenimi, son yıllarda dnya apında matematik eēitimi arařtırmalarında artan bir ilgi grmektedir (Kazak vd., 2023). İstatistik, ērencilerin nicel bilgileri iřleyerek dnyayı anlamalarına yardımcı olur (Ben-Zvi ve Garfield, 2008). İstatistik eēitimi, ērencilerin "veri daēılımlarını okumasını, analiz etmesini, eleřtirmesini ve bunlardan ıkarımlar yapmasını" amalamaktadır (Shaughnessy, 2007, s.968). Bu nedenle, dnya apındaki istatistik mfredatı, ērencilerin istatistiksel dřnme, okuryazarlık ve muhakemelerini geliřtirmek iin ok nemli olan temel fikir ve kavramlara ve istatistiksel arařtırmalara odaklanmaktadır (Ben-Zvi ve Garfield, 2008; MacGillivray ve Pereira-Mendoza, 2011). Son arařtırmalar, istatistik ēretme ve ērenmede istatistiksel arařtırmaların ve temel fikirlerin kullanılmasının, p-deēerleri ve gven aralıkları gibi teorik istatistiklerden ziyade istatistik eēitimi yoluyla dnyayı anlamlandırmaya katkıda bulunduēunu gstermektedir (Watson & Smith, 2022). Bu nedenle, bir istatistik mfredatı, ērencilerin ilk yıllardan itibaren temel kavramlarla istatistiksel uygulamalara katılmalarına izin vermelidir (Bargagliotti vd., 2020).

Yazılı mfredat, ēretmenlerin niyetleri ve uygulamaları yoluyla sınıflarda uygulanır ve ērencilerin ērenmesine yol aar (Stein vd., 2007). Ancak, matematik ēretmenleri matematik sınıflarında istatistik ēretiminde sıklıkla zorluklarla karřılařmaktadır (Litke & Hill, 2020; Eichler, 2011). İstatistiksel arařtırmalara artan vurguya raēmen, birok istatistik mfredatı ve uygulaması hala prosedrlere ve matematiksel ynlere ncelik vermekte, bu da hem ēretmenler hem de ērenciler iin

zorluklar yaratmaktadır (Makar ve Fielding-Wells, 2011; Leavy, 2015). İşlemsel becerilerde ustalaşmak, matematiksel ilişkileri ve olasılığı kavramak istatistik öğretiminin temel bileşenleri olsa da müfredat hedefleri öncelikle istatistiksel araştırmaları amaçları ve döngüsel doğaları bağlamında anlamaya odaklanmalıdır (Newton vd., 2011). Öğrenciler, istatistiksel araştırmalara katılarak istatistiksel düşünme ve muhakeme becerilerini geliştirebilir ve nihayetinde istatistik okuryazarlığına ulaşabilirler (Bargagliotti vd., 2020; Garfield ve Ben-Zvi, 2008). Bu nedenle, istatistik eğitiminde müfredat hedefleriyle uyumlu etkili öğretim uygulamalarının belirlenmesi çok önemlidir.

Sınıf içi uygulamalar, öğretmenlerin inanç ve bilgilerinden etkilenen amaçlanan müfredattan kaynaklanır (Eichler, 2011). Öğretmen yeterliliğinin bir parçası olarak öğretmenlerin inançları ve bilgileri, etkili öğretim uygulamalarını içeren yüksek kaliteli öğretimi teşvik etmede kritik faktörlerdir (Blömeke vd., 2020). Birçok çalışma, öğretmenlerin matematik inançları ve bilgileri ile sınıf uygulamalarına atıfta bulunarak yürürlüğe konan müfredat arasındaki ilişkiyi araştırmıştır (örn. Charalambous vd., 2012; Charalambous ve Hill, 2012; Hill vd., 2008; Hill ve Charalambous, 2012b). Sınıf içi uygulamalar öğretmenlerin inançları ve bilgileri tarafından şekillendirilir (Eichler, 2011). Son yıllarda, matematik eğitimi araştırmalarına benzer şekilde, istatistik eğitiminde öğretmenlerin bilgi ve inançlarını incelemeye yönelik artan bir ilgi vardır (örn. Burgess, 2011; Groth, 2007, 2013; Leavy, 2015). İlgili çalışmalar istatistik eğitimi araştırmalarına ışık tutsa da bu tür inanç ve bilgilerin istatistik sınıflarındaki öğretim uygulamalarına nasıl dönüştüğünü araştırmamaktadır ve daha iyi istatistiksel öğrenme deneyimlerini teşvik etmek için araştırılması gerekmektedir.

Sınıf uygulamalarını şekillendiren faktörlerin yanı sıra, birçok çalışma matematik öğretimi uygulamalarını keşfetmeye ve geliştirmeye odaklanmıştır (örn. Boaler ve Staples, 2008; Charalambous ve Litke, 2018; Schoenfeld, 2013). Bununla birlikte, öğrencilerin daha iyi öğrenmesini teşvik eden istatistik öğretimi uygulamaları konusunda literatür eksikliği bulunmaktadır (Eichler, 2011). Gerstenschlager (2019), yüksek kaliteli öğretimi teşvik etmek için NCTM (2014) tarafından önerilen matematik öğretim uygulamalarını tamamlayan istatistik öğretim uygulamaları önermiştir. Ancak, matematik ve istatistik öğretim uygulamalarını karşılaştırmadan

önce, istatistik eğitiminde sınıf içi uygulamaları geliştirmek için iki disiplin arasındaki benzerlik ve farklılıkları incelemek önemlidir.

İstatistik ve matematik, İlkokuldan 12. sınıfa kadar okul matematiği kapsamında ele alınsa bile ayrı disiplinlerdir. Her iki disiplinin gerekliliklerini ve özelliklerini tanımak, istatistiği etkili bir şekilde öğrenmek ve öğretmek için gereklidir (Gattuso ve Ottaviani, 2011). Temel olarak, her iki alan da bağlamın rolü, problem çözme süreci, değişkenliğin varlığı ve akıl yürütme türleri açısından farklılık göstermektedir (Burrill ve Biehler, 2011; Groth ve Meletiou-Mavrotheris, 2018). Tüm farklılıkları tanımak ve her iki disiplinin birbirine nasıl katkıda bulunduğunu göz önünde bulundurmak, istatistik ve matematikte daha iyi öğrenme çıktıları için çok önemlidir.

Gattuso ve Ottaviani (2011) okul matematiğinde istatistiğe yer vermenin avantajlarını vurgulamıştır. İlk olarak, istatistik bağlamsallaştırılır ve öğrencilere çözmeleri için gerçek problemler sunarak öğrencilerin matematik sınıflarındaki motivasyonlarını artırır. Bu nedenle öğretmenler, öğrencilerin matematiğe yönelik bu tür tutumlarını geliştirmek için istatistiğe yönelik olumlu tutumlarından faydalanabilirler. Ayrıca, süreç içerisinde toplanan veriler kullanılarak istatistiksel araştırma sorularının yanıtlanması için verilerin analiz edilmesi bulguların yorumlanması istatistiksel problem çözme sürecinin önemli bir parçasıdır. Matematik müfredat standartlarının bir parçası olarak iletişim, bulguların yorumlanmasını ve iletişimini içeren istatistiğin doğasından faydalanabilir. Ayrıca, istatistiksel problem çözme sürecinde "aritmetik işlemler, ölçüm becerileri ve kesirler, ondalık sayılar, oranlar ve yüzdelerle ilgili orantısal akıl yürütme" dahil olmak üzere belirli matematik becerileri ve anlayışları gereklidir (Gattuso, 2006; Gattuso ve Ottaviani, 2011; Watson, 2016, s.133).

Gattuso ve Ottaviani (2011) de istatistiksel süreçlerin öğrencilerin matematiksel becerilerini (örn. sınıflandırma) geliştirmedeki katkılarına odaklanmıştır. Her iki disiplinde de temsillerin kullanımı çok önemlidir. Özellikle grafiksel temsiller, istatistikte analiz için verileri temsil etmek (örn. dağılımları anlamak) ve problemlerin analizi için problemleri modellemek (örn. problemleri anlamak ve çözüme götürmek) için gereklidir. Buna ek olarak, Wild ve diğerleri (2018) istatistikte matematiksel kavramların ve soyutlamanın varlığı nedeniyle özellikle veri analizi yaparken

matematiksel düşünmenin gerekli olduğunu tartışmıştır. İstatistik, verilerdeki örüntüler ve genellemeler veya grafiklerde temsil edilen veriler gibi matematiksel uygulamalardan yararlanır (Delmas, 2004). Sonuç olarak, öğretmenler matematik derslerinde istatistiğe yer vermenin avantajlarını kabul etmeli, ancak her bir konunun kendine has özelliklerini de anlamalıdır (Gattuso & Ottaviani, 2011).

1.1. Matematiksel ve İstatistiksel Öğretim Uygulamaları

Araştırmalar, matematik için etkili öğretim uygulamalarını belirlemiştir ve öğrencilerin öğrenme çıktılarını iyileştirmek için bu uygulamaları anlamak çok önemlidir (Copur-Gencturk, 2012; Charalambous & Kyriakides, 2017). Bununla birlikte, muhtemelen istatistiğin matematiğe kıyasla benzersiz doğası nedeniyle istatistik öğretim uygulamaları üzerine araştırma eksikliği vardır (Eichler, 2011). Bu nedenle, matematik sınıflarındaki istatistik öğretimi uygulamalarının daha iyi anlaşılmasına ihtiyaç vardır. İstatistik öğretim uygulamaları, görev hedefleri, problem çözme yaklaşımları, temsiller, söylem sorunları, sorgulama yöntemleri, akıcılık oluşturma, üretken mücadele ve kanıt kullanma gibi matematik eğitimi uygulamalarından farklıdır (Gerstenschlager, 2019). İstatistikteki temel fikirleri ve istatistiksel problem çözme sürecini tanımlamak da dahil olmak üzere, matematik ve istatistik öğretim uygulamaları arasındaki benzerlik ve farklılıkları anlamak, yüksek kaliteli öğretim sağlamak için çok önemlidir.

1.1.1. İstatistikte Temel Fikirler

Okullardaki istatistik çalışmaları, verileri anlamak ve bunlardan sonuçlar çıkarmak için kavramsal bir çerçeve sağlayan, literatürde tanımlanmış ve tartışılmış bir dizi temel fikre odaklanmaktadır (Garfield ve Ben-Zvi, 2004, 2008; Shaughnessy, 2019a; Watson vd., 2013). Garfield ve Ben-Zvi (2004, 2008), öğrencilerin K-12 eğitimleri sırasında tipik olarak karşılaştıkları, veri, dağılım, eğilim, değişkenlik, modeller, ilişkilendirme, örnekler ve örnekleme ve çıkarım dahil olmak üzere istatistikteki sekiz büyük fikrin bir listesini sunmuşlardır. Watson ve diğerleri (2013) de birbiriyle ilişkili beş temel istatistiksel fikir tanımlamıştır: beklenti, varyasyon, dağılım, informal çıkarım ve rastgelelik.

Bu temel fikirlerin anlaşılması, öğrencilere istatistik alanında daha iyi öğrenme ve öğretme deneyimleri sağlamak için çok önemlidir (Watson vd., 2013). Öğrencilerin istatistiksel araştırmalar yapmayı ve anlamlı sonuçlar çıkarmayı öğrenmeleri için yalnızca prosedürler yerine büyük fikirlerin gerekli olduğu savunulmaktadır (Callingham vd., 2021). Shaughnessy (2019a), istatistikteki temel fikirlerin bir listesini oluşturmanın, öğrencilerin istatistiksel kavram ve süreçlere ilişkin anlayışlarını düzenlemelerine yardımcı olabileceğini öne sürmektedir. Callingham ve diğerleri (2021) de istatistik öğretimi ve öğrenimini geliştirmek için akıcılık, muhakeme, problem çözme ve anlama gibi matematiksel yeterlilikleri kullanarak anahtar kavramları veya fikirleri dikkate alarak istatistikte bir öğrenme ilerlemesine dayalı bir müfredatı işaret etmektedir. Sonuç olarak, öğrencilerin istatistik okuryazarı olabilmeleri için istatistikteki büyük fikirleri anlamaları şarttır. Bu nedenle, öğretim uygulamalarının istatistikteki temel fikirlere göre şekillendirilmesi gerekir.

1.1.2. İstatistiksel Problem Çözme Süreci

İstatistik eğitiminde araştırma yaklaşımına verilen önem giderek daha fazla vurgu yapılmaktadır. Bu yaklaşım, istatistikteki temel fikirleri geliştirmek ve çevrelerindeki dünyayı daha derinlemesine anlamak için öğrencileri eksiksiz istatistiksel araştırmalara veya istatistiksel araştırmalardaki belirli adımları dahil etmeyi içerir (Watson & Callingham, 2015; Wild & Pfannkuch, 1999; Pfannkuch, 2018). Öğrencileri soru sormaya, veri toplamaya, sonuçları analiz edip yorumlamaya ve sonuç çıkarmaya teşvik eden araştırma yaklaşımı, öğrencilerin istatistiksel düşünme, muhakeme ve okuryazarlık becerilerini geliştirmelerine yardımcı olur.

Ampirik Araştırmada İstatistiksel Düşünme için Bir Çerçeve, Wild ve Pfannkuch (1999) tarafından geliştirilmiştir ve Araştırma Döngüsü bu çerçevenin bir parçasıdır. Araştırma Döngüsü, PPDAC döngüsünü (problem, plan, veri, analiz ve sonuçlar) içerir ve istatistiksel problem çözme sürecine sistematik bir yaklaşım sağlar. Amerikan İstatistik Derneği (ASA) tarafından geliştirilen GAISE (İstatistik Eğitiminde Değerlendirme ve Öğretim Rehberi) çerçevesi iki boyutlu bir müfredat çerçevesi içermektedir: ilk boyut istatistiksel problem çözme sürecinin aşamalarını içermektedir (Bargagliotti vd., 2020, s.13):

- (1) İstatistiksel olarak araştırma sorularının formüle edilmesi
- (2) Verilerin toplanması/düşünülmesi
- (3) Verilerin analiz edilmesi
- (4) Sonuçların yorumlanması

İkinci boyut, istatistik okuryazarlığının önemini vurgulayarak, öğrencilerin istatistiksel problem çözme süreci boyunca temel istatistik becerilerini geliştirmelerine odaklanan üç gelişim düzeyinden (A, B ve C) oluşmaktadır. GAISE çerçevesi, K-12 ortamlarında istatistik eğitiminin öğretilmesi ve değerlendirilmesine rehberlik etmek üzere tasarlanmış bir müfredat belgesidir. Litke ve Hill (2020), istatistik öğretiminde disiplinle ilgili uygulamaları araştırmak için GAISE çerçevesini kullanarak bir protokol geliştirmiştir. Protokol, GAISE çerçevesi tarafından önerilen uygulamaların artan varlığına sahip üç seviye ile araştırma sürecinin dört bileşenini içerir. Protokol özellikle ilkökul istatistik içeriğine uygundur ve gelişimsel seviye A'ya odaklanmaktadır.

Araştırmalar, sınıf içi uygulamalarda istatistiksel problem çözmesi sürecine yaklaşmanın iki olası yolunu göstermiştir (Burgess, 2011). Birincisi, öğrencilere keşfetmeleri için veri vermeyi ve veri analizi ve yorumlama adımlarına odaklanmayı içerir (örneğin, Akkoç ve Selçuk, 2017; Chick ve Watson, 2002; Makar, 2014; Makar ve Confrey, 2005; Watson ve Moritz, 2000). İkincisi ise istatistiksel araştırma sorularıyla başlayıp araştırma döngüsünün diğer aşamalarını takip etmeyi içerir (örn. Lavigne ve Lajoie, 2007; Öz ve Işıksal Bostan, 2021; Silva vd., 2022; Watson ve English, 2015; Watson ve English, 2017b). Her iki yaklaşım da öğrencileri istatistiksel düşünme, akıl yürütme ve okuryazarlıklarını geliştirmek için istatistiksel problem çözme sürecine dahil eder (Burgess, 2011; Litke ve Hill, 2020). Genel olarak, araştırmacı bir yaklaşım, öğrencilerin istatistiksel kavramlar ve bunların gerçek dünyadaki uygulamaları hakkında daha derin bir anlayış geliştirmelerine yardımcı olur.

1.1.3. Matematik Eğitiminde Öğretim Kalitesi

Matematik eğitiminde öğretim kalitesi, öğrencilerin matematik öğrenimi için çok önemlidir (Griffin vd., 2009) ve öğretmenin sınıftaki gözlemlenebilir davranışları ile tanımlanır (Blömeke vd., 2022). İyi bir matematik öğretimi, akıl yürütme, anlama,

akıcılık ve problem çözme gibi matematiksel yeterliklerin gelişimini destekleyen çeşitli bileşenlerden oluşur (Callingham vd., 2021). Bununla birlikte, iyi öğretim uygulamalarını neyin oluşturduğu ve bunlara nasıl ulaşılacağı konusunda literatürde tartışılmaktadır (Fraenke vd., 2007; Schoenfeld, 2014). Öğretim kalitesini anlamamanın temel amacı, daha iyi öğrenci öğrenimi için fırsatlar yaratmak üzere öğretimi geliştirmektir. Dolayısıyla, öğretim uygulamalarını ölçmek ve anlamak bu hedefe ulaşmak için gereklidir (Blömeke vd., 2022; Bell vd., 2019; Hill ve Grossman, 2013; Schlesinger ve Jentsch, 2016; LMT, 2011; Schoenfeld, 2014; Stein ve Lane, 1996).

Matematiksel Öğretim Kalitesi (MQI) Çerçevesi, matematikte içeriğe özgü öğretim uygulamalarının kalitesini anlayarak matematik öğretimini geliştirmek için geliştirilmiştir. MQI Çerçevesi, öğretimi Cohen ve diğerleri (2003) tarafından öğretim üçgeni ile tanımlanan öğretmen, öğrenci ve içerik arasındaki etkileşim olarak görmektedir. Bu çerçeve, dört boyutta belirli öğretim uygulamalarını tanımlayan 21 bölüm düzeyinde kod aracılığıyla matematik öğretimine kapsamlı bir bakış sağlar: matematiğin zenginliği, öğrencilerle ve matematikle çalışma, hatalar ve yanlışlıklar ve öğrencilerin anlam oluşturma ve akıl yürütmeye katılımı. MQI çerçevesi, matematik eğitimcilerinin okul matematiğindeki sınıf uygulamalarını değerlendirmelerine yardımcı olarak matematik öğretimini anlamak ve geliştirmek için kapsamlı bir yaklaşım sağlar (Charalambous & Litke, 2018; Charalambous & Praetorius, 2018).

Sonuç olarak, literatür istatistik eğitiminde gerçek sınıf uygulamalarını inceleyen sınırlı sayıda çalışma göstermektedir (Eichler, 2011; Litke & Hill, 2020). İstatistiğin okul matematiğindeki yeri ve iki alanın iş birliği nedeniyle istatistik öğretimi hem disiplinle ilgili uygulamaları (Litke ve Hill, 2020) hem de daha geniş matematik içinde konuya özgü uygulamaları gerektirmektedir (Delmas, 2004; Gattuso, 2006; Gattuso ve Ottaviani, 2011; Moore ve Cobb, 2000; Wild vd., 2018). İstatistik ve matematik arasındaki benzerlikler ve farklılıklar, öğretim uygulamaları ve matematik eğitiminde öğretim kalitesi de dahil olmak üzere istatistik eğitimindeki güncel konuların gözden geçirilmesine dayanarak, aşağıdaki bölüm mevcut araştırmanın amacını ve önemini özetleyecektir.

1.2.Çalışmanın amacı

Bu nitel durum çalışması, matematik öğretmenlerinin matematik sınıflarındaki istatistik öğretimi uygulamalarını anlamayı amaçlamaktadır. Bu çalışma, matematik sınıflarındaki istatistik öğretimi uygulamalarını araştırmak için iki çerçeve kullanmaktadır. İlk çerçeve, GAISE Çerçevesini kullanan Transkript Kodlama Protokolü ile öğretmenlerin istatistiksel problem çözme süreçlerinin dört aşamalı modelini kullanıp kullanmadıklarını araştırmaktadır. İkinci çerçeve ise, daha geniş kapsamlı matematik eğitiminde istatistik öğretim uygulamalarının kalitesini ortaya koymak için Matematiksel Öğretim Kalitesi (MQI) aracını kullanmaktadır.

Çalışma, matematik öğretmenlerinin istatistiksel uygulamaları ve matematik sınıflarındaki öğretimlerinin kalitesi hakkında kapsamlı bir anlayış sağlamak için her iki çerçeveyi de kullanmaktadır. Çalışma özellikle, öğretmenlerin uyguladıkları istatistiksel içeriği ve istatistiksel problem çözme süreçlerinin dört aşamalı modeline nasıl yaklaştıklarını inceleyecektir. Buna ek olarak, öğretimin kalitesi dört boyutta incelenecektir: matematik ve istatistiğin zenginliği, öğretmenlerin matematik ve istatistiği anlamalarını desteklemek için öğrencilerle nasıl çalıştıkları ve öğretmenler tarafından hangi matematiksel ve istatistiksel hataların ve yanlışlıkların gösterildiği. Çalışma ayrıca, öğrencilerin istatistik öğretimi sırasında daha geniş matematik bağlamında anlam oluşturma ve akıl yürütmeye nasıl katıldıklarını da araştıracaktır.

1.3. Araştırma Soruları

Çalışmanın amacı ışığında, bu çalışmaya rehberlik eden araştırma soruları, matematik öğretmenleri tarafından sınıfta kullanılan istatistiksel uygulamaların kapsamlı bir şekilde anlaşılmasını sağlamak üzere tasarlanmıştır. Bu nedenle, nitel durum çalışması aşağıdaki araştırma sorularına yanıt aramaktadır:

1. Ortaokul matematik öğretmenlerinin sınıftaki istatistiksel uygulamaları nelerdir?

1.1 Matematik öğretmenleri öğretimlerinde hangi istatistiksel içeriği ele almaktadır?

1.2. Matematik öğretmenleri öğretimlerinde istatistiksel problem çözme süreçlerinin dört aşamalı modeline nasıl yaklaşmaktadır?

2. Matematik öğretmenlerinin daha geniş kapsamlı matematik eğitimi içinde istatistik eğitimi kalitesi nedir?

2.1 Sınıf içi eğitim hem matematik hem de istatistik ile ne ölçüde bağlantılıdır?

2.2 Derslerde işlenen matematik ve istatistik içeriği ne kadar zengindir?

2.3 Matematik öğretmenleri, öğrencilerin hem matematiği hem de istatistiği anlamalarını desteklemek için onlarla nasıl çalışmaktadır?

2.4 Matematik öğretmenleri tarafından sergilenen matematiksel ve istatistiksel hatalar ve yanlışlıklar nelerdir?

2.5 Öğrenciler daha geniş matematik bağlamında istatistik öğretimi sırasında anlam oluşturma ve akıl yürütmeye nasıl katılırlar?

Bu araştırma soruları, matematik öğretmenlerinin sınıftaki istatistiksel uygulamalarını ve istatistik öğretiminin kalitesinin daha geniş matematik eğitimi içinde nasıl olduğunu açıklamaya yardımcı olmaktadır. İşlenen istatistiksel içeriğin ve istatistiksel içeriğin müfredat beklentileri ve istatistiksel problem çözme süreçlerinin dört aşamalı modeline yaklaşımla nasıl uyumlu olduğunun incelenmesiyle, matematik öğretmenlerinin istatistiksel uygulamalarının daha iyi anlaşılması sağlanabilir. Daha geniş kapsamlı matematik öğretiminde istatistik öğretiminin kalitesinin incelenmesi, bir bütün olarak matematik öğretiminin güçlü ve zayıf yönleri hakkında fikir verecektir.

1.4. Çalışmanın Önemi

İstatistik eğitimi, matematik eğitimi araştırmalarında giderek artan bir ilgi görmektedir. İstatistik Eğitiminde Değerlendirme ve Öğretim Rehberi (GAISE), Ulusal Matematik Öğretmenleri Konseyi (NCTM), Matematik için Ortak Çekirdek Devlet Standartları (CCSSM) ve Avustralya ve Yeni Zelanda Matematik ve İstatistik Müfredatları da dahil olmak üzere dünyanın dört bir yanındaki müfredat belgeleri, istatistiksel problem çözme sürecini vurgulamaktadır. Türkiye'deki matematik müfredatının öğrenme hedefleri, istatistiksel problem çözmeye giderek daha fazla

vurgu yapan diğerküreselmüfredatlarla uyumludur (Ader, 2016, 2018; Frischemeier vd., 2022; MEB, 2018, Özmen ve Baki, 2019). Bununla birlikte, müfredat belgeleri, öğretmenlere öğretim uygulamalarını önermek için kolektif bir çalışmayla hızla güncellenmiş ve öğretmenlerin bunları ne kadar uygulayabilecekleri dikkate alınmamıştır. Sınıf içi uygulamaların müfredat belgeleriyle nasıl uyumlu olduğuna dair çok az araştırma yapılmıştır (Litke & Hill, 2020). İstatistik öğretimini iyileştirmek için sahadan bilgi toplamak gereklidir.

İstatistiksel problem çözme süreçlerinin dört aşamalı modelinin kullanılması, matematik öğretmenlerinin istatistiksel uygulamalarını incelemek için kapsamlı bir yaklaşım sağlar (Watson vd., 2018). Öğretmenlerin istatistiksel problem çözme sürecinin dört unsurunu da kullanıp kullanmadıklarını incelemek, öğretimlerindeki potansiyel boşlukları tespit etmeyi mümkün kılar. Bu, istatistiğı ayrı bir disiplin olarak öğretilip öğretilmediklerini değerlendirmenin ötesinde, matematik öğretmenlerinin uygulamalarının daha incelikli bir şekilde anlaşılmasını sağlayabilir (Litke & Hill, 2020; Visnovska & Cobb, 2019). Ayrıca, istatistik öğretiminin uygulamada nasıl görüldüğüne dair kapsamlı bir anlayış sağlar. Bu nedenle, matematik öğretmenlerinin istatistik öğretimine nasıl yaklaştıklarını anlamak önemlidir, çünkü uygulamaları öğrencilerin istatistiksel düşünme, akıl yürütme ve okuryazarlıklarını önemli ölçüde etkileyebilir (Bargagliotti vd., 2020; Ben-Zvi ve Garfield, 2008; MacGillivray ve Pereira-Mendoza, 2011).

Bu nedenle, bu araştırma, ortaokul matematik öğretmenlerinin istatistik öğretim uygulamalarını, Türkiye Ortaokul Matematik Dersi Öğretim Programı'ndaki müfredat hedefleri ve istatistiksel problem çözme süreçlerine genel vurgu yapan beklentiler ile uyumları açısından incelemeyi amaçlamaktadır. Çalışma, yürürlükteki müfredatının uygulanmasını inceleyerek, öğrencilere istatistiksel anlayış için gerekli öğrenme fırsatlarının ne ölçüde sağlandığına ışık tutmaktadır (Litke & Hill, 2020). Mevcut çalışmanın teorik katkılarından biri, matematik sınıflarında istatistik öğretimi bağlamında yazılı ve uygulamalı müfredat arasındaki ilişkinin anlaşılmasını iletmeğdir. Bu çalışma, öğretim uygulamalarını doğal ortamlarda uzun süre gözlemleyerek matematik öğretmenlerinin istatistik uygulamalarının kapsamlı bir şekilde anlaşılmasını sağladığı için de önemlidir. Bu çalışmanın bulguları, müfredat

ve müfredat materyallerinde iyileştirilmesi gereken alanlar konusunda politika yapıcıları potansiyel olarak bilgilendirebilir.

Ayrıca, bu çalışmanın önemi, matematik sınıflarında istatistik öğretiminin kalitesinin anlaşılmasına yaptığı katkıdır. Daha önce de belirtildiği gibi, istatistik birçok ülkede matematik müfredatının bir parçasıdır. Her iki disiplin arasındaki benzerlik ve farklılıkları keşfetmek, birbirlerine nasıl katkıda bulunabileceklerini anlamaya yardımcı olur (Gattuso & Ottaviani, 2011). Bu nedenle, istatistikle ilgili öğretim uygulamalarına daha geniş matematiksel uygulamalar bağlamında bakmak gereklidir.

Mevcut çalışmada, öğretimin kalitesini araştırmak için Öğretimin Matematiksel Kalitesi (MQI) çerçevesi kullanılmıştır. İstatistiksel problem çözme sürecine ilişkin protokol ile birlikte MQI aracı kullanılarak, istatistik öğretiminin kalitesi, öğretmenlerin istatistik uygulamalarının genel öğretim uygulamalarıyla nasıl ilişkili olduğuna dair içgörü kazanmak için daha geniş bir şekilde incelenebilir (Hannigan vd., 2013). Bu çalışma, bir disiplin olarak istatistiğin doğasına ilişkin tartışmanın ötesinde, okul matematiğindeki yeri nedeniyle istatistiği bir disiplin olarak ele alması bakımından önemlidir. Bu nedenle, doğal ortam olan matematik sınıflarında istatistik öğretimi uygulamalarını araştırmaktadır. İstatistiksel araştırmalarda hem disiplinle ilgili uygulamalara hem de matematik daha geniş bir şemsiye konu olarak kabul edildiğinden, matematikte konuya özgü öğretim uygulamalarına bakmaktadır. Bu çalışma, matematik öğretmenlerinin istatistiksel uygulamaları öğretimlerine entegre etmelerini ve matematik müfredatının daha genel taleplerini göz önünde bulundurmaktadır. Matematiğin geniş bağlamı içinde istatistik öğretimi uygulamalarına dair kapsamlı bir anlayış geliştirmek, mevcut çalışmanın diğer teorik katkısı olarak mevcut koşullar altında istatistiğin okul matematiğine nasıl etkili bir şekilde entegre edilebileceği konusunda literatüre bilgi sağlayacaktır. Sonuç olarak, bu çalışma matematik sınıflarında etkili istatistik öğretimi ile ilgili literatüre katkı sağlama potansiyeline sahiptir.

İstatistik eğitiminde öğretim uygulamaları üzerine yapılan çalışmalar sınırlıdır (Litke ve Hill, 2020; Tarr vd., 2006). Bu çalışmada, video analizi üzerinden öğretmen, öğrenciler ve içeriğin etkileşimi de dahil olmak üzere sınıf bağlamında gerçek zamanlı

uygulamaları içeren ve uygulamayla tekrar tekrar etkileşime girmeye olanak tanıyan gözlensel bir çalışma yaklaşımı kullanılmaktadır. Bu nedenle, bu çalışma öğretmenlerin disiplinle ilgili uygulamalar açısından ne yaptıklarına ve matematik sınıflarındaki uygulamalarının kalitesine ilişkin değerli bilgiler sağlayabilir ve matematik sınıflarında istatistik öğretiminin nerede ve nasıl geliştirileceğini anlamaya yardımcı olabilir. Ayrıca, MQI aracı, her bir uygulamanın kalite düzeyini inceleyen bölüm düzeyindeki kodlarda kapsamlı içerik-özel öğretim uygulamalarından oluşmaktadır. Kalite göstergeleri aracılığıyla matematik ve istatistik öğretimi uygulamalarının belirlenmesi, öğretmen eğitimcilerini bu uygulamaların güçlü ve zayıf yönlerinin yanı sıra matematik sınıflarında istatistik öğretimi sırasında ne ölçüde mevcut oldukları konusunda bilgilendirir. Öğretmen eğitimcileri, mevcut öğretim uygulamalarındaki güçlü ve zayıf alanları belirleyerek, okullardaki matematik ve istatistik öğretiminin kalitesini artırmak için hedefe yönelik müdahaleler geliştirebilir ve sonuçta öğrencilerin öğrenme çıktılarını iyileştirebilir.

Genel olarak, öğretmenlerin uyguladıkları müfredattaki hedefleri ne kapsamda ele aldıkları, bu hedeflerle hangi temel fikirlere odaklandıkları ve istatistiksel problem çözme sürecinin hangi adımlarına odaklandıkları da dahil olmak üzere sınıf uygulamalarını araştırmak, daha geniş matematik eğitimi içinde istatistiksel öğretimin kalitesini daha iyi anlamak için çok önemlidir. Bu sayede, öğrencileri etrafımızdaki dünyayı anlamlandırmak için gerekli olan istatistiksel düşünme, okuryazarlık ve muhakeme becerileriyle daha iyi donatmak için istatistik eğitiminde iyileştirmeler yapılabilecek alanları tespit edebiliriz.

2. Alan Yazın Taraması

Alan yazın istatistik öğretme ve öğrenmenin, öğrencilerin bilinçli vatandaşlar olmaları için gerekli olan istatistik okuryazarlığı, akıl yürütme ve düşünme becerilerini geliştirmeye odaklanması gerektiğini göstermektedir (Gomez-Blancarte & Chavez-Aguilar, 2022). Temel fikirler veya istatistiksel kavramlar ve süreçler, istatistik öğretme ve öğrenmenin kritik bileşenleridir (Garfield ve Ben-Zvi, 2004; 2008; Shaughnessy, 2007; Shaughnessy, 2019a; 2019b; Watson vd., 2013) ve öğrencilerin

sadece prosedürleri öğrenmek yerine bu fikirleri anlamaları gerekir (Callingham vd., 2021; Shaughnessy, 2019a; Watson vd., 2018).

İstatistiksel problem çözme süreçleri, istatistik öğretmek ve öğrenmek için çok önemlidir (Wild & Pfannkuch, 1999). Öğrenciler anlamlı sonuçlar çıkarmak için eksiksiz istatistiksel araştırmalara ya da bu araştırmaların belirli adımlarına katılmalıdır (Watson & Callingham, 2015). Müfredat belgeleri olarak GAISE raporları, istatistik okuryazarlığını ve verilere dayalı bilinçli karar vermeyi teşvik etmek için K-12 düzeyinde istatistik eğitimi için öneriler sunmaktadır (Bargagliotti vd., 2020; Franklin vd., 2007). İstatistik eğitiminde istatistiksel araştırmaların rolü, Türk müfredatı da dahil olmak üzere farklı müfredatlarda tartışılmıştır (Ader, 2016; 2018). Çeşitli çalışmalarda araştırma sorularının oluşturulması, veri toplama, veri analizi ve tüm istatistiksel problem çözme süreci de dahil olmak üzere istatistiğin çeşitli yönleri vurgulanmıştır. Ancak, müfredat ve literatürde tartışılan önemine rağmen, sınıf içi uygulamalara yansımamaktadır (Litke ve Hill, 2020). Bu nedenle, istatistiksel öğretim uygulamalarını iyileştirmek için uygulamadan kanıt toplamak önemlidir.

Literatür taraması, özellikle matematik ve istatistik eğitiminde etkili öğretim uygulamalarını teşvik eden olumlu bir öğrenme ortamı geliştirmede öğretmenlerin önemli rolünü vurgulamaktadır. İstatistik eğitimi geleneksel matematik eğitiminden farklı olduğu için farklı öğretim yaklaşımları gerektirmektedir (Gerthenschlager, 2019). İstatistik eğitiminde sınıf uygulamalarının iyileştirilmesine yönelik artan taleple birlikte, istatistiksel problem çözme süreci ve kapsanan büyük fikirler de dahil olmak üzere istatistiğe özgü uygulamaları kapsamlı bir şekilde ele almak önemlidir. Alan yazında, istatistik okuryazarlığı, akıl yürütme, düşünme, temel fikirler ve istatistik eğitimindeki müfredat yaklaşımları ile birlikte istatistiksel problem çözme süreci ve bu konuda yapılan araştırmalar derinlemesine tartışılmaktadır.

Ayrıca alanyazın matematikte öğrenci öğrenimini geliştirmek için öğretim kalitesinin ve öğretim uygulamalarını anlamının önemini vurgulamaktadır. Etkili öğretim, öğretim kalitesi çerçeveleri ve bileşenleri aracılığıyla ölçülebilen ve anlaşılabilen yüksek kaliteli öğretimi gerektirir. Bu çerçevelerden biri olan Matematiksel Öğretim

Kalitesi (MQI), içeriğe özgü bakış açısı, kapsamlılığı ve müfredatla uyumlu uygulamaları nedeniyle bu çalışmanın odak noktasıdır. MQI geçerlilik ve güvenilirlik çalışmalarıyla desteklenmekte, puanlayıcılar için çevrimiçi bir eğitim sunmakta ve Öğretim için Matematiksel Bilgi (MKT) gerektirmektedir. Bu yönler, çerçevenin doğruluğunu artırmakta ve istatistik öğretimi daha geniş matematik öğretimi bağlamında incelemek için kullanışlı hale getirmektedir. MKT gibi matematikte öğretim kalitesini etkileyen potansiyel faktörler de MQI ile ilgili alan yazında tartışılmaktadır. Bu çalışma, uygulamayı bilgilendirmek ve öğrenci öğrenme çıktılarını iyileştirmek için MQI çerçevesini kullanarak istatistik öğretiminin kalitesine ilişkin içgörü kazanmayı amaçlamaktadır.

Genel olarak, alanyazın istatistik okuryazarlığı, akıl yürütme ve düşünme konularını kapsamlı bir şekilde ele almakta, istatistik öğretimi ve öğreniminde araştırmacı yaklaşımın önemini ve istatistiksel becerilerin geliştirilmesinde istatistiksel problem çözmenin rolünü vurgulamaktadır. Ayrıca, istatistiksel süreçleri dahil etmek için Türk matematik müfredatında yapılan değişikliklere ilişkin iç görüler sunmaktadır. Etkili öğretim uygulamaları, matematik ve istatistik eğitiminde olumlu öğrenme çıktıları için çok önemlidir ve öğretmenler bu konuda önemli bir rol oynamaktadır. Bu alanlardaki öğretim uygulamaları arasındaki benzerlik ve farklılıkları anlamak, yüksek kalitede öğretim sağlamak için gereklidir. Bu nedenle, öğretim uygulamalarının ölçülmesi ve anlaşılması, etkili öğrenci öğrenimi için fırsatlar yaratılması açısından önemlidir.

Çalışmanın amacı göz önüne alındığında, alanyazın istatistiğe özgü uygulamaları ve bunların okul bağlamında matematiksel uygulamalarla birlikte yerini tartışmaktadır. Sonuç olarak, son yıllarda istatistik eğitimi üzerine çok sayıda araştırma yapılmasına rağmen, istatistik uygulamalarının matematik öğretimine nasıl entegre edildiğine dair hala bir anlayış eksikliği olduğunu göstermektedir. Özellikle, matematik öğretiminin daha geniş bağlamı içinde istatistik öğretiminin kalitesi araştırılması gerekmektedir. Bu nedenle, bu çalışma, alan yazındaki bu boşluğu doldurmak ve etkili istatistik öğretimine ilişkin iç görüler sağlamak için "GAISE Çerçevesini Kullanan Transkript Kodlama Protokolü" ve Matematiksel Öğretim Kalitesi çerçevesini kullanarak matematik öğretmenlerinin sınıftaki istatistik uygulamalarını araştıracaktır.

3. Yöntem

3.1. Araştırma Deseni

Bu araştırma, iki ortaokul matematik öğretmenin istatistiksel araştırmaları öğretimlerine nasıl entegre ettiklerini ve istatistik öğretim uygulamalarının kalitesini kapsamlı bir şekilde anlamak için betimsel bir toplu durum çalışması tasarımı kullanmaktadır. İncelenen vakalar Ali ve Esra'nın istatistik öğretimi uygulamalarıdır. Bu çalışma, vaka içi analiz ve çoklu vakaların çapraz vaka analizi yoluyla, ortak temaları ve örüntüleri ve benzersiz özellikleri ve varyasyonları belirlemeyi amaçlamaktadır (Creswell, 2007; Yin, 1981;2018). Yin'e (2018) göre, çoklu vaka çalışması yürütmenin yolu, benzer sonuçları tahmin etmek için tek tek vakaları gerçek replikasyon için kullanmayı ve ardından olası nedenlerle zıt sonuçları tanımak için teorik bir replikasyonu içerir. Ayrıca, her bir vaka bağlamında öğretim kalitesini ve istatistiksel araştırmaları analiz ettiği ve ardından "başka bir vakadan gelen açıklamayı" karşılaştırarak çapraz vaka analizi yaptığı için çoklu vaka gömülü tasarım da bu çalışmada kullanılmaktadır (Creswell, 2013; Yin, 1981, s.108).

3.2. Katılımcılar ve Çalışma Bağlamı

Bu çalışmada, 7. sınıfı okutma, ortaokul matematik eğitimi programında lisans derecesine sahip olma, öğretmenlikte en fazla 12 yıllık deneyime sahip olma ve mevcut okulda en az iki yıldır çalışma gibi belirli ölçütleri karşılayan iki ortaokul matematik öğretmenini seçmek için amaçlı örnekleme kullanılmıştır. Bu çalışma için iki ortaokul matematik öğretmen, Ali ve Esra ve bu öğretmenlerin birer 7. sınıfı katılımcı olarak seçilmiştir.

Katılımcıları ve çalışmanın bağlamını kapsamlı bir şekilde anlamak için çeşitli veri toplama araçları kullanılmıştır. Katılımcıların geçmişleri, matematik ve istatistik öğretme ve öğrenme konusundaki görüşleri, istatistiksel içerik anlayışları ve tipik öğretimleri hakkında bilgi toplamak için görüşmeler yapılmıştır. Ayrıca, öğrenci testleri sınıf ortamı hakkında fikir vermiştir. Öğretmenlerle yapılan görüşmelerde

müfredat materyallerinin kullanımını araştırmak ve çalışma için daha fazla bağlam sağlamak amacıyla gerçekleştirilmiştir.

3.3. Veri toplama süreci

Öncelikle kuruldan ve İstanbul İl Milli Eğitim Müdürlüğünden etik onayı alınmıştır. Bu çalışma kapsamında Sarıyer ilçesinde çalışmaya dahil edilme kriterlerini karşılayan ortaokul matematik öğretmenleri belirlenmiştir. Toplam 11 öğretmenle iletişime geçilmiş ve üç okuldan dört 7. sınıf öğretmeni çalışmaya katılmayı kabul etmiştir. Veri toplama süreci için kuruldan etik izin güncellenmiştir. Toplam yedi sınıftan oluşan bu dört öğretmen araştırma sürecine tam olarak katılmış ve seçim kriterleri karşılanmıştır.

İlk veri toplama döneminden önce, bir 6. sınıf matematik öğretmenini birkaç saat gözlemlediğim bir pilot çalışma yürüttüm. Öğretmenin derslerini istatistiksel incelemeler ve Matematiksel Öğretim Kalitesi çerçevesini kullanarak gayri resmi bir şekilde analiz ettim. Gözlemlere ve video kayıtlarına ek olarak, ana çalışmadaki aynı soruları kullanarak öğretmenle mülakat yaptım. Öğretmenenden geri bildirim aldım ve görüşme sorularını buna göre yeniledim.

Ana çalışma 2017-2018 akademik yılının ikinci yarısında, Mart-Haziran 2018 tarihleri arasında gerçekleştirilmiştir. Bu dönemde gözlemler, görüşmeler ve öğrenci değerlendirmeleri gibi veri toplama faaliyetleri gerçekleştirilmiştir. "Çember" konusunun gözlem dönemi, bu konu 7. sınıf matematik derslerinin yıllık planında yer aldığı için mart ayı sonunda başladı. Yaklaşık 15 adet 40 dakikalık dersten oluşan "istatistik" ünitesini Nisan ayında gözlemledim- Ali'nin sınıflarındaki gözlem süresi ünitelerin sırası değiştiği için biraz farklılık gösterdi ve gözlemler mayıs ayında gerçekleşti.

Esra, X Öğretmen, Y Öğretmen ve Ali olmak üzere dört öğretmenle toplam 104 saat sınıf gözlemi gerçekleştirdim. Bu süre Esra'nın iki sınıfı için 13 saat, X Öğretmen için 21 saat, Y Öğretmen için 15 saat ve Ali'nin üç sınıfı için 14 saatten oluşuyordu. Her ders videoya kaydedilmiş ve bunlara açıklayıcı ve yansıtıcı saha notları eşlik etmiştir.

Bu gözlemler, saha notları ve video kayıtları ile birlikte çalışmanın birincil veri kaynaklarını oluşturmuştur. Ayrıca, öğretmenlerle hem çalışma öncesinde hem de sonrasında görüşmeler yapılmıştır. Öğrenci verileri, incelenen vakaları derinlemesine anlamak için tutum testleri, başarı testleri ve arka plan bilgi formları aracılığıyla toplanmıştır. Çalışmanın vakaları, katılımcı seçimi bölümünde tartışıldığı gibi Ali ve Esra'nın öğretmenlik uygulamalarıdır.

3.4. Veri Toplama Araçları

Veri toplama kapsamlı olarak tanımlanır, bu nedenle bir vaka çalışmasında kullanılmak üzere birden fazla bilgi kaynağına ihtiyaç vardır (Creswell, 2013). Bir vaka çalışmasında kullanılabilecek olası veri toplama araçları, doğrudan gözlem veya katılımcı gözlem dahil olmak üzere gözlemler, görüşmeler ve fiziksel eserler veya arşiv kayıtları dahil olmak üzere belgelerdir (Creswell, 2013; Yin, 2018). Bu çalışmada birincil veri kaynakları, saha notları ve sınıf içi eserlerin eşlik ettiği video kayıtlarından oluşan gözlemler ve görüşmelerdir (bkz. Ek B)

3.1. Veri Analizi

Bu vaka çalışmasında, 7. sınıf istatistik sınıflarındaki istatistik öğretiminin doğasını araştırmak için saha notları, görüşmeler ve sınıf eserleri ile sınıf videoları gibi çoklu veri kaynaklarını analiz ettim. Nitel araştırmada veri analizi, toplanan verilerin anlamlandırma süreci olarak tanımlanmaktadır (Flick, 2014). Nitel veri analizi veri toplama süreciyle başlar çünkü veri toplama, analiz ve bulguların raporlanması birbiriyle ilişkilidir (Creswell, 2007). Flick (2014), anlam oluşturma sürecinin araştırma süreci boyunca verilerin kaba ve ayrıntılı bir şekilde incelenmesini içerdiğini belirtmiştir. Yin (2016) nitel veri analizinin katı bir yapıya sahip olmadığını, bunun yerine derleme, dağıtma, yeniden birleştirme, yorumlama ve sonuçlandırma gibi birbiriyle ilişkili bir dizi adımı izlediğini belirtmiştir.

Bu çalışmadaki her bir vaka için Yin'in (2016) beş aşamalı döngüsünü takip ettim. İlk aşamada, veri kaynaklarını sırayla düzenleyerek verileri derledim. Bu, videoların ve saha notlarının düzenlenmesini, sınıf videolarının izlenmesini, derslerin özetlenmesini

ve saha notlarının yeniden okunmasını içeriyordu. Videoya kaydedilen altı dersin tamamını ve kalan derslerden belirli klipleri yazıya döktüm. Bu derleme aşaması çalışmanın her iki amacı için de aynıydı.

İstatistiksel araştırma süreci ve öğretim kalitesi açısından istatistik öğretiminin doğasını keşfetmek için, sınıf eserlerini ve kaynaklarını video bölümleri ve saha notlarıyla eşleştirerek Şematik Diyagram adı verilen bir istatistiksel görevler listesi hazırladım. Bu, sınıfta kullanılan görevleri belirlemek için kullanıldı. Ardından, görevleri amaçlarına göre yeniden düzenlediğim ayrıştırma sürecine başladım. Diyagramı Litke ve Hill'in (2020) "bölüm" tanımına göre yeniden düzenledim. Litke ve Hill'e (2020) göre, istatistiksel görevler olarak tanımladığım epizotlar, "sınıfın karşılaştığı her bir benzersiz veri kümesiyle sınıfın etkileşimlerini" ifade eder (s.487). Bu benzersizlik tanımına dayanarak şematik diyagrama son halini verdim. Sonuç olarak, şematik diyagram, görevlerin istatistik öğretimi içindeki farklılığını ve amacını daha iyi anlamaya yardımcı oldu.

Örnek vermek gerekirse, şematik diyagramda ilk olarak benzersiz veri setlerine sahip istatistiksel görevlere kodlar atadım. Örneğin, görev grafiksel bir gösterim içeriyorsa, göreve bir "grafik" kodu atadım. Ardından göreve, kullanılan grafiksel gösterimin türünü (örneğin çubuk grafik, çizgi grafik veya daire grafik) belirten daha spesifik bir kod atadım. Yin (2016) bu kodları Düzey 1 (grafik türleri) ve Düzey 2 (grafik) kodları olarak adlandırmıştır. Bu analizde ele alınan ana konu ve içerikler arasında çeşitli grafik türleri (örn. daire, çubuk, çizgi), grafiklerin dönüşümleri ve merkezi eğilim ölçüleri (örn. ortalama, medyan, mod) yer aldı. Bu kodlama süreciyle, istatistiksel görev analizi sırasında ortaya çıkan her bir kodun frekanslarını ve yüzdelerini sağlayarak matematik öğretmenlerinin istatistik öğretimlerinde hangi istatistiksel içeriği kapsadıkları araştırma sorusunu yanıtladım.

Şematik diyagramı oluşturduktan sonra, bu çalışmada veri analizi için Bilgisayar Destekli Nitel Veri Analizi (CAQDAS) programı kullandım. Video analizini kullanmayı seçmemin nedeni, sınıf araştırmasına olan ilgim ve sınıf ortamı ile uygulamaya ilişkin önemli bilgileri yakalamayı amaçladım (Kuckartz ve Rädiker, 2019). Veriler iki çerçeveye dayalı içerik analizi kullanılarak analiz edilmiş ve birden

fazla vakanın varlığı nedeniyle hem vaka içi analiz hem de çapraz vaka analizi yapılmıştır. Creswell (2013) "tipik formatın öncelikle her bir vakanın ve vaka içindeki temaların ayrıntılı bir tanımını sunmak olduğunu, bunun vaka içi analiz olarak bilindiğini" (s.101) belirtmiştir. Bu nedenle, GAISE Çerçevesi ve MQI Kullanarak Transkript Kodlama Protokolünü kullanarak her vakanın ayrıntılı bir tanımını yaptım. Vaka içi analizin ardından, çalışmada kullanılan çerçevelere dayalı olarak "vakalar arasında tematik bir analiz" gerçekleştirdim ve her iki vakayı karşılaştırdım (Creswell, 2013, s.101). Ek olarak, GAISE Çerçevesi ve MQI Kullanarak Transkript Kodlama Protokolünü kullanarak öğretmenlerin istatistiksel öğretim uygulamalarını tanımlamak için nitel verilerden örnekler verdim.

4. Sonuçlar ve Tartışma

Bu bölüm, matematik sınıflarında istatistik öğretimi uygulamalarına ilişkin çalışma bulgularının tartışmasını sunmayı amaçlamaktadır. Bu çalışma, matematik öğretiminde kapsanan istatistiksel içerik, öğretmenlerin istatistiksel problem çözme süreçlerinin dört aşamalı modeline nasıl yaklaştıkları ve daha geniş matematik öğretimi içinde istatistik öğretiminin kalitesi dahil olmak üzere istatistik öğretim uygulamalarına odaklanmıştır.

4.1. İstatistiksel Öğretim Uygulamaları

Bu bölümde, disiplinle ilgili uygulamalara ilişkin bulguları tartışacağım. Disiplinle ilgili uygulamaları değerlendirirken, öncelikle öğretmenlerin ele aldığı içeriği inceledim ve içeriğin müfredat hedefleri veya gereklilikleriyle uyumlu olup olmadığını değerlendirdim. Ayrıca işlenen içerik ile istatistikteki büyük fikirler arasındaki bağlantıyı da araştırdım.

4.1.1. İstatistiksel İçerik

Müfredat ve müfredat materyalleri, öğretmenlerin sınıflarda ne yaptıklarını şekillendirmede çok önemlidir (Batur vd., 2021; Litke ve Hill, 2020). Türkiye'deki devlet okullarında, öğretmenlerin müfredatta belirtilen tüm hedefleri kapsamı

gerekmektedir. Ancak, öğretmenlerin bu müfredat hedeflerine ilişkin anlayışlarını ve belirli hedefleri ne ölçüde kapsadıklarını araştırmaya ihtiyaç vardır. Batur ve diğerlerine (2021) göre, Türkiye'deki müfredatta yer alan istatistik hedefleri, Yeni Zelanda ve Amerika Birleşik Devletleri gibi diğer ülkelerle karşılaştırıldığında nispeten sınırlıdır. Araştırmacılar, istatistik alanındaki müfredat hedeflerinin derinliğinin Türkiye'de daha sığ olduğunu tespit etmiştir. Ader'in (2015, 2018) Türk ulusal müfredatına ilişkin analizi, yıllar içinde istatistik hedeflerinin dahil edilmesinde, hedeflerin ve müfredat yaklaşımlarının sayısındaki artışla birlikte bir iyileşme olduğunu göstermiştir. Bununla birlikte, Batur ve diğerlerinin (2021) de belirttiği gibi, bu hedeflerin daha da geliştirilmesi ve iyileştirilmesi için hala yer vardır.

Bu çalışmanın bulguları, öğretmenlerin genel olarak hedefleri takip ettiklerini ve her bir hedefe verilen ağırlık değişse de (örneğin, daire grafiklerine çizgi grafiklerinden daha fazla zaman harcanmıştır) hepsini kapsadıklarını göstermektedir. Bu farklılığın olası bir nedeni, öğretmenlerin kullandıkları ek ders kitaplarında belirli kavramlara verilen ağırlıktır. Ayrıca, daire grafikleri daha fazla matematiksel hesaplama içermekte ve öğrencilerin daireler ve orantılilik hakkında bilgi edinmeleri için bir bağlam sağlamaktadır. Sınıfa getirilen birçok görev, bir dairenin sektörünü bulmak gibi hesaplamalara odaklanmıştır.

Ayrıca, çalışma öğretmenlerin grafik dönüşümü hedefine daha az zaman ayırdıklarını ortaya koymuştur. Bunun olası bir nedeni, bu kazanımın müfredatta ortalama kazanımlardan sonra yer alan son kazanım olmasıdır. Sonuç olarak, grafikler ve grafik dönüşümü ile ilgili bireysel hedefler arasındaki bağlantı, odak kaymaları nedeniyle azalmıştır. Öğretmenler tarafından belirtilen bir diğer neden de hedefin amacını anlamakta güçlük çekmeleridir. Belirli bağlamlar içinde farklı grafiklerin güçlü ve zayıf yönlerini kavramakta zorlanmışlardır. Bu zorluklara rağmen, sonuçlar öğretmenlerin tüm hedefleri kapsadığını göstermiştir. Watson ve diğerlerinin (2018) daha önce belirttiği gibi, ilkökul ve ortaokul öğrencileri betimsel istatistiklerin bir parçası olarak merkez ölçülerini öğrenmiştir. Sınırlı türde grafikler oluşturabilmiş ve yorumlayabilmişlerdir.

Öğretmenlerin müfredatta belirtilen tüm hedefleri kapsamına rağmen, istatistikteki büyük fikirlere sınırlı bir odaklanma söz konusudur ve bu da istatistiğin kavramsal olarak anlaşılmasında eksikliğe yol açmaktadır (Garfield & Ben-Zvi, 2004). Örneğin, öğretim uygulamalarının çoğu öncelikle merkez kavramı, tam olarak üç merkezi eğilim ölçüsü etrafında dönmüştür. Literatüre göre istatistikte önemli bir büyük fikir olarak kabul edilen merkez (Shaughnessy, 2007; Shaughnessy, 2019; Watson vd., 2007), çoğunlukla ortalama bulmaya yönelik hesaplamalar ve prosedürler aracılığıyla ele alınmıştır (Jacobbe ve Carvalho, 2011). Kavramsal anlayıştan ziyade işlemsel anlayışa yapılan bu vurgu literatürde tartışılmıştır (Shaughnessy, 2007).

Bu çalışmadaki müfredat, beklenti ve varyasyon dışındaki büyük fikirleri açıkça içermemektedir. Ancak Ali, örneklem ve temsil kabiliyeti kavramlarını ele alırken "rastgele" kavramını kısaca tartışmıştır. Yine de bu kavram bir büyük fikir olarak kapsamlı bir şekilde tanıtılmamış ya da vurgulanmamıştır. Bazı durumlarda, öğretim uygulamaları iki veri setini karşılaştırırken, veri dağılımını tanımlamak için daire grafikleri kullanırken veya bir aykırı değer varlığında merkez seçimini tartışırken örtük olarak dağılım kavramına odaklanmıştır. Müfredata farklı grafik türlerinin dahil edilmesi, dağılımın büyük fikrini anlatmak için bir fırsat olabilir.

4.1.2. İstatistiksel Problem Çözme Süreci

İstatistiksel problem çözme süreci, istatistik öğretme ve öğrenmenin temel bir yönü olarak kabul edilir (Wild & Pfannkuch, 1999) ve dünya çapında istatistik müfredatının temel hedeflerinden biridir (Arnold & Franklin, 2021; Gal & Garfield, 1997; Leavy, 2015). Son zamanlarda yapılan araştırmalar, istatistiksel araştırmaların çeşitli unsurlarını ve sürecin döngüsel doğasını vurgulamaktadır (Burgess, 2011). Bu çalışmanın amaçlarından biri, ortaokul öğretmenlerinin istatistiksel problem çözme sürecinin dört aşamalı modeline nasıl yaklaştıklarını betimlemektir. Bu amaca ulaşmak için, öğretmenlerin süreci bir bütün olarak ele almadıklarında süreç içinde hangi belirli unsurlara odaklandıklarını belirlemek için öğretim uygulamaları analiz edilmiştir (örn. Litke & Hill, 2020). Buna ek olarak, öğretmenlerin istatistiksel problem çözme sürecinin tüm döngüsünü kapsayıp kapsamadığını belirlemek için öğretim uygulamaları incelenmiştir.

Çalışmanın sonuçları, istatistik öğretimi uygulamalarının, her bir öğretmenin uygulamalarında gözlemlenen sınırlı örnekler dışında, genellikle istatistiksel olarak araştırmacı soruların sorulmasını içermediğini göstermektedir. Bu unsur üzerine yapılan araştırmalar nispeten yenidir ve istatistiksel olarak araştırmacı sorular sorma konusunda sınırlı sayıda çalışma bulunmaktadır (Frischemier & Leavy, 2020). Yedinci sınıf müfredatı bu unsuru açıkça ele almamaktadır. Ancak beşinci ve altıncı sınıf müfredatlarında araştırmacı sorular oluşturmaya ilişkin kazanımlar yer almaktadır. Esra'nın sınıfında, altıncı sınıf istatistik ünitesi önceki yıllarda işlenmediği için, Esra istatistik ünitesini sınıfa en sevdikleri rengi sorarak başlatmıştır. Bunu sınıftan veri toplama, tablo ve daire grafiği oluşturma ve grafikleri yorumlama takip etmiştir. Görüldüğü üzere, Esra'nın sınıfında bir araştırma sorusuyla başlamak, tüm istatistiksel araştırma döngüsünün tamamlanmasını kolaylaştırmıştır.

Veri toplama unsuruyla ilgili olarak, sonuçlar, her iki öğretmenin de birer bölüm dışında bu unsuru sınıf uygulamalarında kapsamlı bir şekilde kullanmadığını göstermektedir. Bargagliotti ve diğerleri (2020), istatistiksel araştırma sorularının ya birincil verilerden ya da ikincil kaynaklardan elde edilen veriler dikkate alınarak veri toplama yoluyla yanıtlanması gerektiğini vurgulamıştır. Bu unsur, veri toplamanın planlama aşamasını da kapsar. Ancak, veri toplama genellikle zaman alıcı ve öngörülemez olarak algılanmaktadır (Ben-Zvi vd., 2018). Veri toplamanın öğretim uygulamalarına sınırlı bir şekilde dahil edilmesinin olası bir nedeni, yedinci sınıf müfredatındaki öğrenme kazanımlarının öncelikle veri analizi ve yorumlamaya odaklanmış olmasıdır.

Çalışmanın bulguları, öğretmenlerin öncelikle istatistiksel problem çözme sürecinin veri analizi unsuruna odaklandıklarını ve döngünün tamamına eşit derecede önem vermediklerini göstermektedir. Watson ve diğerlerinin (2018) de belirttiği gibi, öğretim ağırlıklı olarak grafik oluşturma ve merkezi eğilim ölçülerini hesaplama etrafında dönmüştür. Ayrıca, istatistik eğitimindeki araştırmalar genellikle "iyi tanımlanmış problemlerle" veri analizine vurgu yapmaktadır (Delmas, 2004; Makar ve Fielding-Wells, 2011) ki bu da bu çalışmanın bulgularıyla örtüşmektedir. Ancak NCTM'ye (2009) göre, hesaplamalar ve prosedürler olarak ele alındığında veri analizi bile istatistiksel araştırmaların dört bileşenini de kapsar çünkü veri analizi, istatistiksel

soruları yanıtlamak ve yorumlara yol açmak için gereken verilerin dikkate alınmasını içerir. Bununla birlikte, istatistiksel problem çözmeden ziyade matematiksel problem çözmeye benzeyen iyi tanımlanmış problemler, istatistiksel problem çözme sürecinin tüm bileşenlerini kapsayan veri analizi perspektifini bozabilir.

İstatistiksel problem çözme sürecindeki son unsur, analiz edilen verilere dayanarak sonuçlar çıkarmayı içeren sonuçların yorumlanmasıdır. Ancak, bu çalışmanın bulguları, her iki öğretmenin öğretim uygulamalarının çoğunun sonuçların yorumlanmasını içermediğini; hesaplamalar, prosedürler ve grafik oluşturma dahil olmak üzere öncelikle veri analizi bölümüne odaklanıldığını ortaya koymuştur. Bununla birlikte, yorumlama unsurunun mevcut olduğu bölümlerde, öğrenciler öncelikle verilerin kendisiyle ilgili soruları yanıtlamak için grafikleri okumakla meşgul olmuşlardır.

Bu çalışmanın bulguları, gözlemlenen öğretim uygulamalarının, bir bölüm dışında, araştırma sorularının formüle edilmesinden verilerin toplanmasına, analiz edilmesine ve yorumlanmasına kadar tüm istatistiksel problem çözme sürecini kapsamadığını göstermektedir. Bu sonuçlar, araştırma ve müfredat belgelerinde bu konuya yapılan vurgunun artmasına rağmen gerçek sınıflarda istatistik öğretiminin istatistiksel araştırma döngüsünün tamamından ziyade hesaplamalar ve grafiklere odaklandığını tespit eden Litke ve Hill'in (2020) bulgularıyla tutarlıdır. Bu tutarsızlığın olası bir nedeni, müfredatın istatistiksel araştırmaları vurgularken, istatistiksel araştırmanın hangi unsurlarının belirli sınıflarda ayrı ayrı vurgulanması gerektiğini açıkça belirtmesidir (MEB, 2018). Ayrıca istatistik eğitimindeki sorunlardan biri, öğretmenlerin genellikle müfredatın döngüsel doğasını dikkate almamaları ve altıncı sınıftaki araştırma soruları oluşturma veya aralık kavramı gibi önceki yıllara ait istatistiksel araştırma unsurlarını ve kavramlarını dahil etmeyi ihmal etmeleridir. İstatistiksel araştırmaların benimsenmesi, müfredat belgelerine (örn. GAISE) dayanan Avustralya, Yeni Zelanda ve ABD müfredatları da dahil olmak üzere dünya çapında müfredatta mevcut olsa da bu yaklaşımların gerçek sınıflarda nasıl uygulandığına dair sınırlı kanıt vardır (Litke ve Hill, 2020).

4.2. Geniş Kapsamlı Matematik Öğretimi İçinde İstatistik Öğretiminin Kalitesi

Öğretim kalitesi, etkili öğretim uygulamalarının belirlenmesinde ve öğrenci öğrenmesinin teşvik edilmesinde çok önemlidir (Schoenfeld, 2014). Araştırmacılar, farklı konu alanlarındaki öğretim uygulamalarını incelemek amacıyla öğretim kalitesini değerlendirmek için çeşitli çerçeveler ve gözlemsel araçlar kullanmışlardır. Bu çalışmada, daha geniş kapsamlı matematik öğretimi içinde istatistik öğretiminin kalitesini değerlendirmek için Öğretimin Matematiksel Kalitesi (MQI) çerçevesi kullanılmıştır. LMT (2011) tarafından geliştirilen MQI çerçevesi, öğrencilere matematiksel içerik sunarken gözlemlenen uygulamalara odaklanmaktadır. Öğretmen, öğrenci ve içerik etkileşimlerini incelemek için öğretim üçgenini kullanır. Çerçeve, aritmetik, cebir, geometri ve istatistik gibi alanlar da dahil olmak üzere konuya özgü uygulamaları kapsar. İstatistik giderek ayrı bir disiplin olarak kabul görse de bu çalışmada okul matematiğinin daha geniş bağlamı içinde ele alınmıştır. Disipline özgü uygulamalara ek olarak, matematik eğitimindeki öğretim uygulamalarının kalitesini daha geniş bir şekilde ele almak önemlidir. MQI çerçevesi, Charalambous ve Praetorius (2018) tarafından vurgulandığı gibi, farklı alanlardaki matematiğe özgü öğretim uygulamalarının kalitesi hakkında kapsamlı bir bakış açısı sağlar.

Çalışmanın bulguları, her iki öğretmenin de öğretim uygulamalarının ağırlıklı olarak matematiksel olmayan ve istatistiksel etkinliklerden ziyade matematik ve istatistik etkinliklerine odaklandığını göstermektedir. Esra'nın dersinin sadece bir bölümünde ödev kontrolü yapılmış, matematiksel ya da istatistiksel bir etkinlik gözlenmemiştir. Ali'nin istatistik dersleri, matematik öğretiminin genel kalitesi incelendiğinde Orta ila Orta/Yüksek kalite olarak değerlendirilmiştir. Buna karşılık, Esra'nın istatistik dersleri matematik dersleri içinde Orta kalite olarak değerlendirilmiştir. Genel ders kalitesindeki bu küçük farklılıklar, MQI çerçevesinin dört boyutuna odaklanarak segment düzeyinde bulguların tartışılma ihtiyacını vurgulamaktadır. Bu boyutlar arasında öğrencilere sunulan matematik ve istatistiğin zenginliği, öğretmenlerin öğrencilerle ve matematik ve istatistikle çalışma yöntemleri, öğretimde mevcut olan matematiksel ve istatistiksel hatalar ve kesinsizlikler ve öğrencilerin anlam oluşturma ve muhakeme sürecine katılımı yer almaktadır.

Genel olarak, MQI ile ilgili tüm bulgular göz önünde bulundurulduğunda, istatistik öğretim uygulamalarının genel olarak yüksek kaliteli göstergelerden yoksun olduğu sonucuna varılabilir. Bunun yerine, zayıf ve güçlü özellikleri içeren tipik olarak tanımlanabilirler. Ders düzeyinde, her iki öğretmenin de derslerinin genel kalitesi çoğunlukla tipiktir, ancak bir öğretmenin orta ila yüksek kaliteli derslerinin oranı daha yüksektir. Matematik ve istatistik zenginliği, her iki öğretmenin öğretim uygulamalarında bölüm düzeyinde düşük ila orta kalite olarak değerlendirilmiştir. Buna karşın, öğrencilerle çalışma ve içerik boyutundaki öğretim uygulamaları düşük kalitelidir. Ayrıca, öğrencilerin anlam oluşturma ve akıl yürütmeye katılımı her iki öğretmenin öğretim uygulamalarında da genel olarak eksik ya da düşük kalitededir. Derslerin çoğu bölümünde matematiksel ve istatistiksel hatalar ve kesinlikten yoksundu. Bu sonuçlar, genel olarak ABD'de matematik öğretiminin kalitesi üzerine yapılan ve öğretimin tipik doğasını vurgulayan çalışmalarla uyumludur (ör. Hill ve Lovison, 2022; Hill vd., 2018; Kane ve Stigler, 2012).

Hill ve diğerleri (2018) çalışmalarında, örneklemelerindeki %50'lik yüzdelik dilimde yer alan öğretmenlerin genel olarak doğrudan öğretim sağlayabildiğini ve iki profilde tanımlandığını belirtmiştir. Birincisi, derslerinde matematiksel hataların olmadığı, öğrencileri matematik üzerinde çalışmaya zorlamayan ve öğrencilerin anlam oluşturma ve akıl yürütmeye katılımını teşvik etmeyen az sayıda zenginlik unsuru bulunmaktadır. Mevcut zenginlik unsurları genellikle rutin uygulamalar olarak yer almıştır. İkinci profil, MQI'deki kalite göstergelerine ilişkin öğretimlerinde hem güçlü hem de zayıf unsurlardan oluşmaktadır. Anlam oluşturma sürecine öğrenci katılımıyla birlikte zenginlik unsurlarından oluşmaktaydı, ancak küçük matematiksel hatalar mevcuttu. Mevcut sonuçlar, Hill ve diğerlerinin (2018) çalışmasındaki %50 oranındaki her iki öğretmen profiliyle de tutarlıdır. Benzer şekilde, Kane ve Stigler (2012) de MET çalışmasında benzer sonuçlar bulmuş, gözlemlerinde matematiğin zenginliği, öğrencilerle çalışma ve öğrenci katılımına ilişkin öğretim uygulamalarının çoğu derste düşük kalitede olduğunu, ancak çoğu dersin matematiksel çalışma içerdiğini belirtmiştir. Hill ve Lovison'ın (2022) sonuçları da çoğu dersin MQI boyutlarındaki uygulamalar açısından tipik olduğunu desteklemektedir.

Hollanda bağlamında, Thurlings ve diğerleri (2019) hem MQI hem de genel pedagojik stratejileri birleştiren Kesir İçin Öğretim Kalitesi (QIFI) Çerçevesine dayalı kesir derslerini analiz etmiştir. Elde ettikleri sonuçlar, QIFI'ye dayalı altı öğretmen kümesi olduğunu göstermiştir; bağlantıcı öğretim, deneysel öğretim, kapsamlı öğretim, otomatik pilot öğretim, kitaba göre öğretim ve etkisiz öğretim. Mevcut çalışmanın sonuçları, her iki katılımcı öğretmenin de "aktarımcı doğaları" (s.181) ve müfredat materyallerine bağımlılıkları nedeniyle otopilot öğretim veya kitaba göre öğretim kümelerinde yer alabileceğini göstermektedir. Ayrıca, her iki kümede de özellikle Otopilot öğretim kümesinde, MQI'deki öğretim uygulamaları, bazı zayıf ve güçlü özelliklerin bir arada bulunduğu düşük kaliteli olarak tanımlanmıştır.

Mevcut çalışma, MQI'deki öğretim uygulamalarının iyileştirilmesi ihtiyacına işaret etmektedir. Park ve Kim (2023) tarafından yapılan küçük ölçekli bir çalışmada, iki yeni başlayan öğretmenin öğretim uygulamaları üç yıl boyunca analiz edilmiştir. Bir öğretmenin son gözlemlerde daha az hata ve yanlışlık yapması dışında, öğretim uygulamalarının yıllar boyunca gelişmediği bulunmuştur. Dolayısıyla, yıllar boyunca edinilen deneyim matematik öğretimi uygulamalarını geliştirmemektedir. Bunun yerine, sonuçlar düşük MQI'nin temel nedenlerinin müfredat materyalleri, MKT, öğretmen inançları ve öğrenciyle ilgili faktörlerin ortak katkılarıyla ele alınması gerektiğini desteklemektedir (Blömeke vd., 2020; Borko vd., 2014; Charalambous ve Hill, 2012; Koehler ve Grouws, 1992).

Ayrıca, mesleki gelişim (MG) programları öğretim kalitesinin iyileştirilmesine katkıda bulunabilir. Borko ve diğerleri (2014), Problem Çözme Döngüsü üzerine video tabanlı bir tartışma içeren bir matematik MG çalıştayının, özellikle öğrencilerle ve matematikle çalışma olmak üzere öğretim kalitesini artırdığını göstermiştir. Buna ek olarak, Gallagher ve diğerleri (2019) matematik öğretmen adaylarının yansıtıcı uygulama yoluyla zenginlik boyutuyla ilgili yüksek kaliteli öğretim uygulamalarına ilişkin farkındalıklarını geliştirebileceklerini vurgulamıştır. Benzer şekilde, istatistikte öğretim kalitesi, mevcut çalışmanın sonuçları dikkate alınarak MQI unsurları ile tasarlanmış bir MG programı veya yansıtıcı uygulama yoluyla geliştirilebilir.

F. THESIS PERMISSION FORM / TEZ İZİN FORMU

(Please fill out this form on computer. Double click on the boxes to fill them)

ENSTİTÜ / INSTITUTE

Fen Bilimleri Enstitüsü / Graduate School of Natural and Applied Sciences ☐

Sosyal Bilimler Enstitüsü / Graduate School of Social Sciences ☒

Uygulamalı Matematik Enstitüsü / Graduate School of Applied Mathematics ☐

Enformatik Enstitüsü / Graduate School of Informatics ☐

Deniz Bilimleri Enstitüsü / Graduate School of Marine Sciences ☐

YAZARIN / AUTHOR

Soyadı / Surname : Şeker Akın

Adı / Name : Vuslat

Bölümü / Department : İlköğretim / Elementary Education

TEZİN ADI / TITLE OF THE THESIS (İngilizce / English):

HOW DO MATHEMATICS TEACHERS TEACH STATISTICS? A COMPREHENSIVE EXAMINATION OF STATISTICAL PROBLEM-SOLVING PROCESS AND INSTRUCTIONAL QUALITY

TEZİN TÜRÜ / DEGREE: Yüksek Lisans / Master ☐ Doktora / PhD ☒

1. Tezin tamamı dünya çapında erişime açılacaktır. / Release the entire work immediately for access worldwide. ☐
2. Tez iki yıl süreyle erişime kapalı olacaktır. / Secure the entire work for patent and/or proprietary purposes for a period of **two years**. * ☐
3. Tez altı ay süreyle erişime kapalı olacaktır. / Secure the entire work for period of **six months**. * ☒

* Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. / A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.

Yazarın imzası / Signature Tarih / Date

(Kütüphaneye teslim ettiğiniz tarih. Elle doldurulacaktır.)

(Library submission date. Please fill out by hand.)

Tezin son sayfasıdır. / This is the last page of the thesis/dissertation.