

AN INVESTIGATION OF 8TH GRADE STUDENTS' STATISTICAL
LITERACY THROUGH PROBLEMS WITH AND WITHOUT CONTEXT

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY



BY
TUĞÇE VURAL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MATHEMATICS EDUCATION IN MATHEMATICS AND SCIENCE
EDUCATION

AUGUST 2020

Approval of the thesis:

**AN INVESTIGATION OF 8TH GRADE STUDENTS' STATISTICAL
LITERACY THROUGH PROBLEMS WITH AND WITHOUT CONTEXT**

submitted by **TUĞÇE VURAL** in partial fulfillment of the requirements for the
degree of **Master of Science in Mathematics Education in Mathematics and
Science Education, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Erdiñ Çakırođlu
Head of the Department, **Math. and Science Education**

Assoc. Prof. Dr. Didem Akyüz
Supervisor, **Math. and Science Education, METU**

Examining Committee Members:

Assoc. Prof. Dr. İffet Elif Yetkin Özdemir
Mathematics and Science Education, Hacettepe University

Assoc. Prof. Dr. Didem Akyüz
Mathematics and Science Education, METU

Assist. Prof. Dr. Işıl İşler Baykal
Mathematics and Science Education, METU

Date: 04.08.2020



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 : Tuğçe Vural

Signature :

ABSTRACT

AN INVESTIGATION OF 8TH GRADE STUDENTS' STATISTICAL LITERACY THROUGH PROBLEMS WITH AND WITHOUT CONTEXT

Vural, Tuğçe

Master of Science, Mathematics Education in Mathematics and Science Education
Supervisor: Assoc. Prof. Dr. Didem Akyüz

August 2020, 270 pages

The aim of the present study was to thoroughly investigate eight grade students' achievement and statistical literacy in statistical literacy problems with and without context. Moreover, the eight grade students' achievement in problems with and without context was compared.

The study was conducted in Çankaya and Yenimahalle districts of Ankara in the spring semester of the 2019-2020 academic year. The participants of the study were selected through purposive sampling as 500 eight grade students from four public schools. In the study, both qualitative and quantitative methodologies were used. The data were collected using the Statistical Literacy Test with Context Problems (SLT-CP) and the Statistical Literacy Test without Context Problems (SLT). While qualitative research methodologies were used to determine the achievement of students in statistical content domains and to compare their achievement in problems with and without context, the quantitative research method was used to investigate the statistical literacy of students according to statistical content domains.

The students' achievement in the tests and their statistical literacy differed across statistical content domains and the students' common misconceptions and solutions

were determined. Additionally, the student responses in open-ended problems were classified under four categories as statistical, pre-statistical, non-statistical and unrelated and the responses were examined in detail.

When the students' achievement in statistical literacy problems with and without context was compared in terms of statistical content domains, no significant difference was observed in the achievement of students in the domains of measures of central tendency, graph interpretation and questioning of statistical data. However, a significant mean difference was found in the sampling domain.

Keywords: Statistical Literacy, Elementary School Students, Measures of Central Tendency, Graphs, Sampling, Questioning of Statistical Data

ÖZ

SEKİZİNCİ SINIF ÖĞRENCİLERİNİN İSTATİSTİKSEL OKURYAZARLIKLARININ BAĞLAM İÇEREN VE İÇERMİYEN PROBLEMLER ÜZERİNDEN İNCELENMESİ

Vural, Tuğçe

Yüksek Lisans, Matematik Eğitimi, Fen ve Matematik Bilimleri Eğitimi

Tez Yöneticisi: Doç. Dr. Didem Akyüz

Ağustos 2020, 270 sayfa

Bu çalışmanın amacı ilköğretim sekizinci sınıf öğrencilerinin bağlam içeren ve bağlam içermeyen istatistiksel okuryazarlık problemlerindeki başarılarını ve istatistiksel okuryazarlıklarını derinlemesine incelemektir. Ayrıca, bu öğrencilerin bağlam içeren ve bağlam içermeyen problemlerdeki başarıları da karşılaştırılmıştır.

Bu çalışma 2019-2020 eğitim-öğretim yılının bahar döneminde, Ankara ilinin Yenimahalle ve Çankaya ilçelerinden amaçlı örnekleme tekniği ile seçilmiş dört adet devlet okulundan toplam 500 adet sekizinci sınıf öğrencisi ile yürütülmüştür. Araştırmada nicel ve nitel araştırma yöntemleri bir arada kullanılmıştır. Çalışmanın verileri Bağlam İçeren İstatistiksel Okuryazarlık Testi (İOT-B) ve Bağlam İçermeyen İstatistiksel Okuryazarlık Testi (İOT) ile toplanmıştır. Öğrencilerin istatistiksel konu alanları bazında başarılarının saptanması ve başarılarının bağlam içeren ve içermeyen problemler üzerinden karşılaştırılması için nitel araştırma yöntemleri kullanılırken, öğrencilerin konu alanlarına göre istatistiksel okuryazarlıklarını incelemek için nicel araştırma yöntemi kullanılmıştır.

Öğrencilerin testlerdeki başarıları ve istatistiksel okuryazarlıkları konu alanlarına göre farklılıklar göstermiştir ve öğrencilerin ortak yanlışları ve çözümleri

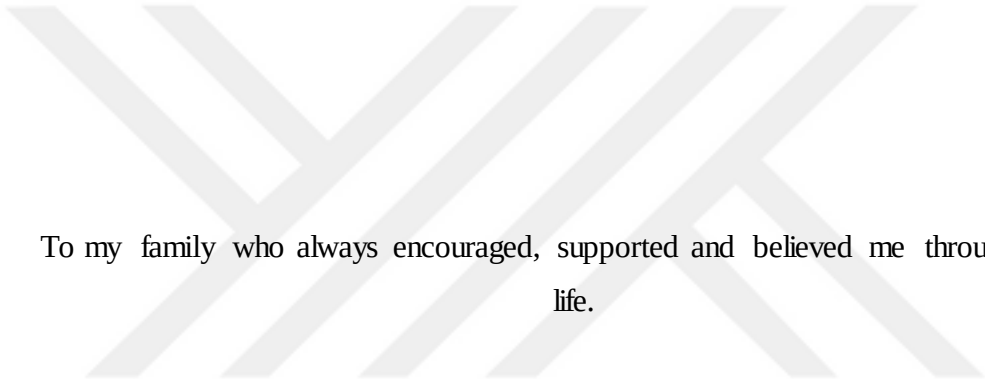
saptanmıştır. Ayrıca, açık uçlu sorulardaki öğrenci yanıtları istatistiksel, ön-istatistiksel, istatistiksel olmayan ve ilgisiz olmak üzere dört kategori altında sınıflandırılmış ve ayrıntılı olarak incelenmiştir.

Öğrencilerin bağlam içeren ve bağlam içermeyen istatistiksel okuryazarlık problemlerindeki başarıları istatistiksel konu alanları bazında karşılaştırıldığında, merkezi eğilim ölçüleri, grafik yorumlama ve istatistiksel verinin sorgulanması konu alanlarında öğrencilerin başarısında anlamlı bir farklılık bulunmamıştır. Ancak, örnekleme konu alanında anlamlı bir fark bulunmuştur.

Öğrencilerin istatistiksel okuryazarlık testlerindeki başarıları ve istatistiksel okuryazarlıkları konu alanlarına göre farklılıklar göstermiştir ve öğrencilerin ortak yanılgıları ve çözümleri saptanmıştır. Açık uçlu sorulardaki öğrenci yanıtları dört kategori altında sınıflandırılmış ve ayrıntılı olarak incelenmiştir.

Öğrencilerin bağlam içeren ve bağlam içermeyen istatistiksel okuryazarlık problemlerindeki başarıları istatistiksel konu alanları bazında karşılaştırıldığında, merkezi eğilim ölçüleri, grafik yorumlama ve istatistiksel verinin sorgulanması konu alanlarında öğrencilerin başarısında anlamlı bir farklılık çıkmazken örnekleme konu alanında öğrencilerin bağlam içeren ve içermeyen problemlerdeki başarıları arasında bağlam içermeyen problemler öncülüğünde anlamlı bir fark bulunmuştur.

Anahtar Kelimeler: İstatistiksel Okuryazarlık, İlköğretim Öğrencileri, Merkezi Eğilim Ölçüleri, Grafikler, Örnekleme, İstatistiksel Verinin Sorgulanması



To my family who always encouraged, supported and believed me throughout my
life.

ACKNOWLEDGMENTS

I would like to extend my gratitude to my advisor Assoc. Prof. Dr. Didem Akyüz, who supervised me in the course of the study by answering my questions despite her busy schedule, and by sharing her precious experiences and constructive feedbacks.

I owe special thanks to Fatih Can for his everlasting patience, courage and support in the face of my endless problems and questions in the course of my study. You supervised me from the beginning to the end of the study. I believe we will overcome all the difficulties in our lives together. Thanks a million my love.

I would also like to thank my mother Kadriye Vural, who always listened to me, supported me, and dealt with my problems throughout my life. You have never given up supporting and loving me. I am proud to be your daughter. I would also like to express my sincere thanks to my father, my brother, my sister, and my grandmother for their love, courage, and support throughout my life. Without your love, tolerance, and patience, I would have never finished this study.

I would like to thank the distinguished thesis committee members, Assist. Prof. Dr. Işıl İşler Baykal and Assoc. Prof. Dr. Elif Yetkin Özdemir, who shared their positive and constructive feedback.

Finally, I would like to extend my appreciation to the teachers and the students who participated in my study for their feedback and patience in the course of the study.

TABLE OF CONTENTS

ABSTRACT.....	v
ÖZ	vii
ACKNOWLEDGMENTS	x
TABLE OF CONTENTS.....	xi
LIST OF TABLES	xv
LIST OF FIGURES	xvii
LIST OF ABBREVIATIONS.....	xxii
CHAPTERS	
1 INTRODUCTION	1
1.1 Aims of the Study.....	4
1.2 Research Questions	4
1.3 Definitions of Important Words	4
1.4 Significance of the Study	5
2 LITERATURE REVIEW	9
2.1 Statistical Literacy	9
2.1.1 Definition of Statistical Literacy.....	9
2.1.2 Theoretical Models and Frameworks for Statistical Literacy	12
2.2 Research Studies on Statistical Literacy	21
2.3 Research Studies on Statistical Literacy in Turkey.....	33
2.4 Context in Problems	36

2.4.1	Context in Statistical Literacy Problems	45
3	METHODOLOGY	49
3.1	Research Design	49
3.2	Population and Sample	51
3.3	Data Collection Instruments	52
3.3.1	Statistical Literacy Test with Context Problems	52
3.3.2	Statistical Literacy Test without Context Problems	72
3.4	Pilot Study	74
3.5	Validity and Reliability of the Instruments	75
3.6	Data Collection Procedures	77
3.7	Data Analysis Procedures	78
3.8	Assumptions and Limitations	82
3.9	Internal and External Validity of the Study	83
3.9.1	Internal Validity	83
3.9.2	External Validity	85
4	RESULTS	87
4.1	Achievement in Problems with and without Context in terms of Statistical Content Domains	88
4.1.1	Achievement in Measures of Central Tendency	89
4.1.2	Achievement in Graph Interpretation	95
4.1.3	Achievement in Sampling	106
4.1.4	Achievement in Questioning of Statistical Data	109
4.2	Statistical Literacy in Problems with and without Context in terms of Statistical Content Domains	118

4.2.1	Statistical Literacy in Measures of Central Tendency	118
4.2.2	Statistical Literacy in Graph Interpretation.....	139
4.2.3	Statistical Literacy in Sampling	163
4.2.4	Statistical Literacy in Questioning of Statistical Data	171
4.3	Comparison between Problems with Context and Problems without Context.....	210
4.3.1	Mean Difference in Measures of Central Tendency	212
4.3.2	Mean Difference in Graph Interpretation.....	213
4.3.3	Mean Difference in Sampling	214
4.3.4	Mean Difference in Questioning of Statistical Data	216
5	DISCUSSION	217
5.1	8 th Grade Students' Achievement and Statistical Literacy in terms of Statistical Content Domains.....	217
5.1.1	Achievement and Statistical Literacy in Measures of Central Tendency.....	218
5.1.2	Achievement and Statistical Literacy in Graph Interpretation.....	221
5.1.3	Achievement and Statistical Literacy in Sampling	224
5.1.4	Achievement and Statistical Literacy in Questioning of Statistical Data.....	225
5.2	Comparison between Problems with and without Context	227
5.3	Implications for Educational Practices.....	230
5.4	Recommendations for Further Research.....	232
	REFERENCES	235
	APPENDICES	249

A.	APPROVAL FROM METU HUMAN SUBJECTS ETHICS COMMITTEE	249
B.	PERMISSION FROM MINISTRY OF NATIONAL EDUCATION	250
C.	STATISTICAL LITERACY TEST WITH CONTEXT PROBLEMS ..	251
D.	STATISTICAL LITERACY TEST WITHOUT CONTEXT PROBLEMS.....	262
E.	TABLE OF SPECIFICATION	270



LIST OF TABLES

TABLES

Table 2.1 Statistical Literacy Construct of Watson and Callingham (2003, p.14) ..	19
Table 2.2 Relationships of models for developing student understanding of variation as a foundation for chance and data (Watson et al., 2003, p.20)	24
Table 2.3 Categories of Statistical Ability (Kimura, 1999; retrieved from Aoyama & Stephens, 2003).....	26
Table 2.4 Frequencies of Research Studies by Years (Tosun & Ünal, 2019)	34
Table 3.1 Table of Content for the Problems in the Statistical Literacy Tests	53
Table 3.2 The Time Schedule for the Data Collection Procedure	78
Table 3.3 Original version of rubric for problem 15 (Watson & Callingham, 2004 ,p.153)	79
Table 3.4 Rubric for problem 15 in the SLT-CP	80
Table 3.5 Rubric for problem 15 in the SLT	80
Table 3.6 Codes and Corresponding Categories for Student Responses	82
Table 4.1 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Measures of Central Tendency Domain	89
Table 4.2 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Graph Interpretation Problems.....	95
Table 4.3 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Sampling Domain	106
Table 4.4 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Questioning of Statistical Data Domain	109
Table 4.5 Student Responses for Problem 8	119
Table 4.6 Students' Responses to Problem 11	124
Table 4.7 Student Responses to Problem 12.....	130
Table 4.8 Student Responses to Problem 13.....	135
Table 4.9 Student Responses to Problem 2.....	140

Table 4.10 Student Responses to Problem 3	147
Table 4.11 Student Responses to Problem 15	155
Table 4.12 Student Responses to Problem 22	164
Table 4.13 Student Responses to Problem 1	172
Table 4.14 Student Responses to Problem 14	178
Table 4.15 Student Responses to Problem 17	188
Table 4.16 Student Responses to the 19 th problem	199
Table 4.17 Descriptive statistics for statistical literacy problems with context and problems without context	211
Table 4.18 Independent sample t-test for total mean scores in problems with context and problems without context.....	211
Table 4.19 Independent sample t-test for mean scores in problems with context and problems without context according to the measures of central tendency domain	213
Table 4.20 Independent sample t-test for mean scores in problems with context and problems without context according to the graph interpretation domain.....	214
Table 4.21 Independent sample t-test for mean scores in problems with context and problems without context according to the sampling domain.....	215
Table 4.22 Independent sample t-test for mean scores in problems with context and problems without context according to the questioning of statistical data domain	216

LIST OF FIGURES

FIGURES

Figure 2.1. Relationship between the Components of Statistical Literacy (Watson, 2006, p.48)	14
Figure 2.2. Gal's Statistical Literacy Model (2002, p.4)	16
Figure 2.3. Five Parts of the Statistical Knowledge Base (Gal, 2002, p.10)	17
Figure 2.4. A sample item from Watson & Callingham (2003, p.35)	20
Figure 2.5. The Distribution of Smokers and Runners by Religion (Schield, 2006, p.2)	28
Figure 2.6. A model of mathematical literacy in practice (OECD, 2013, p.26)	38
Figure 2.7. A sample problem from test (Carraher et al.,1985, p.26)	39
Figure 2.8. A sample problem with penalties context (Boaler ,1993 p.351)	42
Figure 2.9. A sample problem with fertilizers context (Boaler, 1993 p.352)	42
Figure 2.10. A sample question without context (Boaler, 1993, p.352)	43
Figure 2.11. A problem situation with context including unknown variables (Gravemeijer, 1994 p.464)	44
Figure 3.1. The sample size of the study	51
Figure 3.2. The original version of the 1 st problem	55
Figure 3.3. The 1 st problem in the test	55
Figure 3.4. The 2 nd problem in the test	56
Figure 3.5. The 3 rd problem in the test	56
Figure 3.6. The 4 th , 5 th , 6 th and 7 th problems in the test	57
Figure 3.7. The original version of the 8 th problem (Watson & Callingham, 2003)	59
Figure 3.8. The 8 th item in the test	59
Figure 3.9. The 9 th problem in the test	60
Figure 3.10. The 10 th problem in the test	60
Figure 3.11. The bar chart used for the 11 th , 12 th , and the 13 th problems	61
Figure 3.12. The 11 th , 12 th and 13 th problems in the test	61

Figure 3.13. The 14 th problem in the test	62
Figure 3.14. The original version of 15 th problem in the test (Watson & Callingham, 2003)	62
Figure 3.15. The 15 th problem in the test	63
Figure 3.16. The original version of the 16 th problem in the test (PISA, 2011).....	64
Figure 3.17. The 16 th problem in the test	65
Figure 3.18. The original version of the 17 th problem(Sharma, 2017, p.125).....	65
Figure 3.19. The 17 th problem in the test	66
Figure 3.20. The 18 th problem in the test	67
Figure 3.21. The original version of the 19 th problem (Sharma, 2017, p.125).....	68
Figure 3.22. The 19 th problem in the test.....	68
Figure 3.23. The original version of the 20 th problem (delMas et al., 2007).....	69
Figure 3.24. The original version of the 21 st problem (delMas et al., 2007).....	69
Figure 3.25. The 20 th and 21 st problems in the test	70
Figure 3.26. The original version of the newspaper article used in the 22 nd item of the test (Watson & English, 2006)	70
Figure 3.27. The 22 nd problem in the test.....	71
Figure 3.28. The original version of the 23 rd , 24 th and 25 th problems (Sorto &White, 2004).....	71
Figure 3.29. The 23 rd , 24 th and 25 th problems in the test	72
Figure 3.30. The 16 th problem in the SLT-CP.....	73
Figure 3.31. The 16 th problem in the SLT	74
Figure 4.1. The statistical response of participant 120 to P8 in the SLT-CP	120
Figure 4.2. The statistical response of participant 141 to P8 in the SLT-CP	120
Figure 4.3. The statistical response of participant 340 to P8 in the SLT	121
Figure 4.4. The non-statistical response of participant 127 to P8 in the SLT-CP. 122	
Figure 4.5. The unrelated response of participant 65 to P8 in the SLT-CP	123
Figure 4.6. The statistical response of Participant 71 to P11 in the SLT-CP.....	125
Figure 4.7 The statistical response of Participant 319 to P11 in the SLT	126
Figure 4.8. The pre-statistical response of Participant 117 to P11 in SLT-CP	126

Figure 4.9. The unrelated response of participant 234 to P11 in the SLT	128
Figure 4.10. The unrelated response of participant 187 to P11 in the SLT	129
Figure 4.11. The statistical response of participant 233 to P12 in the SLT-CP.....	131
Figure 4.12. The statistical response of participant 234 to P12 in the SLT-CP.....	132
Figure 4.13. The unrelated response of participant 75 to P12 in the SLT-CP.....	133
Figure 4.14. The unrelated response of participant 49 to P12 in SLT-CP.....	133
Figure 4.15. The unrelated response of participant 47 to P12 in the SLT-CP.....	134
Figure 4.16. The statistical response of participant 237 to P13 in the SLT-CP.....	136
Figure 4.17. The unrelated response of participant 208 to P13 in the SLT-CP.....	138
Figure 4.18. The statistical response of Participant 54 to P2 in the SLT-CP	141
Figure 4.19. The statistical response of Participant 107 to P2 in the SLT-CP	142
Figure 4.20. The statistical response of Participant 76 to P2 in the SLT-CP	142
Figure 4.21. The statistical response of Participant 480 to problem 2 in the SLT.	143
Figure 4.22. The statistical response of Participant 236 to P2 in the SLT.....	143
Figure 4.23. The pre-statistical response of Participant 121 to P2 in SLT-CP.....	144
Figure 4.24. The non-statistical response of Participant 92 to P2 in the SLT-CP.	145
Figure 4.25. The statistical response of Participant 89 to P3 in the SLT-CP	148
Figure 4.26. The statistical response of Participant 47 to P3 in the SLT-CP	149
Figure 4.27. The statistical response of Participant 402 to P2 in the SLT.....	150
Figure 4.28. The statistical response of Participant 330 to P2 in the SLT.....	150
Figure 4.29. The pre-statistical response of Participant 194 to P2 in SLT-CP.....	151
Figure 4.30. The pre-statistical response of Participant 158 to P2 in the SLT	151
Figure 4.31. The non-statistical response of Participant 104 to P2 in SLT-CP	152
Figure 4.32. The non-statistical response of Participant 147 to P2 in SLT-CP	153
Figure 4.33. The unrelated response of Participant 375 to P2 in the SLT.....	154
Figure 4.34. The statistical response of participant 141 to P15 in the SLT-CP.....	156
Figure 4.35. The statistical response of participant 334 to P15 in the SLT.....	157
Figure 4.36. The pre-statistical response of participant 120 to P15 in SLT-CP	158
Figure 4.37. The pre-statistical response of participant 394 to P15 in SLT-CP	159
Figure 4.38. The non-statistical response of participant 171 to P15 in SLT-CP ...	160

Figure 4.39. The non-statistical response of participant 407 to P15 in the SLT ...	161
Figure 4.40. The unrelated response of participant 323 to P15 in the SLT.....	162
Figure 4.41. The unrelated response of participant 343 to P15 in the SLT.....	163
Figure 4.42. The statistical response of participant 138 to P22 in the SLT-CP	165
Figure 4.43. The statistical response of participant 120 to P22 in the SLT-CP	166
Figure 4.44. The statistical response of participant 386 to P22 in the SLT	166
Figure 4.45. The pre-statistical response of participant 130 to P22 in SLT-CP....	168
Figure 4.46. The non-statistical response of participant 84 to P22 in SLT-CP.....	169
Figure 4.47. The statistical response of participant 47 to P1 in the SLT-CP	173
Figure 4.48. The pre-statistical response of Participant 180 to P1 in SLT-CP	174
Figure 4.49. The pre-statistical response of Participant 419 to P1 in the SLT.....	175
Figure 4.50. The non-statistical response of Participant 176 to P1 in SLT-CP.....	176
Figure 4.51. The non-statistical response of Participant 434 to P1 in SLT-CP.....	176
Figure 4.52. The statistical response of participant 171 to P14 in the SLT-CP	179
Figure 4.53. The statistical response of participant 234 to P14 in the SLT-CP	180
Figure 4.54. The statistical response of participant 312 to P14 in the SLT	181
Figure 4.55. The statistical response of participant 314 to P14 in the SLT	182
Figure 4.56. The pre-statistical response of participant 95 to P14 in SLT-CP.....	183
Figure 4.57. The pre-statistical response of participant 498 to P14 in the SLT	184
Figure 4.58. The non-statistical response of participant 11 to P14 in SLT-CP.....	185
Figure 4.59. The unrelated response of participant 91 to P14 in the SLT-CP	186
Figure 4.60. The statistical response of participant 114 to P17 in the SLT-CP	189
Figure 4.61. The statistical response of participant 207 to P17 in the SLT-CP	189
Figure 4.62. The statistical response of participant 173 to P17 in the SLT-CP	190
Figure 4.63. The statistical response of Participant 481 to P17 in the SLT	191
Figure 4.64. The statistical response of participant 371 to P17 in the SLT	192
Figure 4.65. The pre-statistical response of participant 180 to P17 in SLT-CP....	193
Figure 4.66. The pre-statistical response of participant 393 to problem 17 in SLT	194
Figure 4.67. The non-statistical response of participant 177 to P17 in SLT-CP...	195

Figure 4.68. The non-statistical response of participant 280 to P17 in the SLT....	196
Figure 4.69. The unrelated response of participant 129 to P17 in the SLT-CP.....	197
Figure 4.70. The unrelated response of participant 396 to P17 in the SLT	197
Figure 4.71. The statistical response of participant 91 to P19 in the SLT-CP.....	200
Figure 4.72. The statistical response of participant 174 to P19 in the SLT-CP.....	201
Figure 4.73. The statistical response of participant 348 to P19 in the SLT	202
Figure 4.74. The pre-statistical response of participant 299 to P19 in SLT-CP	203
Figure 4.75. The pre-statistical response of participant 417 to P19 in the SLT	204
Figure 4.76. The non-statistical response of participant 129 to P19 in SLT-CP ...	206
Figure 4.77. The non-statistical response of participant 217 to P19 in SLT-CP ...	207
Figure 4.78. The non-statistical response of participant 323 to P19 in SLT-CP ...	208
Figure 4.79. The unrelated response of participant 215 to P19 in the SLT-CP	209
Figure 4.80. The unrelated response of participant 336 to P19 in SLT-CP	210

LIST OF ABBREVIATIONS

ABBREVIATIONS

SLT	Statistical Literacy Test with Context Problems
SLT-CP	Statistical Literacy Test without Context Problems
MoNE	Ministry of National Education
SPSS	Statistical Package for Social Sciences
TIMMS	Trends in International Mathematics and Science Study
NCTM	National Council of Teachers of Mathematics

CHAPTER 1

INTRODUCTION

Statistics is an essential part of today's information-driven world. It takes part in a large proportion of everyday information resources from scientific news such as weather forecasting to political news such as interpreting poll results. (Watson, 1997; Schield, 2010). Put differently, people encounter statistics in every aspect of life, and they have to deal with statistics around them. Therefore, individuals need 'literacy' regarding statistics by comprehending and critically evaluating statistical concepts (Trewin, 2005; Watson, 1997). Otherwise, individuals who lack statistical literacy could not discriminate reliable and unreliable data embedded in statistical messages (English & Watson, 2016; Gal, 2002). In this respect, statistical literacy is the most anticipated ability of individuals in data-laden communities and is also considered as the intended result of schooling and as an indispensable portion of adults' literacy. (Gal, 2002). It is also stated by English and Watson (2016) that "preparing our students to be statistically literate in today's world is paramount" (p. 3). As, being a statistically literate individual brings about reading and critically evaluating statistical data properly.

The importance of statistical literacy is shown by adding statistical literacy to the curricula around the world (AEC, 1994a; MoE, 2001; NCTM, 2000; MoNE, 2018). For instance, the Australian mathematics curriculum (AEC, 1991) involves *data and chance* content domain to educate statistically literate citizens by stating how important it is to "understand the impact of statistics on daily life" (p.178). Additionally, the National Council of Mathematics Teachers (2000) emphasizes the indispensability of statistical understanding in terms of critically evaluating the statistical data encountered in daily life. In Turkey, *data handling* is one of the five

learning domains of the Turkish middle school mathematics curriculum (MoNE, 2018). This domain aims to help students make sense of statistical concepts, including measures of central tendency, graphs, sampling, and inferences regarding statistical data. The objectives of the data handling learning domain in the Turkish middle school mathematics curriculum (MoNE, 2018) aim to equip students with the ability of reading and interpreting statistical data which can appear in daily life. In this respect, the 8th grade students' statistical literacy in terms of investigating their ability to read and interpret the statistical data that they encounter in daily life need to be evaluated and validated. In that regard, the present study aims to examine the 8th grade students' statistical literacy and achievement in statistical literacy problems.

Moreover, some research studies (Aoyama & Stephens, 2003; delMas et al., 2005; Friel, Curcio & Bright, 2001; Henriques & Cruz, 2010; Monteiro & Ainley, 2003; Watson, 1997; Watson & Callingham, 2003; Watson & Kelly, 2008; Watson & Moritz, 2000; Schield, 2010;) were conducted in the field of statistics education as a result of the increasing interest in statistical literacy. Some researchers attempted to derive a definition for statistical literacy (e.g., Gal, 2002; Garfield et al., 2003; Wallham, 1993), while other researchers sought to establish frameworks to determine the construct of statistical literacy (Gal, 2002; Watson, 1997; Watson & Callingham, 2003). They examined statistical literacy in terms of affective and cognitive domains.

Watson (1997), who is the pioneer of the research regarding statistical literacy, attempted to define statistical literacy with a three-tiered hierarchy framework. The first tier of Watson (1997) was related to the basic understanding of statistical terminology; the second tier was about the comprehension and interpretation of statistical concepts within a specific context; and the third tier was related to the questioning of statistical arguments presented in media. With the help of these tiers, Watson (1997) defined statistical literacy as the capacity to understand, interpret and analyze statistical messages in various contexts. Following Watson's (1997) hierarchy, many researchers investigated students' statistical literacy (Watson & Moritz, 2000; Watson & Callingham, 2003). Building on Watson's (1997) hierarchy,

Watson and Callingham (2003) developed six levels of hierarchy to conduct in-depth analysis for students' statistical literacy. The hierarchy of Watson and Callingham (2003) became an essential milestone in the field of statistical literacy.

The place of real-life context including social and scientific situations in developing and investigating statistical literacy is an ongoing discussion in the literature (Begg, 1997; Ben-Zvi & Chick et al., 2005; English & Watson, 2016; Gal, 2002; Garfield, 2004; Kelly, Sloane & Whittaker, 1997; Lesh, Amit & Schorr, 1997; Watson, 1997). The researchers pointed out that it is important to use real-world contexts from various media channels which students encounter in daily life. For instance, Chick and colleagues (2005) associated statistical literacy with "transnumerative thinking" by encouraging students to collect, interpret, represent, and analyze data in real-world contexts. In this respect, examining students' statistical literacy through problems, including real-life context, is an essential aspect of statistical literacy research. Compared to this, in the field of research on mathematics education, there are two distinct views regarding the utilization of context in problems. According to the first view, the utilization of context provides benefits for students by engaging them in the problem solving process by activating their daily life experiences and motivating students to solve problems (Chapman, 2006; De Lange, 2003; Kliebard & Franklin, 2003; Lappan & Phillips, 2009; Steen, 2001; Van-den Heuvel-Panhuizen, 2005; Verschaffel, 2000). According to the second view, involving a context in problems is a complex construct which requires sequential and complicated cognitive processes as well as arithmetic and algorithmic operations. (Fusch et al., 2006; Schnotz, 2002; Schnotz et al., 2010). Thus, the utilization of context in problems complicates the problem situation as it requires extra effort and information regarding the context of the problem (Boaler, 1993; Gravemeijer & Doorman, 1999; Roth, 1996; Treffers, 1987) and differentiates students in terms of their interests, socio-economic status and life experiences (Cooper & Dunne, 1999).

Following the previously discussed ideas and studies around statistical literacy and the effects of context on the understanding of its concepts, this study aims to investigate 8th grade students' achievement and statistical literacy through statistical

literacy problems with and without context in terms of the statistical content domains in the curriculum. Subsequently, the students' achievement and statistical literacy were compared and contrasted in terms of the existence of the context in the problems.

1.1 Aims of the Study

The first aim of the present study is to investigate 8th grade students' achievement and statistical literacy through problems with and without context in terms of statistical content domains. Another aim of the present study is to compare 8th grade students' achievement in problems with and without context.

1.2 Research Questions

The present study addresses the following research questions:

1. How successful are the 8th grade students in problems with and without context in terms of statistical content domains (*measures of central tendency, graph interpretation, sampling, questioning of statistical data*)?
2. How is the statistical literacy of 8th grade students in problems with and without context in terms of statistical content domains (*measures of central tendency, graph interpretation, sampling, questioning of statistical data*)?
3. Is there any significant mean difference in students' achievement in statistical literacy problems with and without context?

1.3 Definitions of Important Words

Statistical Literacy: As Watson (1997) stated, statistical literacy is the ability to understand, interpret, and analyze statistical messages in various contexts.

Social-Scientific Context: Scientific or social contexts correspond to printed media such as journal excerpts, scientific reports and magazines people encounter. In the current study, newspapers excerpts, scientific reports and infographics were presented with real data from reliable sources compiled by the researcher.

Problems with Context: These are the problems in which the problem situation including social and scientific context is experientially real to the student. (Doorman & Gravemeijer, 1999)

Problems without Context: These are the problems which do not include social and scientific context.

1.4 Significance of the Study

The data is ubiquitous in the data-driven era as it has permeated into every aspects of individuals' lives. To be a part of the society effectively, people need proficiency in understanding statistics. In addition, people should be able to differentiate misleading information from reliable information through powerful literacy in statistical concepts (Gal, 2002; Wallman, 1993). Therefore, around the world, many curriculum documents (AEC, 1994a; MoE, 2001; MoNE, 2018; NCTM, 2000) have acknowledged the importance of developing students as statistically literate individuals. More specifically, in Turkey, the elementary mathematics curriculum aims to develop conscious citizens who can read and interpret different forms of statistical data encountered in daily life. Thus, it can be said that educating statistically literate and informed citizens is a crucial objective for the curriculum documents around the world. In accordance with this objective, the researchers attempted to define statistical literacy (Gal, 2002; Garfield et al., 2003; Wallham, 1993), derived models for statistical literacy (Gal, 2002; Sharma, 2017; Watson, 1997; Watson & Callingham, 2003) and conducted research studies to investigate students' statistical literacy (e.g., Friel, Curcio & Bright, 2001; Gal, 1995; Mokros & Russell, 1995; Monteiro & Ainley, 2003; Pollatsek et al., 1981; Watson & Moritz,

2000; Watson & Callingham, 2003) in terms of different content domains regarding data and chance curriculum. However, the number of research studies in the field of statistical literacy is very limited in Turkey. As research on statistical literacy is a contemporary topic with many new and exciting research directions, the present study aimed to investigate students' achievement and statistical literacy in statistical literacy problems. With the help of the present study, researchers, curriculum developers, and educators can gain an insight into 8th grade students' achievement and statistical literacy through statistical literacy problems.

On the other hand, the use of context in problems is a disputable issue in mathematics education research. Several researchers (Carragher et al., 1985; Chapman, 2006; Lappan & Phillips, 2009; Mitchell & Carbone, 2011; Nunes et al., 1993; Van-den Heuvel-Panhuizen, 2005) promote the utilization of context in problems to enhance students' motivation and interest in mathematical tasks by concentrating on mathematics (Thompson et al., 1994) and by increasing the authenticity of the task (Palm, 2009). However, some researchers (Boaler, 1993; Fusch et al., 2006; Gravemeijer, 1994; Schnotz, 2002; Schnotz et al., 2010; Walkerdine, 1990) do not endorse the use of context in problems. According to these researchers, the utilization of context in problems complicates the problem by requiring extra information about the context of the problem. On the other hand, some researchers (Ben-Zvi & Garfield, 2004; Callingham, 2003; English & Watson, 2016; Gal, 2002; Watson, 1997; Watson & Gal, 2005) working in the field of statistics education emphasize the importance of context in teaching, learning and examining statistical literacy. However, the number of studies on students' statistical literacy with regard to the existence of context is limited. In the literature, there are no research studies investigating students' statistical literacy through problems with and without context simultaneously. At this point, the present study sheds light on the effect of the existence of context in statistical literacy problems on students' achievement and statistical literacy.

Furthermore, the present study investigates 8th grade students' statistical literacy with open-ended, multiple choice, and True-False items. In the Turkish context, students'

statistical literacy was investigated through open-ended problems. In fact, open-ended tasks enable students to explain their understanding of statistical problems in detail (Gal, 1995; Watson, 2013). From this perspective, the use of various forms of problems while examining the students' statistical literacy makes the current study significant as it enables the students to explain their understanding in detail.

Moreover, the present study gives insight to pre-service and in-service teachers about eight grade students' statistical literacy with regard to content domains of measures of central tendency, graph interpretation, sampling, and questioning of statistical data. This study is also significant in that it provides curriculum developers with evidence to make adjustments and reconsider the objectives regarding the data and chance unit in the elementary mathematics curriculum. Furthermore, this study informs teacher educators about the problematic situations regarding the development of statistical literacy. Lastly, the present study provides insight into the selection of statistical tasks to add context in problems.

CHAPTER 2

LITERATURE REVIEW

2.1 Statistical Literacy

2.1.1 Definition of Statistical Literacy

Today, statistical literacy has become a foremost issue to access information in education (Steen, 1997). Around the world, statistical literacy is the ultimate objective in mathematics curriculum (AEC, 1994; MOE, 2001; MoNE, 2018; NCTM, 2000). According to the MoNE (2018), at the end of elementary school, students should be able to present, interpret, and evaluate statistical data properly. Correspondingly, many researchers attempted to provide a definition for statistical literacy. The definition of statistical literacy is not mature yet. More specifically, there is no agreed-upon definition of statistical literacy, which indicates that it is still under development (Chance, 2002; Sharma, 2017; Shaugnessy, 2007; Rumsey, 2002). In international mathematics education research, studies were conducted to define statistical literacy (e.g., Callingham, 2007; Gal, 2002; Garfield et al., 2003; Wallham, 1993). The current definitions of statistical literacy are discussed below.

Gal (2000) defined statistical literacy as people's capacity to understand and investigate statistical data from new media channels and their capability of expressing their ideas over this kind of information. Correspondingly, Gal (2002) stated that statistical literacy could be described with two intertwined parts. The capability of interpreting and critically evaluating statistical messages is the first part. The second part is the capability of discussing or connecting people's feedback for statistical messages. The researcher indicated that statistical literacy is essential to

acquire social and personal skills. Gal (2002) indicated that these two parts are intertwined and built on various knowledge bases and dispositions. In parallel with Gal's statements, Garfield and colleagues (2003) defined statistical literacy as the comprehension of statistical information or research findings including concepts, symbols, and language by utilizing basic and substantial abilities. These abilities are organizing data, creating and displaying tables, and studying various representations of data. In this way, Garfield and colleagues (2003) provided the ways for Gal's (2000) definition to understand, investigate and express ideas about statistical data gathered from new media channels.

Wallham (1993, p.1) defined statistical literacy as follows:

“Statistical literacy is the ability to understand and critically evaluate statistical results that permeate our daily lives coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions”.

Similar to the definition of Gal (2002), the definition of Wallman (1993) pointed to individual and social need for statistical literacy. Accordingly, Watson (1997) described statistical literacy as comprehension of text, and the meaning and implications of the statistical information included in it in the context of the topic to which it belongs. In other words, Wallman (1993) and Watson (1997) focused on the importance of social context in statistical literacy. In addition to Wallman (1993) and Watson (1997), Callingham (2007) asserted that such a definition is needed to enhance students' mathematical skills not only to comprehend statistical information but also to understand the social context where the data was obtained.

Ben-Zvi and Garfield (2004) distinguished between statistical literacy, statistical thinking, and statistical reasoning by utilizing the definitions discussed above and the definitions of Garfield and colleagues (2003), who defined statistical literacy as the ability to organize and represent data by constructing and displaying tables, whereas statistical reasoning is the interpretation of these tables and summaries of the data. In addition to these, statistical thinking includes the conceptual

understanding of statistical data, such as criticizing and analyzing statistical information. While reading, identifying, and describing statistical information is related to statistical literacy, explaining the process about statistical information is related to statistical reasoning. Furthermore, applying, criticizing and evaluating statistical information is related to statistical thinking. Similar to Ben-Zvi and Garfield (2004), delMas (2002) made a distinction between statistical literacy, statistical thinking and statistical reasoning. delMas (2002) distinguished these three domains by asking critical questions related to the statistical content. More precisely, delMas (2002) suggested an investigation for the nature of the items to determine which domain or domains the task is relevant to.

Rumsey (2002) preferred to use two distinct expressions as “*statistical competence*” and “*statistical citizenship*” instead of the term “*statistical literacy*”. Statistical competence is described as the knowledge which is compulsory to think and reason statistically. It includes data awareness, basic statistical concepts, data generation, interpretation and communication abilities. Additionally, statistical citizenship is described as the main aim to enhance and become a trained person in the data-driven world (Rumsey, 2002). In other words, statistical literacy is being a citizen who successfully obtains, criticizes and determines to utilize statistical information. Rumsey (2002) added that integrating context in statistical problems might strengthen students’ statistical understanding, reasoning, and thinking skills.

As a consequence, many definitions were provided by the researchers for statistical literacy. Statistical literacy is interpreting and evaluating statistical messages critically, and it involves an intense engagement with media (Gal, 2000; Watson, 1997). Additionally, statistical literacy includes the ability to understand concepts, symbols and language to organize, create and display data (Gal, 2002; Garfield et al., 2003). Statistical literacy is related but different from statistical thinking and statistical reasoning. To make a distinction between these constructs, the nature of an item should be investigated in detail by posing some critical questions (Ben-Zvi & Garfield, 2004). Furthermore, statistical literacy can also be expressed with the terms “*statistical competence*” and “*statistical citizenship*” (Rumsey, 2002). Those

terms include the knowledge of statistical thinking-reasoning and development to be an educated individual in an information-driven society. On the other hand, the significance of the use of context in statistical literacy was also emphasized by researchers (Callingham, 2007; Rumsey, 2002; Watson, 1997). Within the scope of all these definitions and assertions, the present study aims to investigate students' achievement and statistical literacy through statistical literacy problems with context and statistical literacy problems without context.

Snell (1999) stated that statistical literacy is the capability of comprehending and reasoning about statistical concepts at the most basic level. Ben-Zvi and Makar (2011) argued that this ability does not include traditional approaches that focus mainly on abilities, procedures, and computations, which do not lead students to reason or think statistically. This ability should include the capability to interpret, evaluate, and criticize statistical data (Gal, 2002). From this perspective, researchers have sought to derive models and frameworks to conceptualize the components or levels of statistical literacy (e.g., Gal, 2003; Watson, 1997; Watson & Callingham, 2004). These models and frameworks are discussed in the following section.

2.1.2 Theoretical Models and Frameworks for Statistical Literacy

In the literature, a number of models and frameworks have been introduced by researchers (e.g., Gal, 2002; Sharma, 2017; Watson, 1997; Watson & Callingham, 2003). In this section, the current models and theoretical frameworks for statistical literacy are discussed.

2.1.2.1 Watson's Three-Tiered Model and the Components of Statistical Literacy:

In terms of describing the related skills and levels of complexity, Watson (1997) emphasized that statistical thinking in social contexts is a crucial part of statistics education. Watson (1997) offered a three-tiered statistical literacy model by

adjusting models from developmental psychology (e.g., Biggs & Collis, 1982; Case, 1985). In this framework, the researcher mostly focused on the cognitive dimension of statistical literacy rather than the affective dimension (Yolcu, 2012).

The first tier of Watson's (1997) model includes abilities related to particular themes from curriculum such as measures of central tendency, graphing, and probability. Context is not required as long as the student questions to learn the present topic. Watson (1997) suggested encouraging students to find an intrinsic reason themselves to attract their attention.

The second tier of Watson's (1997) model requires the identification of different contexts, such as social and scientific context. This stage requires the comprehension and interpretation of statistical information to obtain results and to make statistical inferences about these results.

In the third tier of Watson's (1997) model, a basic understanding of context is not sufficient for statistical literacy. At this stage, the students are expected to question the statistical claims, in addition to a comprehensive understanding of the context.

Watson's three-tiered model was utilized in various domains of statistics in research studies (e.g., Watson & Callingham, 2003; Watson & Moritz, 2000a; Watson & Moritz, 2000b). Based on these studies, Watson (2006) offered a model examining the structure of statistical literacy. The researcher illustrated the main components of statistical literacy by forming association with variation, task context, literacy skills, mathematical/statistical skills, task motivation, and task format. The model is presented in Figure 2.1 as follows:

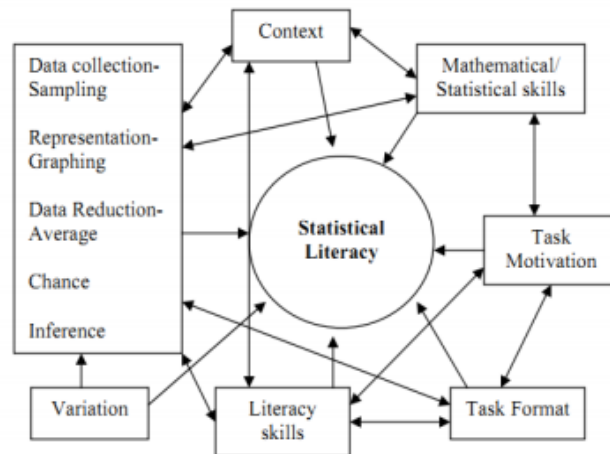


Figure 2.1. Relationship between the Components of Statistical Literacy (Watson, 2006, p.48)

The first component of the model is *statistical skills*. According to the middle grade curriculum, various mathematical (rates, percent, proportions, part-whole relationship) and statistical (comprehending average and probability) abilities are essential to progress in statistical literacy (Watson, 2006).

The second component is *variation*. According to Watson (2006), to develop statistical literacy, students should comprehend the meaning of variation and recognize the relationship between variation and sample.

The third component of the model is *context*. Similar to Watson (2006), many researchers recognized the importance of context for competency in statistical concepts (e.g., Cobb & Moore, 1997; Gal, 2002; Makar et al., 2011; Wallman, 1993). Cobb and Moore (1997) distinguished statistics from mathematics by indicating that '*data are not just numbers; they are numbers with a context*' (p.801). Researchers laid emphasis on context for literacy in statistics. According to Watson (2006), context is quite an important element of statistical literacy in terms of cognition with the problem situation. Watson (2006) categorized the problem context as *isolated context*, *familiar context*, and *unfamiliar context*. For the problems with *isolated context*, understanding the rule is sufficient such as tossing a dice and reading the table. For the problems with *familiar context*, the integration of students' experiences at school such as conducting a survey is necessary. To integrate *unfamiliar context*

into problems, various parts should be presented from media channels involving the nonconversant context to students. To obtain high competency in statistical literacy, students should develop a proficiency in unfamiliar context (Watson, 2006).

The fourth component is *literacy skills*. Watson (2006) used a framework by Luke and Freebody (1997). In this framework, four elements of essential practices for readers are explained. The first element of the framework is related to the decoding of data such as grasping the aim of using graphs in a text. The second element of the framework emphasizes the effect of context on social and cultural dimensions, such as changes in the meaning of average within the scope of context. The third element emphasizes generating meanings about statistical concepts by using graphs, samples, data, and results to attribute a meaning to the text. The fourth element comprises critical questioning by utilizing the reading text. Luke and Freebody (1997) stated that these elements are not in hierarchical order. Watson (2006) argued that the elements of Luke and Freebody (1997) and three-tiered hierarchy have common aspects.

The fifth component of the model is the *task format*. Multiple choice questions and open-ended questions are two of the possible types of questions to assess statistical literacy. The use of multiple-choice items might be engaging for students by providing various alternatives.

The last component of the model is *task motivation*, which includes dispositions required for statistical literacy tasks. By involving the component of task motivation in statistical literacy, Watson (2006) emphasized the affective aspect of statistical literacy in parallel with the perspectives in the existing literature (Gal, 2002; Wild & Pfannkuch, 1999). Correspondingly, Gal (2002) proposed a model for statistical literacy as discussed below.

2.1.2.2 Gal's (2002) Model

Gal (2002) proposed a model for adults aiming at understanding and critically evaluating statistical concepts in various contexts. The model involves knowledge elements and dispositional elements as presented in Figure 2.2.

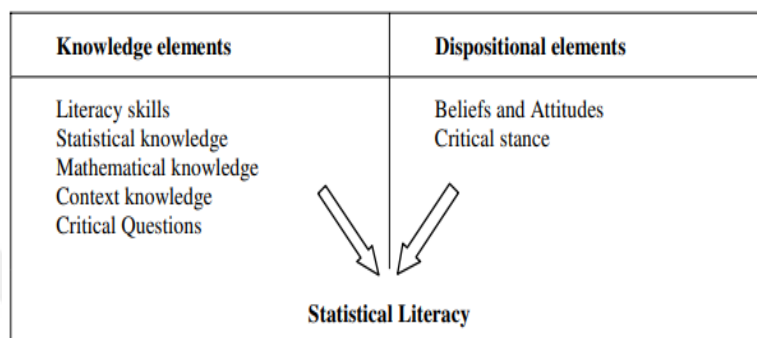


Figure 2.2. Gal's Statistical Literacy Model (2002, p.4)

As can be observed in the model of Gal (2002), knowledge elements involve literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions. Respectively, dispositional elements include beliefs and attitudes and critical stance.

The elements of Gal's (2002) model are not isolated. On the contrary, they have common and different points.

2.1.2.2.1 Knowledge Elements

To be statistically literate, the recipient should be acquainted with the meaning of statistical concepts in media excerpts such as probability, random, variation and chance except for daily life usage. Additionally, statistical literacy necessitates *prose text literacy* and *non-prose text literacy*. *Prose text literacy* requires the comprehension and interpretation of texts excerpted from media (Kirsch & Mosenthal, 1990). *Non-prose text literacy*, on the other hand, requires understanding graphical and tabular representations.

Statistical knowledge is the other element of knowledge elements. This element involves the basic skills required to comprehend statistical information that appear in media. As a result of studies by many researchers (e.g., Garfield, 1997; Moore, 1990; Shaughnessy, 1992), Gal (2002) proposed five parts of statistical knowledge base which are presented in Figure 2.3.

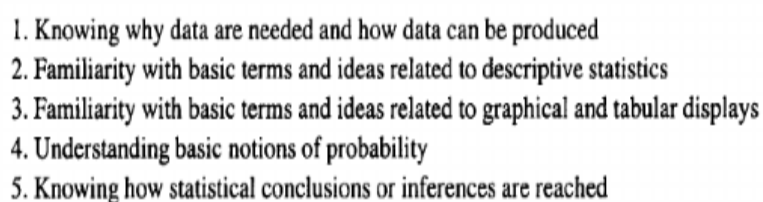
- 
1. Knowing why data are needed and how data can be produced
 2. Familiarity with basic terms and ideas related to descriptive statistics
 3. Familiarity with basic terms and ideas related to graphical and tabular displays
 4. Understanding basic notions of probability
 5. Knowing how statistical conclusions or inferences are reached

Figure 2.3. Five Parts of the Statistical Knowledge Base (Gal, 2002, p.10)

According to five parts of statistical knowledge base, adults should be aware of the need for data by being familiar with the main statistical concepts and terms such as mode, median, mean and graphs. In addition to simple probabilistic concepts, adults should also be aware of the statistical operations such as sampling, the steps for making inferences and recognizing the characteristics of measures of central tendency, graphs and charts. (Gal, 2002)

Mathematical knowledge is another element of knowledge base. Adults need basic mathematical knowledge to build a connection between statistical knowledge and mathematical knowledge. However, Gal (2002) could not specify the amount of mathematics required for statistics. This is still discussed by a number of researchers.

Context knowledge is one of the elements of knowledge base which is essential for thinking critically about statistical messages. More specifically, knowledge of the context where data is produced is important for readers to create a link between the problem situation and statistical measures.

The last element of knowledge base is *critical skills*. Gal (2002) maintains that readers need critical skills to evaluate the acceptability and reliability of the information included in media excerpts. They should question the validity of the

information presented in media. The questions on readers' mind are called '*worry questions*' by Gal (2002). These questions could be exemplified as "Is that claim valid for this data set?", "Can this conclusion be really inferred from that table?" and "Is that graph appropriate to illustrate this data set? Worry questions lead to more reasonable interpretations of statistical situations.

2.1.2.2.2 Dispositional Elements

Dispositional elements of Gal's (2002) model are critical stance, beliefs and attitudes which are intertwined, i.e., not isolated from each other. In contrast to knowledge elements, dispositional elements cannot be categorized into sections. If adults are exposed to a statistical message, they question the message critically. They might be mistaken or prejudiced about statistical messages, which is related to adults' critical stance. Adults' critical stance might be affected by various factors such as beliefs and attitudes (Gal, 2002) as proposed by McLeod (2002). Although attitudes are steady and include strong emotions, beliefs involve cognitive constituents on a large scale.

2.1.2.3 Watson and Callingham's (2003) Construct of Statistical Literacy

Watson and Callingham (2003) collected data from 3852 students from grades 3 to 9 between the years of 1993 and 2000. The researchers specified six hierarchical levels from idiosyncratic to critical mathematics by conducting Rasch analysis. The statistical literacy construct of Watson and Callingham (2003) is in parallel with Watson's (1997) three tiered framework. The six levels of the construct are presented in Table 2.1.

Table 2.1 Statistical Literacy Construct of Watson and Callingham (2003, p.14)

Level	Brief characterization of step levels of tasks
6. Critical Mathematical	Task-steps at this level demand critical, questioning engagement with context, using proportional reasoning particularly in media or chance contexts, showing appreciation of the need for uncertainty in making predictions, and interpreting subtle aspects of language.
5. Critical	Task-steps require critical, questioning engagement in familiar and unfamiliar contexts that do not involve proportional reasoning, but which do involve appropriate use of terminology, qualitative interpretation of chance, and appreciation of variation.
4. Consistent Non-critical	Task-steps require appropriate but non-critical engagement with context, multiple aspects of terminology usage, appreciation of variation in chance settings only, and statistical skills associated with the mean, simple probabilities, and graph characteristics.
3. Inconsistent	Task-steps at this level, often in supportive formats, expect selective engagement with context, appropriate recognition of conclusions but without justification, and qualitative rather than quantitative use of statistical ideas.
2. Informal	Task-steps require only colloquial or informal engagement with context often reflecting intuitive non-statistical beliefs, single elements of complex terminology and settings, and basic one-step straightforward table, graph, and chance calculations.
1. Idiosyncratic	Task-steps at this level suggest idiosyncratic engagement with context, tautological use of terminology, and basic mathematical skills associated with one-to-one counting and reading cell values in tables.

The first level of the construct is the *idiosyncratic level*, which includes interpretation based on personal experiences apart from context. At this level, students basically read data from a table and interpret statistical messages based on their individual beliefs and experiences (Watson & Callingham, 2003). For instance, when statistics about environmental pollution is presented and questions are posed, the student might respond to the questions as “We should not pollute the environment” based on her/his personal beliefs by not taking into consideration the statistical constructs in the question.

The second level of the construct is *the informal level* in which students are involved in the context heuristically without statistical interpretation. They focus on unrelated points in statistical tasks (Watson & Callingham, 2003). For instance, when students are asked about the meaning of arithmetic mean, they provide only an example or one-word explanation for arithmetic mean. More specifically, students use idiosyncratic strategies to provide an answer.

The third level of the construct is *inconsistent level*. Compared to the first two levels, involvement in the context is necessary at this level. Students’ involvement changes

according to item format. The format of the item should encourage students to engage in the context of the item. When students are asked about the generalizability of a small sample in the context, they provide responses such as “There should be more people” or “The people are too old” at inconsistent level. As could be inferred, instead of focusing on remarkable aspects, students are concerned more about superficial characteristics of the statistical task.

The fourth level of the construct is *consistent-noncritical level*. Although students establish a connection with the context, they exhibit non-critical involvement in the context included in the statistical task. At this level, students can identify and illustrate tabular and graphical tasks obtained from media partially. For instance, when a chart is given in a problem context, students at consistent-noncritical level could make interpretations about the chart to some extent.

The fifth level is *critical level*. Similar to the tier-three of Watson’s (1997) hierarchy, students at this level have critical thinking skills. Although proportional reasoning is not necessary, critical thinking, proper use of terminology and engagement with statistical concepts are essential at this level. Students should be able to conduct statistical procedures with two variables simultaneously. They should also be able to interpret the graphs and detect errors in them. A sample from the study of Watson and Callingham (2003) is displayed in Figure 2.4.

M2PI. Explain the meaning of this pie chart. Is there anything unusual about it?

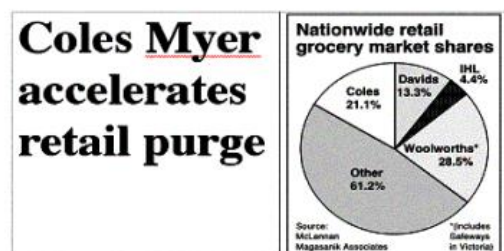


Figure 2.4. A sample item from Watson & Callingham (2003, p.35)

In this item, students at critical level are expected to provide an interpretation about the chart in context and detect the error about the percentages of the parts of the whole in the chart. Students who notice that the whole exceeds 100% and interpret the national grocery market sales are considered to be at the critical level of the construct.

Critical mathematical level is the sixth level. At this level, reasoning skills different from those at the fifth level are required. In addition to reading and interpreting data, students should be able to summarize, make predictions and inferences about statistical expressions in context. This level is similar to tier-three of Watson's (1997) hierarchy.

2.2 Research Studies on Statistical Literacy

In this section, the research studies conducted so far on statistical literacy are discussed in detail.

In the literature, although the number of research studies related to statistical literacy is limited, various studies have been conducted by many researchers (Friel, Curcio & Bright, 2001; Monteiro & Ainley, 2003; Watson, 1997; Watson & Callingham, 2003) in the field of statistics education. The research studies were conducted on many aspects of statistical literacy such as statistical inference (Watson et al., 2003; Watson & Kelly, 2008; Watson & Moritz, 2000), statistical literacy hierarchies (Aoyama & Stephens, 2003; Watson, 1997; Watson & Callingham, 2003), graphs (Aoyama & Stephens, 2003; Curcio, 1987; delMas et al., 2005; Henriques & Cruz, 2010; Schield, 2000), measures of central tendency (Cai et al., 1999; Gal, 1995; Mevarech, 1983; Mokros & Russell, 1995; Pollatsek et al., 1981; Strauss & Bichler, 1988; Watson & Moritz, 2000), and college level statistical literacy (Cimpoeru & Roman, 2018; Cooper, 2018; Yotongyos et al., 2015).

One of these studies belongs to Watson (1997), which is a milestone for most of the research studies on statistical literacy. The researcher worked with 670 students from

grades 6 to 9 and investigated students' statistical literacy based on the three-tiered hierarchy to reveal whether students understand probabilistic and statistical terminology and related concepts when they are involved in the context of a social situation. To achieve this aim, Watson (1997) presented various statistical situations to the students and requested them to interpret these situations including graph interpretation, sampling and questioning of claims. Watson (1997) classified student responses into three tiers. The results of the study indicated that the student responses exhibiting basic comprehension of statistical displays in context corresponded to tier-one and-tier two of Watson's (1997) hierarchy. The students with tier-three went beyond the statistical display and provided higher order responses.

In the next step of Watson's (1997) study, Watson and Moritz (2000) worked with students from grades 3, 6 and 9 on the concept of sampling. The researchers used the data obtained from a longitudinal survey. The student responses were classified as pre-structural, uni-structural, multi-structural and relational levels of Biggs and Collis (1982). Unanswered questions were regarded as pre-structural. For tier-one questions, the sequence of uni-structural- multi-structural-relational was useful to reinforce the notion of sample. The researchers found that the use of a familiar context helps to reinforce the concept of sample at early ages. In tier-two questions which include performing sampling in context, the number of correct responses increased as grade level increased. The researchers stated that since the context of the tier-two question was related to cars, the female students were less involved in the question compared to the male students. In tier-three questions, the researchers indicated that context might distract students' attention away from statistical concepts and direct it to different aspects of the question context. The content of the context might create disequilibrium with respect to gender, grade level or socio-economic levels of students. To observe students' longitudinal improvement, the study was repeated after two years and four years with the same students. After two years, 15% of the students performed at a lower level, 48% performed at a higher level, and 37% remained the same. After four years, 7% students performed at a lower level, 24% remained the same, 60% reached a higher level. The reason for

performing at a lower level over the years might be motivational issues. Additionally, it was concluded that the students might not have reached the higher level instantly, but gradually by experiencing the levels between lower level and higher level. Being aware of this fact might help teachers plan teaching according to students' existing level. In conclusion, Watson and Moritz (2000) investigated students' cognitive development without any intervention during the two-year and four-year period.

Many curriculum reports have emphasized the importance of the components of statistical inference and the importance of experiencing these components (AEC, 1991; MoNE, 2018; NCTM, 1989). For instance, NCTM (1989) indicated the significance of the sampling, randomness and variation concepts in the data collection process by recommending to involve all students in the questioning process for data collection. Correspondingly, Australian Education Council (1991) emphasized the importance of encouraging students to comprehend the concepts such as sampling, random selection and making inferences from sample. In this context, Watson and Kelly (2008) conducted a study with 738 students from grades 3, 5, 7 and 9 to investigate students' definition and examples of 'sample', 'random' and 'variation'. The researchers identified four levels regarding the comprehension of graphs, tables and prose texts based on the basic statistical concepts (variation, random and sample). Watson and Kelly (2008) replicated the study of Watson and Moritz (2000) after a two-year period to observe the longitudinal changes in students' performances. Two years later, it was observed that in addition to significant developments in variation concept, grade 7 students achieved the most considerable progress regarding the sample concept among all the students. On the other hand, grade 9 students' performance of the random concept did not change after two years, which might have originated from the decrease in their interest due to the confusion about literacy skills. In grades 3, 5 and 7, the concepts of random, sample and variation were consolidated with class activities. Their study was parallel to Watson et al. (2003)'s study. It could even be considered as an extension of Watson et al. (2003)'s study. The authors conducted the study with 746 students from the same grade levels. Watson et al. (2003) examined levels and the relationships of

models to develop students' understanding of variation, chance and data by utilizing Watson (1997) and Luke-Freebody's (1997) frameworks as presented in Table 2.2.

Table 2.2 Relationships of models for developing student understanding of variation as a foundation for chance and data (Watson et al., 2003, p.20)

Variable mapping (figure 1)	Luke and Freebody [41]	Statistical literacy hierarchy [40]
Level 4 – Critical aspects of variation: Employing complex justification or critical reasoning	Text analyst Critical practice	Tier 3 – Ability to question claims
Level 3 – Applications of variation: Consolidating and using ideas in context, inconsistent in picking salient features	Text user Pragmatic practice	Tier 2 – Application in context
Level 2 – Partial recognition of variation: Putting ideas in context, tendency to focus on single aspects and neglect others	Text participant Text-meaning practice	
Level 1 – Prerequisites for variation: Working out the environment, table/simple graph reading, intuitive reasoning for chance	Code-breaker Coding practice	Tier 1 – Language, definitions/processes

As can be inferred from the table, the students' responses were divided into four levels under the categories of variation concept. At the first level, the students provided idiosyncratic examples from their daily life. They provided limited response in terms of interpreting variation in the tables and graphs. At the second level, although the students were familiar with the concepts of sample, random and variation, they could not interpret variation in tabular and graphical representations. At level three, the students provided more comprehensive definitions for the concepts of sample, random and variation. Besides, they critically question the process of sample selection. However, the student responses were not compatible with the context of the questions. At level four, the students provided conceptual responses and statistically appropriate solutions for the variation concept by utilizing the concept of mean. The students noticed bias and non-repressiveness in the sampling questions. Overall, the students demonstrated comprehensive cognizance about the concepts of sample, random and variation. In the study, 15% of the fifth

graders and 19% of the seventh graders used the correct randomization technique. The overall percentage for providing the correct randomization technique was 7%.

As a requirement of the data-driven world, people need to be literate about graphical and tabular displays they are intensively exposed to through media channels. More specifically, people should be able to read, analyze and critically evaluate statistical data as a requirement of daily life in the information era. As Shaughnessy et al. (1996) stated, the analysis of data and graphical displays is intertwined and they are strongly correlated with each other. Therefore, people need to comprehend graphical displays to analyze statistical data appropriately. Additionally, graph comprehension includes the construction of graph and choosing the most appropriate one between graphs, in addition to reading and interpreting the data embedded in graphs (Friel, Curcio & Bright, 2001; Monteiro & Ainley, 2003) Thus, as indicated by Shaughnessy (2007), many studies were conducted on different types of graphs. Curcio (1987), who is one of the pioneers of research on graph comprehension, conducted a study with 389 students from grades 4 and 7. In the study, Curcio (1987) investigated the influence of prior knowledge, mathematical content and form of the graph on graph comprehension. The researcher also investigated the correlations between those concepts. Besides, regression analyses were conducted to investigate the predictors of graph comprehension. The items in the measurement test included circle graph, bar graph, line graph and infographics. Additionally, literacy and comparison tasks were included in the measurement test parallel with Friel et al.'s (2001) ideas. Curcio (1987) classified the responses according to the *read the data*, *read between data* and *read beyond data* framework. Although there was no significant difference in graph comprehension in terms of gender, achievement in graph comprehension was affected by gender or grade level. Not surprisingly, the seventh graders' prior knowledge was much more than that of the fifth graders. Curcio (1987) found that young students are required to concretize and visualize graphs compared to elderly students. At that point, young students might be presented with data from real life situations with visual representations. Besides, prior mathematical knowledge is required for the comprehensive understanding about graphs.

In another research study, Aoyama and Stephens (2003) categorized the statistical skills by using Curcio's (1987) studies and Kimura's (1999) categories. The researchers categorized statistical skills from A level to F level by not indicating a strict sequential order. The levels are presented in Table 2.3.

Table 2.3 Categories of Statistical Ability (Kimura, 1999; retrieved from Aoyama & Stephens, 2003)

Level	Features
Level A1	Basic reading of tables and graphs
Level A2	Reading key features from graphs
Level A3	Comparing information from two graphs
Level A4	Reading a simple trend in graphs
Level B	Knowing what constitutes an appropriate source of data for a given question
Level C	Statistical computation skills
Level D	Reading global trends in graphs
Level E	Extracting qualitative information from quantitative information
Level F	Creating new dimensional information

According to Aoyama and Stephens (2003), at level A, students are required to have basic abilities to read graphs and tables. At level F, students are expected to synthesize statistical information to construct new information. Aoyama and Stephens (2003) conducted a study with 55 students from grade 5 (N=17) and grade 8 (N=38). In addition to graph questions, the students responded to Level F tasks, which required students to engage in the information presented in the graph. For two graph reading items, although 16 five graders provided 'read beyond data' responses for at least one of these items; 12 five graders provided the same level of response to both items. In the same way, 36 eighth graders provided 'read beyond data' responses at least to one item, whereas 32 eight graders provided the same level response to both items. In Level F tasks, the majority of the five graders could not provide a statistical explanation for their responses, which corresponds to the idiosyncratic level in Watson and Callingham's (2003) framework. Additionally, the majority of the eight graders provided only one answer from the graph by not

justifying their response, which corresponds to the inconsistent level of Watson and Callingham's (2003) framework.

Henriques and Cruz (2010) aimed to investigate 25 third graders' progress in statistical literacy and to identify their difficulties regarding constructing, reading and evaluating statistical data presented in tables and graphs. The researchers aimed to develop teaching and learning in terms of statistical literacy. They employed both qualitative and quantitative methods by utilizing seven exploratory tasks prepared within the framework of Curcio's (1989) conception levels. The students worked on these exploratory tasks in pairs and they presented their work by sharing their understanding and inferences regarding the given graphs and tables. The results of the study revealed that the third graders made errors regarding the selection of a suitable graph for the condition posed in the tasks and the key elements in their drawing. Additionally, the students were unable to explain their reasoning about their solutions in statistical literacy tasks. However, the third graders' involvement in statistical tasks, attention in drawing graph and their statistical interpretations improved as the unit proceeded.

DelMas, Garfield and Ooms (2005) carried out a study to investigate students' understanding and errors regarding bar graphs. The researchers utilized the items in the ARTIST Project, which includes a variety of statistical literacy problems applied to 909 high school and college students. In terms of graph reading, most of the students were confused about the vertical and horizontal axes. They had difficulty in finding the value of the variables in the histogram and bar graph. As far as the graph interpretation items were concerned, 75% of students interpreted the data given in the histogram successfully. However, a vast majority of these students only summarized the statistics in the graph instead of interpreting the data in the problem context. More specifically, these students ignored the problem context by just focusing on the statistics given in the problem. As for matching different graph representations of identical data, most of the students matched the appropriate histogram and bar graph representations of the same data. The students were also quite successful in matching graphs to variable definitions when they were asked to

find the bar graphs corresponding to particular situations in the problems, e.g., expressing the measurements obtained across a period with a bar graph.

Schild (2006) conducted a survey analysis to investigate the reading of graphs and tables involving ratio and percentage. The students were presented a pie chart and a bar chart illustrating the data concerning the percentage of smokers according to religion and the distribution of runners in a competition by the factors of religion and gender. The pie chart and bar graph are given in Figure 2.5 below:



Figure 2.5. The Distribution of Smokers and Runners by Religion (Schild, 2006, p.2)

In the study of Schild (2006), 19% of the students could not read the information given the pie chart by stating that “20% of the smokers are Catholic”. 62% of the students misinterpreted the relationship between the percentages of the Protestant smokers and Catholic smokers presented in the pie chart. Moreover, more than half of the students misinterpreted the bar graph by stating that “20% of Protestant males are runners.” instead of the appropriate interpretation of “20% of all runners are Protestant males.” Moreover, 44% of the students could not read the data presented in the frequency table showing the percentages of smokers and non-smokers. Schild (2006) emphasized the requirement of more focus on statistical literacy in statistics education.

Similar to the study of Schild (2006), Copeer (2018) investigated 23 undergraduate students’ conception regarding the interpretation of variability in histograms and bar charts. The researcher prepared a nine-item measurement tool to observe students’

understanding about variability in the data set represented with graphs. In the first part, the students were presented graphs including multiple and triple data sets, and they were expected to make a comparison about the variability of these data sets. In the following part of the assessment, the researcher aimed to observe students' data reading abilities with regard to bar graphs. Copeer (2018) used the Mann-Whitney test to compare the scores of the students who took statistics course before the current course and who did not take the statistics course. Besides, Levene's test was utilized to observe the significant difference in variability of scores of those students. In the first part of the assessment, 74% of the students provided between two and four correct responses. In addition, a significant difference could not be found between the groups who took the statistics course prior to current course and who did not take the course ($p=0.45$).

Overall, the students had difficulty in comparing the variability between the same graph types. Although they successfully examined the data unique graph (between 43.5% and 65.2%), they had difficulty in examining the data in multiple and triple graphs (4.3%). At this point, Cooper (2018) recommends instructors to encourage students to compare and contrast data in various graphs.

In the literature, a number of studies were conducted including measures of central tendency. The concept of "average" is the most commonly employed statistical concept in individuals' daily life (Gal, 1995). In this respect, the problem context including daily life experiences might encourage students to communicate with statistics to investigate the concept of average and to reveal students' conception of average more comprehensively (Gal, 1995; Watson, 1997). Otherwise, if the problem requires additional knowledge, the context might complicate students' communication with statistics to examine the concept of average (Gal, 1995).

The concept of "average" expresses the representative value of the data set in different contexts not only proceeding with algorithm-based operations (Gal, 1995; Watson & Moritz, 2000). The concept is the way to represent the central tendency in a given data set by utilizing mean, mode and median of the data set. Mokros and

Russell (1995) investigated students' conception of "average" in terms of the representativeness of the "average" to illustrate the given statistics for students from grade 4, 6 and 8. The researchers examined students' definitions of "average" (mean, mode and median) in a given situation to examine their perception of representativeness. Initially, the researchers investigated the students' conception of mean by stating that although most individuals are able to calculate the numerical value of the mean of the data set, they cannot interpret and comprehend mean as an indicator of relationship, balance point, mode, median and fair share (Mevarech, 1983; Mokros & Russell, 1995; Pollatsek, Lima, & Well, 1981; Strauss & Bichler, 1988). The students' errors regarding the average problems stem from deficiencies in conceptual knowledge of average rather than procedural knowledge (Cai, Moyer & Grochowski, 1999; Mokros & Russell, 1995). The majority of the students' conception of average is restricted with the computational procedure of '*add-them-all-up-and-divide*' (Shaugnessy, 1992). More specifically, most of the individuals' understanding of average is nothing more than '*what you get from adding and dividing*' (Gal, 2005, p.3).

In the study of Mokros and Russell (1995), the students used the terms of "usual" and "typical" to refer to "average" and the majority of the students considered the concept of "average" as the numerical calculation of the mean of the data set. In their study, five widespread approaches to the concept of average were balance point, mode, middle point, computational procedure and representativeness. Correspondingly, some of the students provided answers including their personal beliefs regarding the data. Among the 8th graders, 57% interpreted average as midpoint; 14% as mode; 14% as algorithm; and 14% as the balance point. On the other hand, none of the 8th graders interpreted average in terms of representativeness. The study also revealed that the students' conception of average was varied according to grade level. For instance, while most of the 4th graders interpreted average as the mode of the data set, most of the 6th graders considered average as an algorithmic procedure and a representative value of the data set and the majority of the 8th graders interpreted average as the midpoint of the data set.

In the study of Mokros and Russell (1995), the students perceived average as the *balance point* of a data set. In other words, they adopted the '*leveling out*' approach of Strauss (1987), considering that the values are aligned in an order so that the sum of the deviation of these values from the mean could be the same (Cai et al., 1999; Gal, Rothschild, & Wagner, 1990; Hardiman et al., 1984).

Furthermore, the students interpreted average as the *middle point* of the data set thinking that the data in the middle of the data set best represents the whole data. The students benefitted from their personal experiences while deciding on the middle value of the data set as the representative of the overall data (Mokros & Russell, 1995). Some of the students who interpreted average as *fair share* considered average as a reasonable distribution by giving everyone the same amount of a something. Some students conceived average as the *mode* of the data set, specifying the most frequent value of the given data set (Mokros & Russell, 1995).

Building on the study of Mokros and Russell (1995), Watson and Moritz (2000) conducted a longitudinal study with 94 students between grade 3 and grade 9 to examine students' understanding of average, including mean, median and mode of a data set, within various contexts and to investigate the bilateral relationship between representativeness and the concept of average. The researchers categorized the students' levels from pre-average to application of average. The students at pre-average level could not use the term of average and made idiosyncratic interpretations regarding the concept. The students at the application of average level conducted the computational and conceptual procedures of average successfully. These students mostly used mean to compare more than one data set (Watson & Moritz, 2000). Furthermore, the researchers examined the students' preferences concerning the use of median, mode and mean to make inferences about the concept of average. In the study, most of the grade levels preferred *most* and *middle* to interpret average. The use of *most* was categorized as mode and the use of *middle* was categorized as median by the researchers. These findings were consistent with the findings of Mokros and Russell (1995). Additionally, the majority of the 8th grades utilized mean to make inferences about the average of data set (Watson &

Moritz, 2000). The word '*most*' was commonly used by the students especially while they were mentioning the mode of the data set such as *most repetitive value* and *most frequent value*. When the students were expected to interpret average in context, their preference for median decreased (Watson & Moritz, 2000). They mostly used the algorithmic procedures of mean instead of balancing the data set (Watson & Moritz, 2000). From a longitudinal aspect, similar to the findings of the study of Mokros and Russell (1995), it was revealed that the students' conceptual understanding of average developed in higher grades (Watson & Moritz, 2000).

Research studies were conducted regarding statistical literacy at college level. Yotonyos and colleagues (2015) conducted a survey study to investigate 103 undergraduate students' level of statistical literacy by utilizing the two elements model of Gal (2004), which includes the knowledge element and the dispositional element. The items in the measurement instrument are related to the subcategories of these elements. The survey questions were presented as press clippings by integrating worry questions of Gal (2002) to engage students in critical questioning. The results of the study indicated that because of not integrating statistical concepts within the context through statistics course, the undergraduate students revealed low performance in context knowledge questions. The researchers suggested that teachers should involve more context in statistics courses. In this way, they can direct statistics courses by helping students to experience and practice statistical literacy. In line with Yotonyos and colleagues' (2015) study, Cimpoeu and Roman (2018) studied with second year undergraduate students and administered a 30-item questionnaire. Similar to Yotonyos and colleagues (2015), the researchers prepared questionnaire items by utilizing the two element model of Gal (2002). The student performance in histogram questions differed according to the type of the question. Although the students read the data from the table and graphs successfully, they had difficulty in combining techniques to make inferences. The results of the study revealed that students displayed 87% performance in statistical literacy; 44% performance in statistical knowledge and 38% performance in mathematical knowledge components. Moreover, 17% of the students were posed Gal's (2002)

worry questions to investigate statistical information in question. Schau and Mattern (1997) explained the reason for the students' failure and difficulty in statistics as memorization of statistical formulas when they encounter word problems. The students tried to remember the formula rather than analyzing the statistical situation in the problem context (Schau & Mattern, 1997).

In another study, Doyle (2008) conducted an action research with grade 9 and grade 10 students whose ages ranged from 12 to 15 and three mathematics teachers. Within the scope of the study, mathematics teachers attempted to differ their classroom practices to develop students' statistical literacy by creating critical discourse during mathematics lesson. The researcher and the teachers designed and prepared the lessons together by aiming to improve students' statistical literacy. The teachers enriched statistics instruction with activities, context, classroom discussion, critical questioning tasks, and reading texts from media. The students were encouraged to conduct their own statistical investigations. In this way, they progressed in statistical literacy components requiring language and literacy.

2.3 Research Studies on Statistical Literacy in Turkey

A number of studies have so far been conducted in Turkey concerning statistical literacy (Koparan & Güven, 2015; Şeker, 2018; Tosun & Ünal, 2019; Yolcu, 2012). Most recently, Tosun and Ünal (2019) carried out a content analysis which analyzed the context of studies on statistics and probability learning domain. In the study, 61 open access studies performed at elementary school level, including 34 papers, 24 theses and 3 dissertations were reviewed and they were subjected to content analysis. The research findings were divided into themes including keywords, research method-design, sample size, data collection instruments, data analysis methods and objectives, aims and results of the research. Table 2.4 illustrates the frequency of studies related to statistics including statistical literacy according to years.

Table 2.4 Frequencies of Research Studies by Years (Tosun & Ünal, 2019)

Years	frequency
2003-2005	2
2006-2008	7
2009-2011	17
2012-2014	20
2015-2017	15
Total	61

According to the findings of Tosun and Ünal (2019), the demand for statistical literacy demonstrates an increasing trend in mathematics education research. By following this trend, Kaynar and Halat (2011) investigated 490 eight graders' literacy about frequencies, measures of central tendency and graphs using 10 problems. The researchers benefitted from descriptive statistics and independent sample t- test to analyze student responses. Their results revealed that the most inaccurately answered statistical content was standard deviation with 93% wrong answers, while range problems was the most correctly answered content with 34% failure. Additionally, the students had difficulty in constructing graphs with 70% incorrect responses, whereas range problems were incorrectly answered by 34% of the students. Also, 60% of the students had difficulty in questioning about statistical data; 66% in arithmetic mean; 63% in median and 48% in mode. On the other hand, the researchers found that students' interest and achievement in mathematics were correlated with achievement in students' statistical literacy (Kaynar & Halat, 2011). In line with the study of Kaynar and Halat (2011), 143 7th grade students' reading and construction abilities of line graphs were investigated by Memnun (2013). The students were posed three problems which necessitated reading, interpreting and constructing line graphs. Although 7th graders successfully read the line graphs, most of them failed in constructing line graph. Besides, a positive correlation was found between students' mathematics achievement and competency in reading and

constructing line graphs. The findings of Oruç and Akgün (2010) coincide with those of Memnun's (2011) study as both studies revealed that 7th graders had difficulty in the construction of line graphs. With the same purpose, Kranda (2018) investigated 1229 7th grade students' graphical literacy about the data gathered from social science class by conducting a mixed method study. 42.2% of the 7th graders made correct drawings of pie graph. The incorrect drawings stemmed from confusion about percentages and inappropriate placement of variables. The students were confused about percentages and central angle for the parts of the pie graph. 50.5% of the students made appropriate bar graph drawing. Most of the rest of the students could not place the numerical values in bar graph appropriately or preferred line graph instead of bar graph. Besides, 37.3% of the students made an appropriate drawing of line graph. Most of the rest of the students could not match the values and variables or confused line graph with bar graph. Additionally, female students revealed better performance than male students in terms of reading and drawing graphs (Kranda, 2018).

Yolcu (2012) investigated 1074 8th grade students' statistical literacy according to statistical content domains by utilizing Watson's (1997) three-tiered hierarchy. The researcher integrated social context in statistical literacy questions. Yolcu (2012) indicated that the 8th grade students' performance was highest in Tier 2, whereas it was lowest in Tier 3. The results were in parallel to the findings of Watson and Callingham (2003). In the current study, most of the students provided a correct response to the chance and graphs questions (Yolcu, 2012). Moreover, Yolcu (2012) found a positive correlation between students' attitude towards statistics and their statistical literacy levels.

A number of researchers investigated interventions on statistical literacy (Koparan & Güven, 2015; Şeker, 2018). Şeker (2018) conducted a case study with two classes of 7th graders and their mathematics teachers by using cognitively demanding tasks (CDT) for the concept of average by engaging students in higher order thinking. The researcher observed both the teacher and the students during the application of the tasks. Şeker (2018) applied pretest and posttest to observe the effect of cognitively

demanding tasks (CDT) on the teaching of average concept. After eight lessons, in the first class, the students' correct explanations about the average concept increased by 7.7% to 28.5% as the comparison of pre-test and post-test results revealed, whereas this change was from 16.6% to 20.8% in the second class, meaning that the use of CDTs affected students' learning positively (Şeker, 2018).

Similar to Yolcu (2012), Koparan and Güven (2015) investigated 8th grade students' statistical literacy levels by conducting a quasi-experimental study with 70 students in four-week interval. The researchers observed the effect of project-based learning on students' statistical literacy. The items in the instruments were adapted from Watson and Callingham's (2003) six level hierarchy. It was found that as far as statistics instruction is concerned, project based learning is better than traditional approaches by engaging students in work collaboratively (Koparan & Güven, 2015). Additionally, in the intervention group, students' statistical literacy levels increased from Level 2 to Level 3 after the intervention (Koparan & Güven, 2015).

2.4 Context in Problems

The assessment of statistics involves the blend of statistical content and process with the elaboration on open-ended and meaningful problems to encourage students to investigate, comprehend and communicate with statistical data (Begg, 1997) in real life context. (Kelly et al., 1997; Lesh et al., 1997). Therefore, the use of problems with context helps to conduct in-depth examination of statistical literacy. In the literature, there are various definitions, approaches, perspectives and research studies about the use of context in problems. In this section, these definitions, approaches, perspectives and research studies are discussed.

“Problem solving should be a process actively engages students in making conjectures, discussing and questioning their own thinking, validating results and making convincing arguments.” (NCTM, 1987, p.54)

Therefore, the use of familiar context in problems is key to engage students in the problem solving process by activating their life experiences and motivating students to solve problems (Chapman, 2006; Verschaffel, 2002). To be an informed citizen, students need literacy in various fields with familiar contexts (De Lange, 2003; Steen, 2001). Thus, it can be said that problem context is the key to solve daily life problems by requiring reasoning for the situations (Kliebard & Franklin, 2003).

Borasi (1986) defined the problem context as '*the situation in which the problem is embedded*' (p.128) and indicated that the problem context leads problem solvers to reach solution by presenting information about the problem. Meyer, Dekker and Querelle (2001) extended Borasi's (1986) definition by examining context in five mathematics curricula. The researchers reported that context does not merely include plain texts, but also diagrams, illustrations, graphs and tables (De Lange, 2003; Meyer, Dekker & Querelle, 2001). However, Meyer, Dekker and Querelle (2001) indicated that many mathematics textbooks include context only at the beginning or at the end of the books shortly while the rest of the problems are just problems without context. For this reason, students have difficulty in solving problems with context and they attempt to eliminate the context in the problem. In fact, in mathematics curricula, context has various missions such as increasing motivation and forming a link between mathematical practices, new mathematics solution techniques and mathematical understanding (Meyer, Dekker & Querelle, 2001; Putnam, Lampert, & Peterson, 1990). The researchers recommend these roles of context as a checklist for students' achievement in problems with context in the field of mathematics education. On the other hand, the ability to solve problems in context requires various features. In the literature, several models and frameworks exist explaining the context in the problem solving process (Verschaffel et al., 2000; OECD, 2013). OECD (2013) proposed a framework (Figure 2.6) to illustrate knowledge, abilities and mathematical notions required to cope with the problem solving process.

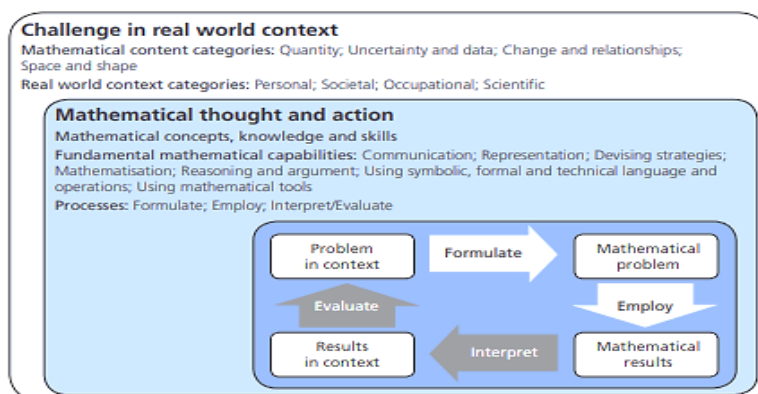


Figure 2.6. A model of mathematical literacy in practice (OECD, 2013, p.26)

The continuum of solving problems with context involves basic mathematical skills and processes such as reasoning, mathematization, communication and representation. In the model of OECD (2013), problem situation and results were considered in context. However, most of the students concentrate only on the right hand side of the model by ignoring the context in problem (Hoogland et al., 2016)

The use of context in problems is perceived as transferring and using mathematics in daily life situations. The data could be interpreted as numbers in context by activating individuals' knowledge of context to critically evaluate the statistics instead of conducting algorithmic procedures (Moore, 1990). Context problems have potential to encourage students to indulge in mathematics, yet they might also lead to certain challenges (Widjaja, 2013). In the literature, there is no consensus on the utilization of context in mathematics education. From this perspective, importance is given to the use of context in research on mathematics education. There exist two distinct views regarding the use of context in the literature.

According to the first view, the utilization of context enables students to maintain unprescribed mental operations, concretizes problem and motivates students (Carraher et al., 1985; Chapman, 2006; Lappan & Phillips, 2009; Mitchell & Carbone, 2011; Nunes et al., 1993; Van-den Heuvel-Panhuizen, 2005; Widjaja, 2013). Context problems encourage students to concentrate on mathematics

(Thompson et al., 1994) by increasing the authenticity (Palm, 2009) of tasks and eliminating students' suspicion concerning the problem situation in the task. These claims were supported by some research studies in the literature (Carragher et al.; 1985; Nunes et al, 1993; Sullivan, Zevenbergen, & Mousley, 2003; Widjaja, 2013)

In this context, Carragher and colleagues (1985) conducted a study with five students from grades 1 to 8. After administering an informal test (including 64 questions) and a formal test (including 99 questions), the students were interviewed. The problems were posed to the students in the streets or at the market, and they were asked in context-free form in the formal test and in context-bounded form in the informal test. A sample question from the interview is presented in Figure 2.7 below:

(I) First example (M, 12 years)

Informal test

Customer: I'm going to take four coconuts. How much is that?

Child: Three will be 105, plus 30, that's 135 ... one coconut is 35 ... that is ... 140!

Formal test

Child resolves the item 35×4 explaining out loud:

4 times 5 is 20, carry the 2; 2 plus 3 is 5, times 4 is 20.

Answer written: 200.

Figure 2.7. A sample problem from test (Carragher et al.,1985, p.26)

As seen in the example in Figure 2.7, the researchers posed the problem at the market. The researchers asked one student how much 4 coconuts are. The student added the price of one coconut on the price of three coconuts by using his own algorithm and technique to obtain the price of four coconuts. The student was more involved in the problem in the informal test involving the context-bounded form of the problem and felt himself free to choose a technique to reach the correct answer. It means that the use of context involving a real life situation assisted students in reaching the correct answer by concretizing the problem situation. The same problem in the formal test involves the context-free form of the problem. This time the student felt that he should use an algorithm or a series of mathematical procedures to reach the correct answer and therefore provided an erroneous response. Thus, it was concluded that the students had less difficulty in solving context problems compared to without context problems. In the test including the context problems, the students'

achievement was 73.7%, while it was 36.8% in the test including the problems without context. Although the arithmetic operations required to reach the correct answer were the same, the use of real life excerpts such as selling coconuts and lemons helped students to solve problems correctly by concretizing the problem situation. The reason was that thanks to the context in the problems, the students were open to mental operations and strategies rather than prescribed complicated rules and procedures (Carragher et al., 1985). Building on that study, Nunes, Carragher and Schliemann (1993) conducted a series of studies in various contexts (i.e., agricultural context, fishermen context, worker context and seafood context) on different content domains of mathematics such as ratio and proportion. More precisely, the researchers asked ratio and proportion problems within the agriculture, fishermen, worker and seafood contexts separately. Then, they investigated the problem solving performances of students in these various contexts. As a result of these consecutive studies, the researchers concluded that the notion of proportion could be enhanced through daily life experiences and situations without being thought at school and could be applied on unfamiliar contexts (Nunes et al., 1993; Millroy, 1994). Therefore, the use of context based on everyday life experiences and situations enhances students' learning of mathematical concepts.

Widjaja (2013) analyzed the effect of using context in problems on Indonesian students' mathematical learning. The sample consisted of 4th, 5th and 6th graders. The students were asked problems with context and they worked alone and in a group. The 4th grade students had difficulty in the problems since the number of students in the problem did not represent their class size. The teacher illustrated the problem by making a drawing related to the context of the problem. Then, the students were encouraged to discuss and share their ideas about the context of the problem. The fifth graders tried to apply algorithm-based solutions despite the teacher's effort regarding the involvement of context. The students made erroneous operations. Then, the teacher made an effort to engage students in the context of the problem by using probing questions benefitting from the problem context. In this way, the students expressed their thoughts and incomprehensible points comfortably. Thanks

to the context, the teacher and students constructed a productive discourse in the classroom. In the 6th grade, the students built a relationship between context and different representations of data. Although students have different life experiences and background, context could be used for meaningful mathematical learning with a strong discussion environment (Widjaja, 2013)

On the other hand, it is widely acknowledged that problems with context are perceived by students as '*the hated word problems*' (Thomas & Gerofsky, 1997, p.21). Thus, the researchers have been seeking the reasons behind this fact. According to another view, the utilization of context is a complex construct, requiring sequential and complicated cognitive processes besides arithmetic and algorithmic operations (Schnotz, 2002; Fusch et al., 2006; Schnotz et al., 2010). The bilateral relationship between the use of context and enhanced conception has been oversimplified by the researchers as they have focused on a single aspect of cognitive development by neglecting the individual differences in thinking and interpreting (Walkerdine, 1990; Boaler, 1993). Solving context problems involves both 'horizontal mathematization' and 'vertical mathematization' processes by including context into mathematical procedures, while context-free problems merely require vertical mathematization focusing only on mathematical constructs (Treffers, 1987; Doorman & Gravemeijer, 1999). Therefore, the utilization of context renders the problem solving process more complicated.

Boaler (1993) investigated the responses of 50 second and eight graders from two different schools to six problems with context and without context including the same content. The researcher also examined the student responses by benefitting from various problem contexts. In the problems with context, some real life situations such as workshop, penalty and cutting wood were employed in different mathematical contents. The problems were classified as fraction questions and number questions. Three questions were asked related to fractions. The first item was a fraction problem including the context of penalty as presented in Figure 2.8 below:

PENALTIES

The school football team is looking for its champion penalty kicker.



Each month Brandon took 10 penalties.
In January he scored 5, in February he scored 7 and in March he scored 6.

Each month Mick took 15 penalties.
In January he scored 7, in February he scored 6 and in March he scored 8.

Who is better at taking penalties?
How did you decide who was best?

Figure 2.8. A sample problem with penalties context (Boaler ,1993 p.351)

In the second fraction problem, the context was differentiated into fertilizer context as given in Figure 2.9 below:

FERTILIZERS



Sam planted some marigold seeds in two different plant tubs.
Each tub contained a different fertilizer, Sam wanted to see which fertilizer was best.
Sam did an experiment over a few months, she kept planting seeds in the tubs and seeing how many plants she got.
She always planted 15 seeds in tub A and 10 seeds in tub B.

These are her results:

	Tub A	Tub B
June	5 plants	3 plants
July	6 plants	4 plants
August	3 plants	2 plants

Which Tub has the best fertilizer?
How did you decide which was best?

Figure 2.9. A sample problem with fertilizers context (Boaler, 1993 p.352)

The third item was a fraction problem without context as presented in Figure 2.10 below:

FRACTIONS			
For each pair of fractions state which one is biggest and explain why:			
1.	$\frac{3}{10}$	or	$\frac{6}{15}$
2.	$\frac{2}{5}$	or	$\frac{7}{15}$
3.	$\frac{8}{20}$	or	$\frac{14}{30}$

Figure 2.10. A sample question without context (Boaler, 1993, p.352)

The context was eliminated from the problems with context and problems without context were obtained. For instance, in the penalties problem, the students were expected to obtain two ratios for two players and make a comparison between two ratios after finding the ratio of successful penalties to overall penalties. Likewise, in the fraction question, the fractions were presented to make a comparison. As a result of Boaler's (1993) study, the students from the first school did not provide distinct responses as the context changed. More specifically, the use of a different context did not make a significant difference in students' solutions. As far as the second school was concerned, however, the students' solutions and achievement were affected by the context of the task. Additionally, the researcher did not observe the significant difference between problems with context and problems without context. Thus, it could be said that the students' performance and context are not directly related. To be more precise, context could not be considered to be a direct motivator and incentive for mathematical tasks (Cobb, 1986; Lave, 1988 as cited in Boaler, 1993). To make sense of the context, students should produce their own meaningful context by reflecting their goals, conceptions, perceptions and mathematics related strategies (Boaler, 1993).

In a similar way, Gravemeijer (1994) investigated whether students' solution process is affected mostly by the context of the problem or the numerical values included in the problem. The researcher presented various problems including number line and unknown variables in arithmetic operations to the students. A sample problem from the study is presented in Figure 2.11:

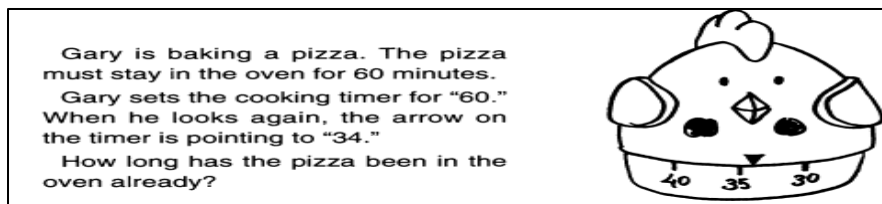


Figure 2.11. A problem situation with context including unknown variables (Gravemeijer, 1994 p.464)

The problem represents the operation as $60 - 34 = ?$ However, the interpretation of the situation in context derived the operation as $34 + ? = 60$. The context converted the subtraction operation into addition with an unknown variable. The researcher observed both strategies in students' solutions. In the problem, the students were affected mostly by the numbers in the problem rather than the situation within the context (Gravemeijer, 1994). More specifically, the students focused on the numbers rather than the context; the context and numbers were apart from each other. Therefore, the context did not provide a convenience for students to solve the problem.

Roth (1996) conducted a case study with 8th grade students to investigate their approaches to context problems and their performance in contextual mathematical situations in a ten-week period. The aim was to observe students' mathematization process by utilizing the ecology context. Firstly, the students collected, analyzed and represented their own data with tables, graphs and descriptive statistics. Then, they prepared written reports to support their work by using context as much as possible. Secondly, word problems and students' data were utilized to observe students' mathematical practices in these situations. Although story problems included a familiar context, this did not reflect on students' mathematical practice. More specifically, towards the end of study, the students did not involve more context in their practices (Roth, 1996). Consequently, they did not make sense of the context in the course of Roth's (1996) study.

Cooper and Dunne (1999) conducted a study with students with low, moderate and high socio-economic status by asking them problems with context. Although the students' performance in problems without context was similar, the researchers observed significant differences between the responses of the students with low socioeconomic level and those with moderate socioeconomic level in problems with context since the problems required the use of daily life knowledge and experience. Obviously, the students' sociocultural background was different from each other. Thus, it can be stated that context had a negative effect on students' performances and discriminated the students.

2.4.1 Context in Statistical Literacy Problems

Watson (2006) interpreted statistical literacy as the intersection of statistics curriculum and real-life including improvisatorial context and instantaneous judgement by using statistical constructs, knowledge of context and critical skills. According to some researchers, students learn statistics as prescribed rules by not engaging with meaningful context where the statistics is applied (Kelly, Sloane & Whittaker, 1997). The use of context and real-life situations for proficiency in statistical literacy is appreciated and investigated by a number of researchers (Ben-Zvi & Garfield, 2004; English & Watson, 2016; Gal, 2002; Gal, 2005; Watson, 1997; Watson & Callingham, 2003). Chick and colleagues (2005) associated statistical literacy with "transnumerative thinking" by engaging students in collecting, interpreting, representing and analyzing data in real world context.

Lesh, Amit and Schorr (1997) conducted a model eliciting activity by integrating real-life context to observe students' interaction with statistical data without any instruction. The researchers presented a context-rich newspaper article and created a discussion environment with context-rich problems. The students used computer-based graphing, average, and trends in data with their group-mates to analyze the data. Lesh, Amit and Schorr (1997) observed that the students' modeling cycle moved from idiosyncratic informal interpretations to formal interpretations

including measures of central tendency, graphs and questioning of data. The use of real world context with model eliciting activities in data analysis encouraged students to produce more productive discourse in terms of literacy about statistical features (Lesh, Amit & Schorr, 1997).

Similarly, Monteiro and Ainley (2003) classified students' interpretations regarding graphs by utilizing Watson's (1997) and Curcio's (1987) hierarchy and integrating context. The student teachers were presented with familiar data in unfamiliar forms with out of school context and an interview was conducted with those students. In parallel with Cooper and Dunne's (1999) findings, the use of context did not provide an advantage for student teachers in Monteiro and Ainley's (2003) study. The researchers pointed out that the integration of context itself does not provide convenience for interpreting and reading data as the knowledge of the problem context requires extra attention. To help students to enhance critical sense, equilibrium between the context knowledge and statistical knowledge should be established by the teachers not the students (Monteiro & Ainley, 2003).

Nisbet, Langrall and Mooney (2007) emphasized both possible positive and negative effects of the use of problem context. More specifically, Nisbet, Langrall and Mooney (2007) investigated the effect of students' knowledge about real-life contexts while analyzing and interpreting statistical data. The researchers worked with 6th, 11th, and 12th graders by employing qualitative methods to deeply examine students' statistical knowledge and its relationship with problem context. The students worked in groups and the problems were assigned to each group. The groups discussed their solutions at the end of the study. The researchers observed three widespread effect of problem context. Some students benefitted from the problem context to promote their statistical interpretations of the data. Some students did not engage in the problem context. The rest were not productive in the statistical task. Nisbet, Langrall and Mooney (2007) concluded that while the problem context might engage some students in statistical tasks, conversely, it might cause the other students to lose attention towards the task. Moreover, it was also concluded that encouraging students to involve in statistical tasks by connecting the statistical data and the

problem context is crucial to improve statistical literacy (Gal, 2002; Nisbet, Langrall & Mooney, 2007; Pfannkuch & Wild, 2004).

Consequently, the related literature indicated that research on statistical literacy was a current issue in mathematics education research and there was an increasing interest for studying on statistical literacy (Shaughnessy, 2007) since statistical literacy is the ultimate objective in mathematics curriculum. Although many researchers sought to derive definition for statistical literacy, there is no common definition of statistical literacy, which indicates that it is still under development (Chance, 2002; Rumsey, 2002; Sharma, 2017; Shaughnessy, 2007). In mathematics education research, studies were conducted to define statistical literacy (e.g., Callingham, 2007; Gal, 2002; Garfield et al., 2003; Wallham, 1993) and to derive frameworks for statistical literacy (e.g., Gal, 2002; Sharma, 2017; Watson, 1997; Watson & Callingham, 2003). On the other hand, the use of context in problems was a disputable issue for research on mathematics education. There were two distinct views regarding the existence of context in problems. According to first view, the utilization of context enables students to maintain unprescribed mental operations, concretizes problem and motivates students (Carragher et al., 1985; Chapman, 2006; Lappan & Phillips, 2009; Mitchell & Carbone, 2011; Nunes et al., 1993; Van-den Heuvel-Panhuizen, 2005; Widjaja, 2013). Respectively, according to another view, the utilization of context is a complex construct, requiring sequential and complicated cognitive processes besides arithmetic and algorithmic operations (Fusch et al., 2006; Schnotz, 2002; Schnotz et al., 2010). As for existence of context in statistical literacy problems, the use of context and real-life situations for proficiency in statistical literacy is appreciated and investigated by a number of researchers (Ben-Zvi & Garfield, 2004; English & Watson, 2016; Gal, 2002; Gal, 2005; Watson, 1997; Watson & Callingham, 2003). In the light of existing literature, the present study intended to investigate the students' statistical literacy through problems with and without context.



CHAPTER 3

METHODOLOGY

The aim of this chapter is to elaborate on the design of the study, population and sample, data collection instruments, pilot study, validity and reliability of the instruments, data collection procedures, data analysis procedures, assumptions and limitations, and the internal and external validity of the study.

3.1 Research Design

The aim of the present study was to investigate 8th grade students' achievement and statistical literacy through problems with and without context. Comparing the 8th grade students' achievement and statistical literacy in statistical literacy problems with and without context was another aim of the study. Therefore, the present study intended to address the following research questions:

1. How successful are the 8th grade students in problems with and without context in terms of statistical content domains (*measures of central tendency, graph interpretation, sampling, questioning of statistical data*)?
2. How is the 8th grade students' statistical literacy in problems with and without context in terms of statistical content domains (*measures of central tendency, graph interpretation, sampling, questioning of statistical data*)?
3. Is there any significant mean difference in students' achievement in statistical literacy problems with and without context?

Initially, cross-sectional survey, which is one of the quantitative research methods, was employed to investigate 8th grade students' achievement in problems with and without context in terms of statistical content domains. According to Fraenkel and Wallen (2006), during the cross-sectional survey research, the data is gathered at one time from a sample selected from the population. For the present study, the data was gathered from the 8th grade public school students in Ankara district. The students responded to the instruments and their responses were analyzed within the framework of the rubric prepared by researcher. The analysis was made out of 3 points for each problem.

In order to address the first research question, the students' responses were analyzed and coded from 0 to 3 according to the accuracy of the responses by using the SPSS program. The frequencies and percentages of achievement were specified for each problem according to statistical content domains.

To address the second research question, the qualitative research methodology was employed to analyze the content of the student responses. The students' statistical literacy was examined and interpreted under categories of student responses by providing sample responses for each of the categories. The categories were determined according to the result of the pilot study and the student responses to the instruments.

Lastly, the students' achievement in both tests were compared with respect to statistical content domains by using statistical comparison methods to address the third research question.

In conclusion, the data were collected with the statistical literacy tests to address three research questions of the study by using both qualitative and quantitative research methods. Therefore, the study was mixed-method research. Mixed method is a blend of qualitative and quantitative methods in a unique study, which enables comprehensive investigation of research questions rather than using a single method alone. (Fraenkel, Wallen & Hyun, 2011)

3.2 Population and Sample

The target population of the present study was all 8th grade students in Ankara. The reason why 8th grade students were chosen was that grade 8 covers all the data and chance topics in the Turkish Ministry of National Education curriculum. Thus, the test was prepared more comprehensively in terms of statistical content. The accessible population of this study was 500 8th grade middle school students from four public schools in Yenimahalle and Çankaya districts of Ankara. Since reaching all the students in Ankara might require considerable time and effort, the convenience sampling method was employed in the current study (Fraenkel & Wallen, 2006). The researcher chose 500 8th grade students from four public schools from Çankaya and Yenimahalle districts in Ankara since the mathematics teachers of those schools were graduate students at METU. These teachers and school administrators kindly approved the application of tests. Therefore, data was collected from 500 8th grade students. The size of sample for the study is presented in Figure 3.1 below.

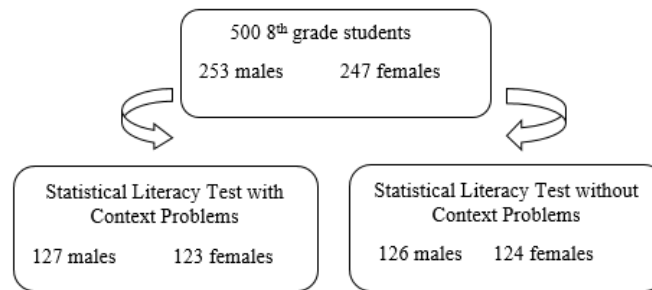


Figure 3.1. The sample size of the study

500 8th grade students were divided into two groups based on their average grade point in mathematics in the first semester of 2019-2020 academic year. After forming two equal groups in terms of mathematics achievement level, the first group of students took the Statistical Literacy Test with Context Problems (SLT-CP), whereas the others took Statistical Literacy Test without Context Problems (SLT). Consequently, the students' achievement level in both groups was equalized to

observe the equal number of students' statistical literacy from each achievement level for the problems with and without context. For instance, one student having 90 average grade point was assigned to SLT where the other student having 90 average grade point was assigned to SLT-CP.

3.3 Data Collection Instruments

The data collection instruments used in the present study were the Statistical Literacy Test with Context Problems (SLT-CP) and Statistical Literacy Test without Context Problems (SLT). To obtain the SLT, the context of the problems in the SLT-CP was eliminated. More precisely, the items in both tests were the same. The only difference was that the problems in the SLT-CP were asked to the students with a social and scientific context. On the other hand, in the SLT the problems were asked directly by not using a social and scientific context. The instruments in the present study are discussed in detail below.

3.3.1 Statistical Literacy Test with Context Problems

The Statistical Literacy Test with Context Problems (SLT-CP) aimed to cover middle school data and chance objectives in the National Mathematics Education Curriculum (MoNE, 2018). The test items were prepared according to the statistical content domains in the middle school data and chance curriculum. For this reason, Table of Specification (See Appendix E) was prepared by the researcher to ensure the comprehensiveness of test in terms of statistical content domains in the curriculum. Within the framework of Table of Specification, the problems were determined by the researcher according to the National Mathematics Education Curriculum (MoNE, 2018) based on the specified statistical content domains. The statistical content domains and the corresponding problems in each domain are illustrated in the Table 3.1 below:

Table 3.1 Table of Content for the Problems in the Statistical Literacy Tests

Content Domain	Related Problems
Measures of Central Tendency	P8, P9, P11, P12, P13
Graph Interpretation	P2, P3, P4, P5, P6, P7, P15, P16, P18, P20, P21
Sampling	P10, P22
Questioning of Statistical Data	P1, P14, P17, P19, P23, P24, P25

In the test, P8, P9, P11, P12, P13 was measures of central tendency problems where P8, P11, P12 and P13 were in open-ended format and P9 was asked as a multiple choice format. Moreover, the students were asked eleven problems related to the graph interpretation content domain. These problems were P2, P3, P4, P5, P6, P7, P15, P16, P18, P20 and P21. Among these problems, P2, P3, P12 and P15 were asked in open-ended form, whereas P16, P18, P20 and P21 were multiple-choice items and P4, P5, P6 and P7 were True-False items. Furthermore, P2, P3 and P15 were related to a pie chart; P4, P5, P6, P7 and P18 belonged to a line graph; P20 and P21 were related to a bar graph, and lastly P16 was about the conversion of the data presented in a pie chart to a bar graph. Furthermore, the students were asked two problems regarding the sampling content domain. In the instruments, P10 and P22 were the sampling problems. The reason for asking two problems regarding the sampling domain was that in elementary mathematics curriculum, the sampling domain was not comprehensive when compared to the measures of central tendency and graph interpretation domains. P10 was related to the techniques of random sampling and P22 was related to the generalizability of the sample to the entire population. Additionally, the students were asked seven problems concerning the questioning of statistical data content domain. These problems were P1, P14, P17, P19, P23, P24 and P25. While P1, P14, P17 and P19 were open-ended items, P23, P24 and P25 were True-False items.

Most of the problems were adapted from the literature appropriate with the Turkish context for 8th grade students. On the other hand, the literature review revealed that the use of context in statistics problems was emphasized by many researchers. In this respect, one of the aims of the present study was to observe the students' achievement and statistical literacy according to the existence of context in statistics problems. Therefore, the context was integrated carefully for the items in the Statistical Literacy Test with Context Problems (SLT-CP) to comprehensively examine students' statistical literacy. The items were designed based on sustainability context to attract students' attention and to raise the students' awareness about environmental issues. To translate statistical terms from the initial versions of the adapted problems, the researcher benefitted from the International Statistical Institute (ISI) multilingual glossary of statistical terms (ISI, 2011). As far as the test format is concerned, the test composed of True-False items, multiple choice items and open ended items. In the literature, multiple choice problems are used for probabilistic and statistical concepts (Joliffe, 1997). Multiple choice format is convenient to use contexts, while others are impractical, particularly to assess large samples (Wild, Triggs & Pfannkuch, 1997). Additionally, open-ended questions are appropriate to address 'real questions' with 'real data' by transferring statistical ideas in written form (Joliffe, 1997). Therefore, those problem formats were included in the Statistical Literacy Test with Context Problems (SLT-CP). The problems in the Statistical Literacy Test with Context Problems (SLT-CP) are explained in detail below. For the Turkish version of the test, see Appendix C.

The context of the first three problems was adapted from the NASA Education (2019) activity called *Ocean World: Earth Globe Toss Game*. In this activity, the steps followed the context of oceans, water, earth and energy saving. To attract students' attention to the context of sustainability and energy sources, the test started with a conversation between most of the students' favorite cartoon characters Sponge Bob and Patrick, who live in ocean. All the items in the test were explained through the conversation between Sponge Bob and Patrick in terms of consistency of the problem contexts in the test.

The first item in the SLT-CP was designed as a warm up problem to observe students' estimation skills within the context of water around the world. The original version of the first item is presented in Figure 3.2



The Blue Marble photograph is the most detailed true color photograph of the world so far. Scientists made observations with the satellite, connecting the oceans, sea, land, ice floes and clouds over the world to each kilometer without any error, and created this perfect photograph. As seen in the photo, most of the world consists of water.

- What types of surface water are there on Earth?
- Does surface water give an accurate representation of the amount of water on Earth?
- Where else is water other than on the surface?

Figure 3.2. The original version of the 1st problem

The items included in the NASA Education (2019) activity were translated into a statistical estimation problem as the 1st problem of the Statistical Literacy Test with Context Problems (SLT-CP) as in Figure 3.3. In the problem, the students were expected to estimate the approximate amount of surface water on earth.



The Blue Marble photograph is the most detailed true color photograph of the world so far. Scientists made observations with the satellite, connecting the oceans, sea, land, ice floes and clouds over the world to each kilometer without any error, and created this perfect photograph. As seen in the photo, most of the world consists of water.

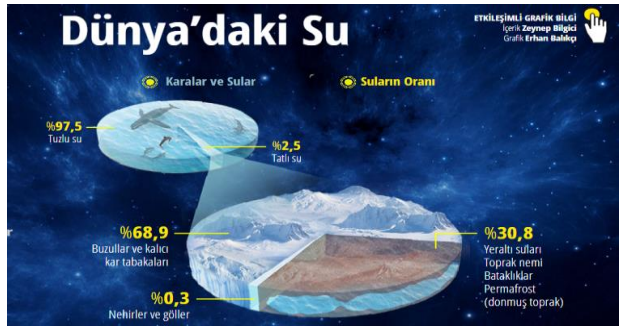
1)What types of surface waters do you think exist on earth? Can you estimate the approximate amount of the water on earth? Explain how you decided.

Figure 3.3. The 1st problem in the test

The second problem is represented in Figure 3.4. It was developed by the researcher. To provide realistic data from the media that students encounter in their daily lives, an infographic prepared by TUBITAK (2016) was used.



On the side, the distribution of the water in the world according to its sources is shown.



Kategori	Oran (%)
Tuzlu su	97,9
Tatlı su	2,5
Buzullar ve kalıcı kar tabakaları	68,9
Nehirler ve göller	0,3
Yeraltı suları	30,8
Toprak nemi	
Bataklıklar	
Permafrost (donmuş toprak)	

2) Which type of the graph best represents the distribution of the water in the world in the most appropriate way by using the information in the figure given above? Why? Explain your answer.

Figure 3.4. The 2nd problem in the test

The problem was related to the distribution of the amount water around the world according to water sources. In the 2nd problem, the students were expected to decide on the most appropriate graph type to represent the given data regarding water sources on earth as presented in the infographic.

In the 3rd problem, the students were asked to represent the given data regarding the distribution of the amount of the water around the world with the chosen chart type in the previous problem. More specifically, the students were expected to draw the graph which they selected in the 2nd problem to illustrate the amount of water in the world. The 3rd problem is represented in Figure 3.5 below:

3) Express the distribution of waters in the world according to their sources with the type of the graph you choose above.

Figure 3.5. The 3rd problem in the test

With the 3rd problem, the researcher aimed to observe students' literacy in terms of reading the given data and constructing the most representative graph that best represents the given data.

The 4th, 5th, 6th and 7th problems were True-False items which require determining the accuracy of the statements by reading and interpreting the data presented in a line graph. These problems were written by the researcher as a result of a comprehensive curriculum, textbook and literature review. The statements include misconceptions and common errors made by students regarding line graph interpretation. Additionally, the authentic data related to the changes in first priority problems by years was parallel with the context of the whole test. The problems and the line graph are given below:

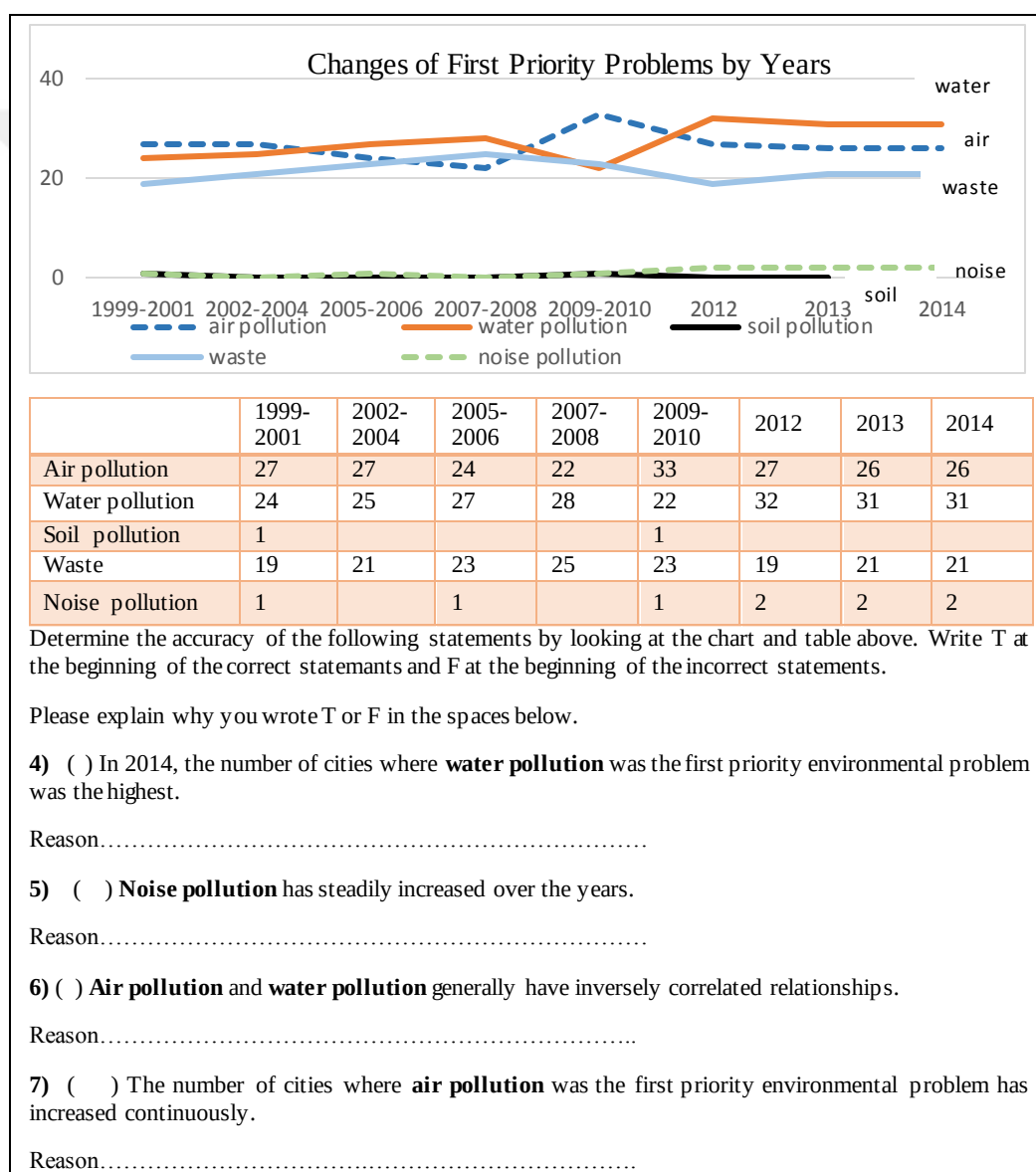


Figure 3.6. The 4th, 5th, 6th and 7th problems in the test

In the line graph, authentic data regarding the number of cities having an environmental problem as the first priority problem was presented according to years. The given table represented the same information with the line graph. The students could utilize the table when they were confused while reading line graph. As an informed citizen, the students should be a decision maker about which of the representations they should use to interpret the given data.

In the 4th problem, the students were expected to examine the statistics about environmental problems in the year 2014 and decide whether water pollution was the most experienced first priority environmental problem. In the 5th problem, whether noise pollution was an over-growing problem according to years was asked. The students should observe the changes in the number of cities having the noise pollution as the first priority environmental problem. In the 6th problem, the students were requested to observe the relationship between the number of cities having air pollution and water pollution as the first priority environmental problem. The students should decide on whether the statistics regarding water pollution and air pollution had an inverse relationship according to the line graph. In the 7th problem, the increase in the number of cities suffering from air pollution as the first priority environmental problem was asked. In other words, it was asked whether the number of cities where air pollution was the first priority environmental problem has increased continuously.

The 8th problem was adapted from the study of Watson and Callingham (2003). This problem was asked to investigate students' understanding of the 'average' concept. The original version of the item is presented in Figure 3.7 below:

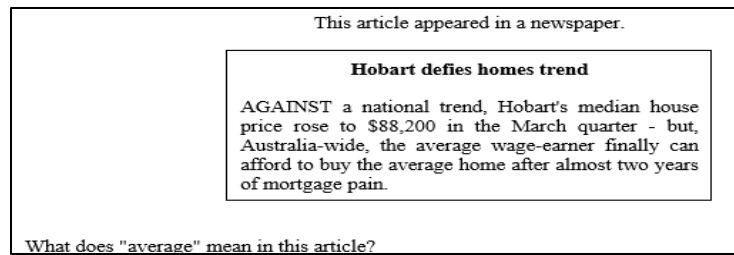


Figure 3.7. The original version of the 8th problem (Watson & Callingham, 2003)

The context of the original problem was changed while adapting the problem as the problem context should be compatible with all the problems in the test. The initial version of the problem was related to home prices in Australia. The context was transformed to the average amount of garbage thrown in Turkey. The final version of the 8th problem in the test is displayed below:

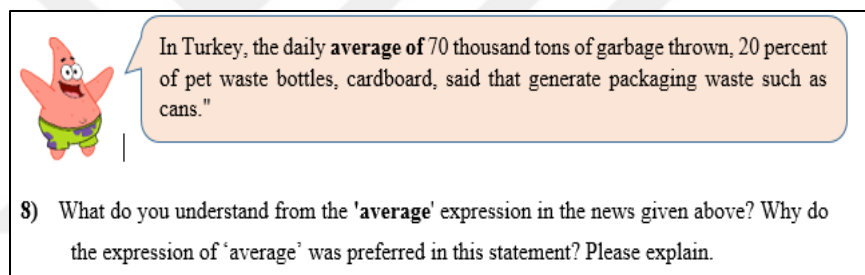


Figure 3.8. The 8th item in the test

The 9th problem was related to finding the average of a data set by using measures of central tendency. The students were asked about the ways for calculating the amount of garbage disposed every day in a week. The students were expected to consider the measures of central tendency to calculate the amount of garbage disposed every day in a week. The problem was asked in a multiple choice format. In option A, the range of the amount of garbage produced for a week was emphasized. In option B, one of the measures of central tendency, median, was calculated to reveal the amount of garbage. In option C, one of the measures of central tendency, arithmetic mean, was calculated to find the amount of garbage by dividing the total amount of garbage by 7. In the last option, one of the measures of

central tendency, mode, was emphasized to find the average amount of garbage thrown daily in Turkey in a week. The problem is presented in Figure 3.9 below:

9)



I want to calculate the **average** amount of garbage that daily thrown in Turkey in a week. For which one of the following methods can we find the closest amount of garbage disposed every day for a week?

A) By subtracting the least amount of garbage from the most amount of garbage produced for a week

B) Sorting the total amount of garbage produced for a week from small to large and selecting the middle number

C) By dividing the total amount of garbage produced daily for a week by 7

D) By writing the total amount of garbage produced daily for a week and determining the most recurring amount.

Figure 3.9. The 9th problem in the test

The 10th problem was adapted from the study of Watson and Callingham (2003). However, the students' interpretation of random was directly asked to students in Watson and Callingham's (2003) study. In the present study, the context of water pollution was integrated to the problem in order to investigate students' understanding about random selection. In the problem, three options were presented for the sampling of the province to examine water pollution. The first and second options did not indicate random selection, but purposive sampling. The third option was an example of random selection. The students were asked which statement(s) was appropriate for random selection. The final form of the 10th item is as follows:



A researcher wants to choose a random province to examine the water pollution in our country to examine the factors that cause water pollution in that province. Accordingly, which of the below choice (s) do I make **random selection**?

I) Since it is easy and cheap to reach, it should go to the nearest seaside.

II) Since the industrial activities in the Marmara Sea are intense, the water pollution of that region will be more intense and it selects a province from the Marmara Region.

III) Writing the plaque numbers of all provinces that have a coast to the sea in a bag and choosing a random provincial plaque number from the bag.

A) Only I B) I and II C) Only III D) II and III

Figure 3.10. The 10th problem in the test

The 11th, 12th and 13th problems were developed by the researcher to examine students' statistical literacy in measures of central tendency which are mode, median and arithmetic mean. The problems required students to read and interpret the data presented in the bar graph which illustrates the frequency of water pollution incidents in different cities of Turkey. The bar graph is presented in Figure 3.11 below:

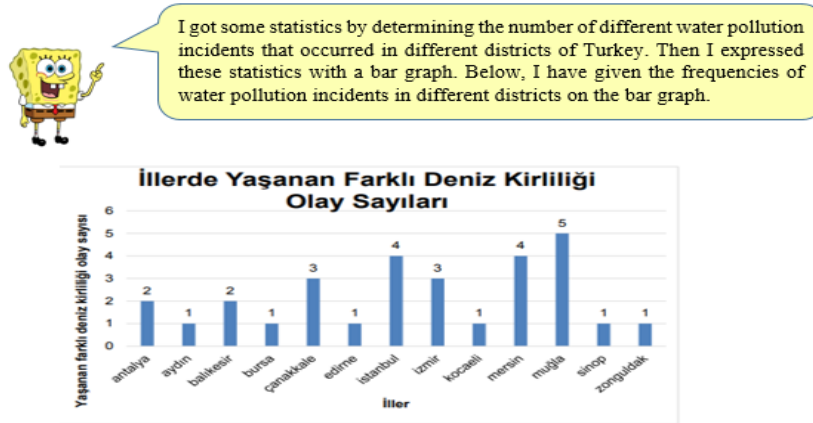


Figure 3.11. The bar chart used for the 11th, 12th, and the 13th problems

In the 11th, 12th and 13th problems, the students were asked to calculate and interpret Measures of Central Tendency within context of water pollution. For these three problems, the students were expected to read the data in the bar graph and interpret mean, mode and median in context and make sense of statistics about the number of water pollution incidents in different cities of Turkey. The items are presented as follows:

According to the given bar graph;

- 11)** What is the mode of the number of water pollution incidents in Turkey? Describe the reason for your answer and how you found it.
- 12)** What is the median of the number of the water pollution incidents in Turkey? Describe the reason for your answer and how you found it.
- 13)** What is the arithmetic mean of the number of water pollution incidents in Turkey? Describe the reason for your answer and how you found it?

Figure 3.12. The 11th, 12th and 13th problems in the test

In the 14th problem, a media report which was prepared by the Chamber of Environmental Engineers was given to the students. The report illustrates the amount of rainfall in different regions of Turkey for the year 2016. The region or regions with less rainfall than normal in 2016 was asked to students. In the 14th problem, the students were expected to read and interpret the data they encounter in daily life presented in the media report. The 14th problem is presented in Figure 3.13 below:

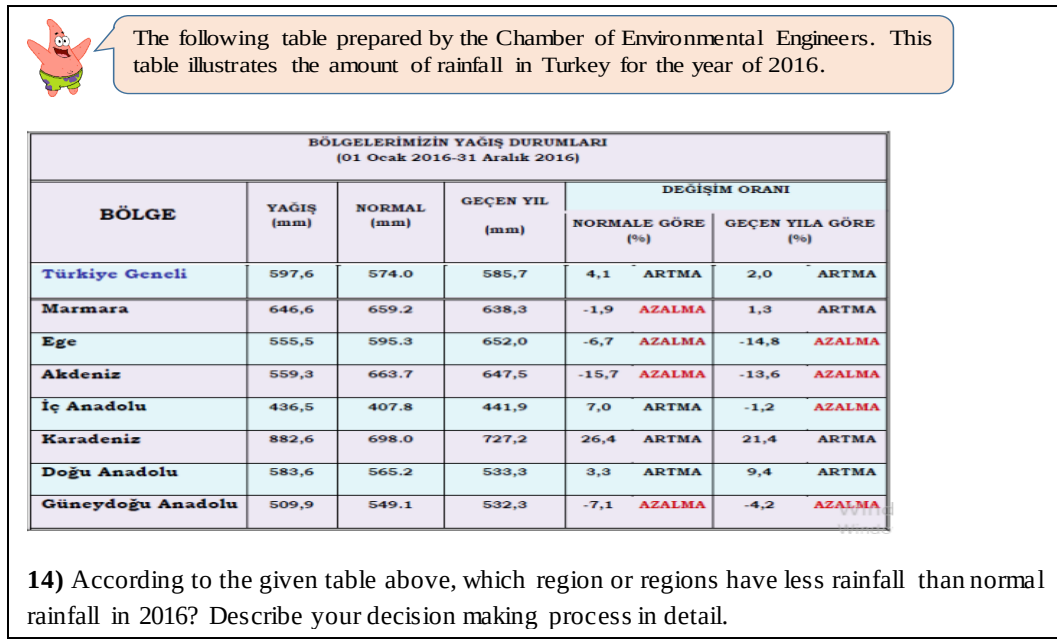


Figure 3.13. The 14th problem in the test

The 15th problem in the SLT-CP was adapted from the study of Watson and Callingham (2003). The original version of the problem is as follows:

M2PI. Explain the meaning of this pie chart. Is there anything unusual about it?

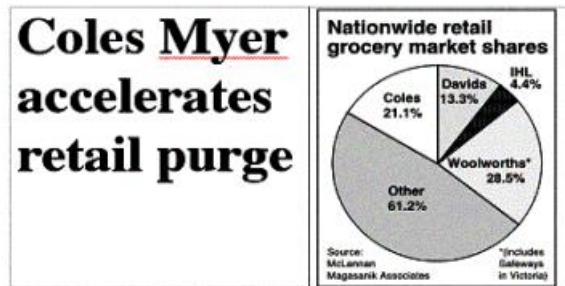


Figure 3.14. The original version of 15th problem in the test (Watson & Callingham, 2003)

In the problem, a pie chart which illustrates the sources of water pollution in a country is given. The students were requested to read and interpret the data presented with a pie chart regarding the grocery market shares. Since the problem context was not compatible with the context of the SLT-CP, the data regarding grocery market shares in the original problem was converted into the data regarding the sources of water pollution in a country.

In the 15th problem, the students were required to express their implications about the pie chart. Then, the students were expected to notice the unusual situation about the pie chart since the chart exceeds 100%. The researcher made several changes in the question. Initially, the context of the question was adjusted into sustainability and environment context. In that regard, the parts of the whole were changed as the sources of water pollution. To make the item realistic, the percentages were retrieved from the official sources. Additionally, the students' interpretations about the chart were asked. The final version of the question is presented below:

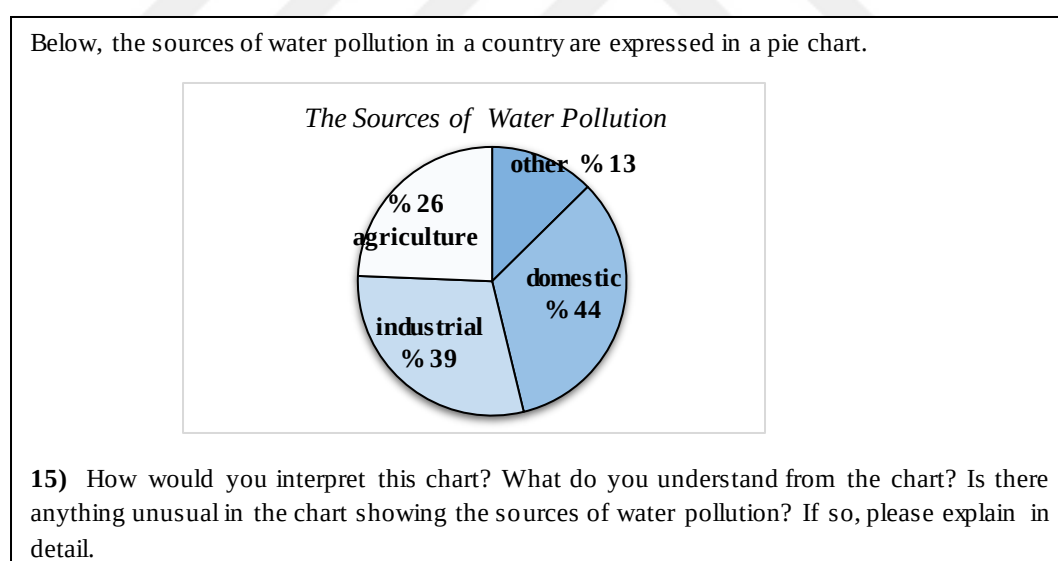


Figure 3.15. The 15th problem in the test

The 16th problem was adapted from a TIMMS (2011) problem. In the original version of the problem, the pie chart illustrates Mr. Johnson's students' favorite subjects. By reading the data presented in the pie chart, the students were requested

to choose the bar graph which shows the same information as the pie chart. The initial version of the item was as follows:

Item label: Comparing pie chart with bar chart

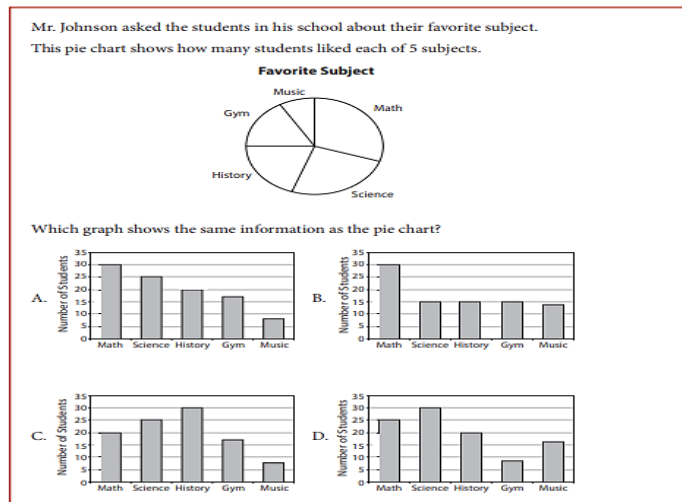


Figure 3.16. The original version of the 16th problem in the test (PISA, 2011)

To obtain the 16th problem in the test, several changes were made in the original version of the problem. Initially, the problem started with a UNICEF report to attract students' attention to the problem context. Then, statistics about the number of water-scarce countries was represented with a pie chart. To engage student into the context of the problem, authentic data was retrieved from the General Directorate of Meteorology (2014). Subsequently, the students were requested to choose the bar chart which gives the same information as the pie chart. The final version of the 16th problem is presented below:

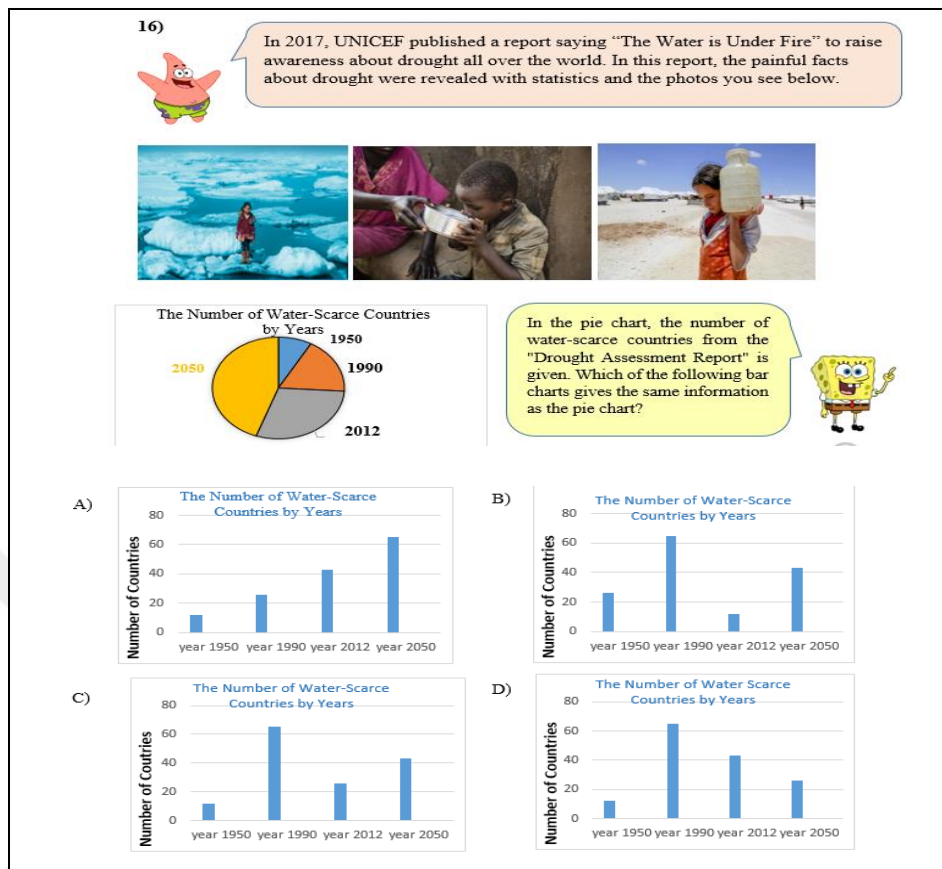


Figure 3.17. The 16th problem in the test

The 17th problem was adapted from the study of Sharma (2017). In the problem, the air temperatures of twelve consecutive days were presented for two cities and it was asked to choose the warmer city by providing a justification. The initial version of the problem is presented in Figure 2.18 below:

Day	1	2	3	4	5	6
Wellington temperature °C	15	16	12	10	13	10
Auckland temperature °C	18	13	16	15	19	20

Day	7	8	9	10	11	12
Wellington temperature °C	13	15	9	13	8	12
Auckland temperature °C	13	11	15	15	16	15

(a) Is Auckland warmer than Wellington? How do you know?

(b) Have you got any questions about the information presented in the tables? Explain your thinking.

Figure 3.18. The original version of the 17th problem(Sharma, 2017, p.125)

Several changes were made regarding the 17th problem of the test. Initially, the problem context was converted into global warming context for the integrity of the whole context of the test. The students were given a table which shows the nine-day temperatures of a city measured in two different intervals. The students were expected to choose the warmer interval between two different intervals. To obtain the correct answer, the students were expected to use the information given in the table. The final version of the 17th problem is as follows:



Global warming, which is one of the causes of drought, brought about imbalances in climate change and air temperatures. In the table below, nine-day temperatures of a city measured in two different intervals are given. *(Answer 17th and 18th questions according to the table)*

	1 st day	2 nd day	3 rd day	4 th day	5 th day	6 th day	7 th day	8 th day	9 th day
1 st interval	15	16	12	10	13	10	13	15	9
2 nd interval	18	13	16	15	19	20	13	11	15



According to the given table, in which interval is this city warmer? How did you decide? Explain the reason for your answer.

Figure 3.19. The 17th problem in the test

The same table was used for the 18th problem. In this problem, the students were asked to choose the best line graph to represent the data in table. The researcher inserted distractors in the options including student misconceptions regarding the construction of the line graph. The 18th item was presented in Figure 3.20 as follows :

	1 st day	2 nd day	3 rd day	4 th day	5 th day	6 th day	7 th day	8 th day	9 th day
1 st interval	15	16	12	10	13	10	13	15	9
2 nd interval	18	13	16	15	19	20	13	11	15



According to this table, which of the following is the correct form of the temperatures in the 1st and 2nd intervals given in the table, shown in line graph?

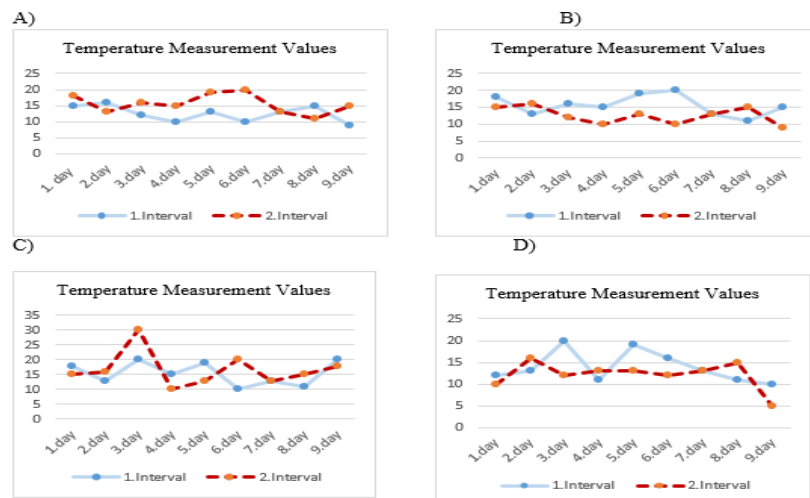


Figure 3.20. The 18th problem in the test

The 19th problem was adapted from the study of Sharma (2017). In the original version of the problem, the students were given three competitors' time scores and they were expected to select one of the three competitors as the most appropriate competitor by providing justification about their choice. The original version of the 19th problem is given in Figure 3.21.

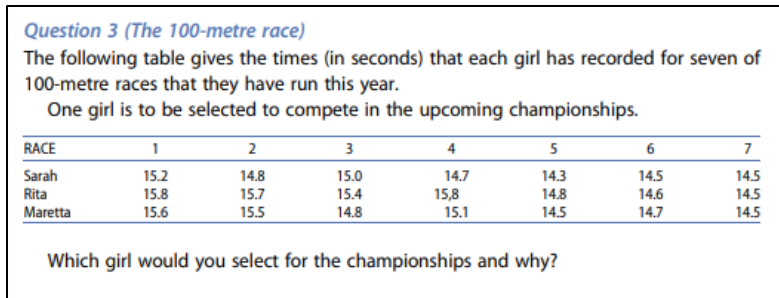


Figure 3.21. The original version of the 19th problem (Sharma, 2017, p.125)

In order to adapt the 19th problem to the current study, several changes were made in terms of context and content of the problem. In the adjusted version, as presented in Figure 3.22, the students were presented a table taken from a newspaper. The table illustrated the water shortage values between the years 2010-2040 for several countries. According to the problem, a researcher wanted to make an investigation in the country where water shortage is experienced the most. The students were expected to help the researcher by determining the most appropriate country which experiences water shortage the most.

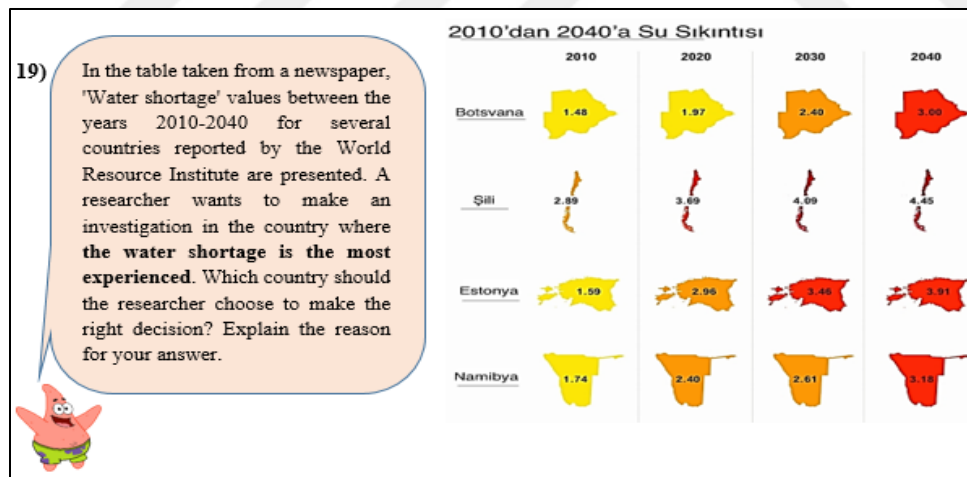


Figure 3.22. The 19th problem in the test

The 20th and 21st items were adapted from the ARTIST Project of delMas et al. (2007). The researchers prepared a set of problems in a project to assess students' statistical literacy. Both the 20th and 21st items were retrieved from the problems

included in the ARTIST Project. The original version of the problems were presented below:

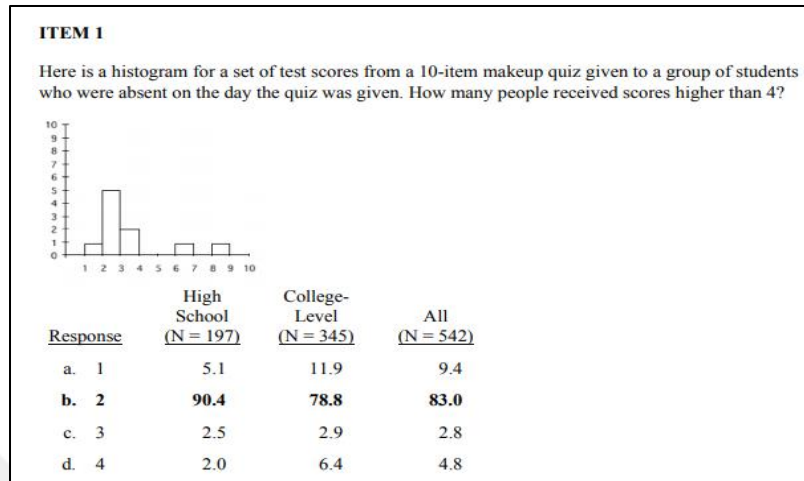


Figure 3.23. The original version of the 20th problem (delMas et al., 2007)

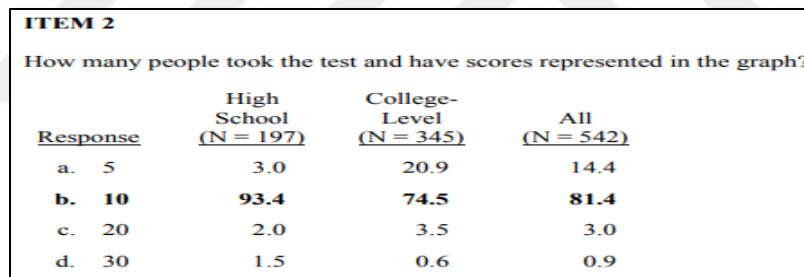


Figure 3.24. The original version of the 21st problem (delMas et al., 2007)

Several changes were made on both questions. Firstly, the context of the problems was changed into statistics about countries experiencing water scarcity for the integrity to the context of the whole test. Additionally, the values used in the problem were altered according to the context. The final versions of the problems were presented below:

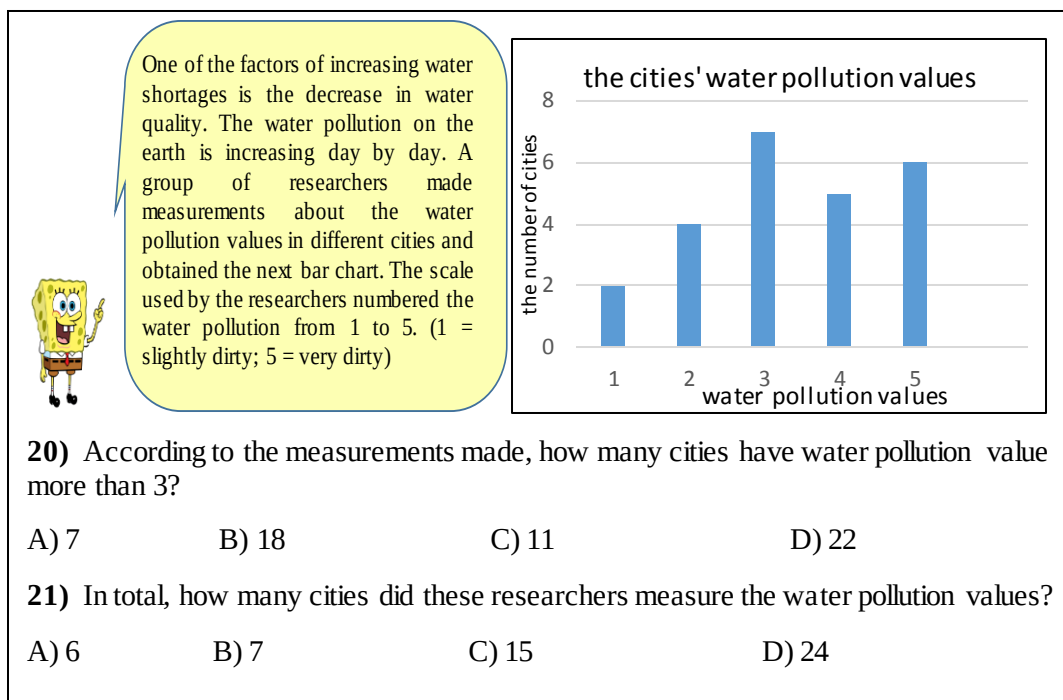


Figure 3.25. The 20th and 21st problems in the test

The 22nd problem was adapted from the study of English and Watson (2016). The researchers presented a newspaper article and encouraged students to consider the set of problems within the scope of a sustainability activity to learn about students' statistical literacy and raise students' interest in sustainability issues. The original version of the newspaper article is as follows:

Friday 17 May 2013

Forestdale Times

Children not environmentally friendly after all

"Australian school children are not as environmentally friendly as we thought," said Mr Plant from the Down-to-Earth Watchers Group. "We surveyed a class of Year 5 students in Tasmania and we were shocked with the results.

The survey found that children do not do enough in their home to conserve the environment. Many students admitted to having long baths, tossing soft-drink cans in with the food rubbish, or worse, just leaving them on benches in parks. No efforts were made to conserve electricity either with most moving from room to room leaving lights, TVs and iPods on.

"Children need to take more care of the environment. They simply don't care at all", Mr Plant said.

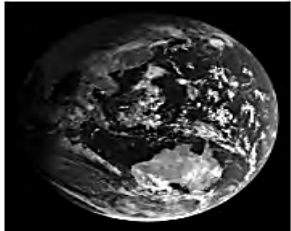


Figure 3.26. The original version of the newspaper article used in the 22nd item of the test (Watson & English, 2006)

The researcher translated the newspaper article into Turkish by summarizing the main ideas. Then, the students were asked their ideas about the generalizability of the sample to the entire population in the newspaper article. The final version of the problem is as follows:

Children Are Not Environmentally Friendly

We think of the children of schoolage in Turkey appeared to be so environmentally friendly. In an interview with an environmental scientist, he said, "The result of the study with a group of 5th grade students in Ankara shocked us." The study carried out showed that the students did not make efforts to protect the nature, waste a long time in the shower and waste water, and do not try to save electricity to save the waste.

The environmental scientist underlined that students should be more careful about this issue because they are not sensitive about the environment.

22) Is that claims of environmental scientist can be generalized to all children in Turkey? It is that claims may be true for every child in Turkey? Why? Please explain the reason for your answer in detail with detailed statistical expressions

Figure 3.27. The 22nd problem in the test

The 23rd, 24th and 25th items were adjusted from the Connected Mathematics Project (Lappan et al., 2002) and the study of Sorto and White (2004). These problems were designed to examine students' conceptual understanding of range, mode and median. Sorto and White (2004) indicated that students even teachers are confused whether measures of central tendency are numerical data or categorical data. Therefore, the researchers designed this kind of problems to investigate the interpretation of mode, median and range of a given data set by making some claims about the given data set. The original version of the problems is as follows:

Item 1

One middle school class generated data about their pets shown below. Students were talking about the data and one said:

“ The mode is dogs, the median is duck, and the range is 1 to 7.”

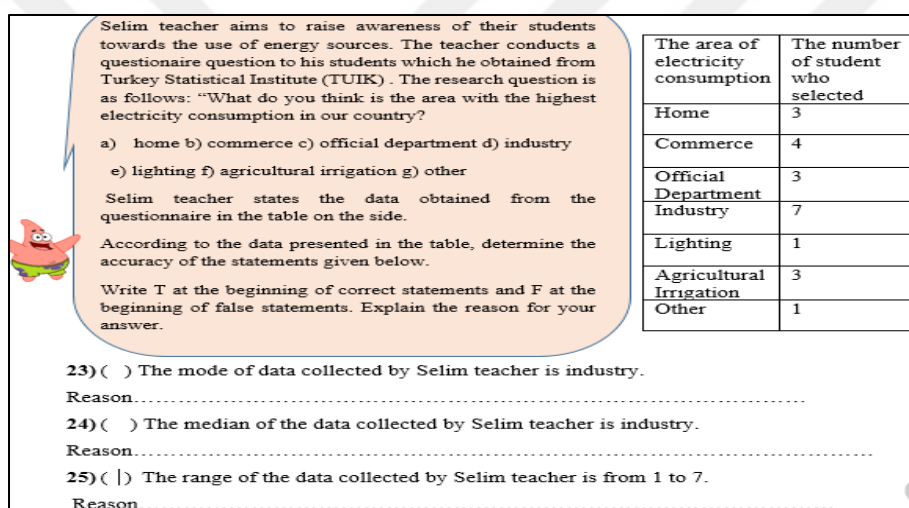
If you think the student is right, explain why.

If you think the student is wrong, identify the mistake(s).

Pet	Frequency
bird	2
cat	4
cow	2
dog	7
duck	1
fish	2
goat	1
horse	3
rabbit	3

Figure 3.28. The original version of the 23rd, 24th and 25th problems (Sorto &White, 2004)

Several changes were made on the original problems. Initially, the problems were converted into True-False items. Subsequently, the context of the problems was made convenient with the context of energy saving by collecting real data from a classroom. According to the problem context, Selim teacher collected data from her classroom by using a research question obtained from the Turkey Statistical Institute (TUIK) and illustrated the data that he collected with a table. Then, the teacher wrote several statements regarding mode, median and range of the collected data set. The students were required to determine the accuracy of the statements by providing explanations about the reason for their decisions. The 23rd, 24th and 25th problems in the test are displayed in Figure 3.29.



Selim teacher aims to raise awareness of their students towards the use of energy sources. The teacher conducts a questionnaire question to his students which he obtained from Turkey Statistical Institute (TUIK). The research question is as follows: "What do you think is the area with the highest electricity consumption in our country?"

a) home b) commerce c) official department d) industry
e) lighting f) agricultural irrigation g) other

Selim teacher states the data obtained from the questionnaire in the table on the side.

According to the data presented in the table, determine the accuracy of the statements given below.

Write T at the beginning of correct statements and F at the beginning of false statements. Explain the reason for your answer.

The area of electricity consumption	The number of student who selected
Home	3
Commerce	4
Official Department	3
Industry	7
Lighting	1
Agricultural Irrigation	3
Other	1

23) () The mode of data collected by Selim teacher is industry.
Reason.....

24) () The median of the data collected by Selim teacher is industry.
Reason.....

25) () The range of the data collected by Selim teacher is from 1 to 7.
Reason.....

Figure 3.29. The 23rd, 24th and 25th problems in the test

3.3.2 Statistical Literacy Test without Context Problems

The Statistical Literacy Test without Context Problems (SLT) was constructed by eliminating the problem context in the Statistical Literacy Test with Context Problems (SLT-CP). The items in both tests were the same in terms of the order of the problems, numerical values, problem situations and statistical terms. The only difference was that the Statistical Literacy Test with Context Problems (SLT-CP) were presented to the students as context-bounded problems, whereas the Statistical Literacy Test without Context Problems (SLT) were presented as the context-free

forms of the problems in the SLT-CP. To illustrate the conversion between the two tests, the 16th problem from both tests are presented below:

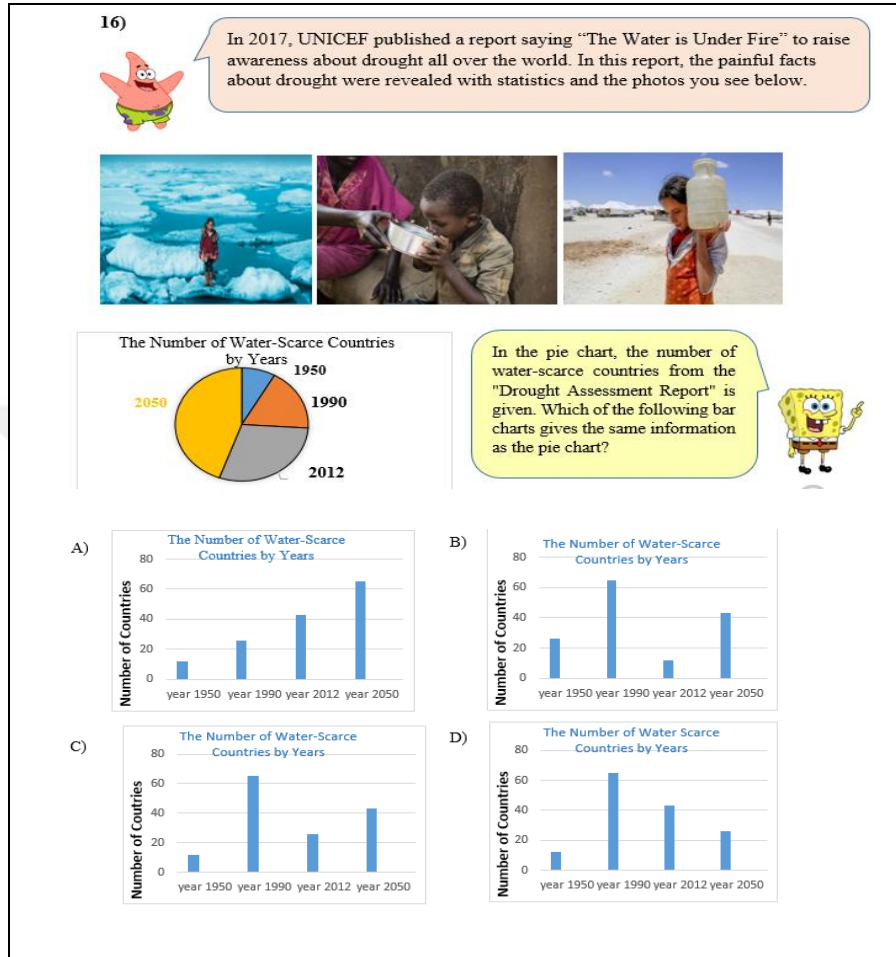


Figure 3.30. The 16th problem in the SLT-CP

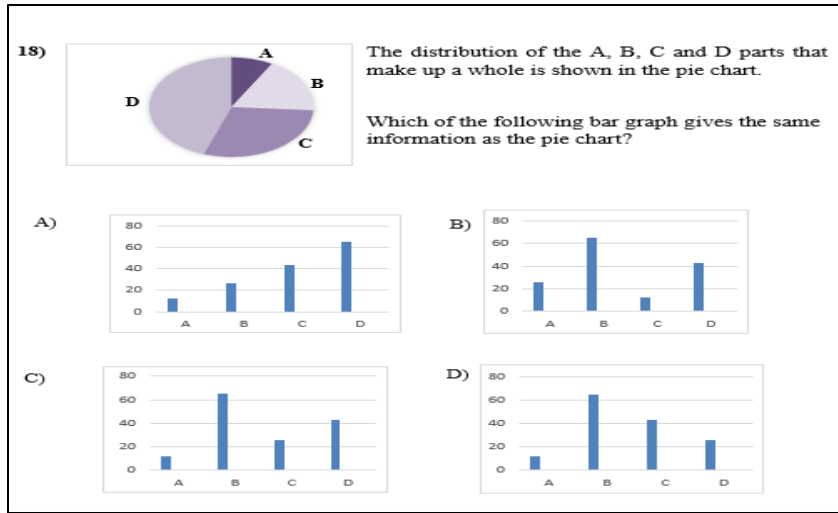


Figure 3.31. The 16th problem in the SLT

As can be seen, the problem context was eliminated from the 16th problem included in the Statistical Literacy Test with Context Problems (SLT-CP) and the 16th item in the Statistical Literacy Test without Context Problems (SLT) was obtained. The same procedures were applied for all 25 problems of the tests. For full Turkish version of the SLT, please see Appendix D.

3.4 Pilot Study

To ensure the duration of the implementation, the validity and reliability of the statistical literacy tests, the tests were piloted with 74 students from a public school in Yenimahalle district of Ankara at the end of the 2019-2020 fall semester. Both tests were composed of 33 items including multiple choice, True-False and open-ended problems. The tests were divided into two parts to observe an individual students' performance on both problems with and without context. More precisely, in the first group, the first half of the test was composed of context problems, whereas the second half was without context problems. Similarly, in second group, the first half was composed of without context problems, while the second half included context problems. Therefore, individual differences in achievement were eliminated in both problem types. To complete the tests, two class hours (80 minutes) were

allocated. According to the students' responses and feedback to the items in the statistical literacy tests, some changes and corrections were made in the statistical literacy tests and the scoring rubric. Some of the questions were eliminated from the tests. Additionally, some of the problems were revised in parallel with students' feedback and the researchers' observations. The problems that most of the students left unanswered or answered incorrectly were specified and these questions were revised. The sentence structures of the problems were reviewed by the researcher and the experts in the field of mathematics education. The final versions of the tests can be seen in Appendix C and Appendix D. Additionally, the scoring rubric was reviewed within the scope of the revisions made in both tests. These revisions were related to the validity of the statistical literacy tests. As for reliability of the instruments in the pilot study, Cronbach alpha coefficient was calculated as .82, which is greater than .70 and thus the instrument was reliable (Pallant, 2007).

3.5 Validity and Reliability of the Instruments

Recently, validity is referred to as 'appropriateness, meaningfulness, correctness, and usefulness of inferences a researcher makes' (Fraenkel, Wallen & Hyun, 2011, p.147). That is, validity is the consistency between the research purpose and the totality of test outcomes. In other words, validity is the collection of evidence to establish that the test interpretation scores relate to its proposed implementation in the curriculum (Creswell, 2012). The content-related evidence is one certain evidence which refers to the instrument's content and format. (Fraenkel, Wallen & Hyun, 2011, p.148). To provide content-related validity, data and chance objectives in the National Mathematics Education Curriculum (MoNE, 2018) related to content domains of measures of central tendency, graph interpretation, sampling and questioning of statistical data were examined. The objectives stated in the National Mathematics Education Curriculum (MoNE, 2018) were specified and the items were matched with those objectives in terms of statistical content domains to ensure the compatibility of the problems with the objectives in mathematics curriculum.

Thus, the table of specification (see Appendix E) was developed according to the objectives in curriculum to ensure the comprehensiveness and appropriateness of the instruments.

To ensure content-related validity of the study, two research assistants, four experienced mathematics teachers and one academician from university revised the problems and they provided feedback related to item format, content and consistence with the objectives in the National Mathematics Education Curriculum related to data and chance statistical content domains. Besides, the adequacy of sample was discussed with these experts. After necessary changes were made based on the feedback, the pilot study was conducted at a school in Yenimahalle District of Ankara. The tests were administered to two classes, with a total of 74 students. Two class hours were allocated for the implementation of the pilot study. During the implementation, the students provided valuable feedback regarding the content and format of the tests by asking questions and stating their opinions regarding the problems included in the tests. The items were revised according to the results and feedback obtained from the pilot study.

The reliability of an instrument refers to ‘the consistency of the scores obtained’ (Fraenkel, Wallen & Hyun, 2011, p.154). Put differently, if an instrument is reliable, then scores from an instrument are stable and consistent. (Creswell, 2012, p.159). Initially, internal consistency methods were utilized to ensure reliability. To do this, Cronbach alpha coefficient (Cronbach, 1984) was calculated with SPSS 20 program because the items were coded from 0 to 3 continuously rather than as correct and incorrect (Creswell, 2012). The Cronbach alpha coefficient was obtained as .85 for the Statistical Literacy Test with Context Problems (SLT-CP) and .84 for the Statistical Literacy Test without Context Problems (SLT). As Pallant (2007) stated, since coefficients were greater .70, the scores on both tests were reliable.

3.6 Data Collection Procedures

The aim of the present study was to investigate 8th grade students' achievement and statistical literacy through problems with and without context in terms of statistical content domains. Additionally, the students' achievement and statistical literacy were compared through problems with and without context. In parallel with these aims, data collection instruments were developed by taking into consideration the related literature during the fall semester of 2019-2020. The necessary revisions were made based on feedback from experts, experienced teachers and the 8th grade students. Before starting with the data collection process, permission was received from Middle East Technical University Human Subjects Ethics Committee (See Appendix A). Afterwards, the necessary permissions were received from the Ministry of National Education (See Appendix B). To ensure the reliability and validity for the instruments, a pilot study was conducted with 74 students at the beginning of January. Based on student responses in the pilot study, several corrections and revisions were made. Then, the researcher visited the schools one by one. The researcher informed the school administrations and mathematics teachers about the content and aim of the study. The administrators and mathematics teachers provided a schedule indicating the appropriate time period for the education program. The appropriate time periods were noted by the researcher. Then, the researcher scheduled the data collection dates for schools based on the appointments. Therefore, the exam dates or school activities did not coincide with the data gathering process. Before the administration of the tests, the researcher matched the students based on their average grade points from mathematics lesson in the 2019 -2020 fall semester. In this way, the threat of subject characteristics to internal validity was controlled. Therefore, both tests were administered to students with similar achievement level. In a class, half of the students were administered the Statistical Literacy Test with Context Problems and half of the class were administered the Statistical Literacy Test without Context Problems. More precisely, the researcher obtained students' average grade points regarding mathematics belonging to the

previous semester. Then, the researcher paired the students having the same grade points in mathematics. For instance, the students having the average mathematics grade point of 98 were assigned the same number. Then, one of the students solved the Statistical Literacy Test without Context Problems, while the other solved the Statistical Literacy Test with Context Problems. Therefore, both tests were solved by the students from each achievement level. By doing this, the researcher eliminated the results that stem from teacher difference across the classes for both test takers. The tests were administered to students during school time by the researcher. The students were informed about the tests and the nature of the study. The researcher indicated that participation in the study was voluntary and they could withdraw from the study any time they want. The researcher also informed students about the confidentiality of the study. Afterwards, two class hours were given to the students to complete the instrument. The students had a break between the two class hours and their papers were collected during the break. A schedule for the data collection process is presented in Table 3.2.

Table 3.2 The Time Schedule for the Data Collection Procedure

Date	Action
September 2019-January 2020	Development of instruments
January 2020	Pilot Study and revisions in the instruments
February-March 2020	Data Collection
March-April 2020	Data Analysis

3.7 Data Analysis Procedures

The aim of the present study was to investigate 8th grade students' achievement and statistical literacy through problems with context and problems without context in terms of statistical content domains. Then, the students' achievement and statistical

literacy were compared through problems with context and problems without context.

Initially, quantitative research methodology was used to specify the 8th grade students' achievement in problems with and without context. A holistic rubric was prepared by the researcher as a result of a comprehensive literature review. While developing the rubric, the researcher utilized Watson and Callingham's (2004) response code examples since some of the problems were adopted from the study of Watson and Callingham (2004). These codes were adapted according to the content of each item in the tests. To make an in depth analysis, a special code example was prepared for each item. Firstly, the rubric for the Statistical Literacy Test with Context Problems (SLT-CP) was prepared. Then, the context was eliminated, and the rubric for the Statistical Literacy Test without Context Problems (SLT) was obtained. A sample from the original version of rubric for 15th problem is provided below.

Table 3.3 Original version of rubric for problem 15 (Watson & Callingham, 2004 ,p.153)

Code	Description	Examples
3	Appropriately interprets the information on the chart and understands proportions	"People have most shares in 'other' companies" "It shows you what companies people invest most in"
2	Inappropriately interprets the chart but understand the proportions	"How many people shop there"
1	Draws extremely basic and non-central conclusions about the intent of the graph	"Shows different stores" "Woolworth's is going well"
	Tautological response or reads percentages direct from chart	"Grocery market shares" "IHL has 4.4%, Woolworth's has ..." etc.
0	Idiosyncratic responses and literal interpretations about pies	"Other stores sell more pies"
	No response	
Code	Description	Examples
2	Detects error	"There is more than 100% there's 128.5%"
1	Focuses on non-central issues	"I thought Coles and Woolworths were bigger than that"
0	Focuses on presentation or fails to give an example	"It has got other shaded but IHL is black and its writing is on the outside of the graph"
	Does not notice anything unusual	"No it looks fine"
	No response	

The context of the rubric was changed according to the context of the 15th problem in the SLT-CP and the rubric was obtained for the 15th problem of SLT-CP as provided below.

Table 3.4 Rubric for problem 15 in the SLT-CP

Classification of Responses	Student Response
Statistical	Statistical interpretation about the chart and recognition of error about the chart. <i>-The biggest source of water pollution is domestic. Total percentage is 122% while it should be 100%</i> <i>- By looking at the chart, we can find the biggest source of water pollution. Total percentage is 122% while it should be 100%</i>
Pre-statistical	Make sense of percentages, but improper comments / but cannot notice the error. <i>- Percentage of water pollution sources</i> <i>-The most industrial and the least other sources constitute the water pollution.</i>
Non-statistical	Idiosyncratic interpretations Unawareness of error <i>-Shows different sources</i> <i>-Domestic pollution should be reduced</i>
Unrelated/Blank	

Moreover, the context of the rubric for 15th problem of SLT-CP was removed and the rubric for 15th problem of SLT was obtained as presented in Table 3.5.

Table 3.5 Rubric for problem 15 in the SLT

Classification of Responses	Student Response
Statistical	Statistical interpretation about the chart and recognition of error about the chart. <i>-The biggest part belongs to 44%. Total percentage is 122% while it should be 100%</i> <i>- By looking at the chart, we can find the biggest and the smallest parts of the whole. Total percentage is 122% while it should be 100%</i>
Pre-statistical	Make sense of percentages, but improper comments / but cannot notice the error. <i>- Percentage of parts of a whole</i> <i>-The most 44% and the least 13%</i>
Non-statistical	Idiosyncratic interpretations Unawareness of error <i>-Shows different percentages</i> <i>-The part belonging 44% should be reduced</i>
Unrelated/Blank	

By using the rubric, the students' responses were coded from 0 to 3. Then, code 0 and code 1 responses were regarded as incorrect answer, and code 2 and code 3 responses were regarded as correct answer. For code 0 responses the students provided totally incorrect response or left the problem blank. For code 1 responses the students provided only an answer irrelevant from statistical responses. Code 2 responses were still statistically appropriate although they were including minor operation mistakes or lack of attention. Code 3 responses were statistically appropriate responses. The students' scores for each problem were entered to SPSS 20 Program. In this way, the frequencies of correct responses and the percentage of achievement for each problem in terms of statistical content domains which were measures of central tendency, graph interpretation, sampling and questioning of statistical data were obtained.

In order to investigate students' statistical literacy in terms of statistical content domains, the content of student responses was analyzed by the researcher. Content analysis helps researchers to analyze written responses indirectly (Fraenkel, Wallen & Hyun, 2011). Thus, the students' responses were categorized under four categories of student responses as statistical, pre-statistical, non-statistical and unrelated or unanswered. These categories were determined by the researcher as a result of a comprehensive literature review and the findings of the pilot study. For each category, the frequencies and percentages together with the sample student responses were presented in a table to present a detailed examination. To specify categories of student responses, code 3 responses were regarded as statistical, code 2 responses were regarded as pre-statistical, code 1 responses were regarded as non-statistical and code 0 responses were regarded as unrelated or blank responses. The codes and corresponding categories were presented in Table 3.6.

Table 3.6 Codes and Corresponding Categories for Student Responses

Code	Category
3	Statistical
2	Pre-statistical
1	Non-statistical
0	Unrelated / Blank

Lastly, in order to compare the students' achievement and statistical literacy through problems with context and problems without context, an independent sample t-test was employed. SPSS 20 Program was used to reveal whether there was a significant mean difference between students' achievement in problems with and without context. The scores of the students taking the SLT-CP were assigned as group 1 and those of the students taking the SLT were specified as group 2. Then, the mean difference in students' achievement in the SLT-CP and SLT was investigated, compared and contrasted. Additionally, descriptive statistics for overall test achievement and for statistical content domains were examined to compare the students' achievement and statistical literacy in the SLT-CP and SLT.

3.8 Assumptions and Limitations

In the present study, there were several assumptions and limitations. In this part, these assumptions and limitations are discussed.

First of all, it was assumed that all the students responded to the problems in the instruments by not being affected by others' ideas. They responded to the items by themselves with no intervention of external factors. The students focused on the instruments. Furthermore, the items in the instruments investigated students' existing status about statistical literacy. Also, it was assumed that the students remember the terms of mode, median, average, range, graph construction and graph interpretation from the 5th, 6th and 7th grades.

The use of convenience sampling might bring about some limitations to the study since the sample might not represent the population. This was the major limitation of the study. In addition, the students' statistical literacy was investigated with limited numbers of questions. More precisely, the findings of the study was limited with the student responses only for 25 problems. The study was limited to the students from four public school. On the other hand, in each class, two different instruments were distributed to the students. Half of the students solved the Statistical Literacy Test with Context Problems and half of the students solved the Statistical Literacy Test without Context Problems. Although the researcher made the required explanations, the students might have had tendencies about the other test. The students might have wondered the problems in others' tests. At this point, the researcher indicated that both tests contained the same problems, and the content is the same. The only difference was that one test included a context and the other did not include context. Nevertheless, this might be a limitation for this study. Moreover, the students were matched according to their academic achievement levels while distributing the instruments including problems with and without context. This brings about the problems related to the individual differences in student responses.

3.9 Internal and External Validity of the Study

Both internal validity and external validity issues are indispensable for any kind of research study. In this section, internal and external validity issues for the present study are discussed.

3.9.1 Internal Validity

Internal validity can be explained as the monitored changes in dependent variable which only results from the independent variable (Fraenkel, Wallen & Hyun, 2011). Put differently, Merriam (2009) describes internal validity as the suitability between

research findings and the reality. That is, internal validity is concerned about whether the researcher measures what is intended to measure.

According to Fraenkel, Wallen and Hyun (2011), there exists three main threats to internal validity in a survey research as mortality (loss of subjects), location, and instrumentation. To control these threats, some precautions should be taken.

Mortality (Loss of subjects): Since the present study is a cross-sectional survey which collects data at one point of a time, loss of subjects was not an internal validity threat.

Location: Fraenkel, Wallen and Hyun (2011) stated that location threat occurs when the data is collected from the participants who have different location conditions such as lighting, conditioning and wideness. In this study, data were collected from public schools in Ankara. The researcher observed the classrooms before collecting data. The conditions of the classrooms were similar to each other. There was not a gap between the location conditions. Hence, location threat was eliminated on a vast scale.

Instrumentation: Instrumentation threat is connected to the way that data are gathered and utilized. Instrumentation threat includes instrument decay, data collector characteristics and data collector bias (Fraenkel, Wallen & Hyun, 2011).

Instrument decay threat: Instrumentation threat may occur when the function of the instrument is altered and the scoring procedure is changed. (Fraenkel, Wallen & Hyun, 2011). This threat comes into question when the instruments allow for various comments of the results or when the scorer gets tired and scores the test disparately. The instruments were applied by the researcher one point at a time by following the same data collection schedule to standardize the conditions during the data gathering process. To control the changes in scoring procedures, a different rater scored the responses by using the same scoring key. The researcher and co-coder scored the 50 students' papers which is 10% of the participants up to obtaining 95% agreement in scoring. Consequently, the raters came to an agreement about the scoring of

instruments. Moreover, the researcher scheduled the scoring process to minimize the changes in scoring procedure. Therefore, instrumentation threat was eliminated on a large scale.

Data Collector Characteristics: During the implementation of the instruments, the participants might change their responses due to the data collector. In the present study, the data were collected only by the researcher. In other words, the data collector was the same for all the participants. Additionally, the researcher did not interact with the students except for the explanations of the expectations from the students before starting the implementation of the instruments. Therefore, data collector characteristics was not a threat to the study.

Data Collector Bias: During the implementation of the instruments, the data collector might manipulate the conditions according the tendencies about the study. To control the data collector bias, the researcher did not interact with the students not to affect their responses.

3.9.2 External Validity

External validity “is the extent to which the results of a study can be generalized” (Fraenkel, Wallen & Hyun, 2011, p. 103). External validity includes two dimensions as population generalizability and ecological generalizability.

Population generalizability is “the degree to which a sample represents the population of interest” (Fraenkel, Wallen & Hyun, 2011, p. 103). In other words, representativeness of the sample to the population is a crucial aspect of population generalizability. In the present study, the target population was all 8th grade students in Ankara; the accessible population was the 8th grade public school students from four different public schools in Ankara. Since the accessible population was specified by the researcher’s convenience, the generalization of results to the population might be limited in this study. Additionally, since the students are public school students, the results of the study might not be extended to private school

students. Therefore, the sample of the present study might not represent all the 8th grade students in Ankara. On the other hand, the researcher tried to keep sample size of the study large as much as possible to generalize the results of the study to the population.

Ecological generalizability “is the degree to which the results of the study can be extended to other setting and conditions.” (Fraenkel, Wallen & Hyun, 2011, p. 105). The researcher must explain the setting of the study in detail. The results of the study might be generalized to other public schools which have similar conditions.



CHAPTER 4

RESULTS

The aim of the present study was to investigate 8th grade students' achievement and statistical literacy in problems with and without context in terms of statistical content domains. Subsequently, the students' achievement in problems with and without context was compared in terms of statistical content domains.

In this section, the results obtained from the Statistical Literacy Test with Context Problems (SLT-CP) and the Statistical Literacy Test without Context Problems (SLT) are presented in parallel with the research questions of the present study. The items in the tests were investigated under four statistical content domains, which are measures of central tendency, graph interpretation, sampling, and questioning of statistical data. The items were grouped and analyzed in terms of these statistical content domains. The Statistical Literacy Test with Context Problems (SLT-CP) and the Statistical Literacy Test without Context Problems (SLT) both include 25 problems, which are basically the same. The difference is that while the problems in the SLT-CP include context, the problems were posed by removing the context in the SLT. The problems include open-ended, True-False, and multiple-choice items. Most of the problems were in the open-ended form to examine students' statistical literacy and thinking in detail.

In the first part, frequencies and percentages of achievement belonging to the 8th grade students' in statistical literacy problems in terms of four statistical content domains were presented. The students' achievement in statistical literacy problems with and without context was compared and contrasted. Then, the 8th grade students' statistical literacy was investigated and categorized by conducting a comprehensive examination of the students' responses towards the instruments. The examination of

8th grade students' current statistical literacy was supported by sample student responses. Lastly, statistical methods were utilized to investigate whether there is a significant mean difference in 8th grade students' achievement in statistical literacy problems with and without context.

4.1 Achievement in Problems with and without Context in terms of Statistical Content Domains

The Statistical Literacy Test with Context Problems (SLT-CP) includes problems with context, while the Statistical Literacy Test without Context Problems (SLT) includes the same problems in context-free form. The problems were categorized and analyzed under four statistical content domains which are *measures of central tendency, graph interpretation, sampling, and questioning of statistical data*. Additionally, these problems were analyzed based on the rubric prepared by the researcher. The data analysis process was explained under the related part. During the analysis of data, code 0 and code 1 responses were regarded as incorrect answers in open-ended problems, while code 2 and code 3 responses were regarded as correct answers. For code 0 responses the students provided totally incorrect response or left the problem blank. For code 1 responses the students provided only an answer irrelevant from statistical responses. Code 2 responses were still statistically appropriate although they were including minor operation mistakes or lack of attention. Code 3 responses were statistically appropriate responses. No response and unrelated responses were coded as 0. In multiple-choice problems, correct responses were coded as 3 and they were considered as correct; incorrect responses were coded as 0 and they were regarded as incorrect. The students' achievement was specified according to their correct answers.

In this section, 8th grade students' achievement in problems with context and problems without context is discussed according to content domains. The descriptive

analysis of the correct responses of 500 eight grade students in the Statistical Literacy Test with Context Problems (SLT-CP) and the Statistical Literacy Test without Context Problems (SLT) in terms of the four statistical content domains is presented in Table 4.1, Table 4.2, Table 4.3 and Table 4.4.

4.1.1 Achievement in Measures of Central Tendency

In both tests, P8, P9, P11, P12 and P13 were measures of central tendency problems. The number of students who provided a correct answer and the percentage of achievement for each measures of central tendency problem in both tests are presented in Table 4.1.

Table 4.1 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Measures of Central Tendency Domain

	Measures of Central Tendency				
	P8	P9	P11	P12	P13
Problems with Context	79 31.6%	91 36.4%	64 25.6%	47 18.8%	147 58.8%
Problems without Context	100 40%	91 36.4%	96 38.4%	48 19.2%	117 46.8%

As Table 4.1 illustrates, 8th grade students' achievement in the measures of central tendency problems with context varied between 18.8% and 58.8%, whereas their achievement in problems without context ranged from 19.2% to 46.8%. This means that the 8th grade students' achievement in context-bounded measures of central tendency problems showed more variability than the context-free form of these problems.

In problem 8, the 8th grade students' understanding of the "average" concept was investigated. In the Statistical Literacy Test with Context Problems (SLT-CP), the students were given a newspaper quote which included statistics about the amount of garbage types thrown in Turkey in a day. The students were requested to explain their understanding and make an interpretation about the reason for the use of the term "average" in the newspaper quote.

79 (31.6%) students made statistically appropriate interpretations about the use of the term "average" in the newspaper quote. Those students described the central tendency in the data set of the amount of garbage thrown in Turkey in a day by using the "concept". Additionally, some of the correct responses included the use of the measures of central tendency such as mean, mode, and median to interpret the term "average". Most of the incorrect responses were composed of answers that were away from statistical explanations such as *approximately, more or less, and about*. These students interpreted the term "average" with non-statistical explanations.

On the other hand, in the Statistical Literacy Test without Context Problems (SLT), the statistics regarding the amount of garbage thrown in Turkey in a day were removed from the problem, and the problem was asked directly to examine the students' understanding of the term "average". 100 (40%) of the 8th grade students made statistically appropriate interpretations about the use of the term. These students explained the central tendency in the data set by using the concept of "average". They interpreted the term "average" as the "*representative value*" of a data set as a measure of central tendency. Similar to the context-bounded form, the inappropriate responses included idiosyncratic and non-statistical explanations such as '*because he is not completely sure of the exact amount, more or less than, about and approximately*'. As can be seen, the students were more successful in the context-free form of problem 8.

The difference in 8th grade students' achievement in the context-bounded and context-free forms of problem 8 might have resulted from the statistics about the amount of garbage thrown in Turkey. Students were required to elaborate on their statistical thinking through extra information related to the context of the problem. In this respect, the students might have performed better in problem 8 in the SLT rather than problem 8 in the SLT-CP.

Similar to problem 8, problem 9 was related to the concept of "average". The problem was asked in multiple-choice format. In problem 9, the techniques for calculating the "average" of a data set were questioned, while problem 8 investigated students' interpretations regarding the use of the term "average". More precisely, in the context-bounded form of problem 8, the students were asked about the methods for calculating the average amount of garbage thrown daily in Turkey in a week by attracting students' attention with environmental waste issues. The students were requested to choose the option which is not appropriate to find the average amount of garbage thrown in Turkey in a week. In the context-free form, the problem asked about the methods to calculate the average of seven numbers. The students were asked to choose the inappropriate option to find the "average" of the given data set. The results revealed that the students showed the same achievement in the context-bounded and context-free forms of problem 9.

For both forms of problem 9, the students' achievement was 91 (36.4%). It might be inferred that attracting students' interests with environmental waste issues did not affect their achievement positively or negatively in terms of evaluating the methods to obtain the "average" of a given data set. Besides, the inclusion of environmental waste issues accompanied by the statistics problem might have complicated the problem situation, and therefore, it might have affected students' statistical thinking in terms of comprehending the written explanations regarding the amount of the waste thrown in Turkey. However, the students should be able to make sense of the written part to determine the inappropriate way of finding the average amount of the

waste thrown in Turkey in a week. On the other hand, the existence of environmental waste issues accompanying the statistics problem might have motivated some students to solve the problem. This might be due to the fact that the use of real-life situations such as environmental waste increases the students' motivation to solve problem.

Problem 11 was asked with problem 12 and problem 13 subsequently with the same context and the same bar graph in the SLT-CP. In these problems, some statistics about the number of different water pollution incidents that occurred in different districts of Turkey was presented in a bar graph. The data was obtained from the report of the Chamber of Environmental Engineers.

In problem 11 in the SLT-CP, the students were asked to find and interpret the “mode” of the number of water pollution incidents that occurred in different districts of Turkey. 64 (25.6%) of the 8th grade students found and interpreted the “mode” of the number of different water pollution incidents by presenting statistically appropriate explanations. These students made correct explanations about the “mode” of the number of incidents that occurred in various districts of Turkey, and they performed the appropriate operation to calculate the “mode” of the given data set.

The majority of the students (74.4%) provided an incorrect response in problem 11. Most of these students determined the “peak value” rather than the “mode” of the number of incidents in various districts of Turkey. Put differently, these students confused the “mode” of data set with the “peak value” of the data set. Moreover, some of the students wrote the name of the city with the highest water pollution instead of finding the numerical value of the “mode” of the given data. The reason for this type of answer was that the students could not decide whether the given data was numerical or categorical. For the context-free form of the problem, this situation did not create a problem as the statistics regarding the water pollution incidents that

occurred in different districts of Turkey was eliminated and only numerical values were presented to the students to find and interpret the “mode” of the given data set.

In problem 11 in the SLT, 96 (38.4%) students provided a correct response. As can be seen, the students’ achievement was better in the context-free form of problem 11. The difference in achievement might be attributed to not using the water pollution incidents context. The existence of an additional written part in the context-bounded form of the problem might have required the interpretation of the written part as well as finding and interpreting the “mode” of the given data set. This might be the reason for lower achievement in the context-bounded form of the 11th problem.

Problem 12 was the most incorrectly answered measures of central tendency problem in both the SLT-CP and the SLT. The problem was related to finding the “median” of the given data set, which is one of the central tendency measures. The students were asked to find the “median” of the number of water pollution incidents that occurred in different districts of Turkey by presenting a bar graph. The students were expected to read the data in the bar graph and find the median of the given data set. 47 (18.8%) students found and interpreted the median of the number of water pollution incidents that occurred in different districts of Turkey successfully. To obtain the 12th problem in the SLT, the situation regarding the number of water pollution incidents that occurred in different districts of Turkey was removed, and the data in the problem were attained. In other words, the problem was asked by not using a social or scientific context but using the same data with SLT-CP.

In the 12th problem of the SLT, 48 (19.2 %) of the students found and interpreted the median of the given data set. The students showed similar achievement in the 12th problem of the context-bounded and context-free form of the problem. Probably, the use of the data obtained from different districts of Turkey with water pollution

incidents did not make a vast difference in 8th grade students' achievement in problem 12.

Furthermore, when the student responses were analyzed for both forms of the problem, it was observed that the students' difficulties were similar. Initially, the students confused the term "median" with "arithmetic mean". More specifically, the students calculated arithmetic mean instead of finding the median. Moreover, some of the students did not order the data in ascending or descending order to find the "median" of a given data set. They wrote the data from left to right in the given order and chose the value in the middle of the data set. A vast majority of the students wrote the values once although the values were repeated more than once. This indicates that the 8th grade students did not make sense of the steps of finding the "median" for a given data set, and they could not understand median conceptually.

The 13th problem was the most correctly answered measures of central tendency problem in both SLT-CP and SLT. The problem context and bar graph were the same with problem 11 and problem 12. More precisely, problem 11, problem 12 and problem 13 were successive questions with the same context and the same bar graph in the SLT-CP. In problem 13, the students were asked to find and interpret the "arithmetic mean" of the number of water pollution incidents in different districts of Turkey.

147 (58.8%) students provided a correct answer for the 13th problem in the SLT-CP, whereas the number of students who provided the correct answer for the 13th problem of the SLT was 117 (46.8%). The students showed better achievement in the 13th problem with context. It can be inferred that the use of real-world data and real-world situation by integrating the number of water pollution incidents that occurred in different districts of Turkey helped students to comprehend the problem situation and thus find and interpret the "arithmetic mean" of the given data set. Probably, the students considered the number of incidents in the problem as a real-life situation

encountered through media channels such as newspapers, magazines, billboards, and televisions. Thus, the students might have built a relationship between real-life situations and the problem situation, and might have revealed better performance in interpreting the arithmetic mean of the given data set in the context-bounded form of the 13th problem.

4.1.2 Achievement in Graph Interpretation

In the tests, P2, P3, P4, P5, P6, P7, P15, P16, P18, P20, and P21 were graph interpretation problem. The distribution of students' correct answers and the percentage of achievement for each graph interpretation problem in both tests are presented in Table 4.2.

Table 4.2 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Graph Interpretation Problems

	Graph Interpretation										
	P2	P3	P4	P5	P6	P7	P15	P16	P18	P20	P21
Problems with Context	212 84.8 %	133 53.2 %	85 34 %	162 64.8 %	110 44 %	157 62.8 %	121 48.4 %	230 92 %	204 81.6 %	151 60.4 %	200 80 %
Problems without Context	221 88.4 %	190 76 %	85 34 %	174 69.6 %	106 42.4 %	177 70.8 %	131 52.4 %	242 96.8 %	194 77.6 %	112 44.8 %	210 84 %

As Table 4.2 illustrates, the 8th grade students' achievement in graph interpretation problems with context varied between 34% and 92%, whereas the achievement in the problems without context ranged from 34% to 96.8%. It can be inferred that 8th grade students' achievement showed more variability in the context-free form of the

graph interpretation problems compared to the context-bounded measures of central tendency problems.

In the 2nd problem of the SLT-CP, an infographic titled ‘Water on Earth’ was given. The graph illustrates authentic data retrieved from TUBITAK. The students were asked to select the most appropriate graph to represent the distribution of water according to the resources in the world presented with the infographic. More specifically, the students were expected to choose a graph to illustrate the data and to provide justification for choosing that graph.

212 (84.2%) students selected bar graph and pie chart, which were statistically appropriate ways for representing categorical data. Besides, these students explained the reason for their selection as ‘*With the help of the pie chart, we can better observe the distribution of percentages.*’ or ‘*We observe the rates better with the pie chart.*’

The rest of the students who gave an incorrect answer chose line graph to represent the distribution of water. Probably, these students were not able to differentiate continuous data from categorical data. This might stem from the fact that the distinction between continuous data and categorical data is not emphasized adequately in textbooks. Students encounter various forms of data representations in textbooks; however, they cannot select the most appropriate and practical ways to illustrate the given data. Moreover, even if they choose the right way to represent data, they cannot explain the reason behind choosing that representation.

In the 2nd problem of the SLT, the problem was asked directly by giving the percentage of A, B, and C parts of a whole by not presenting problem in context. 221 (88.4%) students gave statistically appropriate responses. Similar to the correct responses for the context-bounded form of the problem, these students successfully chose bar graph and pie chart, which are statistically appropriate ways to illustrate categorical data. Likewise, the incorrect responses included the construction of line

graph, which is inappropriate to represent categorical data. As could be seen, the 8th grade students' achievement in context-bounded form and context-free form of the problem was very close. This means that the use of an infographic representing water resources on earth did not influence students' graph selection to illustrate data. The use of context could have helped students to comprehend the problem and the data given in the infographic. However, it seems that the elimination of context and asking the problem directly might have simplified the problem situation in terms of understanding and representing the given data in the 2nd problem.

In the 3rd problem, the students were expected to construct the graph that they selected in 2nd problem. 133 (53.2%) students constructed the selected graph successfully in the context-bounded form of the 3rd problem, whereas that percentage was 190 (76%) for context-free form. As could be seen, the students had better achievement in the 3rd problem without context than the 3rd problem with context. In the context-bounded form of the problem, the students might not have built a connection between the problem situations in the 2nd and 3rd problems. Some of the students could not read the data on the infographic, which illustrates the amount of water according to the resources in the world. More precisely, the students might not have made sense of the percentages of water resources written in the infographic. However, in the context-free form, the problem was asked directly to draw an appropriate graph with the given parts of the whole. Therefore, the context-free form of the 3rd problem might have been considered to be simpler than the context-bounded form by the 8th grade students.

In the tests, problem 4, problem 5, problem 6 and problem 7 were connected with a line graph and a table that were presented at the beginning of these problems. In the SLT-CP, before asking the problems, a written part was included to attract students' attention to the problems. The written part mentioned the importance of water for living creatures. Then, the students' attention was attracted to the unconscious

attitudes of human-beings that cause water pollution and thus climate changes posing a threat to the water we need. To support the written part of the problem, real-world data showing statistics about the changes in the first priority environmental problems such as air pollution, water pollution, soil pollution, waste, and noise pollution was presented. The statistics related to these problems was represented in a line graph and a table. The students were expected to determine the accuracy of the statements in problems 4-5-6-7 and to provide statistical reasoning for their decision about these statements by reading the data from the line graph and the table. To obtain problems 4-5-6-7 in the SLT, the written part involving the necessity of water and the environmental issues related to water was removed from the problem. A line graph and a table including the change in the values of the variables A, B, C, D and E were presented instead of the variables of first priority environmental problems. Additionally, the years in the SLT-CP were replaced with intervals in the SLT. Then, the statements were given to the students to determine the accuracy of these statements and they were asked to provide statistical reasoning for their decision. The students' achievement in problem 4, problem 5, problem 6 and problem 7 was discussed as follows.

The statement included in the 4th problem of the SLT-CP was *"In 2014, the number of provinces where water pollution was the first priority environmental problem was the highest."* This statement was converted into *"The variable B has the highest value in the 8th interval"* to obtain the 4th problem in the SLT. As could be seen, the variable water pollution was replaced with variable B, and the year of 2014 was replaced with the 8th interval. When the students' responses were examined, it was seen that problem 4 was the least correctly answered graph interpretation problem in the SLT-CP and SLT. Only 85 (34%) students provided a statistically appropriate response to the 4th problem in both the SLT-CP and SLT. Probably, the integration of a written part mentioning the necessity of water for living creatures and the use of real-world data regarding the changes in the first priority environmental problems

did not influence the students' performance in problem 4 positively or negatively. Besides, the majority of the students who gave incorrect responses investigated water pollution or variable B horizontally. To be more precise, these students examined the values of water pollution or variable B by years. However, they should have examined the values of the variables in the year 2014. This means that the students confused the axes in the table and line graph. Instead of looking at the values of the variables in 2014 and specifying the variable having the highest value, they investigated the values of variables by years and determined the year when the variable got the highest value.

The statement in the 5th problem of the SLT-CP was as follows: "*Noise pollution has steadily increased over the years.*" The students were expected to observe the progression of the number of provinces having noise pollution as the first priority environmental problem by interpreting the line graph or reading data from the table.

162 (64.8%) students noticed the mistake in the statement and marked the statement as False by providing statistical justifications from the line graph or the table because the number of provinces having noise pollution as the first priority environmental problem was 0 at the intervals 2002-2004 and 2007-2008. The students who incorrectly determined the accuracy of the statement thought that the blank cells in the table indicated the stability in the number of provinces. However, it can be easily observed from the line graph that the blank cells indicated the number of provinces having noise pollution as 0 rather than no change in the number of provinces.

In the 5th problem in the SLT, the statement was converted into "*The variable E has steadily increased over the intervals*" by eliminating the issues about changes in the first priority environmental problems. 174 (69.6%) students marked the statement as false by providing a statistical explanation from the line graph or the table. As can be seen, the 8th grade students' achievement did not differ according to the use of statistics about the changes in first priority environmental problems such as air pollution, water pollution, soil pollution, waste and noise pollution that they could

encounter in the news, journals or textbooks. Although the use of statistics in the context of environmental problems might have helped the students to interpret the problem situation and decide on the accuracy of the statements easily, these statistics might have confused students by providing additional information which was not required to decide on the accuracy of the statements.

The statement in problem 6 in the SLT-CP is as follows: “*Air pollution and water pollution generally have an inversely correlated relationship*” The students were expected to examine the relationship between the number of provinces with air pollution and water pollution as the first priority environmental problem. 110 (44%) of the students examined the relationship between the number of provinces with air pollution and water pollution as the first priority environmental problem successfully and decided on the correctness of the statements. In general, when the number of provinces with air pollution as the first priority environmental problem increased, the number of provinces with water pollution as the first priority environmental problem decreased. Similarly, when the number of provinces with water pollution as the first priority environmental problem increased, the number of provinces with air pollution as the first priority environmental problem decreased. It means that these variables are inversely related.

For the 6th problem in the SLT, the statement was converted into “*The variables A and B have an inversely correlated relationship*”. 106 (42.4%) students determined the accuracy of the statement correctly. When variable A increased, variable B decreased or vice versa. As could be seen, similar to problem 5, the use of statistics about the changes in first priority environmental problems such as air pollution, water pollution, soil pollution, waste and noise pollution did not make a difference in 8th grade students’ achievement in the 6th problem significantly.

The statement in the 7th problem in the SLT-CP was as follows: “*The number of provinces where air pollution is the first priority environmental problem has*

increased continuously”. The students were expected to evaluate the accuracy of the statement by examining the line graph and the table.

157 (62.8%) students determined that the statement was incorrect, which is the statistically correct answer. These students noticed that the number of provinces where air pollution is the first priority environmental problem fluctuated. In other words, the number of provinces increased, decreased or remained constant over the years.

In problem 7 in the SLT, the statement was changed into “*The variable A has increased continuously*” by not using the number of provinces with first priority environmental problems. 177 (70.8%) students noticed the incorrectness of the statement because of the fluctuations in the values of variable A. As could be seen, the 8th grade students showed better achievement in the 7th problem in the SLT than in the SLT-CP. This might be because the use of statistics related to air pollution and other environmental problems might have complicated the students’ statistical decision making while examining the change in the number of provinces where air pollution is the first priority environmental problem. In this respect, the 7th problem without context might have been comprehended more easily by the students than the 7th problem with context.

In problem 15 in the SLT-CP, a pie chart which illustrates the sources of water pollution in a country was given. According to the chart, the source of 26% of water pollution was agricultural, 39% was industrial, 44% was domestic and 13% was other. The students were asked to interpret the pie chart and express their inferences related to the distribution of water pollution sources. Then, the students were asked whether there is an unusual situation about the pie chart. In the chart, the sum of the percentages of the sources of water pollution exceeds 100%.

121 (48.4%) students interpreted the pie chart and explained their inferences related to the distribution of the sources of water pollution in a country. Some of those

students noticed the unusual situation about the pie chart, and they interpreted the pie chart referring to agricultural, industrial, domestic and other sources of water pollution. However, some of these students could not notice that the sum of the percentages of the sources exceeded 100%. Additionally, some students noticed an unusual situation about the percentages on the pie chart; however, they could not interpret the percentages in the graph. In that case, the pie chart representing the percentages of the parts of a whole was given, and the students were expected to explain and interpret the percentages on the pie chart. To obtain the 15th problem in the SLT, the sources of water pollution in a country were removed from the problem. More precisely, the parts of a whole were presented with a pie chart, and the students were requested to interpret the pie chart and detect the error regarding the pie chart.

131 (52.4%) students gave the correct answer to problem 15 in the SLT. The achievement students showed in problem 15 in the SLT-CP and SLT was close. The students produced similar solutions in the SLT-CP and SLT. When the student answers were analyzed, it was observed that the students confused the percentages of parts with the central angles of these parts. More specifically, the students asserted that the whole was required to be 360% rather than 100%. This might have stemmed from students' confusion about the different expressions of the whole as in textbooks, the whole in pie charts is represented with 360° and the parts in the pie chart are expressed as parts of 360° instead of 100%. Accordingly, the students are not accustomed to percentages on pie charts in textbooks.

Problem 16 was the most correctly answered graph interpretation problem in both tests. The 16th problem in the SLT-CP started with a UNICEF report titled "The Water is Under Fire" to raise awareness about drought. In this report, some disturbing facts about drought were revealed with statistics and photos to attract students' attention in the problem. Then, the statistics about the number of water-scarce countries from the "Drought Assessment Report" was represented with a pie chart. The problem was asked in multiple-choice format. The students were

requested to choose the appropriate bar graph which gives the same information with the pie chart. Option A was the most appropriate choice.

230 (92%) students selected the most appropriate bar graph to represent the statistics about the number of water-scarce countries. On the other hand, in the SLT, instead of using statistics from the “Drought Assessment Report”, the parts of a whole were presented in a pie chart and the students were asked to select the most appropriate bar graph to represent the given statistics. 242 (96.8%) students selected the most appropriate bar graph correctly. The students’ achievement was high in both forms of the problem. It can be inferred that the 16th problem in the SLT-CP and SLT was understandable, the bar graphs in the options were clear, and the sentences were explicit. Additionally, the students showed slightly better performance in the context-free form of the 16th problem. This difference could be due to the fact that the context might have complicated the problem situation for the students as it involves extra information which might have distracted students’ attention away from the aim of the problem situation.

Problem 18 was asked to the students in multiple-choice format. In the SLT-CP, a table was presented at the beginning of the problem. The table showed the 9-day temperature for a city in two different intervals. In the problem, the students were requested to choose the best option to represent the data in the table with a line graph. 204 (81.6%) students chose the most appropriate line graph correctly.

Problem 18 in the SLT included the change in the values of variables A and B for 9-day intervals instead of the 9-day long temperature for a city in two different intervals and the students were asked to select the best option to represent the data in the table with a line graph. 194 (77.6%) students selected the correct line graph representing the data for variables A and B in the table. As can be inferred, the students’ achievement was close to each other in context-bounded form and context-free form of the 18th problem. The students showed slightly better achievement in the context-bounded form of the problem. It could be inferred that the use of realistic contexts

such as temperature in a time interval might have helped students to be involved in the problem situation and thus to make statistically appropriate decisions.

Problem 20 and Problem 21 were in multiple-choice format, and they were asked consecutively with the same context in the tests. In the SLT-CP, these problems were asked according to a bar graph presented at the beginning. In the problem, the decrease in water quality was shown as the factor leading to an increase in water shortage. The problem drew attention to the increase in water pollution on earth day by day. According to the problem, a group of researchers measured the water pollution values in different cities. The scale used by the researchers ordered the water pollution values from 1 to 5. The data was obtained as a result of the measurements illustrated with a bar graph representing the cities' water pollution values. In problem 20 and problem 21, the students were expected to read and interpret the bar graph showing water pollution values in different cities. To obtain the 20th and 21st problems in the SLT, the data were presented with a bar graph directly as a result of the measurement of the levels of a group and the students were requested to read and interpret the results of the measurements.

In problem 20 in the SLT-CP, the students were expected to find the number of cities whose water pollution value is above 3. 151 (60.4%) students specified the number of cities with water pollution values of 4 and 5 and added the number of these cities to reach the correct answer. When the students' incorrect responses were examined, it was seen that the students included the cities with water pollution level of 3 in their calculations although the problem asked about the cities whose water pollution levels were above 3. Therefore, these students reached an incorrect answer by adding the cities whose water pollution levels were 3, 4 and 5. Some of the incorrect responses indicated that some students considered only the cities with water pollution level of 3. Thus, they marked an incorrect option.

In problem 20 in the SLT, 112 (44.8%) students specified the number of people whose levels were above 3. These students specified the number people with level 4

and level 5, and then added the number of people corresponding to these levels. The students' correct solutions in problem 20 were the same in the SLT-CP and SLT. It can be inferred from the number of students who provided a correct response for problem 20 that the 8th grade students showed better achievement in problem 20 in the SLT-CP compared to the 20th problem in the SLT. This could be due to the fact that the students might have built a connection between real-life situations and the problem situation with the help of real-life context embedded in the problem. The authenticity of the task might have affected the students' understanding and motivation positively. More specifically, the students might have established empathy with the group of researchers introduced in the problem, and they might have been involved in the statistics related to water pollution levels of the cities. In the problem, the use of classroom context might engage students in the problem situation. Therefore, the students might visualize the problem better in context-bounded form of the 20th problem.

Problem 21 was the follow-up problem of the problem 20 with the same content. In the problem, the students were asked to find the total number of cities in which the researchers measured water pollution. 200 (80%) students found the total number of cities correctly in problem 21 in the SLT-CP, while 210 (84%) students reached the correct answer in problem 21 in the SLT. The students' achievement in problem 21 was very close to each other in both tests. The integration of authentic task in the problem seems not to have created a huge difference in students' achievement. Additionally, in both forms of the problem, most of the incorrect responses involved the addition of water pollution levels rather than the number of cities. To be more precise, the students were confused about the values on the x-axis (water pollution levels) and y-axes (number of cities) written on the bar graph although the axes were labeled with the related variables.

4.1.3 Achievement in Sampling

In the tests, P10 and P22 were sampling problems. The distribution of students' correct answers and the percentage of achievement for each problem in both tests are presented in Table 4.3.

Table 4.3 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Sampling Domain

	Sampling	
	P10	P22
Problems with Context	164 (65.6%)	136 (54.4%)
Problems without Context	180 (72%)	152 (60.8%)

As can be seen in Table 4.3, the 8th grade students' achievement in sampling problems with context varied between 54.4% and 65.6%, whereas their achievement in sampling problems without context ranged from 60.8% to 72%. The students' achievement differed significantly according to the existence of context in problems. More precisely, the 8th grade students showed better achievement in sampling problems without context.

In problem 10 in the SLT-CP, a researcher wants to choose a random province to examine water pollution in a country and the factors that cause water pollution in that province. Accordingly, the students were asked to help the researcher choose one of the options to make a random selection. Three different statements were given for the researcher's sampling process of the province to examine water pollution. The first and second statements did not indicate random selection because they were

an example of purposive sampling. The third statement was an example of random selection. The students were asked which statement(s) was appropriate for the random selection of province.

164 (65.6%) students selected the correct option, which is putting the license plate numbers of all the coastal provinces in a bag and choosing randomly from the bag. In the SLT, the statement(s) were asked directly by not involving the researcher's study. 180 (72%) students specified the statement representing random selection correctly. The number of students who provided a correct response in the context-free form of problem 10 was nearly twice as the number of students who provided a correct response in the context-bounded form of the problem. This means that the students determined random sampling in the 10th problem more easily without context than the 10th problem with context. This may be because the context might have complicated the students' statistical decision-making process by including the problem context in addition to statistical thinking required for the problem. More specifically, the involvement of the researcher's study and the narrative part related to water pollution in the problem might have distracted the students' attention from the investigation of the statements indicating random selection.

In problem 22 in the SLT-CP, the students were given an excerpt from a newspaper titled "Children Are Not Environmentally Friendly". This excerpt presented the results of a study carried out with the 5th grade students in Ankara. In the excerpt, a part was included to explain the details of the study. According to the excerpt, the study results revealed that the students do not make efforts to protect the nature, waste a huge amount of water in the shower, and do not try to save electricity. The problem examined 8th grade students' ideas about the generalizability of the results of that study for all the children in Turkey.

136 (54.4%) students made statistically meaningful interpretations about the generalizability of the results of the study conducted at a school in Ankara to all the

children in Turkey. The students stated that the sample size of the school in Ankara is insufficient to generalize the study's results to all the school children in Turkey. For instance, some students stated that *"It could not be interpreted only by looking at a group of students"* or *"The size of the sample is insufficient to generalize a sample to a population"*. To obtain problem 22 in the SLT, the context of the problem in the SLT-CP was reduced. The results of a study were presented and the generalizability of the results of that study was investigated in the SLT.

152 (60.8%) students made statistically appropriate interpretations about the generalization of the results of the study conducted with a group of students in Ankara to all the students in Turkey. They indicated the insufficiency of the sample size to generalize the population. As could be seen, similar to the other sampling problem (problem 10), the students showed better achievement in the context-free form of the 22nd problem compared to the context bounded form. This may be attributed to the fact that the problem context might have made the problem confusing as it requires understanding the problem context in addition to the statistical thinking and decision-making process. Not to use extra information in asking the problem might have facilitated the understanding of the problem and thus the frequency of providing correct response.

4.1.4 Achievement in Questioning of Statistical Data

In the tests, P1, P14, P17, P19, P23, P24 and P25 belonged to the questioning of statistical data domain. The distribution of students' correct answers and the percentage of achievement for each questioning of statistical data problem in both tests are presented in Table 4.4.

Table 4.4 Distribution of Participants' Correct Answers in Problems with and without Context regarding the Questioning of Statistical Data Domain

	Questioning of Statistical Data						
	P1	P14	P17	P19	P23	P24	P25
Problems with Context	139 55.6%	128 51.2%	144 57.6%	81 32.4%	39 15.6%	94 37.6%	60 24%
Problems without Context	139 55.6%	126 50.4%	145 58%	142 56.8%	36 14.4%	61 24.4%	43 17.2%

Table 4.4 shows that the 8th grade students' achievement in the questioning of statistical data problems with context was between 15.6% and 57.6%. The most correctly answered problem was the 1st problem, and the least correctly answered problem was the 23rd problem. Whereas, achievement in questioning of statistical data problems without context ranged from 14.4% to 58%.

Problem 1 was a warm-up problem for the context and the nature of the SLT-CP. At the beginning of the problem, the marine characters Sponge Bob and Patrick have introduced themselves, and they specified the aim of the context of the SLT-CP. Then, the Blue Marble photograph was introduced to attract students' attention to

the amount and the necessity of the water around the world. In the problem, the students were expected to make reasonable estimations about the type and the approximate amount of surface water sources on earth. The 1st problem aroused students' interest in the context of the test. To obtain the 1st problem in the SLT, the context was removed from the 1st problem in the SLT-CP, and the students were directly requested to make statistical estimation about the number of living species around the world by not involving a narrative part at the beginning of the problem.

In the 1st problem in the SLT-CP, 139 (55.6%) students made logical estimations about the amount of water and the number of living species around the world. The students sampled the water sources on the surface and made statistically reasonable predictions about the amount of water sources on earth. Even if some of these responses were illogical, the students explained their estimation process and their ways of thinking logically.

Similarly, in the 1st problem of the SLT, 139 (55.6 %) students sampled the living species and estimated the number of living species around the world appropriately. As could be seen, the students showed the same achievement in the 1st problem in both tests. It can be deduced that the use of context did not create a difference in the 8th grade students' achievement in the 1st problem.

Some students might have perceived the part including information about the process of taking the Blue Marble photograph as an unnecessary detail. For this reason, their attention might have been distracted. However, asking the problem directly as in the context-free form might be perceived by some students as meaningless. Since the problem was the first problem of the test, the students might not have made sense of the aim of the problem. Thus, the use of context which attracts students' attention to solve problems could be useful for the 1st problem of the test.

In the 14th problem of the SLT-CP, a table prepared by the Chamber of Environmental Engineers was presented to the students. The table illustrates the amount of rainfall in different regions of Turkey for the year 2016. Additionally, the table shows some statistics regarding the amount of rainfall, expected amount of rainfall and the comparison between the amount of rainfall in 2016 and in 2015. The rate of change in the amount of rainfall was also shown in the table. In the problem, the students were expected to find out which region or regions have less rainfall than normal (expected) rainfall in 2016. They were also expected to explain the reason behind their decision by providing statistics from the table.

128 (51.2%) students answered the question correctly by stating that the regions with less rainfall than normal in 2016 were Marmara, Ege, Akdeniz, and Güneydoğu Anadolu regions. The students used different statistics from the table to reach the correct answer. Initially, some of the students compared the amount of rainfall in 2016 with the normal amount of rainfall. Afterwards, they selected the regions with less amount of rainfall than normal. Moreover, some of the students utilized from the statistics regarding the ratio of change in the amount of rainfall. Those students specified the regions which experienced a decrease in the amount of rainfall compared to the normal rainfall. Therefore, they appropriately selected the regions with less rainfall than normal (expected) rainfall in 2016.

For the 14th problem in the SLT, the statistics related to the amount of rainfall in 2016 was eliminated from the problem. The numerical values in the data remained the same. The numerical values were presented in the same way by indicating that the statistics was the result of eight measurements. The regions were not mentioned. Then, the students were asked which measurement or measurements indicate less value than the expected value. 126 (50.4%) students correctly specified the measurements having less value than the expected value as the 2nd, 3rd, 4th and 8th measurements. Similar to the 14th problem in the SLT-CP, the students specified the

measurements having less value than the expected value by using different statistics from the table. Some of the students compared the values at the ‘result of the measurement’ column and the ‘expected value’ column and then they determined the result of measurements that are less than the expected value. Likewise, some students examined the values in the column ‘the result of measurement concerning expected value’. Then, the students determined the measurements with a negative sign in that column because the results of the measurements were less than the expected value in these measurements. As seen, the students’ achievement in the 14th problem in the SLT-CP and SLT was about the same. Although the students’ achievement did not differ greatly based on the use of a table obtained from the media report, the students often encounter that kind of a table in problem 14 in popular media and they should be able to interpret and evaluate that kind of media excerpts as a part of the data-driven society as the media reports might be full of biased, misleading and deceptive information. In this respect, problem 14 was significant in terms of making inferences about the students’ evaluation of the media reports that they face in daily life by presenting a table from media and requesting students to interpret and evaluate information embedded in the table.

In the 17th problem of the SLT-CP, the students were presented information about global warming which is one of the causes of drought. To observe the effect of global warming on a city’s temperature, a table shows the 9-day temperature for a city in two different intervals. The students were asked to decide in which interval the city is warmer. Additionally, the students were expected to provide justification for their answers and explain how they think while answering the problem.

144 (57.6%) students determined the interval at which the city was warmer in the 9-day period by utilizing the statistical concepts. The students mostly utilized the measures of central tendency. Initially, the students calculated *arithmetic mean* for both intervals and they specified the interval having the higher arithmetic mean during the 9-day time period. In addition, some students benefitted from *the sum* of

the 9-day period temperature for both intervals and selected the second interval as the interval when the city was warmer. Furthermore, some of the students investigated the day by day *difference* in the air temperature. Then, they added up the differences and they decided that during the second interval the city was warmer.

In the 17th problem of the SLT, the problem was asked directly by presenting the same statistics with the 17th problem in the SLT-CP. In this test, the statistics was presented as the values of the variables A and B in 9 different intervals and the students were asked which variable has the higher value in nine different intervals. 145 (58%) students found that variable B had higher values than variable A in nine different intervals. Similar to the 17th problem in the SLT-CP, the students mostly benefitted from the concepts of *arithmetic mean*, *sum* and *difference* while investigating the variable with higher values at 9-day intervals. As can be seen, the students showed the same achievement in problem 17 with and without context. It might be inferred that the involvement of students in the problem by presenting information about global warming, which is one of the causes of drought, did not affect the students' achievement in the 17th problem positively or negatively. The information regarding global warming might have helped some students construe the problem situation in problem 17 with context and thus focus on the statistical methods for determining the warmer interval. Besides, not to use of a context in the problem in the SLT might have caused that students not to comprehend and make sense of the problem situation and therefore the number of students who provided a correct response in problem 17 in the SLT fell behind the number of students who provided a correct response in the 17th problem of the SLT-CP.

Problem 19 in the SLT-CP includes 'water shortage' values between the years 2010-2040 for several countries as reported by the World Resource Institute. According to the report, Botswana, Chile, Estonia, and Namibia have been suffering from water shortage. Water shortage values for the years 2010, 2020, 2030 and 2040 were

presented. As stated in the problem, a researcher wants to make an investigation in the country where water shortage is experienced most. It is asked which country the researcher should choose to make the right decision. The students were expected to compare the water shortage values assigned by the World Resource Institute and specify the country where the water shortage is experienced most. 81 (32.4%) students determined the country where water shortage is suffered most by benefitting from the statistical methods. Some students utilized the sum of the water shortage values by adding the water shortage values in the years 2010, 2020, 2030 and 2040. Then, they determined the country with the highest water shortage value as Chile correctly. Moreover, some of the students calculated the arithmetic mean for the water shortage value of each country between 2010-2040 and selected Chile as the country suffering from water scarcity most.

The remaining 169 (67.6%) students could not correctly determine the country suffering from water shortage most. Some of these students calculated the range of the water shortage values of the countries between the years 2010 to 2040 and they selected the country with the highest range. However, the problem did not question the change or difference in the values of water shortage. Additionally, some of the students selected the country with the peak water shortage value. Indeed, an overall evaluation of water shortage in the countries was questioned instead a snapshot in a random time interval. Furthermore, some of the students examined the water shortage values for only the year of 2040 by considering that year as the most up-to-date measurement for water shortage and selected Chile. Even though the selection was the correct answer, the students' method for selection was statistically inappropriate for the solution of the problem. To obtain problem 19 in the SLT, water shortage values for some countries in the years 2010, 2020, 2030 and 2040 were eliminated from the 19th problem in the SLT-CP by keeping the statistics the same. Instead of providing water shortage values of four countries, the data was presented in four rows and the students were requested to specify the row having the highest

value. Besides, the students were required to explain the reason behind their answer. 142 (56.8%) students successfully determined the row having the highest value as the 2nd row. The students' appropriate and inappropriate solution strategies were the same with the 19th problem in the SLT-CP. As could be seen, the students showed better achievement in problem 19 in the SLT when compared to problem 19 in the SLT-CP. This could be originating from the fact that the 8th grade students might have been unwilling to engage with the context including water shortage, the name of various countries and statistics related to these constructs. The name of the countries might have sounded unfamiliar for the students and this situation might have demoralized the students to engage in the problem. Hence, the 19th problem without context might have been conceived as simple, clear and more understandable by the students.

Problems 23, 24, and 25 in the SLT-CP were asked consecutively based on the same context. At the beginning, a questionnaire conducted by Selim teacher was introduced. Selim teacher aims to raise his students' awareness about the conscious use of energy sources. The teacher asked his students a research question, which was obtained from Turkish Statistical Institute (TUIK). The research question was that "What do you think is the area with the highest electricity consumption in our country?" Then, the students chose the best option corresponding to the research question. The options were home, commerce, official departments, industry, lighting, agricultural irrigation and other. Subsequently, the teacher illustrated the data he collected in a table. Then, the statements related to the data on the table were presented in problems 23, 24 and 25. The students were expected to evaluate and determine the accuracy of the given statements critically. To obtain the 23rd, 24th and 25th problems in the SLT, the classroom situation in which the data was collected for a research question was eliminated from the problem. The same data was presented with the same table. However, this time, the table showed the number of recurrence

of the variables A, B, C, E, F and G. The students were requested to determine the accuracy of the given statements.

The statement included in problem 23 of the SLT-CP was “*The mode of the data collected by Selim teacher is the industry*”. 39 (15.6%) students critically evaluated the data and decided on the accuracy of the statements. These students made a distinction between the mode and the peak value. In the statements, the area having the peak value was given as the mode of the data. However, the mode was 1 as the most recurrent value in the data set. On the other hand, these students were able to notice that the mode should be a numerical value rather than a categorical value. They indicated that the numerical value of mode is 1 as a number of students selected the options of lighting and other in question presented by the teacher. The rest of the students could not decide whether the data’s mode is numerical or categorical, or they specified only the accuracy of the statement by not providing a statistical justification. The 23rd problem in the SLT was converted into the statement “The mode of the data is A”.

36 (14.4%) students critically evaluated and appropriately decided on the accuracy of the statements by indicating the mode of the data as the numerical value of 1. Similar to the inappropriate responses for the 23rd problem in the SLT-CP, the rest of the students could not decide whether the mode of the data is numerical or categorical, or they specified only the accuracy of the statement by not providing a statistical justification.

The statement included in the 24th problem of the SLT-CP was “*The median of the data collected by Selim teacher is industry*”. 94 (37.6%) students thought that to obtain the median of the data set, the data should be sorted in ascending or descending order. Therefore, these students ordered the data in ascending or descending order and selected the number of the students in the middle of the data.

Additionally, these students noticed that the median should be numerical rather than categorical. Therefore, they indicated that the median of the data collected by Selim teacher should be 3. Similarly, to obtain problem 24 in the SLT, the context of the research study in Selim teacher's class was eliminated from the problem and the statement was provided directly as "*The median of the data set is E.*" Similarly, 61 (24.4%) students noticed that the median should be numerical rather than categorical and the data should be ordered in an increasing or decreasing order. Hence, the students ordered the data in an increasing or decreasing order and they selected the middle value to obtain the median of the given data set.

The statement included in the 25th problem of the SLT-CP was "*The range of the data collected by Selim teacher is from 1 to 7*". 60 (24%) students critically evaluated the statement by analyzing the data obtained by Selim teacher and noticed the fallacy in the statement. These students indicated that the range should be " $7-1=6$ " and made interpretations based on this idea. To obtain the 25th problem in the SLT, the context was eliminated from the problem and the statement was provided directly as "*The range of the data set is 1 to 7*"

In problem 25 of the SLT, 43 (17.2%) students correctly determined the incorrectness of the statement by stating that the range should be a numerical value; namely, the range of the data set is " $7-1=6$ ".

As could be seen, the students showed better achievement in the 23rd, 24th and 25th problems in the SLT-CP compared to the same problems in the SLT.

4.2 Statistical Literacy in Problems with and without Context in terms of Statistical Content Domains

Another aim of the study was to investigate the 8th grade students' statistical literacy in problems with and without context. The students' responses to each of the problems in the SLT-CP and SLT were examined and categorized as statistical, pre-statistical, non-statistical, and unrelated. To examine students' literacy, the students' responses to open-ended problems were examined because open ended tasks encourage students to explain the depth of their understanding of the statistics problems (Gal; 1995; Watson, 2013). Thus, the students' responses to the open-ended items included in measures of central tendency problems, graph interpretation problems, sampling problems and questioning of statistical data problems were thoroughly examined to make inferences about students' statistical literacy in problems with and without context.

4.2.1 Statistical Literacy in Measures of Central Tendency

The SLT-CP and SLT consist of 25 problems each and the 8th, 11th, 12th and 13th problems were open-ended measures of central tendency problems. The 8th grade students' statistical literacy in these problems was examined separately on problem basis.

4.2.1.1 The 8th Grade Students Statistical Literacy in Problem 8

Problem 8 was a measures of central tendency problem, which investigated the students' understanding of the concept of "average". In the 8th problem of the SLT-CP, the students were given a newspaper excerpt including some statistics about the amount of garbage. The excerpt was as follows: *"In Turkey, an average of 70 thousand tons of garbage thrown daily, and 20 percent of the waste is composed of*

packaging waste such as plastic bottles, cardboard, and cans.” Then, the students were asked to interpret the term “average” in the given newspaper excerpt.

To obtain the 8th problem of the SLT, the problem context was removed from the problem in the SLT-CP. The students’ responses were analyzed and classified by the researcher. The classification of students’ responses to problem 8 in the SLT-CP and SLT is presented below:

Table 4.5 Student Responses for Problem 8

Classification of responses	Student Response	Problem 8 with context		Problem 8 without context	
		f	p	f	p
Statistical	The explanation of central tendency in the data set by using the average concept. <i>- representative value of data set</i>	14	5.6%	15	6%
Pre-statistical	The explanation of average with three measures of central tendency <i>-mean,</i> <i>-mode,</i> <i>-median</i>	65	26%	85	34%
Non-statistical	The use of non-statistical explanations and terms to explain average. <i>-about</i> <i>-more or less</i> <i>-approximately</i>	151	60.4%	113	45.2%
Unrelated/ Blank		20	8%	37	14.8%

As can be inferred from Table 4.5, in problem 8 of the SLT-CP, 14 (5.6%) students successfully described the central tendency in the data set by using the average

concept within the problem context. Participant 120's response belonging to that strategy is explained in Figure 4.1 below:

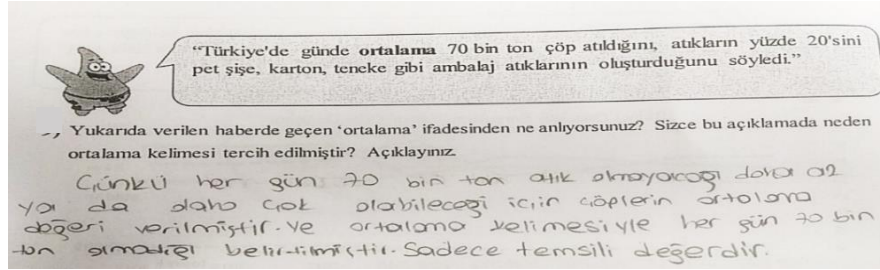


Figure 4.1. The statistical response of participant 120 to P8 in the SLT-CP

As can be seen, Participant 120 stated as follows:

‘because the amount of garbage could be more or less than 70 thousand tons, the average amount of garbage was given. By using average term, it was indicated that the average amount of garbage was not 70 thousand tons every day, it was only representative value.’

Therefore, the student mentioned average's representativeness of the sample. Additionally, it was observed that some of the students using this strategy responded to problems 8 as *‘the average is the representative value of the amount of waste disposed by people’* or *‘the amount of waste disposed in Turkey on any day’*.

For example, Figure 4.2 shows the understanding of Participant 141 of the statistical term ‘average’ within the context of the garbage thrown in Turkey.

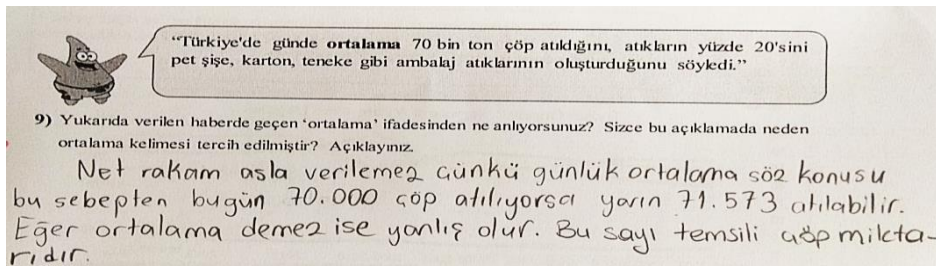


Figure 4.2. The statistical response of participant 141 to P8 in the SLT-CP

As seen, Participant 141 responded as *'an exact amount could not have been provided since an average amount is in question. Therefore, the amount of garbage could be 70.000 tons today, but it could be 71.573 tons for tomorrow. If the writer does not use the term of average, this will be wrong. The given value is representative amount.'*

The findings of the present study revealed that 15 (6%) students used the same strategy as participant 141 for problem 8 of the SLT. Although the problem did not contain context, the number of students who preferred that strategy was almost the same in the context-bounded form of the problem. The students described the central tendency in the data set by using the average concept. As seen in Figure 4.3, Participant 340 mentioned average's representativeness of a group by indicating as *'I think, the term of average represents the overall of the group'*.

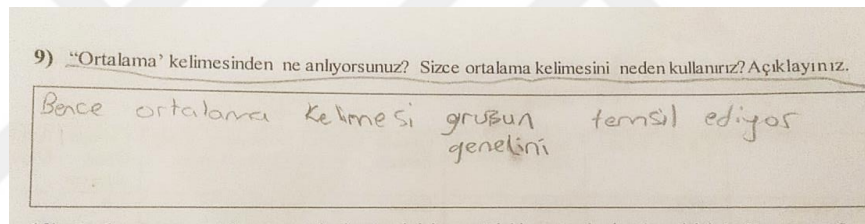


Figure 4.3. The statistical response of participant 340 to P8 in the SLT

For problem 8 in the SLT-CP, 65 (26%) students provided explanations with three measures of central tendency (mode, median or arithmetic mean) to interpret average. Some of the students who used this strategy interpreted the average amount of garbage as the mode of the amount of garbage by writing *'the most repetitive amount of garbage thrown away in Turkey'*. Some of the students interpreted the average amount of garbage as median of the amount of garbage by writing *'the amount of garbage thrown is the number in the middle'*. Additionally, some students interpreted the average amount of garbage as arithmetic mean of the amount of the garbage by writing *'add up and divide all numbers'* and *'add up garbage amounts*

and divide by number'. Similarly, for problem 8 in the SLT, the students provided similar responses to problem 8 in the SLT-CP. For instance, the students wrote:

“Mean is the most recurrent value in a data set, and it could be found by adding up all values and dividing the sum by the number of values in a given data set or middle number of an ascending data set.”

The only difference in the content of the responses is that the students did not use context to interpret the concept of average. At this point, it was observed that some students tried to derive a context to provide an answer to problem 8 in the SLT. This might be due to the fact that the existence of a problem context might help students to build a relationship between the problem situation and real life. Therefore, providing an answer could be easier with the help of problem context.

The results of the study revealed that 151 (60.4%) students provided non-statistical explanations for problem 8 in the SLT-CP. This means that more than half of the students used non-statistical terms to interpret the concept of average in the problem. To be more precise, these students' responses mostly included expressions such as *about, more or less, approximately, rounded form of any amount of tone approximate to 70 thousand tones, more or less than 70 thousand tons*. A sample student response (participant 127) is presented below in Figure 4.4.

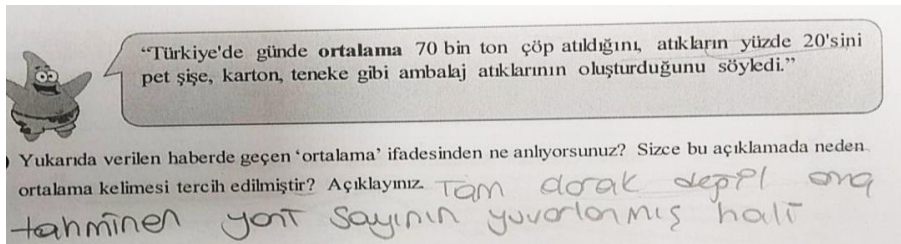


Figure 4.4. The non-statistical response of participant 127 to P8 in the SLT-CP

As can be seen in Figure 4.4, participant 127 stated as *'not exactly, but approximately therefore, the rounded form of the number.'* by using the terms of *'approximately', 'rounded form of a number', 'not exactly'* to express average. Likewise, 113 (45.2%)

students used non-statistical terms to explain and interpret average in problem 8 in the SLT. Similar to the SLT-CP, the students used the terms ‘about’, ‘approximately’ and ‘more or less’ mostly to interpret the concept of average. This solution approach was preferred more in the SLT-CP compared with the SLT.

20 (8%) of 250 students provided unrelated or no solution for problem 8 in the SLT-CP. In these solutions, the students did not give an answer to the problem or the students demonstrated no statistical understanding about problem 8 with context. Most of the students provided idiosyncratic responses far from statistical literacy and the context. One of the student responses belonging to this solution category is explained below:

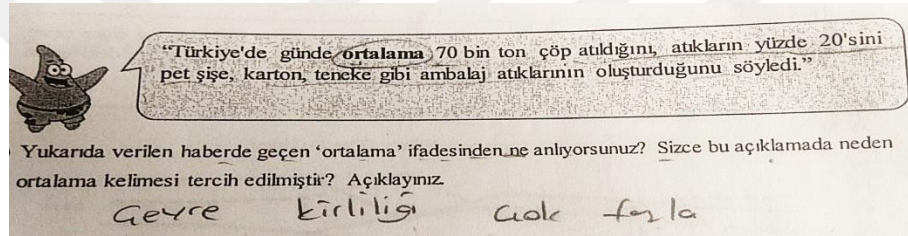


Figure 4.5. The unrelated response of participant 65 to P8 in the SLT-CP

As can be seen in the figure, participant 65 gave an irrelevant response to problem 8 in the SLT-CP by writing that '*environmental pollution is too much*'. The response of participant 65 has no indication of literacy about the concept of average in a social context. For problem 8 in the SLT, 37 (14.8%) students used that strategy by providing answers away from the concept of average. These students provided unrelated personal responses to problem 8. The number of students who provided unrelated or no response to problem 8 was higher in the SLT than in the SLT-CP. It could be inferred that with the help of context, the students comprehended the problem situation better, and thus, the number of students who provided a related response was higher in the SLT-CP.

4.2.1.2 The 8th Grade Students Statistical Literacy in Problem 11

Problem 11 was a “measures of central tendency” problem which investigated the students’ understanding of the “mode” of the data set. In problem 11 of the SLT-CP, the students were asked to calculate and interpret the mode of the data presented in a bar graph. The problem was about water pollution and it was presented in a bar graph showing the water pollution incidents that occurred in different districts of Turkey. The students were expected to read data in the bar graph and find the mode of the given data set. To obtain problem 11 in the SLT, the context of the problem in the SLT-CP was removed from the problem. The 8th grade students’ responses to problem 11 were analyzed and classified by the researcher. The classification of students’ responses in the SLT-CP and SLT is presented in Table 4.6.

Table 4.6 Students’ Responses to Problem 11

Classification of responses	Student Response	Problem 11 with context		Problem 11 without context	
		f	p	f	p
Statistical	Explaining and finding the numerical value of the mode of the data set by using statistical terms.	58	23.2%	93	37.2%
Pre-statistical	-Finding the numerical value of mode, but statistically insufficient explanations. -Statistically appropriate explanations related to mode, but deficiencies in calculating the numerical value.	6	2.4%	3	37.2%
Non-statistical	Non-statistical / no explanation	24	9.6%	12	4.8%
Unrelated/ Blank		162	64.8%	142	56.8%

As seen in Table 4.6, the majority of the students could not provide a correct solution for problem 11 in the SLT and SLT-CP. The students might not have remembered the concept of mode from the 7th grade because of the lapse of time. Additionally, it was seen in student solutions that the students confused the concept of mode with median. More precisely, some of the students found and interpreted the mode of the number of water pollution incidents that occurred in different districts of Turkey. The same situation was valid for problem 11 in the SLT.

The results of the current study showed that 58 (23.2%) students explained the mode of the number of water pollution incidents that occurred in different districts of Turkey and found the numerical value of the mode successfully.

For instance, participant 71 demonstrated a comprehensive understanding of the mode of the number of water pollution incidents that occurred in different districts of Turkey by stating as '*1, the most recurrent value which belongs to Aydın, Bursa, Edirne Kocaeli and Sinop.*' The answer of participant 71 is presented in Figure 4.6 below.

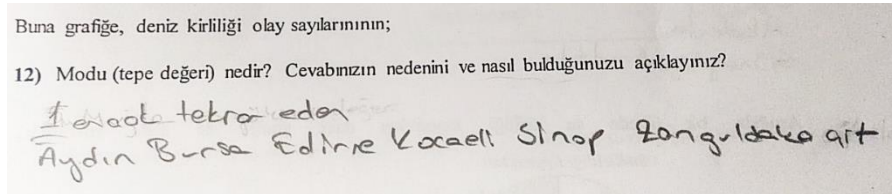


Figure 4.6. The statistical response of Participant 71 to P11 in the SLT-CP

In addition to carefully reading the statistics in the bar graph and finding the mode of the data set, the participant indicated the cities which correspond to the numeric value of mode.

Similarly, 93 (37.2%) students successfully explained the mode and found the numerical value of the mode of the given data set in problem 11 of the SLT. These

students explained the way through which they found the mode and the meaning of the mode in the data set. A sample response from participant 319 providing a statistical response to problem 11 in the SLT is presented below:

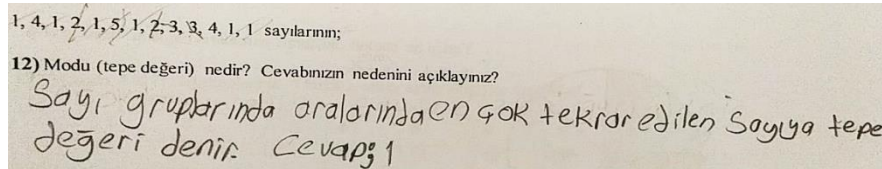


Figure 4.7 The statistical response of Participant 319 to P11 in the SLT

Participant 319 stated as '*mode is the most recurrent value among the number set. Answer is 1.* As can be understood from participant 319's response, the student defined the mode as the most frequent number of the given data set and he selected value "1" which is the most recurring number in the given data set.

According to Table 4.6, 6 (2.4%) students found the numerical value of mode; however, they provided statistically insufficient explanations or conversely they gave statistically appropriate explanations related to mode, but there were trivial deficiencies in calculating numerical value in students' responses. Additionally, some students' responses were free from the context of data representing number of water pollution incidents that occurred in different districts of Turkey. A sample response from participant 117 is presented below:

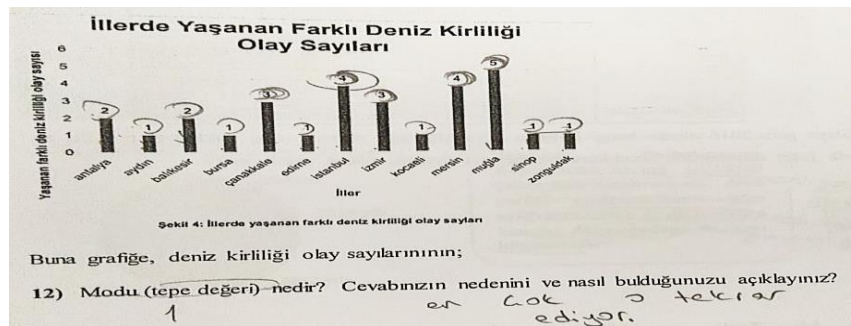


Figure 4.8. The pre-statistical response of Participant 117 to P11 in SLT-CP

As seen participant 117 indicated as '*1 since it recurs the most*'. The student found the numerical value of the data set successfully by providing the reasoning behind finding the mode of the data set. However, the student did not build a connection between the number of water pollution incidents that occurred in different districts of Turkey as the students did not mention the different districts of Turkey presented in the bar graph. Although the students' response related to the mode of the data set was statistically correct, it would have been better if the student had involved the context of the water pollution incidents. The students' response could have been more meaningful by involving the variables in the problem. On the other hand, 3 (1.2%) students found the numerical value of mode; however, they did not provide statistically sufficient explanations, or they gave statistically appropriate explanations related to mode, but there were trivial deficiencies in calculating the numerical value in students' responses. The students' way of thinking for that type of response was the same with the students' response for problem 11 in the SLT-CP.

According to Table 4.6, 24 (9.6%) students gave non-statistical responses to problem 11 in the SLT-CP. These students' responses did not include statistical explanations about the mode of number of water pollution incidents that occurred in different districts of Turkey. Some of these students wrote only the numerical value of the mode by not providing an explanation or justification for their answer. More specifically, these students wrote "*the mode is 1*" or "1" as a response to problem 11. As could be seen, the responses did not involve the context of water pollution incidents.

12 (4.8%) students provided non-statistical responses to problem 11 in the SLT. Similar to students' response to the 11th problem in the SLT-CP, the students' responses included only the numerical value of the mode without a statistical justification. As could be inferred from the result, the number of students who provided a non-statistical response to problem 11 was twice as many as those in the SLT-CP. The existence of the context of water pollution incidents might have

directed the students to provide non-statistical explanations without the context since, in addition to statistical explanation, the students were required to engage in the problem context. That might have made the provision of a statistical response to problem 11 more difficult.

162 (64.8%) students provided statistically inappropriate, unrelated or blank response to problem 11 in the SLT-CP. When these students' incorrect responses to SLT-CP were examined, it was found that the underlying reasons behind the responses were explicit. Even, some students made more than one mistake while answering the problem.

Initially, 105 (42%) students confused the mode of the number of water pollution incidents for different districts of Turkey with the district having the highest number of incidents. To be more precise, these students confused the *mode* of the data set with the *peak value* of data set including the number of incidents in different districts of Turkey. This error could be exemplified with Participant 234's response as follows:

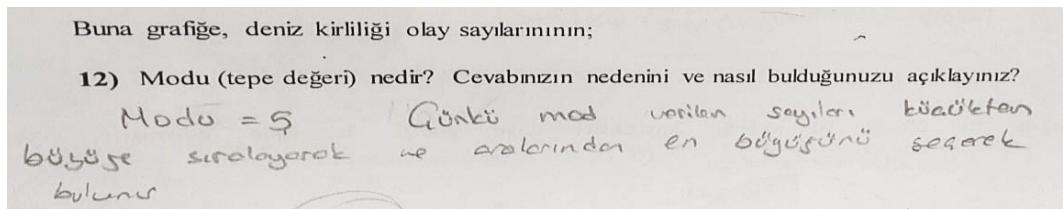


Figure 4.9. The unrelated response of participant 234 to P11 in the SLT

As could be seen, participant 234 wrote as '*mode is 5 since mode could be found by ordering the given numbers from smallest to biggest and selecting the biggest number among these numbers.*' The student specified the *peak value* for the number of incidents in different districts of Turkey as the mode of data set. This confusion might have stemmed from the synonym of mode in Turkish language. The term mode

is called as the peak value in Turkish mathematics textbooks. Therefore, the students might have considered the mode of the data set as the peak value.

9 (3.6%) students wrote the data recurring several times once while ordering the number of water pollution incidents for different districts of Turkey. For instance, the students wrote the data as “1,2,3,4,5” and specified 5 as the mode of the data set. The students did not notice that some values in data set recurred more than once.

43 (17.2%) students were confused whether the mode of the data indicating the number of incidents in different districts of Turkey is numerical data or categorical data. To be more precise, these students wrote the names of the districts with the most frequent water pollution incidents. After ordering the data represented in the bar graph as 1,1,1,1,1,1,2,2,3,3,4,4,5; the students specified 1 as the most recurrent value and selected the district names with the value 1. A sample response for that strategy from participant 187 is presented below:

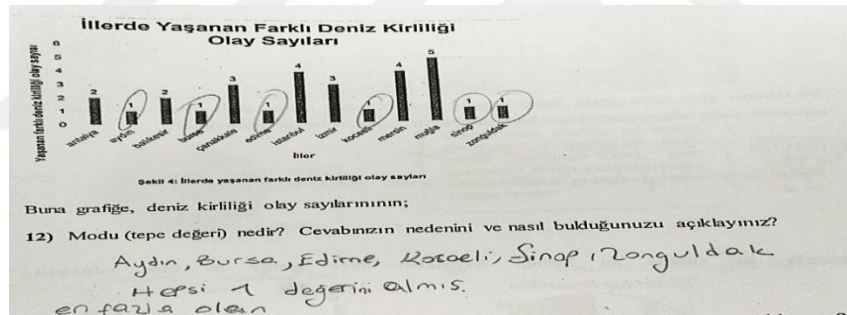


Figure 4.10. The unrelated response of participant 187 to P11 in the SLT

As could be seen, participant 187 stated as ‘All the districts of Aydın, Bursa, Edirne, Koala, Sinop, Zonguldak got the value of 1 which is the most.’ Therefore, the student specified the names of the districts with the most frequent number of water pollution incidents.

142 (56.8%) students provided statistically inappropriate, unrelated or blank responses to problem 11 in the SLT. The students’ errors and the underlying reasons

for these errors were similar with the 11th problem in the SLT-CP. These students wrote the given data once or selected the peak value of the data set instead of selecting the mode as the most frequent value.

4.2.1.3 The 8th Grade Students Statistical Literacy in Problem 12

Problem 12 was a measures of central tendency problem which investigates students' understanding of the median value of the data set representing the number of water pollution incidents that occurred in different districts of Turkey. In problem 12 in the SLT-CP, the students were asked to calculate and interpret the median of the data presented in the bar graph. The problem was asked by presenting a bar graph regarding water pollution incidents that occurred in different districts of Turkey. The students were expected to read the data in the bar graph and find the median of the given data set. To obtain the 12th problem in the SLT, the context of the 12th problem in the SLT-CP was removed from the problem.

The 8th grade students' responses were analyzed by the researcher and the students' responses for problem 12 were classified. The classification of students' responses for the SLT-CP and SLT are presented in Table 4.7.

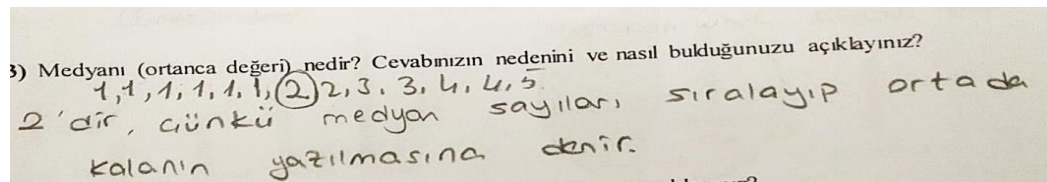
Table 4.7 Student Responses to Problem 12

Classification of Responses	Student Response	Problem 12 with context		Problem 12 without context	
		f	p	f	p
Statistical	Explaining and finding numerical value of median of the data set by using statistical terms.	38	15.2%	43	17.2%

Table 4.7 (continued)

Pre-statistical	-Finding numerical value of median, but statistically insufficient explanations. -Statistically appropriate explanations related to median, but deficiencies in calculating numerical value.	9	3.6%	5	2%
Non-statistical	Non-statistical / no explanation -only the numeric value of median	15	6%	15	6%
Unrelated/Blank		188	75.2%	187	74.8%

The results of the current study revealed that 38 (15.2%) students explained and found the numerical value of median of the data set representing the number of water pollution incidents in different districts of Turkey by using statistically appropriate justifications. These students wrote the number of water pollution incidents in an increasing or decreasing order and selected the value in the middle of the data set as median. A sample response is presented below:

**Figure 4.11.** The statistical response of participant 233 to P12 in the SLT-CP

As seen, participant 233 stated that '2 because median could be found by ordering the numbers and selecting the number in the middle'. The student put the data regarding the number of water pollution incidents in the bar graph in order and chose the value in the middle of the data set.

Additionally, as shown in Participant 234's response in Figure 4.12, some students eliminated the values one by one from the beginning to the end. Participant 234 responded as *'I wrote the given numbers consequently. Then, I started to eliminate numbers one by one from right and left. Lastly, I found the average of the remaining two numbers.'*

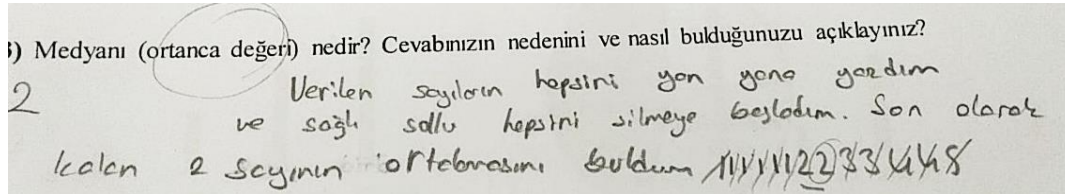


Figure 4.12. The statistical response of participant 234 to P12 in the SLT-CP

On the other hand, in problem 12 in the SLT, 43 (17.2%) students explained and found the numerical value of the median of the given data set by using similar strategies used in problem 12 in the SLT-CP.

9 (3.6%) students found the numerical value of median in problem 12 in the SLT-CP; however, they provided statistically insufficient explanations regarding the median of the number of water pollution incidents in different districts of Turkey. More precisely, although they found the median of the data regarding water pollution incidents as 2, these students did not explain their solution strategy adequately. Moreover, 5 (2%) students provided the correct value of the median of the given data set as 2. However, these students could not elaborate on their thinking and solution strategy sufficiently.

15 (6%) students provided non-statistical explanations or no explanations regarding the median of the number of water pollution incidents in Turkey for the 12th problem in the SLT-CP. These students provided only the numerical value of the median, which is insufficient to understand the students' solution strategy and statistical literacy about the median concept. In the same way, 15 (6%) students gave answers without a statistical explanation. Some of these students did not provide an

explanation regarding their answers by writing the numerical value of the median of the given data set. Unfortunately, the students' literacy about the median of a data set could not be observed with that kind of responses.

188 (75.2%) students gave unrelated responses to problem 12 in the SLT-CP or they left the problem unanswered. When the students' solutions were examined, several distinct errors were noticed by the researcher. Initially, 17 (6.8%) students did not order the data regarding the number of water pollution incidents in Turkey in an ascending or descending order to find the median of the data. They transferred the data with the same order in the bar graph. A sample response from participant 75 is presented below:

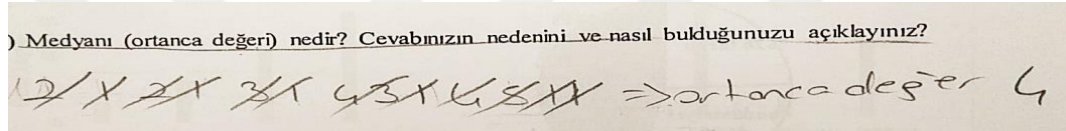


Figure 4.13. The unrelated response of participant 75 to P12 in the SLT-CP

As seen, the student transferred the number of incidents regarding water pollution in the same order as the bar graph by not ordering the data in an increasing or decreasing order.

89 (35.6%) students wrote the number of incidents of water pollution once. To be more precise, these students wrote the numbers as 1, 2, 3, 4, 5 and selected the median of the data set as 3. However, some of the numbers were repeated more than once in the data set. A sample response from participant 49 regarding that strategy is given below:

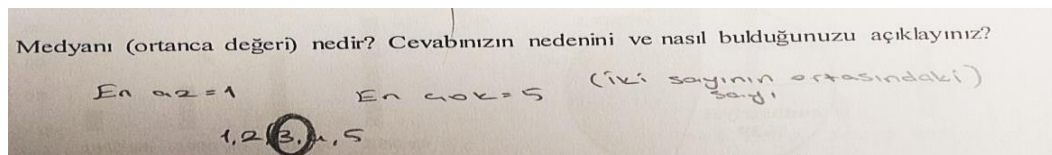


Figure 4.14. The unrelated response of participant 49 to P12 in SLT-CP

As seen, participant 49 wrote as '*1,2,3,4,5. Median is the middle of the two numbers. The least is 1 and the most is 5*'. Therefore, the student did not write the number of incidents having more than one frequency.

Moreover, 43 (17.2%) of students wrote the names of the district instead of writing the number of incidents regarding water pollution. However, the median should be a numerical value. Probably, these students are not able to distinguish the numerical data and categorical data. A sample response regarding that solution from participant 47 is presented as follows:

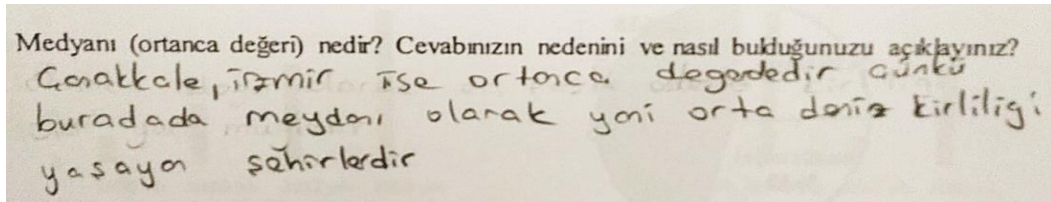


Figure 4.15. The unrelated response of participant 47 to P12 in the SLT-CP

As seen, participant 47 wrote as '*Çanakkale and İzmir is median value. Since, these districts experience median that is middle water pollution.*' The student wrote the names of the cities instead of using numerical values. The student did not notice that the median should be numerical data instead of categorical data. Additionally, the participant wrote the number of incidents once similar to participant 49.

On the other hand, 187 (74.8%) students provided unrelated responses to problem 12 in the SLT or they left the problem unanswered. Except for confusing the numerical data and categorical data, the students' responses were similar under that category of responses since the problem was presented by not using categorical data in the SLT. This might have prevented students from getting confused about numerical and categorical data.

4.2.1.4 The 8th Grade Students' Statistical Literacy in Problem 13

Problem 13 was a “measures of central tendency” problem which investigated students' understanding of the arithmetic mean of the data set representing the number of pollution incidents in different districts of Turkey. In problem 13 of the SLT-CP, the students were asked to calculate and interpret the arithmetic mean for the data presented in the bar graph. The problem was asked in the context of water pollution by presenting a bar graph regarding water pollution incidents that occurred in different districts of Turkey. The students were expected to read the data in the bar graph and find the arithmetic mean of the given data set. To obtain the 13th problem in the SLT, the context in problem 13 in the SLT-CP was removed from the problem.

The students' responses were analyzed and classified by the researcher. The classification of students' responses in the SLT-CP and SLT is presented in Table 4.8.

Table 4.8 Student Responses to Problem 13

		Problem 13 with context		Problem 13 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Explaining and finding numerical value of arithmetic mean for the data set by using statistical terms.	106	42.4%	73	29%
Pre-statistical	-Finding numerical value of arithmetic mean, but statistically insufficient explanations. -Statistically appropriate explanations related to median, but deficiencies in calculating numerical value.	41	16.4%	44	17.6%

Table 4.8 (continued)

Non-statistical	Non-statistical / no explanation -only the numeric value of arithmetic mean	45	18%	31	12.4%
Unrelated/Blank		58	23.2%	102	40.8%

106 (42.4%) students found and interpreted the arithmetic mean of the number of water pollution incidents in different districts of Turkey with statistically appropriate justifications. These students added up the number of water pollution incidents and obtained the sum of the incidents. Then, they divided the sum by the number of districts given in the bar graph and obtained the arithmetic mean of the number of water pollution incidents in Turkey. A sample response from participant 237 is presented below:

15) Aritmetik ortalaması nedir? Cevabınızın nedenini ve nasıl bulduğunuzu açıklayınız?

$$2+2+1+1+1+1+1+1+3+3+4+4+5 = 29$$

$$\begin{array}{r} 29 \\ 13 \overline{) 29} \\ \underline{26} \\ 30 \\ \underline{26} \\ 40 \end{array}$$

2,23...
ben bütün sayıları topladım
ve çıkan sonucu 13 ile sayısına böldüm.

Figure 4.16. The statistical response of participant 237 to P13 in the SLT-CP

These students applied the statistical method to obtain the numerical value of the arithmetic mean correctly. In addition, the students explained the solution process by making connection with the context clearly. As could be seen, the arithmetic mean of the number of water pollution incidents was not a whole number. It was a decimal number. The students hesitated about the value of the arithmetic mean because they thought that the arithmetic mean of the number of water pollution incidents must be a whole number. Some of the students wrote the arithmetic mean as 2, while some of them wrote the arithmetic mean as 2.23. Both two responses were regarded as

correct response to equalize the student responses with 13th problem in the SLT. At this point, the researcher aimed to encourage students to consider the form of the value of the arithmetic mean depending on the context it is used.

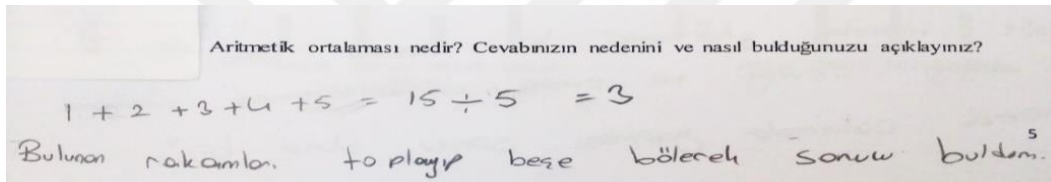
On the other hand, 73 (29%) students found and interpreted the arithmetic mean of the given data set for 13th problem of SLT. The students found the total of the values in the data set and they divided the total by the number of values to obtain the arithmetic mean for the data set. Additionally, these students explained the process of obtaining arithmetic mean in statistically appropriate sentences.

41 (16.4%) students found the numerical value of arithmetic mean; however, they made statistically insufficient explanations regarding their solution in the 13th problem of the SLT-CP. These students wrote what they did; however, they did not interpret the arithmetic mean in the problem context. More specifically, these students did not interpret the arithmetic mean of the number of water pollution incidents. The students should have explained what the sum of these values represent, how arithmetic mean of 2,23 was obtained as the arithmetic mean of the number of water pollution incidents. Moreover, 44 (17.6%) students found the numerical value of arithmetic mean, but they could not provide enough statistical explanation regarding the mean of the given data set. When these students' responses were examined, the sense of arithmetic mean was observed. However, the students were not able to interpret the arithmetic mean and explain the solution path while finding the arithmetic mean for the given data set.

45 (18%) students provided non-statistical answers or they only wrote the value of the arithmetic mean regarding the number of water pollution incidents for problem 13 in the SLT-CP. For instance, these students wrote "2" as the answer. However, to make inferences about the students' literacy about the arithmetic mean, the students should explain their statistical decision making process and arithmetic operations to obtain values of statistical constructs. Therefore, these answers were categorized as non-statistical explanation as there was no indication of statistical literacy. Similarly,

31 (12.4%) students gave a non-statistical response to problem 13 in the SLT. These students provided the numerical value of the mean for the given data set. They did not provide a statistical justification as an indicator of statistical literacy. For this reason, these students' responses were categorized as non-statistical response.

58 (23.2%) students did not answer or gave an unrelated answer to problem 13 in the SLT-CP. When the students' answers were analyzed, it was observed that some students wrote the values in the given data set once. More specifically, some students wrote the number of water pollution incidents as 1,2,3,4 and 5 although the numbers were repeated more than once in the data set. Then, they added the numbers up and divide them by 5 to obtain the arithmetic mean of the data set. A sample response from participant 208 is presented below:



Aritmetik ortalaması nedir? Cevabınızın nedenini ve nasıl bulduğunuzu açıklayınız?

$$1 + 2 + 3 + 4 + 5 = 15 \div 5 = 3$$

Bulunan rakamları toplayıp beşe bölerek sonucu buldum.

Figure 4.17. The unrelated response of participant 208 to P13 in the SLT-CP

As can be seen, participant 208 wrote as *'I added the numbers and divided them into five to find the result'*. The student wrote the values in the data set once. However, the students should have added up all number of water pollution incidents in different districts of Turkey to obtain the total of the number of incidents and divide the total by the number of districts presented in the bar graph.

Moreover, 102 (40.8%) students did not provide an answer or gave unrelated answers to problem 13 in the SLT. Similar to the 13th problem in the SLT-CP, the most frequent error made by the students was that the students wrote the values in the data set only once, even if the values were repeated more than once.

4.2.2 Statistical Literacy in Graph Interpretation

The SLT-CP and SLT consists 25 problems and problems 2, 3, and 15 are open-ended graph interpretation problems. These problems involve choosing the most appropriate graph for the data set, representing the given data with the chosen graph, interpreting different types of graphs, making inferences from the graph and conversion between different types of graphs. The students' statistical literacy in these problems was examined separately on problem basis.

4.2.2.1 The 8th Grade Students' Statistical Literacy in Problem 2

In the 2nd problem of the SLT-CP, an infographic titled 'Water on Earth' was given to the students. The graph illustrates authentic data retrieved from TUBITAK. The students were asked to select the most appropriate graph to represent the distribution of water according to the resources in the world presented in the infographic. More specifically, the students were expected to choose a graph to illustrate the data and provide a justification for choosing that graph. To obtain the 2nd problem in the SLT, the context of the 2nd problem in the SLT-CP was eliminated from the problem, and the most appropriate graph to represent the distribution of the A, B and C parts of a whole was asked.

The students' responses were analyzed and classified by the researcher. The classification of students' responses for the SLT-CP and SLT is presented in Table 4.9.

Table 4.9 Student Responses to Problem 2

		Problem 2 with context		Problem 2 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Statistically appropriate chart selection and statistical explanations. -With the help of the pie chart, we can better observe the distribution of percentages. -We observe the rates better with the pie chart.	103	41.2%	96	38.4%
Pre-statistical	Statistically appropriate but personal, idiosyncratic justifications -.....because it's easier to understand - just the right graphic name - because everything is more clear, understandable	109	43.6%	125	50%
Non-statistical	Statistically inappropriate but reasoned chart selection -line graph selection, but correctly explained	6	2.4%	5	2%
Unrelated/Blank		32	12.8%	24	9.6%

The results of the current study revealed that 103 (41.2%) students made statistically appropriate chart selection by providing statistical justification for choosing that chart in the 2nd problem of the SLT-CP. To provide a statistical answer to the problem, the students could read the authentic data given in the infographic. By doing this, the students could decide whether “the amount of water” is continuous or

categorical, which affects the students' choice of the representation of the amount of water on earth according to water resources. In that case, if the data is continuous, the selection of line graph could be the most appropriate decision. Otherwise, if the amount of water on earth presented in the infographic is categorical, selecting the pie chart or the bar graph could be suitable to represent the distribution of amount of water on earth. A vast majority of the students who provided a statistical response selected the pie chart to represent the amount of water on earth which was presented in the infographic titled "Water on Earth". These students provided responses such as "With the help of the pie chart, we can better observe the distribution of percentages in the infographic" or "We observe the rate of the amount of water better with the pie chart".

A sample student response including statistics is presented as follows:

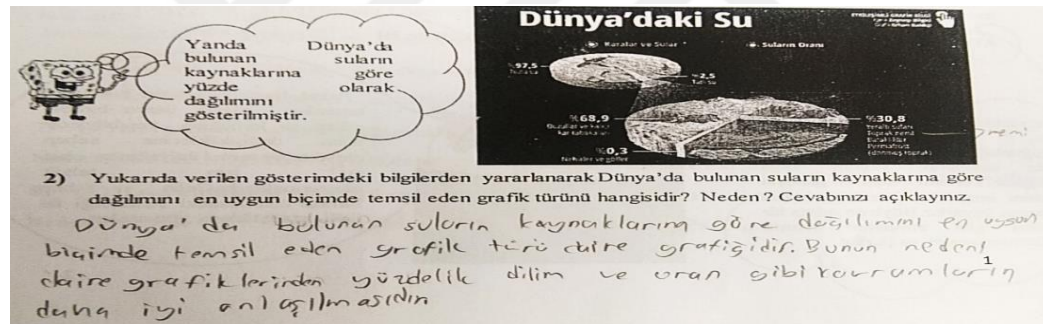


Figure 4.18. The statistical response of Participant 54 to P2 in the SLT-CP

As seen, participant 54 stated as 'Pie chart is the best representative of the distribution of water on earth according to sources. The reason is that the percentages and rates could be more understandable with pie chart.' The student preferred the pie chart to represent the distribution of water on earth according to the water sources given in the infographic. The student provided a statistical justification for choosing the pie chart to represent the data. The student indicated that we can better observe the percentages and rates with a pie chart.

Similarly, participant 107 stated that the better way is the “pie chart because we can see the rate of the amount of water easily.”

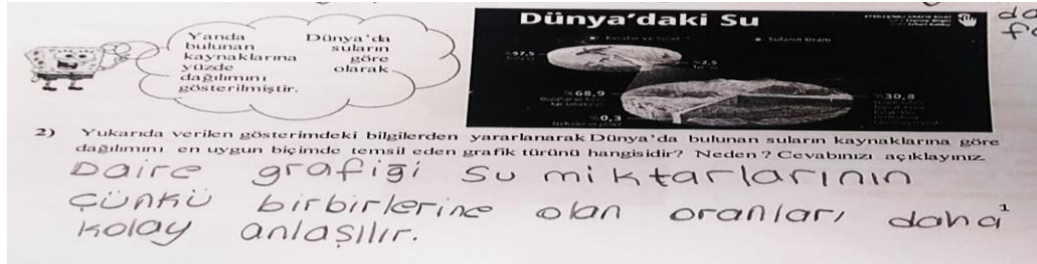


Figure 4.19. The statistical response of Participant 107 to P2 in the SLT-CP

As seen, participant 107 believed that the use of the pie chart helps to see the rate of the water amount according to different resources of water presented in the infographic.

The rest of the students who provided a statistical response selected the bar graph as the most suitable graph to illustrate the distribution of water on earth according to water sources. These students provided statistical justifications for the selection of bar graph to represent the distribution of water on earth according to water sources.

A sample response from participant 76 is presented below:

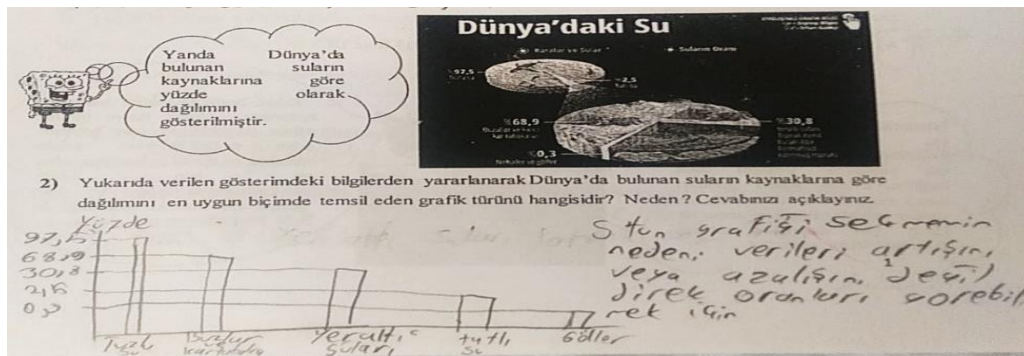


Figure 4.20. The statistical response of Participant 76 to P2 in the SLT-CP

Participant 76 selected the bar graph by indicating that:

“The reason for the selection of bar graph is that instead of observing the increase or decrease in data, we can observe the rates of the data directly”

In fact, participant 76 was aware of the discontinuity of the data presented in the infographic by stating the increase and decrease in data. Therefore, participant 76 provided a statistical response for problem 2 in the SLT-CP.

96 (38.4%) students provided a statistical response for the 2nd problem of the SLT. These students selected bar graph and pie chart to represent the distribution of A, B and C parts of a whole. They justified their selection with statistical expressions. As seen in participant 480's response in Figure 4.21, some of the students selected the pie chart by stating that *“Pie chart is the most appropriate way to represent the distribution of parts of the whole.”*

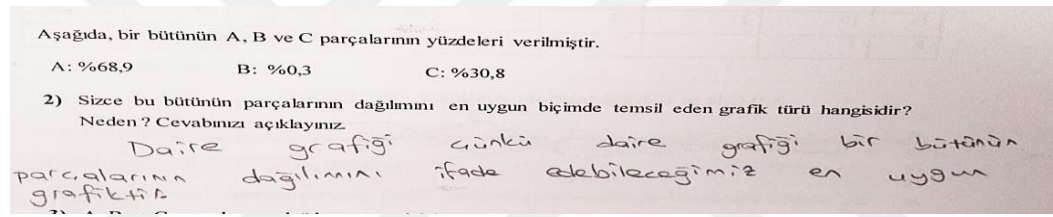


Figure 4.21. The statistical response of Participant 480 to problem 2 in the SLT

Additionally, as seen in participant 236's response in Figure 4.22, some students argued that *“We could use the pie chart because we can better express the percentages of the parts.”*

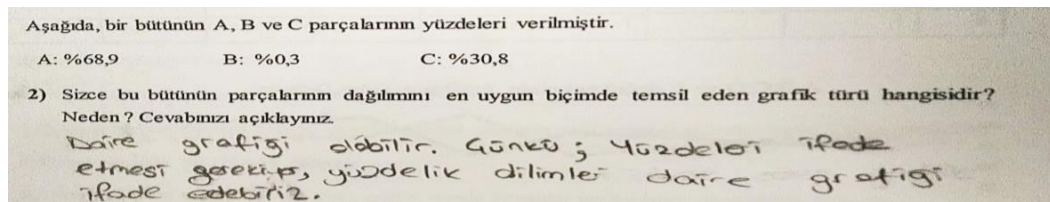


Figure 4.22. The statistical response of Participant 236 to P2 in the SLT

109 (43.6%) students selected the statistically appropriate graph to represent the data regarding the amount of water on earth according to water resources; however, they

justified their selection with personal or idiosyncratic statements such as “...because it's easier to understand with a pie chart” and “...because everything is more clear, understandable with a bar graph”

For instance, the students selected the bar graph or the pie chart to represent the distribution of water according to the resources on earth. They justified their answer by stating that the best way is the “bar chart because it's easier to understand” or “because everything is more clear, understandable with a pie chart”

A sample response from participant 121 is presented below:

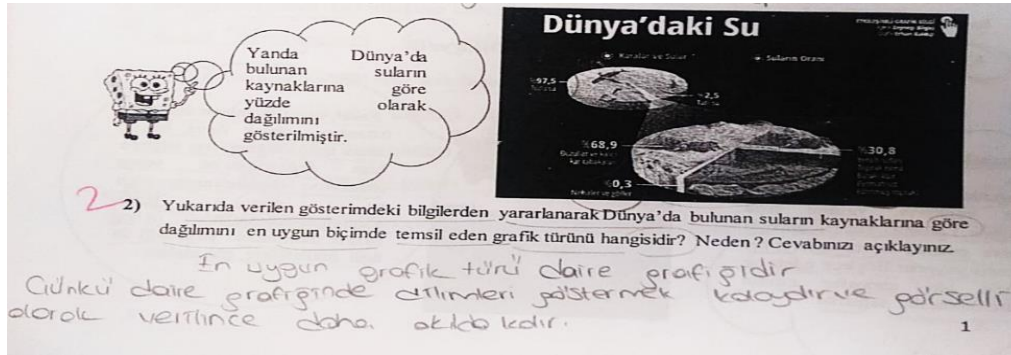


Figure 4.23. The pre-statistical response of Participant 121 to P2 in SLT-CP

Participant 121 stated that:

“Pie chart is the most appropriate way to represent data because it is easy to illustrate the segments of the pie and the pie chart presents data in a visual form which makes the data easy to remember.”

As could be seen in participant 121's response, the student selected the appropriate chart type; however, the student could not provide a statistical justification for selecting the pie chart. Instead, the student provided idiosyncratic reasoning about the selection of the pie chart to represent the data regarding the amount of water according to resources.

6 (2.4%) students chose the statistically inappropriate graph to represent the categorical data regarding the amount of water on earth according to resources;

however, they provided a correct explanation regarding the selected graph in problem 2 of the SLT-CP. To be more precise, these students selected the line graph, which is inadequate to represent categorical data. However, these students explained the use of the line graph based on the usage of line graph in statistics. At this point, these students' responses were categorized as non-statistical response. A sample response from participant 92 is as follows:

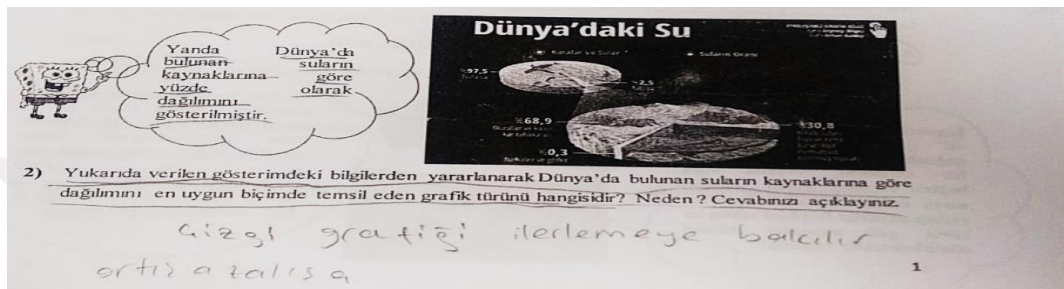


Figure 4.24. The non-statistical response of Participant 92 to P2 in the SLT-CP

As seen, the student chose the line graph which is inadequate to represent the categorical data regarding the amount of water on earth according to resources; however, the student was aware of the relationship between the continuity of the data and the use of the line graph. This could be inferred from the student's response stating that *"With the line graph, the progress in increase and decrease could be observed."*

Similar to the 2nd problem of the SLT-CP, 5 (2%) students selected the statistically inappropriate graph to represent the categorical data regarding the A, B and C parts of a whole; however, they provided a correct explanation regarding the selected graph in the 2nd problem of the SLT. Although the students selected a line graph which is inappropriate for representing categorical data, they made explanations consistent with the properties of the line graph. These students' responses were categorized as a non-statistical response.

58 (23.2%) students provided an unrelated response or left the problem unanswered in the SLT-CP. These students selected the line graph instead of the bar graph or the pie chart to represent the data regarding the amount of water on earth according to water resources, which is categorical data. Similarly, 102 (40.8%) students gave an unrelated response or did not answer the 2nd problem in the SLT.

4.2.2.2 The 8th Grade Students Statistical Literacy in Problem 3

Problem 3 was a follow-up problem of problem 2. The students were expected to construct the graph that they specified in the 2nd problem. More specifically, in the 2nd problem the students were presented an infographic which represents the distribution of the amount of the water according to surface water sources and the students were expected to choose a graph to illustrate the data regarding the amount of water sources. In the 3rd problem of the SLT-CP, the students were expected to make the most appropriate drawing with the graph which they selected in the 2nd problem. To obtain the 3rd problem in the SLT, the infographic which presents the authentic data regarding the amount of the water according to surface water sources was removed from the problem. Then, the A, B and C parts of a whole was given and the students were expected to construct the graph which they selected as the most appropriate graph.

The students' responses were analyzed and classified by the researcher. The classification of students' responses for the SLT-CP and SLT is presented in Table 4.10.

Table 4.10 Student Responses to Problem 3

		Problem 3 with context		Problem 3 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Statistically appropriate construction to represent the given data. - <i>Pie chart</i> - <i>Bar graph</i>	97	38.8%	78	31.2%
Pre-statistical	Statistically appropriate selections and drawings with small deficiencies. - <i>indicating variables, but not indicating percentages on graph</i> - <i>indicating percentages, but not indicating variables on graph</i>	36	14.4%	112	44.8%
Non-statistical	Personal drawing independent from the given data - <i>Selected chart type that is not statistically appropriate but correct drawing</i> - <i>only drawing of graph without including the variables and the percentages</i> - <i>the drawn graph is not 360 'or 100%, it exceeds 100%</i>	70	28%	31	12.4%
Unrelated/Blank		47	18.8%	29	11.6%

97 (38.8%) students made statistically appropriate graph selection and they constructed the graphs representing the given data. More specifically, the students selected the bar graph or pie chart to represent the distribution of the amount of the water according to surface water sources presented in the infographic in problem 2.

The majority of the students constructed a pie chart to represent the data presented in the infographic. This might be due to the fact that the illustration in the infographic was presented in the form of a circle. The students might have been affected by that presentation. It could be inferred that the use of authentic tasks including the media excerpts from real life might affect students' statistical literacy. As a result, most of the students selected pie chart to illustrate the data regarding the distribution of the amount of the water according to surface water sources. A sample response from participant 89 is presented below:

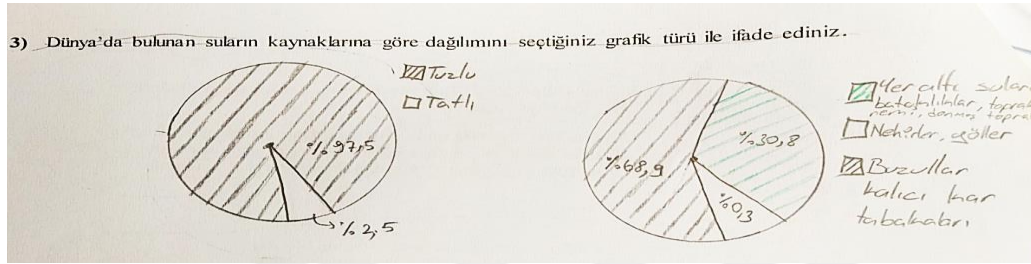


Figure 4.25. The statistical response of Participant 89 to P3 in the SLT-CP

As could be seen in the participant 89's response, the student drew the pie chart successfully by indicating the water resources in terms of salinity (fresh water or salty water) and the sources of water (ground water, underground water and glacier). Additionally, the student labeled all the variables belonging to the pie chart and specified the percentages on the chart. The distribution of the rates of the percentages was correct. For instance, the segment belonging to glacier (68.9%) was almost as twice as the segment for underground water (30.8%). Therefore, participant 89 provided a statistical response for the 3rd problem in the SLT-CP.

Furthermore, some students constructed the bar graph to represent the data regarding the distribution of the amount of the water according to surface water sources presented in the infographic. A sample response from participant 47 is presented below:

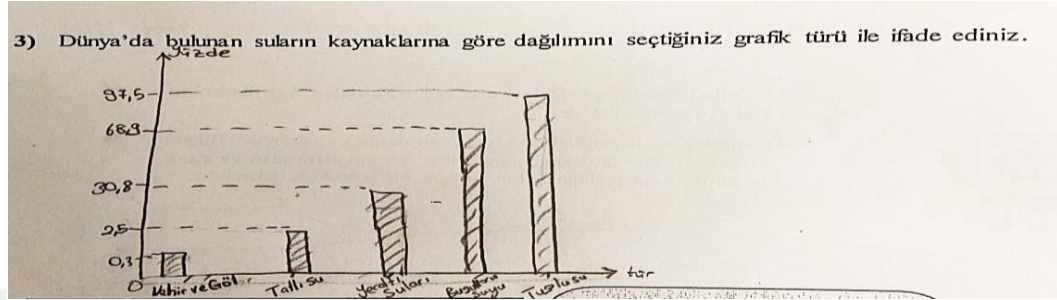


Figure 4.26. The statistical response of Participant 47 to P3 in the SLT-CP

As seen in participant 47's response, the student represented the percentages of fresh water, salty water, ground water, underground water and glacier with a bar graph successfully. The student labeled the axes where x-axis represents the water resources and y-axis represents the percentages of the water resources on earth. The length of the bars belonging to water sources was coherent with the percentages. Therefore, it can be said that participant 47 provided a statistical response to the 3rd problem of the SLT-CP.

78 (31.2%) students made a statistically appropriate graph selection and they constructed the graphs representing the given data. More precisely, the students selected the bar graph or the pie chart to represent the distribution of the A, B and C parts of the whole. Some students constructed a pie chart to represent the A, B and C parts of the whole. A sample response from participant 402 is as follows:

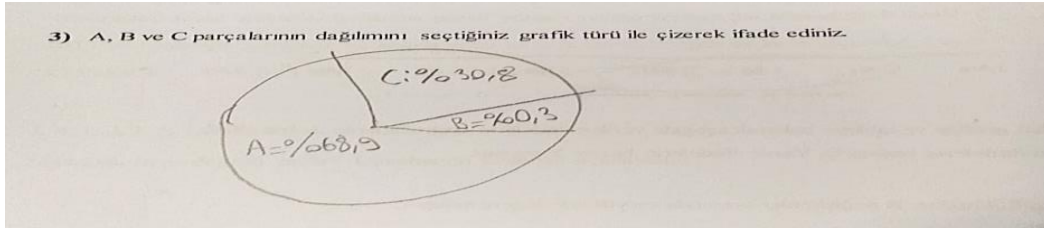


Figure 4.27. The statistical response of Participant 402 to P2 in the SLT

As could be seen, participant 402 labeled the parts of the whole as A, B and C on the pie chart and specified the percentages of these parts placed correctly, together with the labels. Hence, we could better observe the distribution of parts of the whole.

Some students represented the A, B and C parts of the whole by utilizing the bar graph. A sample response from participant 330 is given below:

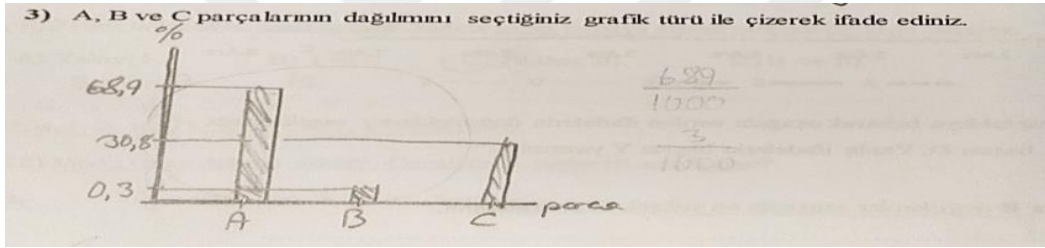


Figure 4.28. The statistical response of Participant 330 to P2 in the SLT

Participant 330 illustrated the A, B and C parts of the whole by benefitting from the bar graph. The student labeled the x-axis as parts and y-axis as percentages. Then, he matched the percentages of the parts appropriately. Thus, it can be said that participant 330 provided a statistical response for the 3rd problem in the SLT.

36 (14.4%) students made a statistically appropriate selection and constructed a graph with small deficiencies. More specifically, these students preferred the bar graph or pie chart to represent the data regarding the distribution of the amount of the water according to surface water sources. However, while constructing these graphs, the students did not place the percentages or variables on the graph. A sample response from participant 194 is given below:

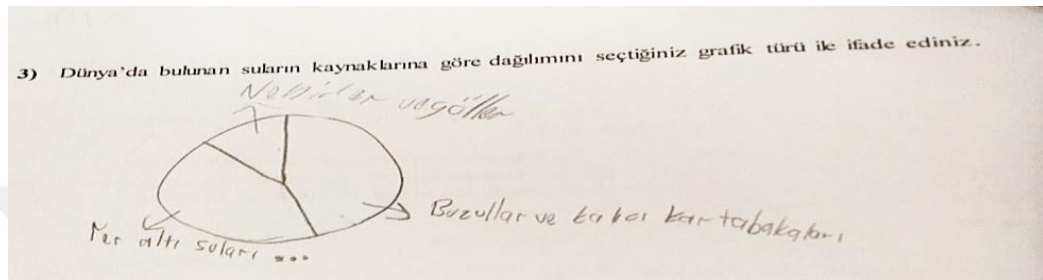


Figure 4.29. The pre-statistical response of Participant 194 to P2 in SLT-CP

As seen in participant 134's response, the student labeled the parts of the whole parallel with the data regarding the distribution of the amount of the water. However, the student did not place the percentages on the segments of the pie chart.

Similarly, 112 (44.8%) students specified the suitable graph and drew the graph correctly with small deficiencies. These students represented the A, B and C parts of a whole with a bar graph or a pie chart. However, the students did not specify the labels or the percentages on the graph. A sample response from participant 158 is given below:

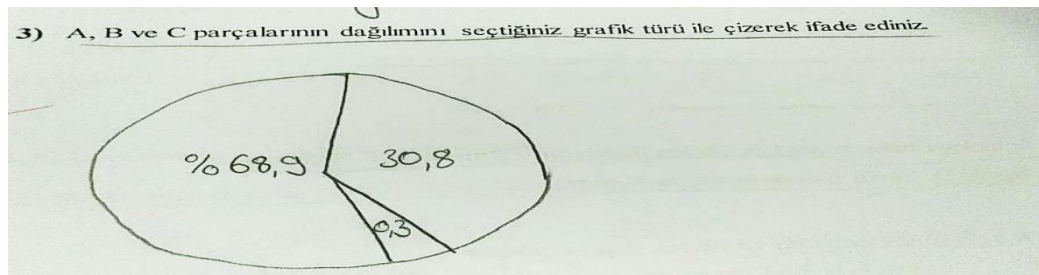


Figure 4.30. The pre-statistical response of Participant 158 to P2 in the SLT

As seen in participant 158's response, the student drew a pie chart to represent the A, B and C parts of the whole. The student placed the percentages of the parts; however, she did not label these percentages. The reader could not infer which percentage belongs to which part. Additionally, the student forgot to use “%” sign in front of the percentages.

70 (28%) students made an idiosyncratic graph drawing independent from the given data in the 3rd problem of the SLT-CP. More precisely, these students selected a line graph to represent the distribution of the amount of the water according to surface water sources. However, the students made a correct drawing with the line graph which is inappropriate to illustrate the distribution of the amount of the water according to surface water sources as the data was categorical. Moreover, some of these students made pie chart drawing in which the sum of the segments exceeds 360° or 100%. A sample response from participant 104 is presented below:

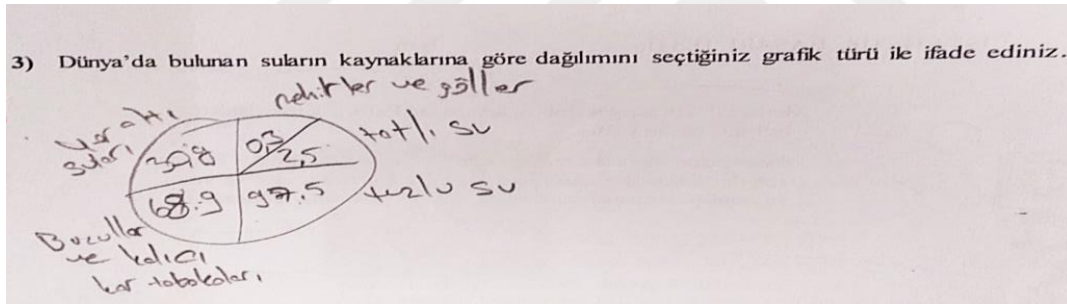


Figure 4.31. The non-statistical response of Participant 104 to P2 in SLT-CP

As seen, participant 104 disregarded the fact that the sum of the percentages of the segments representing the amount of water according to sources should be 100%. This might be because the student might not have interpreted the infographic given in Problem 2 correctly. In the infographic, there were two illustrations. The first illustration gave information about the saltiness of the surface waters and the second illustration showed the distribution of these waters according to sources as ground

water, underground water and glacier. Probably, the student could not distinguish between these two illustrations. In fact, to be an informed citizen, the student should have been able to make a distinction between the given data through media channels such as infographics, newspapers or journals. Additionally, some of these students made a sketchy drawing of the graph without including the variables or the percentages. A sample response from participant 147 is given below:

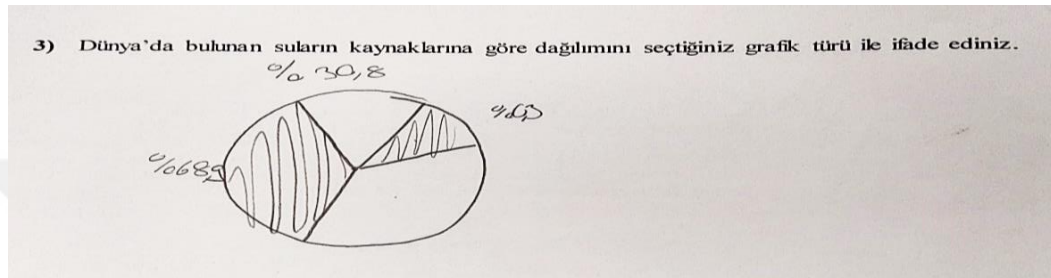


Figure 4.32. The non-statistical response of Participant 147 to P2 in SLT-CP

As seen, participant 147 did not specify the sources of water. The reader could not understand which percentage belongs to which water source. Moreover, one segment of the graph is empty, and the sum of the percentages of the other segments already made 100%. When that segment is included, the sum of the percentages of the segments exceeds 100%. Therefore, we could state that the student drew a non-statistical chart regarding the distribution of the amount of the water according to sources.

Furthermore, 31(12.4%) students drew an idiosyncratic graph independent from the data given in the 3rd problem of the SLT. These students represented the distribution of the A, B and C parts of the whole erroneously. For instance, these students' pie chart exceeded 100% or 360°. Additionally, some of the drawings were sketchy. They did not include percentages and the labels of the variables, but only a circle and three parts.

47(18.8%) students provided an unrelated response or left the 3rd problem of the SLT-CP unanswered. These students sketched only a simple circle by not labeling the variables and the percentages of the water sources around the world.

Similarly, 29 (11.6%) students gave an unrelated answer or left the 3rd problem of the SLT unanswered. A sample response from participant 375 for an unrelated response is as follows:

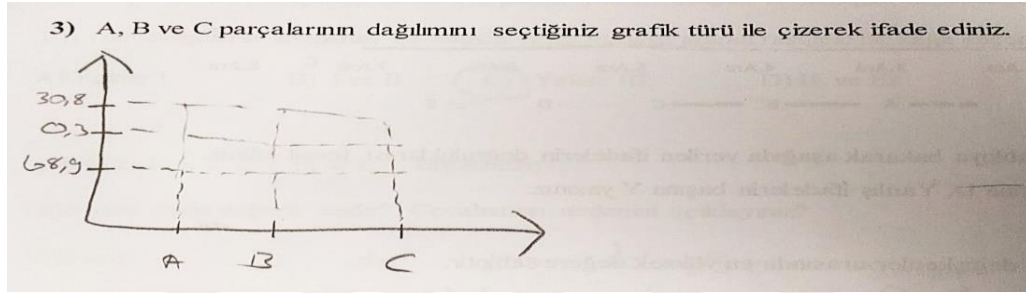


Figure 4.33. The unrelated response of Participant 375 to P2 in the SLT

As could be observed, initially, participant 375 did not label the axes with the parts of the whole and percentages. Additionally, the student could not place the percentages in an ascending order from the zero point of the graph. Furthermore, the student did not match the parts of the whole with their percentages. A reader cannot infer which variable belongs to which percentage. All in all, it could be inferred that participant 375 provided an unrelated response to the 3rd problem of the SLT.

4.2.2.3 The 8th Grade Students Statistical Literacy in Problem 15

Problem 15 is a graph interpretation problem which investigates the students' interpretation of a pie chart. In the 15th problem of the SLT-CP, a pie chart illustrating the sources of water pollution in a country is given. According to the chart, 26% of water pollution stems from agricultural reasons, while 39% from industrial, 44% from domestic and 13% from other factors. The students were requested to interpret the pie chart and express their inferences related to the distribution of the sources of

water pollution. Then, the students were asked whether there is an unusual situation about the pie chart or not because in the chart, the total of the percentages of the sources of water pollution exceeds 100%. To obtain the 15th problem of the SLT, the context of water pollution was eliminated from the problem. The same statistics were presented in the form of parts of a whole. More specifically, the sections of the pie chart illustrated the four parts of a whole instead of four sources of water pollution.

The students' responses were analyzed and classified by the researcher. The classification of students' responses in the SLT-CP and SLT is presented in Table 4.11.

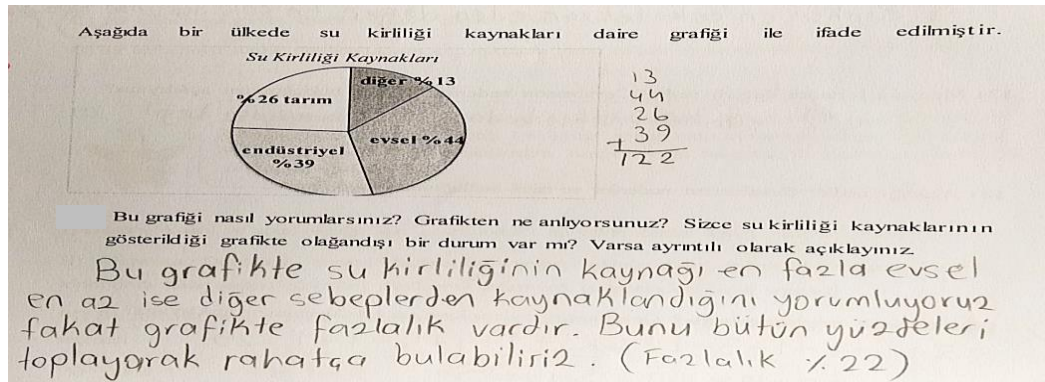
Table 4.11 Student Responses to Problem 15

		Problem 15 with context		Problem 15 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Statistical interpretation about the chart and recognition of error about the chart. -The biggest source of water pollution is domestic. Total percentage is 122% while it should be 100% - By looking at the chart, we can find the biggest source of water pollution. Total percentage is 122% while it should be 100%	28	11.2%	20	8%
Pre-statistical	Make sense of percentages, but improper comments / but cannot notice the error. - Percentage of water pollution sources -The most industrial and the least other sources constitute the water pollution.	93	37.2%	111	44.4%

Table 4.11 (continued)

Non-statistical	Idiosyncratic interpretations Unawareness of error -Shows different sources -Domestic pollution should be reduced	72	28.8%	43	17.2%
Unrelated/Blank		57	22.8%	76	30.4%

28 (11.2%) students made statistically adequate interpretations about the pie chart illustrating the sources of water pollution in a country, in addition to detecting the error about the pie chart. More specifically, these students interpreted the distribution of water pollution sources in terms of their distribution. For instance, the students stated that “*the biggest source of water pollution is domestic and the least important source is the other factors not mentioned in the pie chart*”. In addition to that, these students detected the error with the percentages of the sources of water pollution. They noticed that the sum of the percentages of the water pollution sources exceeds 100%. To clarify the statistical responses to the 3rd problem of the SLT-CP, a sample response from participant 141 is presented below:

**Figure 4.34.** The statistical response of participant 141 to P15 in the SLT-CP

As can be seen, participant 141 responded as ‘In this graph, we interpret that the biggest source of water pollution is domestic and the least source of water pollution is other sources. However, there is an excess in the graph. We can observe that easily by adding the percentages’. The student made statistically adequate interpretations regarding the sources of water pollution. The student indicated that the biggest source of water pollution is domestic pollution and the least common source of water pollution is categorized as “other”. In addition, the student noticed the error about the pie chart by adding the percentages of industrial, domestic, agricultural and other sources of water pollution. Afterwards, the student noticed that the total of the percentages of these sources makes 122% which is 22% more than the whole (100%). Therefore, it can be said that participant 141 provided a statistical response by interpreting the water pollution sources and recognizing the error about the percentages of the water pollution sources.

20 (8%) students interpreted the pie chart with statistically adequate statements. In general, the students ordered the parts of the whole from the smallest to the biggest or from the biggest to the smallest part. Moreover, these students recognized the error by indicating that the total of the percentages of the parts exceeds the whole (100%). A sample response from participant 334 is presented below:

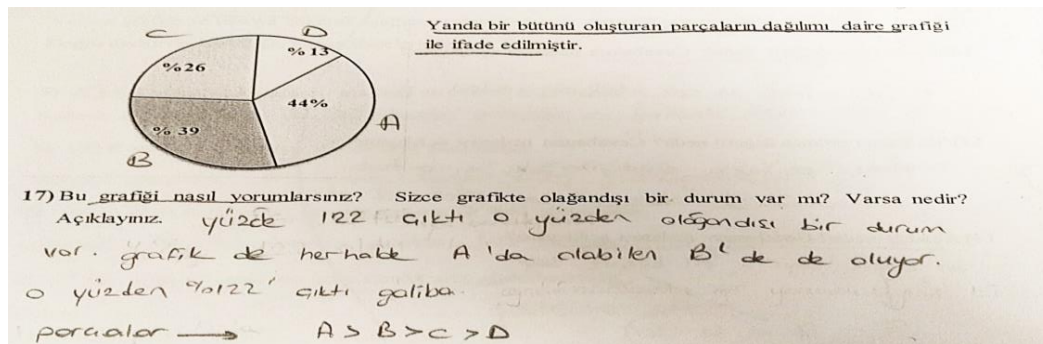


Figure 4.35. The statistical response of participant 334 to P15 in the SLT

Participant 334 stated as ‘The total is 122%. Therefore, there is an unusual situation. I guess A and B could have some joints. Because of that the total is 122%’. The

student noticed the error with the percentages of the parts of the whole. The student reasoned this situation with the intersection areas of the parts of the whole by indicating that part A and part B could include common areas. In addition, the student labeled the parts as A, B, C and D to sort these parts in descending order. Hence, the participant provided a statistical answer to the 15th problem of the SLT-CP. Furthermore, 93 (37.2%) students inappropriately interpreted the pie chart, but noticed the error about the percentages about the water pollution sources. Otherwise, these students could not have noticed the unusual situation regarding the percentages, but they made statistically adequate interpretations about the pie chart in the 15th problem of the SLT-CP. Some of the students made sense of percentages about the water pollution sources, but they could not notice the error about the percentages and focused on different parts of the chart. A sample student response from participant 120 is presented below:

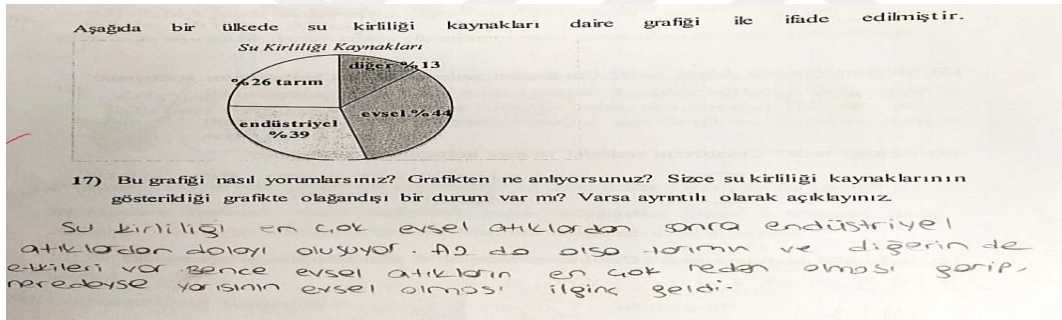


Figure 4.36. The pre-statistical response of participant 120 to P15 in SLT-CP

As can be seen in the figure, participant 120 made inferences from the pie chart as follows:

“The biggest source of water pollution is domestic waste, followed by industrial waste and domestic waste. The effect of the agricultural and other waste is less than domestic and industrial waste. I am surprised that the biggest reason for water pollution is domestic”.

The student made adequate interpretation regarding the sources of water pollution; however, the student did not notice the unusual situation about the percentages of the

sources of the water pollution. Hence, participant 120 made pre-statistical explanations about the 15th problem of the SLT-CP.

111 (44.4%) students either could not notice the error about the percentages or they could not interpret the pie chart appropriately in the 15th problem of the SLT. Some students interpreted the percentages of parts of the whole correctly, but they did not make sense of the total of the percentages of the parts. Some students realized that the total of the percentages exceeds 100%, However, they could not interpret the parts of the whole. A sample response from participant 394 is below:

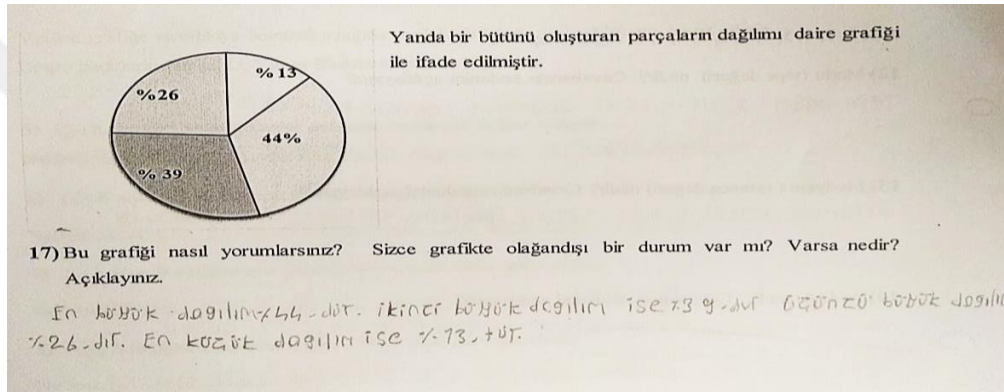


Figure 4.37. The pre-statistical response of participant 394 to P15 in SLT-CP

Participant 394 made statistical comments about the distribution of the segments of the pie chart which represents the parts of a whole by stating that “*The biggest distribution is 44%, the second is 39%, and the third one is 26% and the last one is 13%.*”.

As could be seen, the student preferred to interpret the distribution of the part by sorting the amounts of parts in descending order. However, the student did not mention that the total of the percentages is more than the whole (100%). This means that the student interpreted the pie chart appropriately, but he could not notice the unusual situation regarding the percentages of the segments of the pie chart.

Therefore, it can be said that participant 394 gave a pre-statistical response to the 15th problem of the SLT.

72 (28.8%) students made idiosyncratic interpretations disregarding the purpose of the pie chart in the 15th problem of the SLT-CP. In addition, these students could not recognize the error in the pie chart. These students provided responses such as ‘the graph shows different sources’ and ‘domestic pollution should be reduced.’. A sample response from participant 171 is presented below:

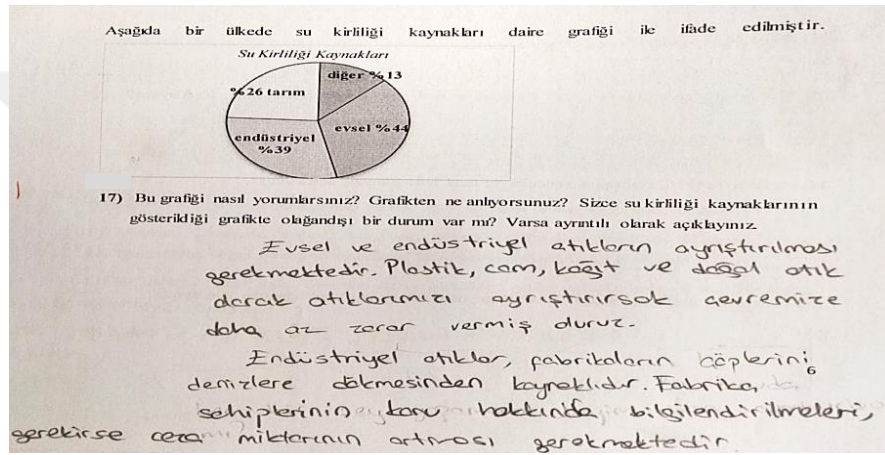


Figure 4.38. The non-statistical response of participant 171 to P15 in SLT-CP

As can be inferred from the figure, participant 171 did not use the data regarding the sources of water pollution. Instead, the student expressed her personal ideas and beliefs regarding the water pollution sources. The student made some suggestions to decrease water pollution. However, the student inferred from the pie chart that domestic and industrial waste have a significant place in terms of water pollution. Accordingly, the student made suggestions to reduce the domestic and industrial waste. Therefore, it could be stated that participant 171 provided a non-statistical response to the 15th problem of the SLT-CP since the student did not use the statistical data regarding the percentages of water sources represented in the bar chart.

43 (17.2%) students made personal interpretations regarding the given data presented in the pie chart. These students could not use the data representing the percentages of the parts of a whole. Additionally, they could not notice that the total of the percentages makes 122% which exceeds 100%. A sample response from participant 407 is presented below:

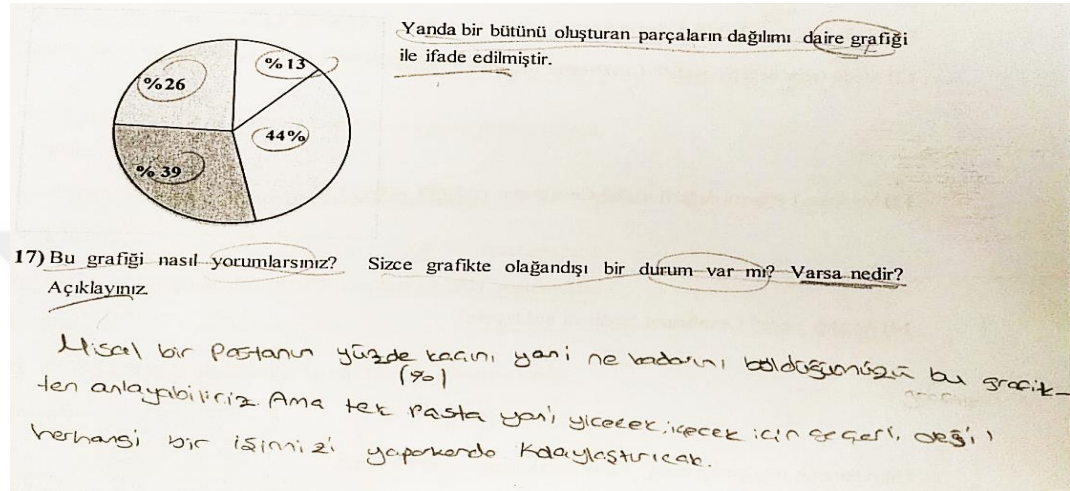


Figure 4.39. The non-statistical response of participant 407 to P15 in the SLT

As seen, participant 407 stated as ‘For instance, we could make inferences about the parts of a cake with the help of this graph. However, not only for eating and drinks we can also benefit from that graph for performing any action’. The student could not interpret the meaning of the segments representing 44%, 39%, 26% and 13% in the pie chart. The student connected the segments of the pie chart with the parts of a pie and she indicated that the use of percentage eases the representation of the parts.

57 (22.8%) students gave an unrelated response to the 15th problem of the SLT or left it unanswered. Most of these students did not attempt to solve the problem. The rest wrote “There is not an unusual situation about the chart” or “No” as an answer.

76 (30.4%) students provided unrelated responses or they did not give an answer to the 15th problem of the SLT-CP. Different from the unrelated responses to the 15th

problem of the SLT-CP, the students made a specific error in the 15th problem of the SLT. Initially, they confused the percentage with the angle and they stated that pie chart could be represented with angles. More specifically, these students calculated the total percentages of the parts as $44\%+39\%+26\%+13\%=122\%$. Then, they stated that the sum of the percentages of the parts should be 360% . In fact, the parts of the whole are represented with percentages not angles. Obviously, these students confused two representations of the whole which are 100% and 360° . To elaborate on this error, a sample response from participant 323 is presented below:

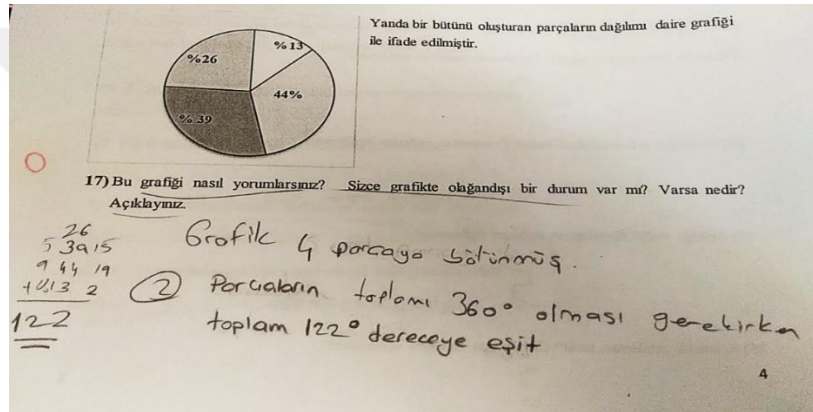


Figure 4.40. The unrelated response of participant 323 to P15 in the SLT

As seen, participant 323 responded as ‘The graph was divided into four parts. The total of the parts should be 360° where the total is equal to 122° ’. The student considered the total of the percentages to detect the error regarding the pie chart and found the total as 122° instead of 122% . In the problem, the four parts of the whole were presented with percentages, not with angles. However, the student confused the central angle of the pie chart with the percentages of the segments of the pie chart. This might result from the fact that in elementary mathematics textbooks, the whole in a pie chart is mostly represented with 360° and the parts in the pie chart are

expressed as parts of 360^0 instead of 100%. Accordingly, the students are not accustomed to percentage on a pie chart. Instead of percentages, they are more likely to use angles to represent the segments of the pie chart. The response of participant 343 consolidates the students' trouble as to whether the segments of a pie chart could be represented with angles, percentages or both.

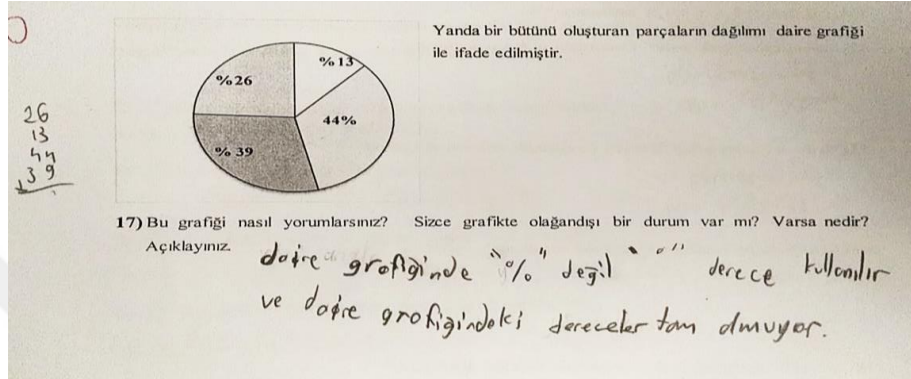


Figure 4.41. The unrelated response of participant 343 to P15 in the SLT

Participant 343 stated that “*In the pie chart, angle should be used, not percentage*”. As could be seen, the student argues that pie chart could not be represented with percentages, but with angles. Consequently, participant 323 and participant 343 provided an incorrect response to the 15th problem of the SLT.

4.2.3 Statistical Literacy in Sampling

Problem 22 was the only open-ended sampling problem in the SLT-CP and SLT. The 8th grade students' statistical literacy in problem 22 was examined in detail.

4.2.3.1 The 8th Grade Students Statistical Literacy in Problem 22

In problem 22 in the SLT-CP, the students were given an excerpt from a newspaper titled “Children Are Not Environmentally Friendly” In this except, there is an

interview conducted with an environmental scientist, for whom the result of a study conducted with a group of 5th grade students in Ankara revealed that the 5th grade students did not make efforts to protect the nature, spend a long time in the shower and waste water, and do not try to save electricity. In the problem, the students' interpretations about the generalizability of the claims of the environmental scientist for all the children in Turkey were investigated. To obtain problem 22 in the SLT-CP, the result of the study was reduced and presented more directly by not making a connection with the environmental issues.

The 8th grade students' responses to problem 22 were analyzed and classified by the researcher. The classification of students' responses in the SLT-CP and SLT is presented in Table 4.12.

Table 4.12 Student Responses to Problem 22

Classification of Responses	Student Response	Problem 22 with context		Problem 22 without context	
		f	p	f	p
Statistical	High level statistical implications from sample to population	55	22%	70	28%
Pre-statistical	Statistically appropriate implication, but insufficient statistical justification <i>-It would be true if there were more people.</i> <i>-Because a group of students says</i> <i>-No, every person is different</i>	81	32.4%	82	32.8%
Non-statistical	Non-statistical, idiosyncratic responses <i>-The garbage should not be thrown around.</i> <i>-Everybody pollutes the nature</i>	47	18.5%	44	17.6%
Unrelated/Blank		67	26.8%	54	21.6%

55 (22%) students indicated in problem 22 of the SLT-CP that the result of a study conducted with a group 5th grade students in Ankara could not be generalized to all the children in Turkey. These students read the claims of the environmental scientist in the newspaper except regarding the unconscious attitudes of the 5th grade students. Afterwards, they critically evaluated the statements they read and they decided that the results of the study conducted with the 5th grade students in Ankara could not be generalized to all the children in Turkey. These students supported their answers with statistical statements such as “It cannot be interpreted only by looking at a group of students” and “The sample is insufficient to generalize to the whole population”.

A sample response from participant 138 is given below:

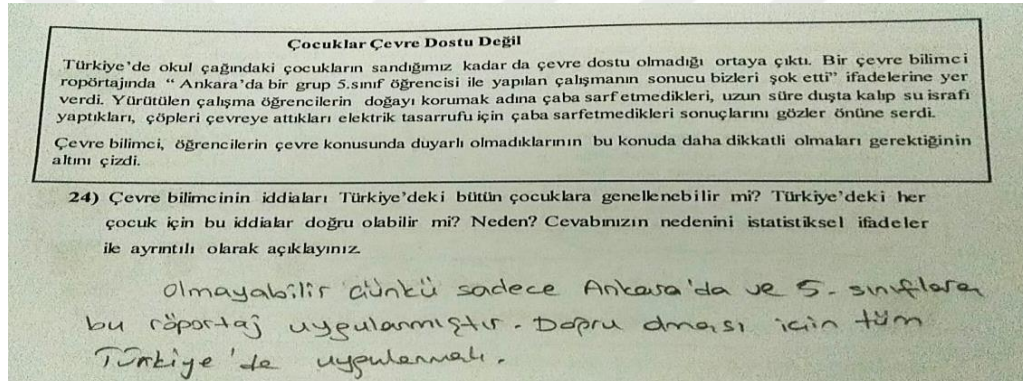


Figure 4.42. The statistical response of participant 138 to P22 in the SLT-CP

As can be seen from participant 138's response, the student indicated that “This study is conducted with the 5th grade students in Ankara. To generalize the results of the study, the study should be applied to all the students in Turkey.”

It could be deduced that the student was aware of the requirements for generalizing the sample of a study to the whole population from which the sample is selected. Similarly, participant 120 stated that “The children in Turkey might be different from the students in the sample”.

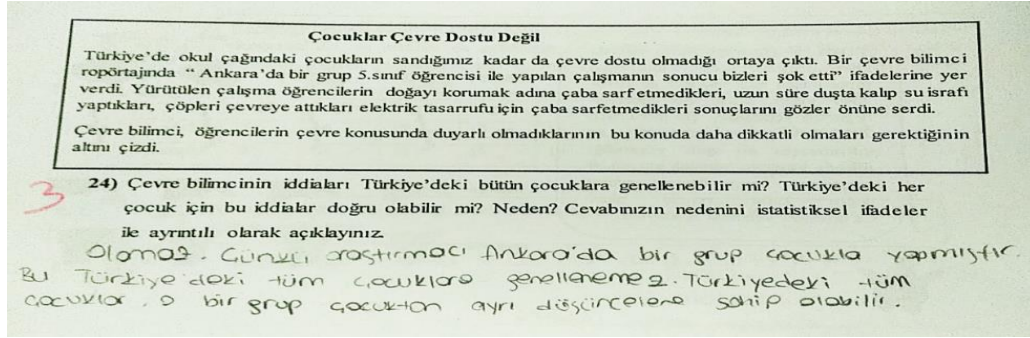


Figure 4.43. The statistical response of participant 120 to P22 in the SLT-CP

As seen, participant 120 responded as *“It cannot be true since the researcher studied with a group of student from Ankara. This could not be generalized to all children in Turkey. They could think differently from a group of student”*. Therefore, the student attracted attention to the personal differences between children making up the whole population and the students in the sample.

70 (28%) students stated that the results of the study conducted with a group of students in Ankara could not be generalized to the population from which the sample was selected. More specifically, the students were aware of the requirement for generalizing the sample to population. A sample response from participant 386 is presented below:

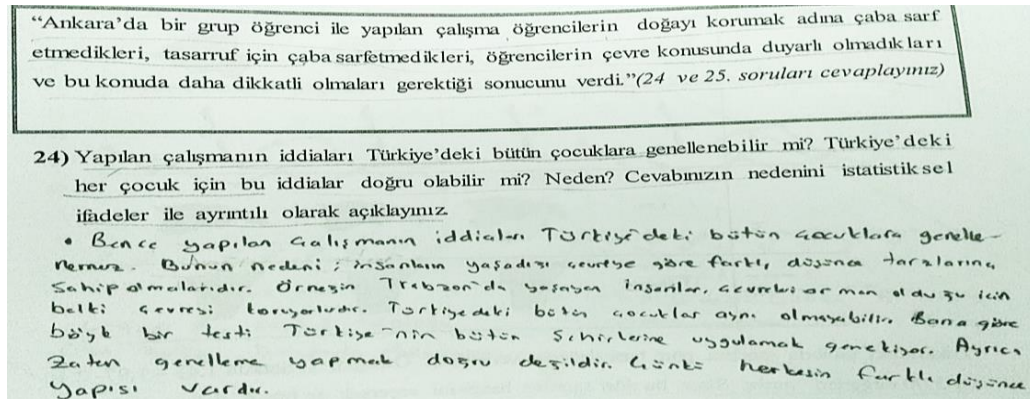


Figure 4.44. The statistical response of participant 386 to P22 in the SLT

Participant 386 responded as “I think the claims could not be generalized to all children in Turkey. The reason is that the people might have different points of view. For instance, the people living in Trabzon could save the environment because Trabzon is a forest land. All the children in Turkey might not be the same. In that regard, the test should be applied to students from all cities in Turkey. Additionally, making generalization is not a good idea. Since, each person has different mentality. As can be inferred, the student stated that the claims of the study could not be generalized to all the children in Turkey. The student mentioned the personal differences between the children in the sample and the whole population. Also, she asserted that to generalize the results of a study, this test should be applied in all cities of Turkey. Hence, participant 386 interpreted the generalizability of the sample to the population with statistical justifications in problem 22 of the SLT.

81 (32.4%) students made appropriate inferences regarding the generalizability of the sample to the whole population in problem 22 of the SLT-CP. More specifically, these students decided that the results of a study conducted with a group of 5th grade student could not be generalized to all the children in Ankara. These students mostly used statements such as “*It would be true if there were more people*”, “*because the problem mentions a group of people*” and “*No, every person is different*”. As seen, the students determined the requirements of the generalization of the results of the study conducted with 5th grades to all the children in Turkey correctly. Additionally, they provided justifications without statistical statements. To clarify these students’ responses, a sample response from participant 130 is presented as follows:

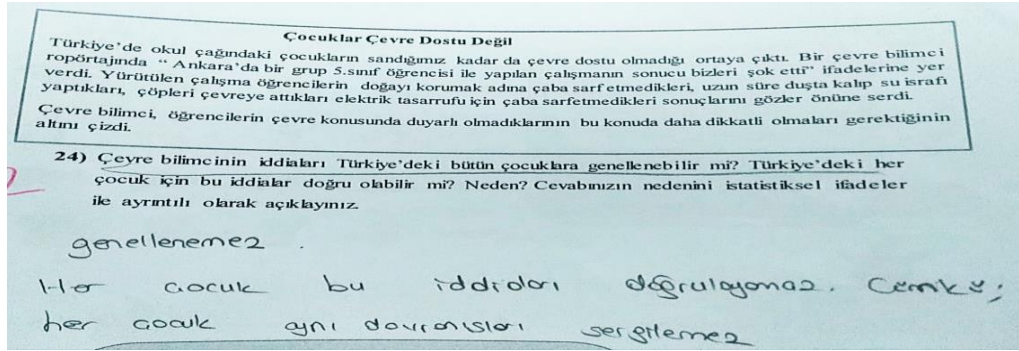


Figure 4.45. The pre-statistical response of participant 130 to P22 in SLT-CP

Participant 130 responded as *"It could not be generalized. Every student could not validate the claims since each student do not act in a particular manner"*. As could be seen, participant 130 mentioned the personal differences between the students in the sample and population. The student realized that the sample could not be generalized to the whole population; however, he could have provided a statistical justification for his decision.

82 (32.8%) students made statistical inferences regarding the generalizability of the sample to population in problem 22 of the SLT. These students stated that the result of a study conducted with a group of students in Ankara could not be generalized to all the children in Turkey. Similar to problem 22 of the SLT-CP, the students supported their decisions with statements such as *"It would be true if there were more people"*, *"because the problem mentions a group of people"* and *"No, every person is different"*.

In fact, the students might have provided more statistical justifications regarding the generalizability of the sample to population. These students made a statistically appropriate decision about the generalizability of the results of the study conducted in Ankara to Turkey, but they made pre-statistical justifications about their decision.

47 (18.5%) students interpreted the environmental scientist's statements presented in the newspaper except with idiosyncratic responses. These students determined the

generalizability of the study conducted with 5th grade students in Ankara by not using a statistical statement. Instead, they made interpretations within the framework of their life experience and personal thoughts and ideas. The students provided responses such as “*The garbage should not be thrown away*” and “*Everybody pollutes the nature*”. A sample response from participant 84 is presented below:

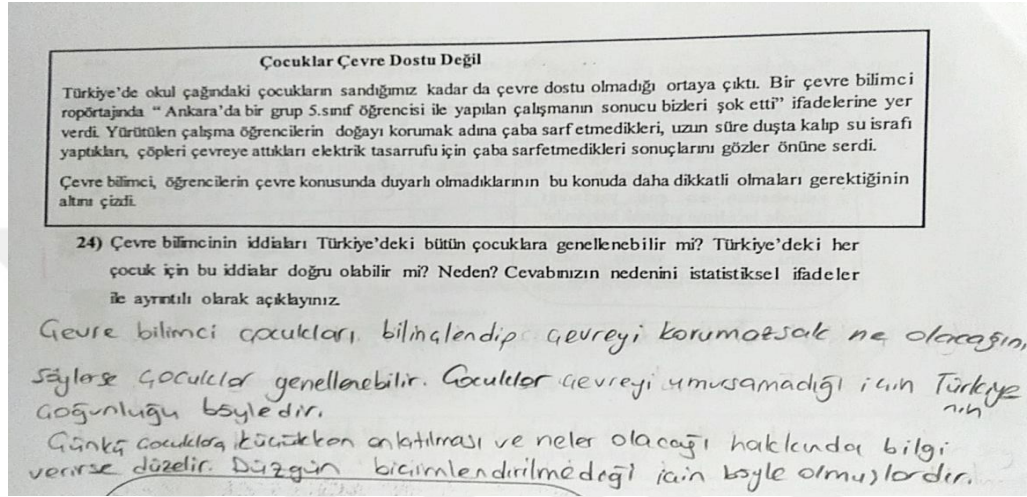


Figure 4.46. The non-statistical response of participant 84 to P22 in SLT-CP

Participant 84 stated as “*If the environmental scientist informs students about the possible outcomes of unconscious behaviors, the students could be generalized. Since the students do not care about the environment, most of the students in Turkey is the same. If the children be informed in early ages, the situation could get better.*” As could be seen, the student made a decision about the environmental scientist’s statements regarding the 5th grade students’ attitudes towards the nature. The student made suggestions to prevent environment from children’s unconscious attitudes. However, the student did not support her ideas with statistical justifications regarding the relationship between the results of the study conducted with 5th grade students in Ankara and the generalizability of the results of that study to all the children in Turkey. As an informed citizen, the students were expected to read the newspaper

except and critically evaluate the information presented in that excerpt. The students who provided idiosyncratic responses should develop their statistical literacy in terms of reading and evaluating the data that they encounter in the media.

Similarly, in problem 22 in the SLT, 44(17.6%) students provided personal justifications about the generalizability of the results of the study carried out in Ankara to all the children in Turkey. Although these students decided that the results of the study could not be generalized to all the children in Turkey, they did not state the requirements of generalization of the sample to the whole population. Instead, they made idiosyncratic comments based on their personal beliefs.

67 (26.8%) students gave unrelated responses to problem 22 in the SLT-CP or left the problem unanswered. These students were not aware of the requirements for the generalizability of the results of the study conducted with 5th grades in Ankara to all the children in Turkey. Therefore, they wrote:

“Yes, we can generalize the results of the study” or “Yes, the claims of the environmental scientist are correct”

The students could not critically evaluate the environmental scientist's claims.

54 (21.6%) students gave an unrelated answer or did not give an answer to problem 22 in the SLT. Similar to the responses to problem 22 of the SLT-CP, the students decided that the results of the study could be generalized to all the children in Turkey by giving responses such as:

“Yes we can generalize these results to all children “or “The results are valid for everybody”.

As seen, the students could not indicate a statistical justification for their responses as an indicator of statistical literacy.

4.2.4 Statistical Literacy in Questioning of Statistical Data

In the test, the 1st, 14th, 17th, 19th, 23rd, 24th and 24th questions were designed to investigate 8th grade students' questioning of statistical data. In these questions, the 8th grade students' statistical literacy in everyday social decision making was investigated by using media, reports, graphs and tables. Additionally, context was employed at all stages of the questioning of statistical data questions. The answers of the 8th grade students were analyzed within the framework of a rubric developed by the researcher. The researcher analyzed and presented the results of the questioning of statistical data items.

4.2.4.1 The 8th Grade Students Statistical Literacy in Problem 1

The 1st problem was a warm up problem for the context and nature of Statistical Literacy Test with Context Problems (SLT-CP). At the beginning of the problem, the marine characters Sponge Bob and Patrick introduced themselves and they specified the aim of the context of the SLT-CP. Then, the Blue Marble photograph was introduced to attract the attention of the students to the amount and necessity of the water around the world. In the problem, the students were expected to make reasonable estimations about the types and the approximate amount of surface water sources on earth. With the help of the 1st problem, students' interest in the context of the test was aroused and they attempted to make logical estimations to answer the problem. To obtain the 1st problem in the SLT, the context was removed from the 1st problem in the SLT-CP and the students were directly requested to make statistical estimations about the number of living species around the world by not involving a narrative part at the beginning of the problem.

The 8th grade students' responses were analyzed and classified by the researcher. The classification of students' responses to the SLT-CP and SLT is presented in Table 4.13.

Table 4.13 Student Responses to Problem 1

		Problem 1 with context		Problem 1 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Reasonable predictions	115	46%	71	28%
Pre-statistical	Appropriate but numerically inappropriate predictions	24	9.6%	68	27.2%
Non-statistical	-Only numerical estimates -Only exemplification	68	27.2%	84	33.6%
Unrelated/Blank		43	17.2%	27	10.8%

115 (46%) students successfully exemplified the species of surface water and they made conceivable estimations regarding the approximate amount of the species of surface water in the 1st problem of the SLT-CP. These students read the narrative part of the problem to be accustomed with the context of the test and to provide a reasonable answer to the first problem. Some of these students evaluated the surface waters in terms of salinity and exemplified the surface waters as fresh water and salt water. Thus, they estimated the amount of fresh water and salty water on earth. The students preferred percentages, surface areas, rates or authentic amounts of water to estimate the amount of water. A sample response from participant 107 is presented as follows:

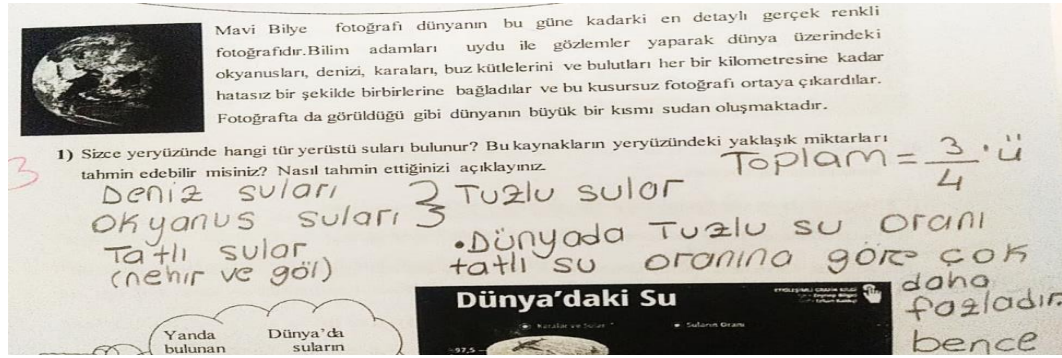


Figure 4.47. The statistical response of participant 47 to P1 in the SLT-CP

As seen, participant 107 wrote as “Sea water, ocean water, freshwater. The rate is $\frac{3}{4}$. On earth, the rate of salty water is more than the rate of freshwater”. The student exemplified the surface water sources as marine waters, ocean waters, rivers and lakes. The student also predicted that the amount of salty water is more than fresh water. The student further estimated that $\frac{3}{4}$ of the earth is composed of water including marine waters, ocean waters, rivers and lakes. As could be seen, participant 107 exemplified the resources of surface waters and evaluated the amount of fresh water and salty water. Afterwards, he estimated the amount of water as $\frac{3}{4}$ of the earth. The numerical value of the students’ estimation and exemplification of sources of the water were reasonable. Therefore, it can be said that participant 107 provided a statistical response to the 1st problem of the SLT-CP.

71 (28%) students made statistical estimations about the number of living species around the world in the 1st problem of the SLT. These students exemplified the living species on earth mostly as animals, plants, viruses and humans. Some of the students classified living species as terrestrials and aquatic living creatures. In addition, the students made their estimation around 1 million, which was the approximate amount of the number of living species. Hence, these students made authentic estimations in the 1st problem of the SLT and thus provided a statistical response to the problem.

When the students’ answers in the 1st problem of the SLT-CP are compared with the 1st problem of the SLT, it is observed that the students made sense of the problem

with the help of the problem context. Thanks to the information regarding the Blue Marble photograph, the students got accustomed to the overall content of the test and they started to consider the content of the SLT-CP. Therefore, the students provided statistical responses to the 1st problem of the SLT-CP compared to the 1st problem of the SLT.

24 (9.6%) students made appropriate but numerically untrue predictions regarding the approximate amount of surface water in the 1st problem of the SLT-CP. The students expressed the species of surface waters in terms of salinity and exemplified the surface waters as fresh water and salt water. Thus, they estimated the amount of fresh water and salty water on earth. Similar to the students who provided statistical responses, these students preferred percentages, surface areas, rates or authentic amounts of water to estimate the amount of water. However, the students' estimations were far from the reality. To clarify the students' pre-statistical answer, a sample response from participant 180 is given below:

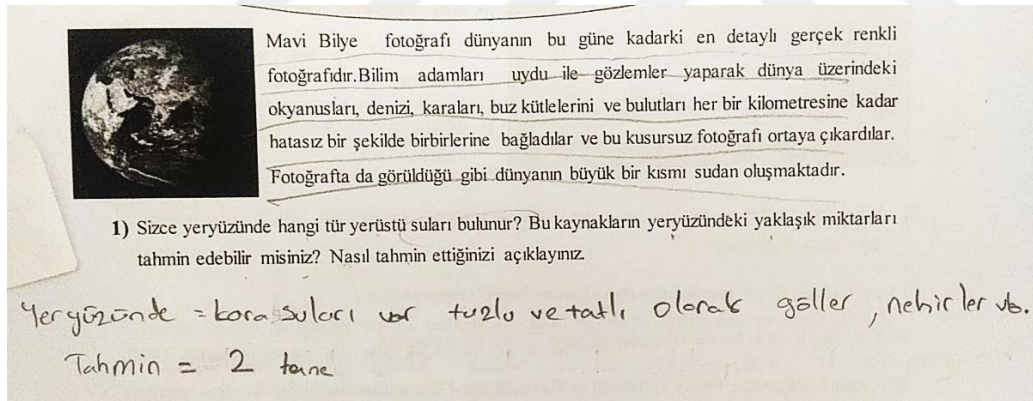


Figure 4.48. The pre-statistical response of Participant 180 to P1 in SLT-CP

As seen, the student exemplified the species of surface water as territorial waters, lakes and rivers by indicating the salinity of the water. It could be inferred that the student engaged with the problem context by giving examples regarding the amount of surface waters. However, the student's numerical estimation was 2. In the problem, the amount of the water was asked. The student did not estimate the amount

with percentages, rates, ratios or surface areas. Although the student attempted to make an estimation regarding the total amount of surface water, her estimation and the examples she provided were far from reality.

68 (27.2%) students made appropriate but numerically insubstantial predictions regarding the approximate amount of living species in the 1st problem of the SLT. Although there was an attempt to estimate the approximate amount of living species, the students' estimations were far from reality. A sample response from participant 419 is presented below:

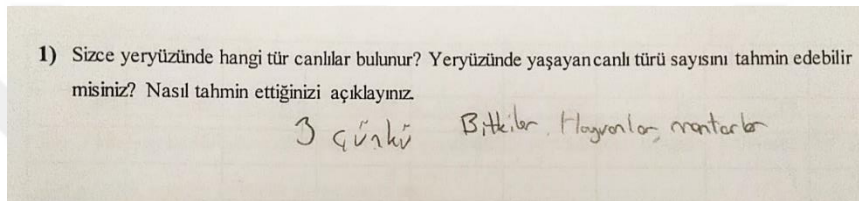


Figure 4.49. The pre-statistical response of Participant 419 to P1 in the SLT

As could be seen, the student exemplified the living species by writing as “*plants, animals and fungi*”. Then, the student estimated the approximate number of the living species as 3. The students' response was somehow irrational. The student even did not regard the humans as a living species. Obviously, participant 419 made non-statistical estimation regarding the approximate number of living species.

68 (27.2%) students either sampled the species of living creatures around the world or they estimated the number of living creatures around the world in the 1st problem of the SLT-CP. On the other hand, some students made only numerical estimation

by not exemplifying or justifying their estimation. A sample response from participant 176 is given below:

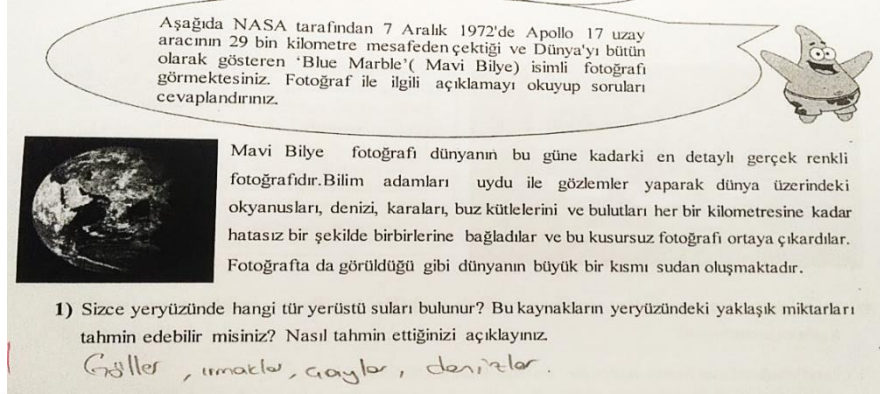


Figure 4.50. The non-statistical response of Participant 176 to P1 in SLT-CP

As seen, participant 176 exemplified the sources of surface waters by writing as “*rivers, streams, brooks and seas*” and by not making an estimation about the amount of surface water on earth. The aim of the 1st problem in the test was to investigate the students’ estimation skills as the students are required to use their estimation skills in interpreting the statistical statements presented through media channels. However, the students did not provide an estimation. Hence, it could be stated that these students provided a non-statistical response to the 1st problem of the SLT.

84 (33.6%) students either made only numerical estimation about the number of living species or they exemplified only the living species in the 1st problem of the SLT-CP. A sample response from participant 434 is as follows:

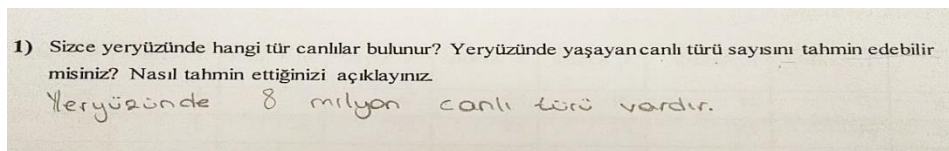


Figure 4.51. The non-statistical response of Participant 434 to P1 in SLT-CP

As seen, participant 434 made only a numerical estimation regarding the number of living species around the world by stating as “*There are eight millions of living species around the world*”. The student did not give examples for the living creatures on earth. He did not indicate which species the estimation includes. In fact, the 8th grade students should be able to provide examples and make reasonable estimations regarding the issues in the newspaper excerpts they read or the news they listen to. This is a requirement for being a part of the society in which people critically evaluate the reading or listening included in daily life.

43 (17.2%) students gave unrelated answers or they left the 1st problem of the SLT-CP unanswered. These students could not give examples about the species of surface water and they could not make reasonable estimations regarding the approximate number of the species of surface water. Some of these students provided responses such as “*The amount of water cannot be estimated*”; however, they did not provide a statistical justification about the reasons for the unpredictability of the approximate amount of surface waters. Hence, these students provided non-statistical responses to the 1st problem in the SLT-CP.

27(10.8%) students gave unrelated answers to the 1st problem of the SLT or they did not answer the problem. These students were not able to exemplify the living species and estimate the number of the living species on earth. Some of these students indicated that the number of the living species cannot be guessed. However, the 8th grade students should be able to make predictions about such issues that they encounter through the media channels.

4.2.4.2 The 8th Grade Students Statistical Literacy in Problem 14

In problem 14 in the SLT-CP, a table taken from a report prepared by the Chamber of Environmental Engineers was presented to the students. The table illustrates the amount of rainfall in different regions of Turkey for the year 2016. Additionally, the

table presented some statistics regarding the amount of rainfall, expected amount of rainfall and the comparison between the amount of rainfall in 2016 and the previous year. The rate of change in the amount of rainfall was also placed on the table. In the problem, the students were expected to find which region or regions received less rainfall than normal (expected) rainfall in 2016. They were also expected to explain the reason behind their decision by providing statistics from the table. To obtain the 14th problem in the SLT, the statistics related to the amount of rainfall in 2016 was eliminated from the problem. The numerical values in the table remained the same. This time, the numerical values were presented in the same way by indicating that the statistics was the results of eight measurements without giving the regions. Then, the students were asked which value or values were below the expected value.

The 8th grade students' responses were analyzed and classified by the researcher. The classification of students' responses to the 14th problem in the SLT-CP and SLT is presented in Table 4.14.

Table 4.14 Student Responses to Problem 14

		Problem 14 with context		Problem 14 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Summarizes data -looking for chance column -comparing the amounts of received rainfall and expected rainfall -Justifications regarding the regions of Marmara, Ege, Akdeniz and Güneydoğu Anadolu -Justifications regarding measurements of 2 nd , 3 rd , 4 th and 8 th measurements	79	31.6%	86	34.4%

Table 4.14 (continued)

Pre-statistical	Basic data reading -Marmara, Ege, Akdeniz and Güneydoğu Anadolu regions - 2 nd , 3 rd , 4 th and 8 th measurements	49	19.6%	40	16%
Non-statistical	Idiosyncratic interpretations -Ege region always receives less rain, I look the percentages. -8 th measurement, I inferred from the table	42	16.8%	53	21.2%
Unrelated/Blank		80	32%	71	28.4%

79 (31.6%) students summarized the data given in the table retrieved from a report prepared by the Chamber of Environmental Engineers in statistically appropriate methods. These students used various techniques to find the regions receiving less rainfall than the expected rainfall in 2016. Initially, some students examined the “change” column of the table to observe the changes in the amount of rainfall. The students examined the rate of change in received rainfall according to normal rainfall. A sample response from participant 171 for this method is as follows:

16)

Aşağıda Çevre Mühendisleri Odası tarafından hazırlanan tabloda Türkiye’deki bölgelere 2016 yılında düşen yağış miktarları verilmiştir.

BÖLGE	YAĞIŞ (mm)	NORMAL (mm)	GEÇEN YIL (mm)	DEĞİŞİM ORANI	
				NORMALE GÖRE (%)	GEÇEN YILA GÖRE (%)
Türkiye Ortan	597.6	574.0	585.7	4.1 ARTMA	2.0 ARTMA
Marmara	646.6	609.2	638.3	-1.9 AZALMA	1.3 ARTMA
Ege	555.5	595.3	652.0	-6.7 AZALMA	14.8 AZALMA
Akdeniz	559.3	663.7	647.5	-15.7 AZALMA	19.6 AZALMA
İç Anadolu	436.5	407.8	441.9	7.0 ARTMA	1.2 AZALMA
Karadeniz	882.6	698.0	727.2	26.4 ARTMA	21.4 ARTMA
Doğu Anadolu	583.6	565.2	533.3	3.3 ARTMA	9.4 ARTMA
Güneydoğu Anadolu	509.9	549.1	532.3	-7.1 AZALMA	4.2 AZALMA

Bu tabloya göre 2016 yılında hangi bölgeye veya bölgelere normal yağış miktarından daha az miktarda yağış düşmüştür? Nasıl karar verdiğinizi ayrıntılı olarak açıklayınız?

Kararlarında yatan artma ve azalma miktarlarını karşılaştırdım. Kararlarımı bu şekilde yaptım.

Marmara, Ege, Akdeniz, Güneydoğu Anadolu bölgelerine normale göre daha az yağış düşmüştür.

Figure 4.52. The statistical response of participant 171 to P14 in the SLT-CP

Participant 171 responded as “The amount of rainfall in Marmara, Ege, Akdeniz, Güneydoğu Anadolu regions was less than the normal amount of rainfall. I looked the amount of increase and decrease writing on the side of the table”. As seen in the participant 171’s response, by examining the change column of the table, the students looked for the increase and decrease in the amount of the rainfall received in different regions of Turkey. The change in the amount of rainfall means there should be a decrease in the received rainfall compared to the expected amount of rainfall. In that case, Marmara, Ege, Akdeniz and Güneydoğu Anadolu regions were the regions with less rainfall than the normal amount of rainfall. Furthermore, some students compared the values in the “rainfall” column and “normal” column in the table. They thought that the amount of received rainfall should be less than the amount of normal rainfall. A sample response from participant 234 is given below:

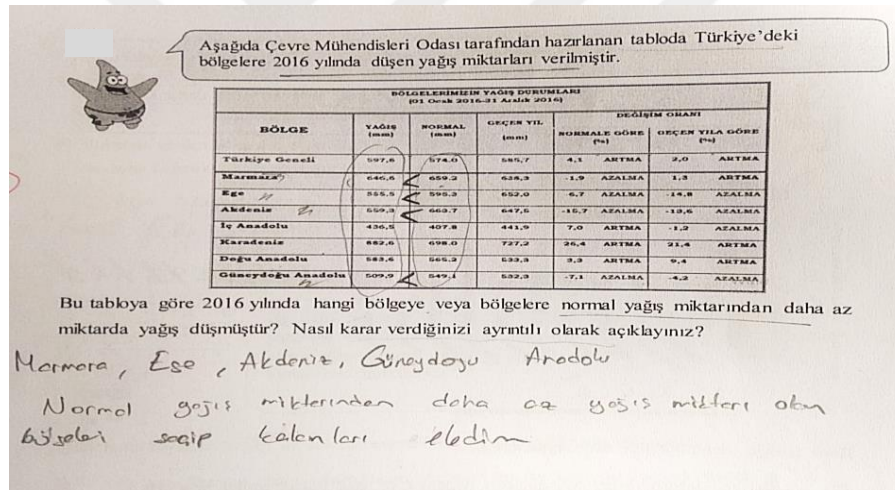


Figure 4.53. The statistical response of participant 234 to P14 in the SLT-CP

Participant 234 responded as “Marmara, Ege, Akdeniz and Güneydoğu Anadolu regions. I selected the regions taking less than the normal amount of rainfall”. As can be inferred from the participant 134’s response, the student compared the received rainfall and the normal rainfall for each of the regions and specified Marmara, Ege, Akdeniz and Güneydoğu regions as the regions where the value of received rainfall is smaller than the value of normal rainfall. In addition, the student

explained her reasoning about the decision making process by writing as “I specified the regions which received less rainfall than the normal amount of rainfall and I eliminated the rest of the regions”

Consequently, these students summarized the data presented in the report with a statistically appropriate method and therefore they provided a statistical response to the 14th problem in the SLT-CP.

86 (34.4%) students interpreted the data presented in the table representing the results of the eight measurements by using the appropriate variables from the table for the 14th problem in the SLT. These students determined the results of the measurements which were less than the expected measurement result. They followed different paths to do so. Many students made a comparison between the ‘result of the measurement’ column and the ‘expected value’ column. The students compared the values of both columns for all eight measurements. Afterwards, they determined the 2nd, 3rd, 4th and 8th measurements with less values than the expected value. A sample response from participant 312 is given below:

16) Aşağıda ekte edilen bazı ölçüm sonuçları verilmiştir.

	Ölçüm Sonucu	Beklenen Değer	Önceki Ölçüm Sonucu	Beklenene Göre Ölçüm Sonucu	Önceki Ölçüm Sonucu
1. Ölçüm	597,6	574,0	585,7	4,1	2,0
2. Ölçüm	646,6	659,2	638,3	-1,9	1,3
3. Ölçüm	555,5	595,3	652,0	-6,7	-14,8
4. Ölçüm	559,3	663,7	647,5	-15,7	-13,6
5. Ölçüm	436,5	407,8	441,9	7,0	-1,2
6. Ölçüm	882,6	698,0	727,2	26,4	21,4
7. Ölçüm	583,6	565,2	533,3	3,3	9,4
8. Ölçüm	509,9	549,1	532,3	-7,1	-4,2

Buna göre hangi değer veya değerler beklenen değer in altında kalmıştır? Nasıl karar verdiniz?

2., 3., 4., 8. ölçümler beklenen değer in altında kalmıştır. Tabloyu inceledim; önce beklenen değere baktım daha sonra ölçüm sonucuna bakıp karşılaştırma yaptım.

Figure 4.54. The statistical response of participant 312 to P14 in the SLT

The student stated that:

“The results of 2nd, 3rd, 4th and 8th measurements are below the expected value. Firstly, I looked for the expected value, and then, I looked at the result of the measurement and I made a comparison.”

It could be inferred that the student utilized the variables and the values in the table to find the appropriate measurements showing less rainfall than the expected rainfall.

Some students examined the result according to the expected column. These students preferred to examine the data presenting a comparison between the result of the measurement and the expected value of measurement. Thereby, the students did not need to compare two columns regarding the result of the measurement and the expected value of the measurement. This was a wise choice to specify the measurements indicating less value than the expected value. A sample response from participant 314 is presented below:

16) Aşağıda elde edilen bazı ölçüm sonuçları verilmiştir.

	Ölçüm Sonucu	Beklenen Değer	Önceki Ölçüm Sonucu	Beklenene Göre Ölçüm Sonucu	Önceki Ölçümüne Göre Ölçüm Sonucu
1. Ölçüm	597,6	574,0	585,7	4,1	2,0
2. Ölçüm	646,6	659,2	638,3	-1,9	1,3
3. Ölçüm	555,5	595,3	652,0	-6,7	-14,8
4. Ölçüm	559,3	663,7	647,5	-15,7	-13,6
5. Ölçüm	436,5	407,8	441,9	7,0	-1,2
6. Ölçüm	882,6	698,0	727,2	26,4	21,4
7. Ölçüm	583,6	565,2	533,3	3,3	9,4
8. Ölçüm	509,9	549,1	532,3	-7,1	-4,2

Buna göre hangi değer veya değerler beklenen değerin altında kalmıştır? Nasıl karar verdiniz?

2.-3.-4.-8. Altında kalmıştır.
Beklenene göre ölçüm sonucu negatif olanlar Beklenene göre altında kalmış demektir

Figure 4.55. The statistical response of participant 314 to P14 in the SLT

The student stated that:

“The 2nd, 3rd, 4th and 8th measurements are below the expected value. The measurement values which have the negative sign indicate rainfall less than expected”.

Obviously, the student specified the appropriate measurements excluding the unnecessary information included in the table and selecting the suitable column to decide on the appropriate measurements.

49 (19.6%) students read the data from the table retrieved from a report prepared by the Chamber of Environmental Engineers, but they could not specify all the regions receiving less rainfall than the expected rainfall in the 14th problem of the SLT-CP. Even if they specified the correct regions by basically reading data from the table, these students could not provide a full statistical justification for their responses. A sample response from participant 95 is given below:

Aşağıda Çevre Mühendisleri Odası tarafından hazırlanan tabloda Türkiye'deki bölgelere 2016 yılında düşen yağış miktarları verilmiştir.

BÖLGE	YAĞIŞ (mm)	NORMAL (mm)	GEÇEN YIL (mm)	DEĞİŞİM ORANI	
				NORMALE GÖRE (%)	GEÇEN YILA GÖRE (%)
Türkiye Geneli	597.6	674.0	585.7	-11.1	ARTMA
Marmara	646.6	659.2	638.3	-1.9	ARTMA
Ege	655.5	655.3	652.0	-0.7	ARTMA
Akdeniz	559.3	663.7	647.5	-15.7	ARTMA
İç Anadolu	426.5	407.8	441.9	7.0	ARTMA
Karadeniz	682.6	698.0	727.2	26.4	ARTMA
Doğu Anadolu	583.6	565.2	533.3	3.3	ARTMA
Güneydoğu Anadolu	509.9	549.1	532.3	-7.1	ARTMA

Bu tabloya göre 2016 yılında hangi bölgeye veya bölgelere normal yağış miktarından daha az miktarda yağış düşmüştür? Nasıl karar verdiğinizizi ayrıntılı olarak açıklayınız?

Marmara, Ege, Akdeniz, Güneydoğu Anadolu

Figure 4.56. The pre-statistical response of participant 95 to P14 in SLT-CP

As seen, the student wrote “Marmara, Ege, Akdeniz and Güneydoğu Anadolu”. Participant 95 read the data from the table included in the report of the Chamber of Environmental Engineers regarding the amount of rainfall in different districts of Turkey. Additionally, the student correctly specified the regions receiving less rainfall than the expected rainfall. However, the student did not make an explanation about the usage of data in the table. For instance, the student did not indicate the variables or values that he used. A reader cannot make an inference about how the student determined the regions with less rainfall than expected. The student could

clarify his thinking with statistical justifications by utilizing the data given in the table. He provided a pre-statistical response to the 14th problem of the SLT-CP.

40 (16%) students read the data pertaining to the results of the eight measurements presented in the table in the 14th problem of the SLT. These students were able to read the data from the table, but they could not fully interpret the data presented in the table. More precisely, these students were aware what the columns of the table represent, but they could not fully determine the regions receiving less rainfall than expected. A sample response from participant 498 is presented below:

	Ölçüm Sonucu	Beklenen Değer	Önceki Ölçüm Sonucu	Beklenene Göre Ölçüm Sonucu	Önceki Ölçüme Göre Ölçüm Sonucu
1. Ölçüm	597,6	574,0	585,7	4,1	2,0
2. Ölçüm	646,6	659,2	638,3	-1,9	1,3
3. Ölçüm	555,5	595,3	652,0	-6,7	-14,8
4. Ölçüm	559,3	663,7	647,5	-15,7	-13,6
5. Ölçüm	436,5	407,8	441,9	7,0	-1,2
6. Ölçüm	882,6	698,0	727,2	26,4	21,4
7. Ölçüm	583,6	565,2	533,3	3,3	9,4
8. Ölçüm	509,9	549,1	532,3	-7,1	-4,2

Buna göre hangi değer veya değerler beklenen değerin altında kalmıştır? Nasıl karar verdiniz?

3. Ölçüm = 555,5 Çünkü beklenenden düşük olmuştur.
8. Ölçüm = 509,9

Figure 4.57. The pre-statistical response of participant 498 to P14 in the SLT

According to the table, the 2nd, 3rd, 4th and 8th measurements indicate less rainfall than the expected amount of rainfall. As seen, participant 498 specified the 3rd and 8th measurements by reading these measurements' results as 555.5 and 509.9 from the table. It could be inferred that participant 498 was able to read the data from the table. Furthermore, the student explained the reason for choosing these measurements. However, the student could not specify all of the measurements indicating less rainfall than expected since the 2nd and 4th measurements indicated less rainfall than the normal amount of rainfall as well. Obviously, the student could not specify all of the measurements indicating less rainfall than normal.

42(53%) students provided idiosyncratic responses about the table taken from a report prepared by the Chamber of Environmental Engineers. These students selected only one of the appropriate regions and they provided a personal justification about their selection by stating:

“It is Marmara region because it is the most appropriate” and “It is Akdeniz region. That is what I read from the table.”

A sample response from participant 11 is presented below:

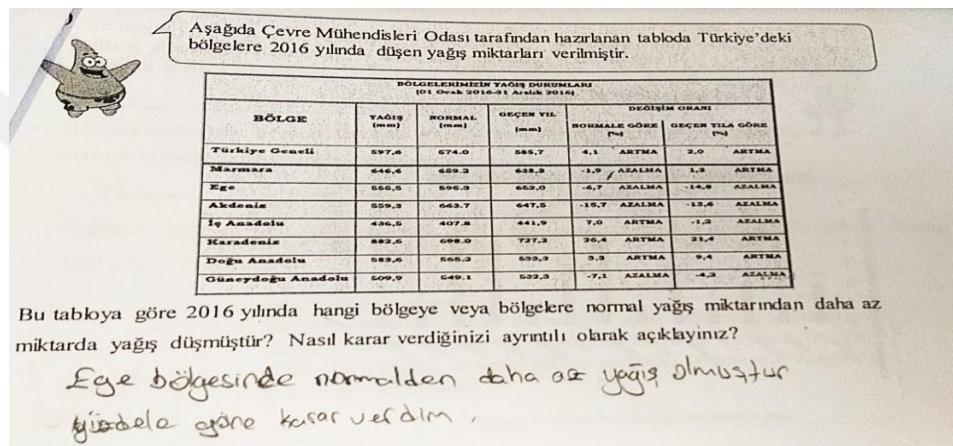


Figure 4.58. The non-statistical response of participant 11 to P14 in SLT-CP

As seen, participant 11 selected Ege region as the region receiving less rainfall than the normal amount of rainfall by stating as:

“Ege region received less rainfall than normal. I determined that by looking at the percentages”

The student selected one of the four appropriate regions. However, the student provided a personal justification since percentages were not used to present the amount of rainfall. The amount of rainfall was represented by the unit of mm. Although the rate of change was presented with percentages, the student did not indicate the way he benefitted from the percentages. Thus, the student’s response was regarded as a non-statistical and idiosyncratic response because there was not

an evidence of statistical literacy in terms of reading and interpreting the table in the student's response.

53 (21.2%) students provided idiosyncratic and non-statistical responses to the 14th problem of the SLT. These students specified one of the four appropriate measurements which indicate less value than the expected value. However, these students could not show full literacy about the data presented in the table. These students provided personal justifications about their answer such as “8th measurement, because I inferred it from the graph”

As seen, the students specified one of the four appropriate measurements, but they did not justify the reason behind their selection. They should have provided a statistical justification from the data presented in the table considering the results of the measurements.

80 (32%) students did not provide an answer or they gave an unrelated answer to the 14th problem of the SLT-CP. These students did not make a selection from among the regions to determine the regions receiving less rainfall than the normal amount of rainfall. They made unrelated comments regardless of the aim of the problem. A sample response for an unrelated answer by participant 91 is as follows:

BÖLGE	YAĞIŞ (mm)	NORMAL (mm)	GEÇER YIL (mm)	DEĞİŞİM ORANI	
				NORMAL GÖRE (%)	GEÇER YILA GÖRE (%)
Türkiye Geneli	597,6	674,0	585,7	4,1 ARTMA	2,0 ARTMA
Marmara	646,6	659,2	638,9	-1,9 AZALMA	1,3 ARTMA
Ege	656,5	696,3	652,0	-4,7 AZALMA	-14,9 AZALMA
Akdeniz	559,3	663,7	647,5	-15,7 AZALMA	-19,6 AZALMA
İç Anadolu	486,5	407,8	441,9	7,0 ARTMA	-1,2 AZALMA
Karadeniz	882,6	698,0	727,2	26,4 ARTMA	21,4 ARTMA
Doğu Anadolu	583,6	565,2	533,9	3,3 ARTMA	9,4 ARTMA
Güneydoğu Anadolu	509,9	549,1	532,9	-7,1 AZALMA	-4,2 AZALMA

Bu tabloya göre 2016 yılında hangi bölgeye veya bölgelere normal yağış miktarından daha az miktarda yağış düşmüştür? Nasıl karar verdiğinizizi ayrıntılı olarak açıklayınız?

Tabloyu okumayı bilen her birsey bu tabloya bakıp karşılaştırma yaparak kolayca cevaba ulaşabiliriz.

Figure 4.59. The unrelated response of participant 91 to P14 in the SLT-CP

Participant 91 made an unrelated comment far from the problem situation and the data by stating:

“Every person who knows to read the table could easily reach the answer by making comparisons. “

As seen, the student's response was not related with the problem.

71 (28.4%) students provided an unrelated response or they left the problem unanswered. Similar to the 14th problem in the SLT-CP, these students' responses were irrelevant with the problem situation.

4.2.4.3 The 8th Grade Students' Statistical Literacy in Problem 17

In problem 17 in the SLT-CP, the students were presented information about global warming which is one of the causes of drought. To observe the effect of global warming on a city's air temperature, a table showing the 9-day temperature for a city in two different intervals was presented. Then, the students were asked in which interval the city is warmer. Additionally, the students were expected to provide a justification about their answer and explain their way of thinking while answering the problem. To obtain the 17th problem in the SLT-CP, the problem was asked directly by using the same statistics. The values of the variables A and B in 9 different intervals were presented by using the same data as the 17th problem of the SLT-CP.

The 8th grade students' responses were analyzed by the researcher and the students' responses to the 17th problem were classified. The classification of students' responses to the problem in the SLT-CP and SLT is presented in Table 4.15.

Table 4.15 Student Responses to Problem 17

		Problem 17 with context		Problem 17 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Summarization of the data and use of the appropriate values in the table. -Mode, -Median, -Arithmetic mean, -Total sum.	114	45.6%	116	46.4%
Pre-statistical	Reading only data from the table. -The numbers in the 2 nd interval is more -The values of variable B is bigger	30	12%	29	11.6%
Non-statistical	Interpreting the data with non-statistical, idiosyncratic expressions. -2 nd interval, I chose the day with the highest temperature -2 nd interval, it seems more -2 nd interval, since the peak value belongs to 2 nd interval -The Variable B, the table shows	51	20.4%	53	21.2%
Unrelated/Blank		55	22%	52	20.8%

114 (45.6%) students questioned the statistical data regarding air temperature appropriately. More precisely, these students examined the values of the 9-day air temperature for both two intervals and they made a decision about the warmer interval by utilizing the adequate statistical constructs. The majority of the students used measures of central tendency such as *mode*, *median*, *arithmetic mean*, *total sum* and *difference*. This means that the students made use of a wide variety of statistical methods to question the statistical data regarding the results of the measurements for the variables A and B as given in the table. Especially, the use of arithmetic mean

and total sum was the most preferred concepts by the students to determine the warmer interval.

Initially, 47 (18.8%) of these students utilized the *total sum* to compare the air temperatures in two intervals. A sample response from participant 114 is presented below:

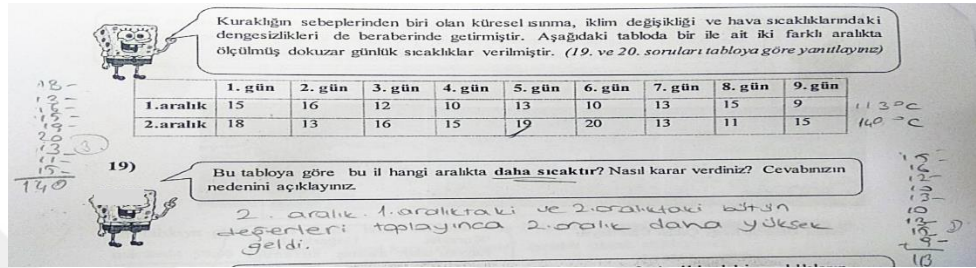


Figure 4.60. The statistical response of participant 114 to P17 in the SLT-CP

As seen, participant 114 calculated the sum of 9-day temperatures for the 1st interval and the 2nd interval by adding up the values of the air temperatures in both two rows of the table. As a result, the student calculated the total air temperature values as 113 °C for the 1st interval and 140 °C for the 2nd interval. Then, the student compared the total sum for both two intervals and decided that the total for the 2nd interval is higher than the total of 1st interval. Hence, the city is warmer in the 2nd interval.

37 (14.8%) students used *arithmetic mean* to determine the warmer interval between the 1st and 2nd interval. A sample response from participant 207 is presented below:

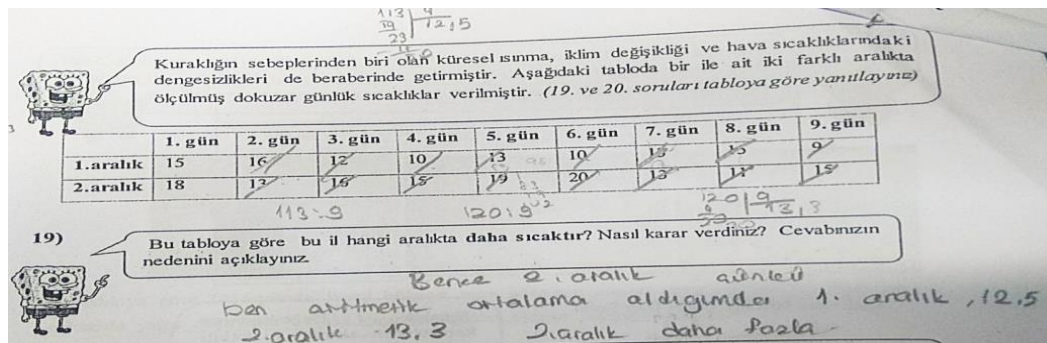


Figure 4.61. The statistical response of participant 207 to P17 in the SLT-CP

As seen, participant 207 responded as “I think 2nd interval since when I calculated the arithmetic mean, it was 12.5 for the 1st interval and 13.3 for the 2nd interval. 2nd interval is greater”. The student calculated the arithmetic mean for air temperatures in the 1st and the second interval by adding up the values of 9-day air temperatures and dividing the total by 9. Consequently, the student calculated the arithmetic mean for the 1st interval as 12.5⁰C and the arithmetic mean for the 2nd interval as 13.3 ⁰C. Since the arithmetic mean for air 9-day temperature is higher in the 2nd interval, the student stated that the city is warmer in the 2nd interval.

Furthermore, the *median value* was utilized by some students to compare the values of air temperature in the first and the second interval. These students ordered the 9-day temperatures in ascending or descending order, and they selected the value in the middle as 13 for the 1st interval and 15 for the 2nd interval. Then, they compared the two median values for both intervals and they decided that the city is warmer in the 2nd interval.

17 (6.8%) students considered the *difference* of between the air temperatures each day. To be more precise, these students looked for the difference in air temperature values by subtracting these values from each other day by day. Afterwards, the students evaluated the differences and decided that the city is warmer in the 2nd interval. To explain that kind of response in a more detailed manner, a sample response from participant 173 is presented below:

Kuraklığın sebeplerinden biri olan küresel ısınma, iklim değişikliği ve hava sıcaklıklarındaki dengesizlikleri de beraberinde getirmiştir. Aşağıdaki tabloda bir ile ait iki farklı aralıkta ölçülmüş dokuz günlük sıcaklıklar verilmiştir. (19. ve 20. soruları tabloya göre yanıtlayınız)

	1. gün	2. gün	3. gün	4. gün	5. gün	6. gün	7. gün	8. gün	9. gün
1. aralık	15	14	12	10	13	10	13	15	9
2. aralık	18	13	16	15	19	20	13	11	15

1. Bu tabloya göre bu il hangi aralıkta daha sıcaktır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

2. aralık daha sıcaktır. Çünkü farkları toplayıp 2. aralıktaki farklardan 1. aralıktaki farkları çıkarırsamda 1. aralıktan 2. aralığın farklarını çıkarmaktan daha büyük bir sayı çıkar.

Figure 4.62. The statistical response of participant 173 to P17 in the SLT-CP

As seen, participant 173 responded as “2nd interval is warmer. I added the differences and I subtracted the difference for 2nd interval from difference for 1st interval. This result was greater than the subtraction of the difference for 1st interval from 2nd interval”. The student preferred to compare the values of air temperature day by day for the 9-day interval. The student wrote the differences alongside the highest temperature with the plus sign. Then, the student added the differences for both intervals. Subsequently, the student subtracted the total for the first interval from the total for the second interval and then subtracted the total for the second interval from the total for the first interval. Finally, the student compared the results and decided that the total for the second interval is highest. Eventually, the student decided that the city is warmer in the second interval.

116 (46.4%) students questioned the statistical data pertaining to the results of the measurements of the variables A and B in 9 different intervals concordantly in the 14th problem of SLT. These students benefitted from *mode*, *median*, *arithmetic mean*, *total sum* and *difference* to find the variable with higher values among all 9 measurements. Similar to the 14th problem of the SLT-CP, the students mostly preferred arithmetic mean and the total sum.

Furthermore, some students utilized *total sum* to determine the variable with higher values in 9 different intervals in the 17th problem of the SLT. These students added up the values of the variables A and B and they obtained the total sum for these variables. Afterwards, they compared the total sum for both variables. A sample response from participant 481 is presented below:

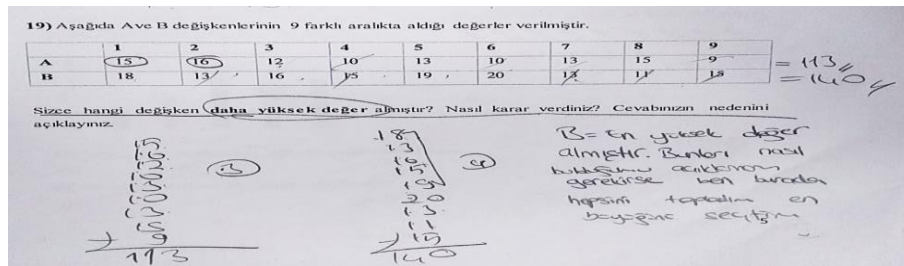


Figure 4.63. The statistical response of Participant 481 to P17 in the SLT

As seen, participant 481 responded as “B got the highest value. I added all and selected the biggest one”. The student preferred to benefit from total sum to find the variable with the higher values in 9 intervals. The student added up the values of A in 9 intervals, and then added up the values of B in 9 intervals. Subsequently, the student compared the total values of the variables A and B in 9 intervals and specified variable B as the variable taking the highest value.

Some of the students calculated the arithmetic mean for both the values of variable A and variable B in 9 intervals and compared these values and determined variable B as the variable taking the highest value. A sample response from participant 371 is presented below:

Aşağıda A ve B değişkenlerinin 9 farklı aralıkta aldığı değerler verilmiştir.

	1	2	3	4	5	6	7	8	9	
A	15	16	12	10	13	10	13	15	9	= 12,5
B	18	13	16	15	19	20	13	11	15	= 15,5

Sizce hangi değişken daha yüksek değer almıştır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

B daha yüksek değer almıştır. Aritmetik ortalamasını alıp karar verdim. B'nin ortalaması daha yüksektir.

$$A = \frac{117}{9} = 12,5$$

$$B = \frac{140}{9} = 15,5$$

Figure 4.64. The statistical response of participant 371 to P17 in the SLT

Participant 371 responded as “B got the higher value. I calculated the arithmetic mean to determine. The arithmetic mean of B is greater”. The student calculated the arithmetic mean for the values of variable A in 9 intervals as 12.5, and for variable B the arithmetic mean was found as 15.5. The participant compared the values of arithmetic mean and determined that variable B got a higher value than B in 9 intervals.

30 (12%) students made a statistically appropriate decision by indicating that the city is warmer in the 2nd interval, but they did not use statistical techniques to support their decision in the 17th problem of the SLT-CP. Instead, the students supported their decision by stating as:

“The numbers in the 2nd interval is higher than the numbers in the 1st interval” or “When I compared the intervals on the basis of each of 9 days, I saw that B has higher values than A”.

A sample response from participant 180 is as follows:

Kuraklığın sebeplerinden biri olan küresel ısınma, iklim değişikliği ve hava sıcaklıklarındaki dengesizlikleri de beraberinde getirmiştir. Aşağıdaki tabloda bir ile ait iki farklı aralıkta ölçülmüş dokuz günlük sıcaklıklar verilmiştir. (19. ve 20. soruları tabloya göre yanıtlayınız)

	1. gün	2. gün	3. gün	4. gün	5. gün	6. gün	7. gün	8. gün	9. gün
1. aralık	15	16	12	10	13	10	13	15	9
2. aralık	18	13	16	15	19	20	13	11	15

Bu tabloya göre bu il hangi aralıkta daha sıcaktır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

2. Aralık daha sıcaktır. Çünkü karşılaştırma yapıldığında 2. Aralığın değerleri daha yüksek.

Figure 4.65. The pre-statistical response of participant 180 to P17 in SLT-CP

As could be seen, participant 180 compared the values of air temperature day by day and made an approximate estimate by just looking at the values roughly. Subsequently, the student selected the second interval as having higher temperature than the first interval by stating as “The 2nd interval, because I made a comparison and I concluded that the values in the 2nd interval is higher”. Obviously, the student determined the interval with the higher temperature correctly and her statements were not erroneous. However, the student did not use some statistical concepts to support her decision such as arithmetic mean, mode, total sum, median or any other statistical concepts.

29 (11.6%) students specified variable B as taking the higher value than variable A, which was the appropriate decision. These students acted based on their anticipation in the decision making process by looking at the data. Instead of using statistical concepts to support their decision, they preferred to take a look at the data. A sample response from participant 393 is as follows:

1) Aşağıda A ve B değişkenlerinin 9 farklı aralıkta aldığı değerler verilmiştir.

	1	2	3	4	5	6	7	8	9	
A	15	16	12	10	13	10	13	15	9	=2
B	18	13	16	15	19	20	13	11	15	=6

Sizce hangi değişken daha yüksek değer almıştır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

B daha yüksek değer almıştır. Çünkü A'ya göre sayıları daha yüksektir.

Figure 4.66. The pre-statistical response of participant 393 to problem 17 in SLT

As can be seen, participant 393 decided that variable B got the higher value. The student justified her decision by stating:

“The numbers are higher in variable B compared to variable A”

Namely, the student compared the values of variable A and variable B in 9 intervals and she marked the variable with higher value in each of these intervals. Subsequently, the student counted the number of intervals that the variables were marked. Variable A was marked 2 times and variable B was marked 6 times. The student noted these numbers at the end of the rows belonging to the variables and she decided that variable B got higher values. Thus, the student correctly specified variable B as getting higher values by counting the number of intervals in which the variables get higher values. However, the student ignored the magnitude of the difference for each of the 9 intervals. The student could note the magnitudes of the differences. She could have made comparisons based on the magnitudes of the differences. Therefore, it could be inferred that the student overlooked the magnitudes of the differences in the values of variable A and variable B.

51 (20.4%) students determined that the city was warmer in 2nd interval by reading the data presented in the table included in the 14th problem of the SLT-CP. However, these students explained their reasoning with non-statistical expressions or they did

not provide a justification regarding the reason for the selection of the 2nd interval. They made explanations as follows:

“2nd interval. I chose the day with the highest temperature”

“2nd interval because it seems more”

“2nd interval because the peak value belongs to the 2nd interval”

Obviously, these responses were not an indication of statistical literacy regarding the table which shows the air temperatures of a city in two different intervals. A sample response from participant 177 is presented below:

Kuraklığın sebeplerinden biri olan küresel ısınma, iklim değişikliği ve hava sıcaklıklarındaki dengesizlikleri de beraberinde getirmiştir. Aşağıdaki tabloda bir ile ait iki farklı aralıkta ölçülmüş dokuz günlük sıcaklıklar verilmiştir. (19. ve 20. soruları tabloya göre yanıtlayınız)

	1. gün	2. gün	3. gün	4. gün	5. gün	6. gün	7. gün	8. gün	9. gün
1. aralık	15	16	12	10	13	10	13	15	9
2. aralık	18	13	16	15	19	20	13	11	15

Bu tabloya göre bu il hangi aralıkta daha sıcaktır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

2. aralıkta daha sıcak. Batarak inceledim. Büyük olanı seçtim.

Figure 4.67. The non-statistical response of participant 177 to P17 in SLT-CP

Participant 177 stated as *“The 2nd interval is warmer. I examined and selected the biggest one”*. Although the student specified the 2nd interval as the warmer interval correctly, the student could not provide a statistical justification about his answer.

53 (21.2%) students specified the variable B as the value taking the higher value, but they could not benefit from statistical constructs to choose between the values. Instead, the students wrote “*I looked at the table, and variable B is the answer*”. A sample response from 280 is presented below:

Aşağıda A ve B değişkenlerinin 9 farklı aralıkta aldığı değerler verilmiştir.

	1	2	3	4	5	6	7	8	9
A	15	16	12	10	13	10	13	15	9
B	18	13	16	15	19	20	13	11	15

Sizce hangi değişken daha yüksek değer almıştır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

Tabloya bakarak karar verdim.
B değişkeni, o daha çok artımdan diğerine göre.

Figure 4.68. The non-statistical response of participant 280 to P17 in the SLT

Participant 280 stated as “*I looked at the table and saw that variable B increases more, so the answer is variable B*”. As seen, the student observed the increase in the values of variables A and B. However, in the problem, the students were asked which variable gets higher value instead of asking which variable increased more. Although the student selected the appropriate variable, the student interpreted the problem erroneously and followed the inappropriate method, which is non-statistical.

55 (22%) students provided an unrelated response or they left the problem unanswered for the 14th problem of the SLT-CP. Some of these students selected the 1st interval as the warmer interval, which is the incorrect answer. Moreover, some students selected a day instead of specifying an interval by misinterpreting the problem. A sample response from participant 129 is presented below:

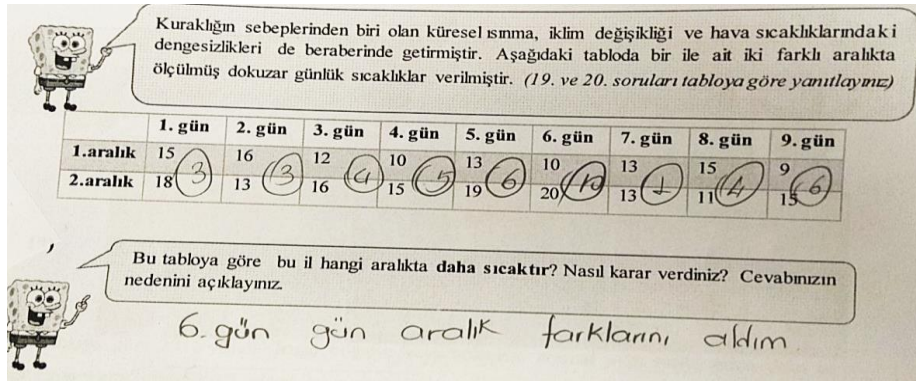


Figure 4.69. The unrelated response of participant 129 to P17 in the SLT-CP

Participant 129 stated as “6th day, I calculated the differences for the intervals”. As could be seen, the student took the difference of the values of air temperatures for each day and decided the 6th day as the answer since the difference was the highest in 6th day. However, instead of the warmer day, the warmer interval was asked. It could be inferred that the student misinterpreted the problem and gave an unrelated answer to the 14th problem of the SLT-CP.

52 (20.8%) students left the 14th problem of the SLT unanswered or gave an unrelated answer to the problem. Some of these students tried to specify the interval with the highest value instead of finding the variable getting the higher value at 9 intervals. Furthermore, some of these students determined variable A as the variable having higher values, which is the incorrect answer. A sample response from participant 396 is presented below:

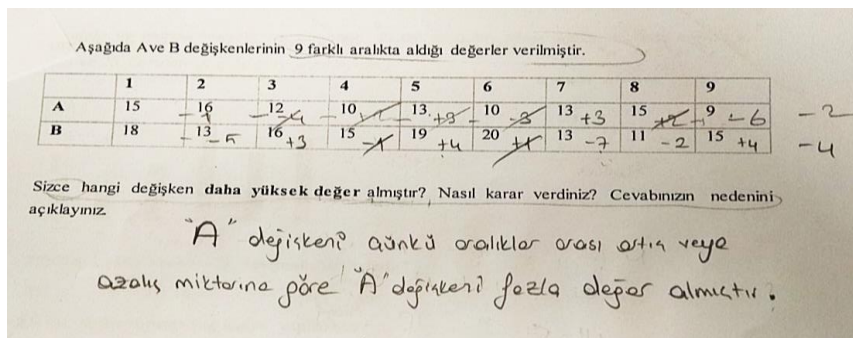


Figure 4.70. The unrelated response of participant 396 to P17 in the SLT

Participant 396 stated as “*Variable A. According to increase and decrease in intervals Variable A got the greater value*”. As could be seen, participant 396 observed the progression of variable A and variable B according to 9 intervals. The student noted the increase and decrease in these variables and at the end the student added these changes. Then, the student compared the total of the changes and decided that variable A got the higher value. To observe the changes in the variables, it was not meaningful to make a comparison between two variables. The student should have compared the values of both variables at each of the 9 intervals. Therefore, it could be deduced that the student provided an unrelated answer to the 14th problem of the SLT.

4.2.4.4 The 8th Grade Students Statistical Literacy in Problem 19

In the 19th problem of the SLT-CP, 'water shortage' values between the years 2010-2040 for several countries reported by the World Resource Institute are presented. According to the report, the countries of Botswana, Chile, Estonia and Namibia have been suffering from thirst. These countries' water shortage values were presented for the years of 2010, 2020, 2030 and 2040. According to the problem, a researcher wants to make an investigation in the country where water shortage is experienced the most. The students were asked which country the researcher should choose to make the right decision. To obtain the 19th problem of the SLT, the problem context of water shortage values for some countries in the years of 2010, 2020, 2030 and 2040 was removed from the 19th problem in the SLT-CP by keeping the statistics the same. Instead of providing water shortage values of four countries, the data was presented in four rows and the students were requested to specify the row with the highest value overall. Besides, the students were required to explain the reason for their answer.

Table 4.16 Student Responses to the 19th problem

		Problem 19 with context		Problem 19 without context	
Classification of Responses	Student Response	f	p	f	p
Statistical	Summarization of the data and use of the appropriate values in the table. <i>-mode</i> <i>median</i> <i>arithmetic mean</i> <i>total sum</i>	45	18%	79	31.6%
Pre-statistical	Reading data from the table. <i>-looking for the increase</i>	36	14.4%	63	25.2%
Non-statistical	Interpreting the data with non-statistical, idiosyncratic expressions. <i>-methods other than central tendency measures</i> <i>- peak value</i>	81	32.4%	51	20.4%
Unrelated/Blank		88	35.2%	57	22.8%

45 (18%) students questioned and critically evaluated the excerpt from the media report prepared by the World Resource Institute to make a suggestion to the researcher. More specifically, these students determined the country where water shortage is most experienced by benefitting from statistical procedures and concepts. They justified their suggestion for the researcher by utilizing statistical constructs such as arithmetic mean, total sum and statistical comparison methods. Initially, some students utilized the total sum to find out the country where water shortage is

experienced most. These students preferred to add up all the water shortage values for each of the four countries. Subsequently, the students compared the total of water shortage values and specified the country with the highest water shortage value. A sample response from participant 91 is presented below:

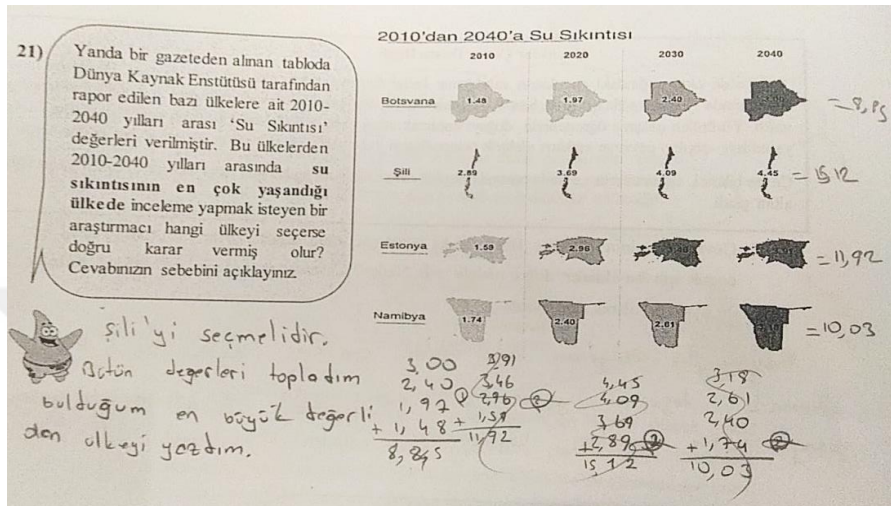


Figure 4.71. The statistical response of participant 91 to P19 in the SLT-CP

Participant 91 stated as “The researcher should select Chile. I added the all values and I selected the country which have the highest value”. The student calculated the sum of water shortage values from 2010 to 2040 for the countries of Botswana, Chile, Estonia and Namibia. The sum was 8.85 for Botswana; 15.12 for Chile, 11.92 for Estonia and 9.93 for Namibia. The student selected Chile as the country where water shortage is experienced the most since the total sum of water shortage values was the highest for Chile among the four countries. As could be seen, the student used a statistical method to help the researcher determine the appropriate country where water shortage is experienced most.

Moreover, some students compared the water shortages of the countries on a yearly basis. Instead of an overall evaluation, these students preferred to compare the water shortage values of the countries for each year. To be more specific, a sample response from participant 174 is presented below:

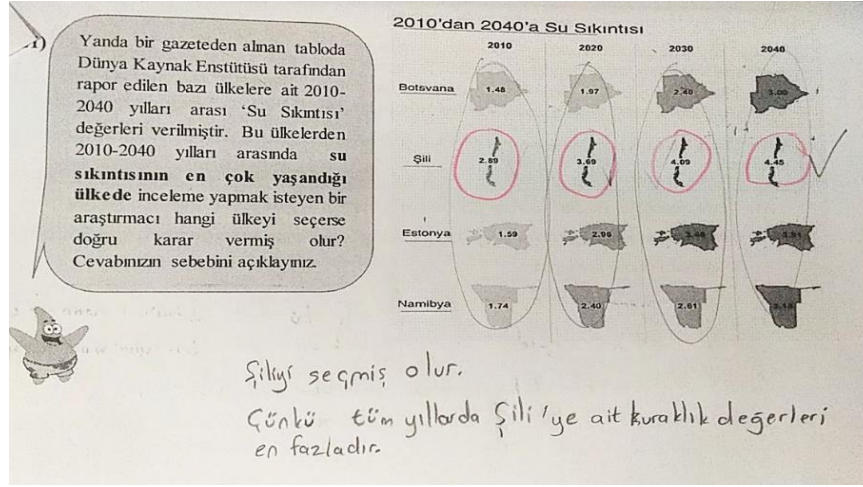


Figure 4.72. The statistical response of participant 174 to P19 in the SLT-CP

Participant 174 stated as “*The researcher should select Chile since for all years the values regarding Chile is highest*”. As can be seen, participant 174 preferred to compare the countries’ water shortage values year by year. Initially, the student looked at the year of 2010 column and he compared the water shortage values. Then, the student marked Chile as having the highest water shortage value. The student went through the same process for 2020, 2030 and 2040. The student marked Chile for all the years from 2010 to 2040. Finally, the student decided that Chile is the country where water shortage values are the highest. The student recommended Chile to the researcher. Obviously, the student was able to make comparisons between the statistics given in the media excerpt and take appropriate decisions within the framework of these comparisons.

The number of students preferring the arithmetic mean was restricted for the 19th problem of the SLT-CP. When the students’ responses were examined, it was seen that the students left the operations unfinished and turned to making interpretations through total sum instead of arithmetic mean. Overall, it was observed that the students mostly preferred to use arithmetic mean to make statistical inferences from the data because the students were accustomed to arithmetic mean in their daily lives. However, for this problem it was seen that the students avoided using arithmetic

mean. Most probably, this situation stems from the fact that the water shortage values were in the form of decimal numbers and making division with decimal numbers compelled students. Therefore, the students might have considered comparing the water shortage values with the total sum of these values for each of these countries.

79 (31.6%) students interpreted and questioned the statistical data regarding the given values of 4 rows in the 19th problem of the SLT. These students utilized arithmetic mean, total sum and statistical comparison methods to determine the row with the highest values. Some students used the total sum to find the row with the highest values. A sample response from participant 348 is as follows:

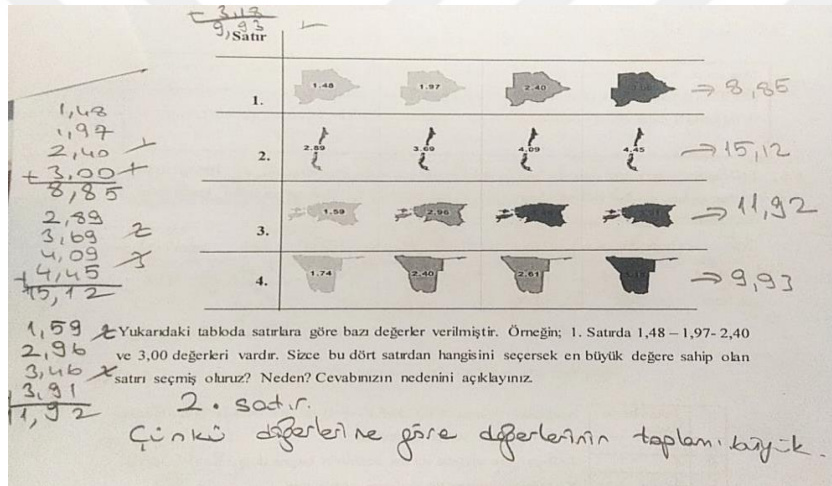


Figure 4.73. The statistical response of participant 348 to P19 in the SLT

Participant 348 responded as "2nd line since the total for this interval is higher than the others". As seen, the student calculated the total sum for each of the four intervals separately. The total sum was 8.85 for the 1st row; 15.12 for the 2nd row, 11.92 for the 3rd row and 9.93 for the 4th row. The student selected the 2nd row as the row having the highest values correctly.

Some of these students preferred to compare the values belonging to rows according to the values in the columns. More specifically, the students looked at the first values

of each row and selected the 2nd row as getting the highest value, and then they observed the second values, third values and the fourth values of the rows and they found that each time the 2nd row got the highest value. Consequently, they decided that the 2nd row got the highest value among all the 4 rows.

36 (14.4%) students read the data presented in the report appropriately, but they did not utilize the certain statistical techniques to find the country where water shortage is experienced most in the 19th problem of the SLT-CP. Some of the students looked for the increase in the water shortage values of the countries Botswana, Chile, Estonia and Namibia. More specifically, these students guessed the country where water shortage is experienced the most as the country where the water shortage values increased mostly without considering the initial water shortage values. These students disregarded the fact that the water shortage values in 2010 is different for each of the countries. To exemplify this approach, a sample response from participant 299 is presented below:

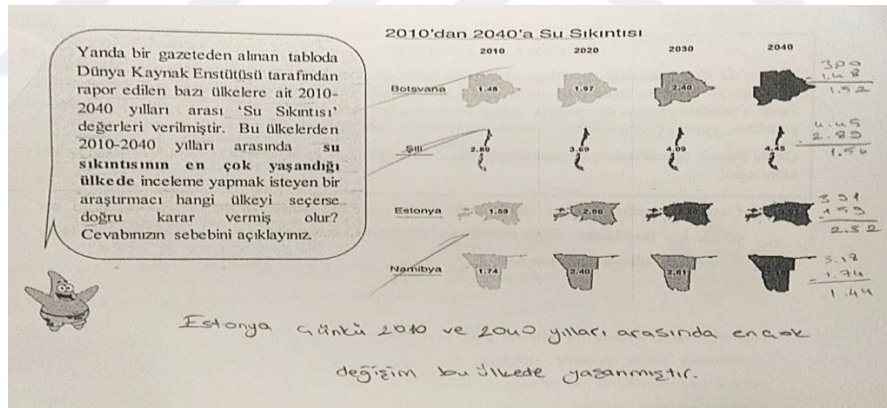


Figure 4.74. The pre-statistical response of participant 299 to P19 in SLT-CP

Participant 299 responded as “Estonia since the biggest change between the years of 2010 and 2040 was experienced in this country”. As could be observed, the student investigated the change in water shortage values between the years 2010 and 2040. In accordance with this purpose, the student subtracted the water shortage value in

2040 from the water shortage value in 2020. For instance, for Estonia the water shortage value in 2040 was 3.91 and the water shortage value for 2010 was 1.59. The student observed the change by subtracting 1.59 from 3.92. By doing this, the student obtained a value representing the change in water shortage values for each of the countries. The change was 1.52 for Botswana; 1.56 for Chile; 2.32 for Estonia and 1.44 for Namibia. Therefore, the student specified Estonia as the country where water shortage is experienced the most. In fact, Estonia was not the country where water shortage is experienced the most. However, the student approached the problem from a different viewpoint. It is obvious that the integration of problem context provided opportunity to students for deriving multiple entries and exit points to solve problems. In the 19th problem of the SLT-CP, the students' answers varied according to their approaches to the situation in the problem.

63 (25.2%) students read the data involving the values of four rows in the 19th problem of the SLT. These students did not benefit from the definite statistical concepts while deciding on a row having the highest values. More specifically, the students specified the 2nd row as the row with the highest value which is the correct answer for the problem. However, they did not use statistical methods such as arithmetic mean or total sum. A sample response from participant 417 is presented below:

Satır				
1.	1.48	1.97	2.40	
2.	2.95	3.00	3.00	3.00
3.	1.59	2.95	2.95	2.95
4.	1.24	2.40	2.91	

Yukarıdaki tabloda satırlara göre bazı değerler verilmiştir. Örneğin; 1. Satırda 1,48 – 1,97- 2,40 ve 3,00 değerleri vardır. Sizce bu dört satırdan hangisini seçersek en büyük değere sahip olan satırı seçmiş oluruz? Neden? Cevabınızın nedenini açıklayınız.

Bence 2. satırı seçersek en büyük değere sahip olan satırı seçmiş oluruz. Çünkü neredeyse en büyük değerler o satırda yer alıyor.

Figure 4.75. The pre-statistical response of participant 417 to P19 in the SLT

As could be seen, the student specified the 2nd row as the row with the highest values which is the appropriate answer to the problem. The student stated as *“I think if we choose the 2nd row, we will choose the line with the greatest value because about all the high numbers are placed in the 2nd row”*

As seen, the student looked at the values in the rows from a broad perspective and decided that the values in the 2nd row are the highest.

81 (32.4%) students interpreted the data with non-statistical methods. For instance, these students selected the country with the peak water shortage value among the countries in any year between 2010 and 2040. More specifically, when the students looked at the table, they observed 16 values in the table regarding four countries and four years. The students looked for the highest water shortage value among these 16 values and specified 4.45 as the highest water shortage value which belongs to Chile. Chile was the correct answer to the problem. However, these students did not make an overall observation of the water shortage values in the course of progression from 2010 to 2020. These students made an instant decision by focusing on only one year among four years. Therefore, although these students specified Chile as the country where water shortage is experienced the most, they did not exhibit a statistical approach to find the correct country.

Moreover, some students who gave non-statistical answers to the 19th problem of the SLT-CP determined Chile as the country where water shortage is experienced the most. However, these students selected Chile by looking at the countries' surface areas. A sample response from participant 129 is as follows:

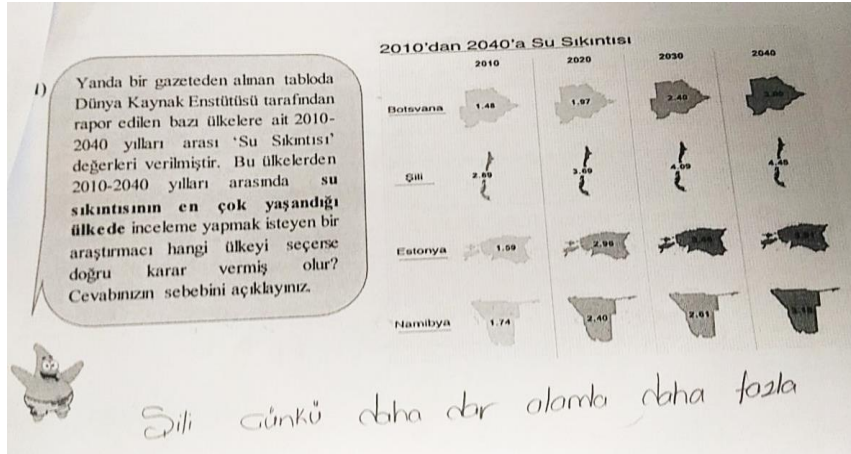


Figure 4.76. The non-statistical response of participant 129 to P19 in SLT-CP

Participant 129 stated as “*Chile since more in the narrower region*”. As could be seen, participant 129 selected Chile as the country where water shortage is experienced the most. The student explained her rationale by stating as “*Chile, because it experienced the highest amount of water shortage in the smallest surface area*”

As seen, the student preferred to decide according to the water shortage value per meter square. Instead, the student should have compared the overall water shortage value.

Additionally, it is observed that some students compared the water shortage values of the countries by considering only the year 2040. Probably, these students considered that the water shortage value in 2040 is the most current value indicating the countries' water shortage situations. A sample response from participant 217 is as follows:

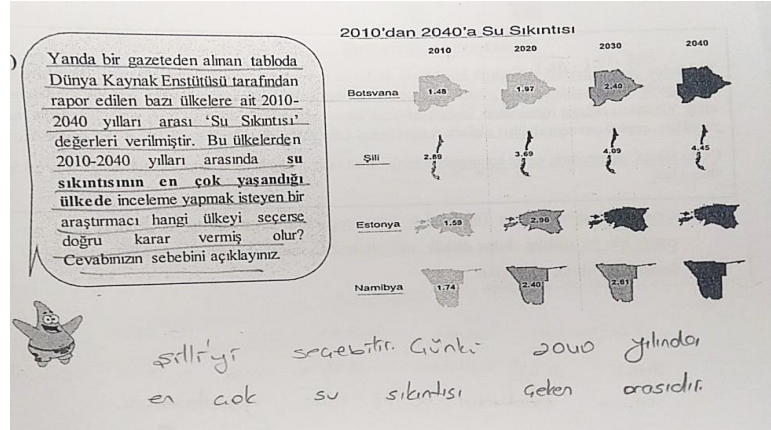


Figure 4.77. The non-statistical response of participant 217 to P19 in SLT-CP

The participant 217 stated that:

“The researcher could select Chile because Chile is the country suffering from water shortage the most in 2040.”

As seen, the student observed the countries' water shortage values only in the year of 2040 by ignoring the values in 2010, 2020 and 2030. In fact, in the problem the cumulative of the water shortage of the countries was questioned. The student should have observed the overall water shortage values from 2010 to 2040.

51(20.4%) students used non-statistical methods to determine the row with the highest value. These students selected the 2nd row which is the appropriate row, but they looked at all the values in the table and selected the row with the highest value by not considering overall values of the rows. The students might have benefitted from arithmetic mean, total sum or the statistical difference to make a comparison between the rows. A sample response from participant 323 is presented below:

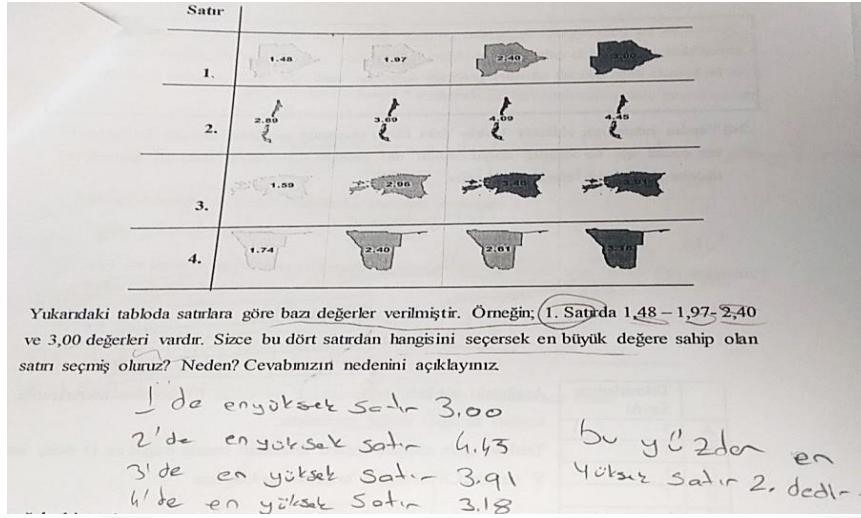


Figure 4.78. The non-statistical response of participant 323 to P19 in SLT-CP

As seen, participant 323 specified the peak values for each of the rows. More specifically, the student looked at the first row and selected the peak value as 3.00. Subsequently, the student proceeded with the second, third and the fourth rows. The peak value for the 2nd row was 4.45; 3.91 for the 3rd row and 3.18 for the 4th row. Among these peak values, 4.45 was the highest and the student chose the 2nd row including the value of 4.45. As could be seen, the student did not benefit from the statistical techniques to compare the overall values of the rows. For one measurement, the value belonging to the 2nd row could be the highest, but the student might have conducted an overall evaluation including all of the four measurements. Instead, the student preferred to use peak values to choose the row with the highest values.

88 (35.2%) students provided an unrelated response or left the 19th problem of the SLT-CP unanswered. Except for the students who gave no response, the others misunderstood or misinterpreted the problem situation. These students focused on irrelevant points in the given media report. For instance, some students thought that they should specify the country with the lowest water shortage values between 2010 and 2040. A sample response from participant 215 is as follows:

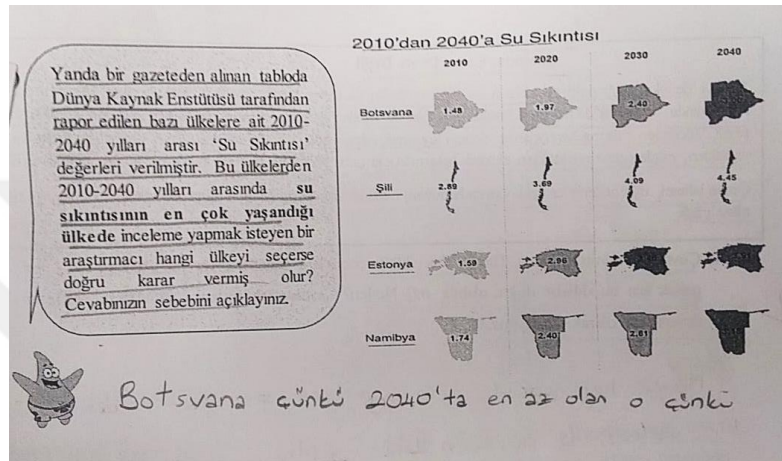


Figure 4.79. The unrelated response of participant 215 to P19 in the SLT-CP

Participant 215 selected Botswana as the country which suffered from water shortage the least in 2040 by stating as “*Botswana, because in 2040 this country suffered the least*”. As seen, the student thought that she should select the country with the lowest water shortage values. Indeed, the student should have compared the cumulative water shortage values of the countries.

57 (22.8%) students did not answer the 19th problem of the SLT or they provided an unrelated answer. These students could not comprehend the problem situation and therefore they gave responses disregarding the problem situation by focusing on the irrelevant aspects of the given data. A sample response from participant 336 is presented below:

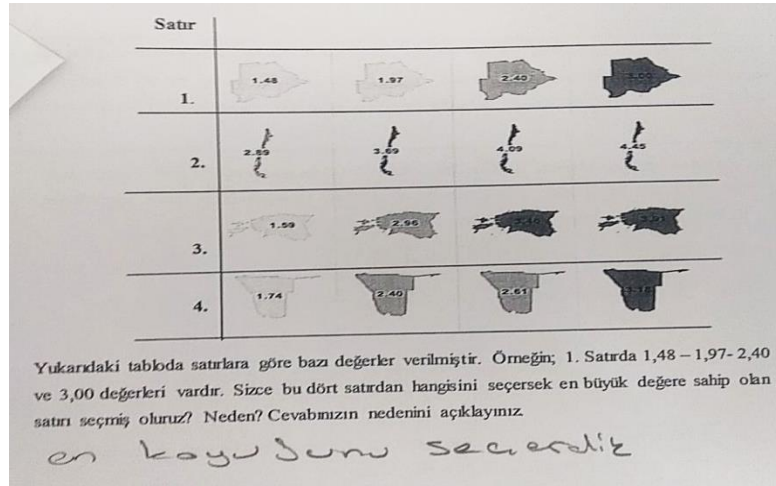


Figure 4.80. The unrelated response of participant 336 to P19 in SLT-CP

Participant 336 stated as “*We could choose the darkest color*”. As could be seen, participant 336 somehow built a relationship between the tone of the color of the areas and the intensity of water shortage. Actually, there was not a relationship between the tone of the color and the water shortage values of the countries. The students’ response was not reasonable in terms of the problem situation. Therefore, we can infer that the student provided a non-statistical response to the 19th problem of the SLT.

4.3 Comparison between Problems with Context and Problems without Context

Descriptive statistics for statistical literacy problems with context and problems without context are presented in Table 4.17. The maximum score in both tests was 75 by including 25 problems over 3 points, and the minimum score was 0. The mean for the problems without context is 40.30, whereas the mean for the problems with context is 39.63.

Table 4.17 Descriptive statistics for statistical literacy problems with context and problems without context

	N	Mean	SD	SE	Skewness	Kurtosis
Without context	250	40.30	12.98	.82	-0.07	-0.45
With context	250	39.63	13.44	.85	-0.15	-0.59

An independent sample t-test was conducted to compare the means for students' achievement in problems with and without context (Pallant, 2007). The results are presented in Table 4.17. Initially, assumptions for the tests were examined in terms of overall scores in SLT and SLT-CP and scores in terms of statistical content domains. For normality, as can be seen in the table, Skewness and Kurtosis values for problems with and without context are between -2 and +2. Additionally, the histograms showed normal distribution. Therefore, the normality assumption is met. The scores for the problems with and without context were independent and the students did not influence each other. The researcher and the teachers ensured the independence of observations. Also, the students' achievement was continuous.

Table 4.18 Independent sample t-test for total mean scores in problems with context and problems without context

	Levene's Test for Equality of Variances		t-test for Equality of Means		
	F	Sig.	t	df	Sig.(two-tailed)
Equal variances assumed	.366	.545	.569	498	.570
Equal variances not assumed			.569	497.395	.570

As a result, there was no significant difference in terms of achievement in problems without context ($M=40.3$, $SD= 12.98$) and problems with context ($M=39.63$, $SD=13.44$); $t(498) = .57$; $p=.57$ (two-tailed). The p value is greater than the critical value (0.05) whether or not the variances of the samples are equal. Therefore, by considering both tests, the students' achievement in statistical literacy did not differ according to the existence of context in the problem.

The study proceeded with in-depth content domain-based and question-based investigation to reveal the possible reasons for the result of the independent sample t -test for total scores.

Next, independent sample t -test was conducted for achievement in statistical literacy problems with and without context according to four statistical content domains. Therefore, the difference in mean scores both tests was questioned under the content domains of measures of central tendency, graph interpretation, sampling, and questioning of statistical data. Assumptions for independent sample t - test were checked together with total scores as explained previously.

4.3.1 Mean Difference in Measures of Central Tendency

According to the measures of central tendency content domain, 500 eighth grade students' mean achievement scores in statistical literacy problems with context and without context were compared with independent sample t -test. The tests included five measures of central tendency problems with 15 points in total. The results are presented in Table 4.19.

Table 4.19 Independent sample t-test for mean scores in problems with context and problems without context according to the measures of central tendency domain

Measures of Central Tendency		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig.(two-tailed)	Mean difference	Std.error difference
	Equal variances assumed	1.939	.164	-.095	498	.924	-.032	.337
	Equal variances not assumed			-.095	495.402	.924	-.032	.337

As can be seen Table 4.19, there was no significant difference in achievement in problems without context ($M=5.56$, $SD= 3.90$) and problems with context ($M=5.59$, $SD=3.63$); $t(498) = -.095$; $p=.92$ (two-tailed). The p-value is greater than the critical value (0.05) whether or not the variances of the samples are equal. Therefore, by considering both tests, the students' achievement in statistical literacy did not differ based on the existence of context in measures of central tendency problems.

4.3.2 Mean Difference in Graph Interpretation

As far as the graph interpretation domain is concerned, it can be said that 500 eight grade students' mean achievement scores in statistical literacy problems with context and without context were compared with independent sample t-test. The tests

included 11 graph interpretation problems with 33 points in total. The results are presented in Table 4.20.

Table 4.20 Independent sample t-test for mean scores in problems with context and problems without context according to the graph interpretation domain

Graph Interpretation		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig.(two-tailed)	Mean difference	Std.error difference
	Equal variances assumed	.034	.853	.689	498	.491	.388	.563
	Equal variances not assumed			.689	498	.491	.388	.563

According to Table 4.20, there was no significant difference in achievement in problems without context ($M=21.98$, $SD= 6.30$) and problems with context ($M=21.60$, $SD=6.30$); $t(498) = -.69$; $p=.49$ (two-tailed). The p-value is bigger than 0.05. Therefore, by considering both tests, the students' achievement in statistical literacy did not differ depending on the existence of context in graph interpretation problems.

4.3.3 Mean Difference in Sampling

As for the sampling domain, 500 eighth grade students' mean achievement scores in statistical literacy problems with context and without context were compared with

independent sample t-test. The tests included 2 graph sampling problems with 6 points in total. The results are presented in Table 4.21.

Table 4.21 Independent sample t-test for mean scores in problems with context and problems without context according to the sampling domain

Sampling		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig.(two-tailed)	Mean difference	Std.error difference
	Equal variances assumed	7.871	.005	2.120	498	.034	.368	.174
	Equal variances not assumed			2.120	490.6670.	.034	.368	.174

As can be seen in Table 4.21, a significant difference was found in terms of achievement in sampling problems without context ($M=3.83$, $SD= 1.82$) and with context ($M=3.46$, $SD=2.06$); $t(498) = 2.12$; $p=.034$ (two-tailed). The magnitude of differences in means was small ($\eta^2=.009$). The p-value is less than 0.05. Therefore, by taking both tests into consideration, the students' achievement in statistical literacy problems differed depending on the existence of context in the sampling content domain in favour of the problems without context. In other words, the students performed better in problems without context in the sampling content domain problems.

4.3.4 Mean Difference in Questioning of Statistical Data

As far as the graph interpretation domain is concerned, 500 eighth grade students' mean achievement scores in statistical literacy problems with context and without context were compared with independent sample t-test. The tests included 7 questioning of statistical data problems with 21 points in total. The results are presented in Table 4.22.

Table 4.22 Independent sample t-test for mean scores in problems with context and problems without context according to the questioning of statistical data domain

Questioning Of Statistical Data		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig.(two-tailed)	Mean difference	Std.error difference
	Equal variances assumed	.762	.383	-1.31	498	.896	-.052	.398
	Equal variances not assumed			-1.31	497.42	.896	-.052	.398

As seen in Table 4.22, a significant difference was not found in terms of achievement in questioning of statistical data problems without context ($M=8.93$, $SD= 4.38$) and with context ($M=8.98$, $SD=4.53$); $t(498) = -1.31$; $p=.90$ (two-tailed). The p-value is bigger than 0.05. Hence, by considering both tests, the students' achievement in statistical literacy did not differ depending on the existence of context in the questioning of statistical data problems.

CHAPTER 5

DISCUSSION

The aims of the current study were to investigate 8th grade students' achievement and statistical literacy in problems with and without context in terms of statistical content domains in elementary mathematics curriculum. Initially, the frequency of the 8th grade students' correct answers and the percentages of achievement for each problem were investigated. Then, these students' statistical literacy in problems with context and problems without context was investigated by examining student solutions and categorizing these solutions. Subsequently, the mean difference in students' achievement in problems with and without context was examined by employing the independent sample t-test.

In this section, the findings of the study were discussed, and compared and contrasted with the research studies in the literature. Additionally, implications for educational practices were explored and several recommendations were made for further research studies considering the findings of the current study.

5.1 8th Grade Students' Achievement and Statistical Literacy in terms of Statistical Content Domains

In this section, the 8th grade students' achievement and statistical literacy are discussed in terms of the statistical content domains, which are measures of central tendency, graph interpretation, sampling and questioning of statistical data problems.

5.1.1 Achievement and Statistical Literacy in Measures of Central Tendency

The students were asked five problems related to the measures of central tendency content domain. These problems were P8, P9, P11, P12 and P13. Initially, the students' achievement in measures of central tendency problems in both tests was examined. Then, the students' statistical literacy was investigated by analyzing and categorizing the content of the student solutions in both statistical literacy tests (SLT and SLT-CP).

When the students' responses were examined, it was found that a few students evaluated the concept of average in terms of *representativeness*. Similar to the findings of the studies of Mokros and Russell (1995) and Yolcu (2012), most of the students could not perceive average as the representative of a data set.

Some students used *mean*, *mode* and *median* to interpret the concept of average. These students most frequently preferred to use *mean* to define and interpret average. This result of the study was parallel with the findings of previous research studies (Çatman-Aksoy, 2018; Enisoğlu, 2014; Uçar & Akdoğan, 2009; Watson, 2006; Yolcu, 2012). The reason behind the demand for the mean could stem from the elementary mathematic curriculum in Turkey. According to the MoNE (2018), the concept of arithmetic mean is included in the curriculum before the concepts of mode and median. More specifically, the concept of arithmetic mean appears in the data and chance curriculum in the 6th grade, while the concepts of mode and median appear in the 7th grade for the first time. The students build the concepts of mode and median on the concept of mean. Therefore, they possibly relate the concept of average with mean because of the reinforcement of mean occurs earlier than mode and median. Moreover, in the textbooks, the concepts of average and arithmetic mean are used interchangeably. This use in textbooks brings about students' anticipation of mean as the same as average, and could be achieved by encouraging

students to use mode and median by providing situations in which mode and median best represent the data (Seftlage, 2007). For instance, if the data includes extreme values and represents skewed distribution, the use of *median* could be suitable to represent the data set. Similarly, if the data involves undetermined values, *mean* could best represent the data. Moreover, if the given data is discrete rather than continuous, *mode* could best represent the given data set. Consequently, students should be presented with this kind of data by the educators to specify the best representative measure of central tendency according to the given data.

Different from the findings of the present study, several research studies concluded that *mode* and *median* are the most frequently employed measures of central tendency to interpret and define the concept of average (Mokros & Russell, 1995; Watson & Moritz, 2000). In the study of Mokros and Russell (1995), the 8th grade students interpreted average as the midpoint of the data set. The diversity in students' preference among the measures of central tendency could be depending on their out-of-school experiences or the context and content of the statistics problems.

As a result of the examination of students' responses, it was found that a vast majority of the students used non-statistical statements of '*approximately*', '*okay*', '*about*', '*normal*' and '*more or less*' to interpret and identify their understanding of average. This finding of the study was consistent with the findings of Watson and Moritz (2000).

Many students' conception of average was limited with the algorithmic procedure of *add-them-all-up-and-divide* (Shaugnessy, 1992), in other words with '*what you get from adding and dividing*' (Gal, 2005, p.3) rather than conceptual understanding and interpretations of average. This finding is consistent with the research studies in the literature (Cai, Moyer & Grochowski, 1999; Konold & Higgins, 2003; Pollatsek, Lima, & Well, 1981; Strauss & Bichler, 1988; Uçar & Akdoğan, 2009; Yıldırım, 2006). This situation might be due to the fact that students are exposed to the algorithmic procedures to find the average in early grades. Even if students cannot

reinforce the comprehension of the concept of average, they are subject to procedural algorithms to obtain the numerical value of average. The same situation applies to all college level students (Randall, 2006). Randall (2006) recommends teachers to pay attention to not to have a tendency towards the use of numerical value of mean to make inferences about average.

Some of the students providing incorrect responses to the problems including average mentioned their personal beliefs instead of providing a statistical response. A similar result was obtained by Mokros and Russell (1995). The students' misconceptions and misinterpretations regarding the "average" problems in the statistical literacy tests might stem from the deficiencies in conceptual knowledge of average rather than procedural knowledge (Cai, Moyer & Grochowski, 1999; Mokros & Russell, 1995). Thus, students should develop a relational understanding towards the concept of average instead of building an instrumental understanding by making connections with the interrelated schemas (Pollatsek, Lima, & Well, 1981; Skemp, 1979).

The findings of the current study indicated that while finding the *median* of a data set, the students mostly utilized algorithmic procedures rather than conceptual procedures. Most of the students attempted to calculate the numerical value of the median rather than interpreting the median of the data set within the problem context. The students placed the values in ascending and descending order and specified the value in the middle of the data set. However, some of the students made errors while finding the median of the data set. Initially, some students considered the repetitive values as repeated once. In the SLT-CP, some students wrote the name of the city instead of writing the numerical value. These students ignored that the median of the data set should be numerical data instead of categorical data. These findings confirm the results of the study by Sorto and White (2004) conducted with prospective teachers.

When the students were asked to find and interpret the *mode* of a given data set presented with the bar graph, the majority of the students provided an incorrect response. Most of the incorrect responses resulted from the confusion between the *peak value* of the data set and the *mode* of the data set. Put differently, the students found the peak value in the data set instead of specifying the mode of the data set. This confusion might have stemmed from using the same expression for mode and peak value in Turkish language. In elementary mathematics textbooks, the mode of a data set is stated as the “peak value”. Therefore, the students perceive mode as the highest value in the data set instead of the most recurrent value.

5.1.2 Achievement and Statistical Literacy in Graph Interpretation

The students were asked eleven problems related to the graph interpretation content domain. These problems were P2, P3, P4, P5, P6, P7, P15, P16, P18, P20 and P21.

When the students’ solutions in problem 2 were examined, it was found that most of the students appropriately determined the most appropriate graph to represent categorical data. The findings of the present study revealed that a vast majority of the students selected the pie chart to represent categorical data, whereas the others selected the bar graph, which is another appropriate graph to represent categorical data. This might be due to the fact that in textbooks, pie charts might have been given more place than bar graphs. Another reason could be that teachers might be inclined towards the use of pie chart during the class. When it comes to the construction of the selected graph in problem 3, more than half of the students successfully constructed the selected graph. While constructing the pie chart, the most common error was exceeding 100%. More specifically, the students did not notice that the sum of the percentages in the segments of the pie chart should be 100%.

When the interpretations and the unusual situation regarding the pie chart were asked in problem 15, a few students both interpreted the data given in the pie chart and

detected the error in the pie chart. These findings including the detection of the error indicate higher level literacy (Watson, 1996). Therefore, these students' responses correspond to the *third-tier of Watson's (1996) hierarchy* of statistical literacy and also *critical level* of statistical literacy in the study of Watson and Callingham (2003).

Some students either interpreted the data in the pie chart or specified the error regarding the total of the percentages of the segments of the pie chart by mentioning the statistics, which corresponds to the *second-tier of Watson's (1996) hierarchy* of statistical literacy and similarly *consistent non-critical level* of statistical literacy in the study Watson and Callingham (2003).

A significant number of students provided non-statistical or unrelated responses including only *wrong* and *right*, which do not indicate statistical literacy towards the reading and interpretation of the pie chart. This finding was consistent with the findings of Watson (1996), who examined students' statistical literacy by asking the same problem as problem 15 in the instruments. Moreover, some of these students confused the percentages of the segments with the central angle of the segments of the pie chart. To be more precise, the students assumed the whole as 360° instead 100%. This finding was in agreement with the findings of Kranda (2018), who investigated students' literacy about graphs including line graphs. It was common among these students that the pie chart should be represented with the central angle not with percentages. This might be due to the fact that in textbooks, pie charts are mostly represented with angle and the segments of the pie chart are expressed by the parts of 360° . To overcome this widespread misinterpretation among the students, the teachers might include different representations of pie chart by involving percentages in classroom practices. In this way, students might be more familiar with the representation of pie chart with percentage.

In the current study, the students appropriately interpreted the data given in the line graph from problem 4 to problem 7. There is not a significant problem regarding the students' interpretation of the line graph. This finding corresponds to the study of

Memnun (2011), who investigated students' achievement in various line graphs. In problem 6, about half of the students interpreted the inverse proportion between the variables by providing statistical responses. In line graph problems, some students were confused about the variables on the x-axis and y-axis. More precisely, when the students were asked whether the most recurrent source of environmental pollution was water pollution in 2014, the students sought to find the year with the highest value of water pollution. This finding was parallel with the study of delMas, Garfield and Ooms (2005) which was conducted with histograms and bar graphs. In a similar way, the students were confused about the variables on x-axis and y-axis. Additionally, in problem 18, the majority of the students were successful about converting the data presented in the table to a line graph. A similar finding was obtained by delMas, Garfield and Ooms (2005) about the conversion of the identical data interchangeably with the bar graph and histogram.

When the students' answers to the bar graph interpretation problems were examined, it was observed that most students could read and interpret the data presented in the bar graph. However, the students made several common errors while reading and interpreting the data given in the bar graph. Initially, the students made interpretations according to the data in the horizontal axis instead of using the data in the vertical axis. For instance, in problem 20 which asks about the number of cities with water pollution level higher than 3, the students added the water pollution levels above 3 instead of adding the number of cities with water pollution levels of 4 and 5. These findings confirm the findings of the study of delMas, Garfield and Ooms (2005) which involved the same problems with problem 20 and problem 21 in the instruments. They also revealed that many students specified the score values instead of the frequency of scores for each value. Moreover, in problem 20, some students specified the number of cities belonging to level 3 instead of adding the number of cities with level 4 and level 5 water pollution. Some of the students also took level 3 into consideration while specifying the levels more than 3. In fact, level 3 should not have been included in the solutions as the problem was about water pollution levels

higher than 3. Additionally, when the total number of cities was asked in problem 21, some students specified the peak value among the number of cities instead of adding up the number of cities belonging to all levels. Furthermore, in problem 21, a common mistake was that the students added up the number of levels instead of adding up number of cities. This might be considered as the confusion about the vertical axis and horizontal axis as mentioned by delMas, Garfield and Ooms (2005).

5.1.3 Achievement and Statistical Literacy in Sampling

The students were asked two problems regarding the sampling content domain. In the instruments, P10 and P22 were the sampling problems. As a result of the study, a significant difference was observed in students' achievement and statistical literacy in sampling problems in the SLT and SLT-CP in favor of SLT. This issue is discussed under the related part of the discussion chapter.

When the students' answers in the sampling problems were analyzed, it was observed that the students showed better performance while determining the appropriate techniques for random sampling than evaluating the representativeness and the generalizability of the sample to population. For both two forms of the problems, this might be due to the fact that the students might make connections with real-life experiences while making a random selection, whereas the evaluation of the generalizability of the sample to population requires a statistical point of view. More than half of the students revealed an understanding related to random sampling and generalizability of the sample to population. Moreover, in problem 10, it was observed that some students could not comprehend the meaning of 'random' for sampling. These students were not able to specify the situation which best fits random sampling. Watson and Kelly (2008) reached similar findings concerning the students' understanding of random sampling. In their study, about half of the grade 9 students could not define the meaning of random for sampling. This could be due

to the confusion resulting from the traditional use of the term of random in students' daily life (Watson & Kelly, 2008). The students might have interpreted random selection as selecting an option randomly by not considering the statistical aspect of random selection. As Watson and Kelly (2008) indicated, this confusion about random selection could be remediated through classroom practices. At this point, teachers should be attentive to checking all students' understanding of random instead of only the willing students. Additionally, teachers should create a healthy classroom discourse where the opinions related to random sampling could be freely discussed rather than attempting to reach a correct response.

On the other hand, in problem 22, more than half of the students interpreted the generalizability of the sample to population successfully. These students noticed that to make a generalization about the result of the study for all the children in Turkey, the survey should be administered to all the children in Turkey. Similar findings were found by Watson and English (2016). The researchers used the same problem as a classroom activity and the majority of the students provided similar responses as the present study (Watson & English, 2016). Moreover, some students could not comprehend the sample in the problem context. The students were not able to make sense of the current sample in the newspaper excerpt. The findings of the study conducted by Watson and Kelly (2008) support the findings of the present study. In the study of Watson and Kelly (2008), one third of the nine graders were not able to define and interpret the meaning of sample.

5.1.4 Achievement and Statistical Literacy in Questioning of Statistical Data

The students were asked seven problems concerning the questioning of statistical data content domain. These problems were P1, P14, P17, P19, P23, P24 and P25.

The students' achievement and statistical literacy were varied in the questioning of statistical data problems. Initially, in problem 1, more than half of the students were able to make statistically appropriate predictions concerning the situations from daily life. The estimation skills have an important place in mathematics education and daily life. The students' statistical estimation skills might be developed with classroom discussions at the beginning of the data and chance unit. The teachers might spare time for excerpts from the media channels to reinforce students' estimation skills. When the media excerpts were presented to the students and they were asked to read and interpret the statistics belonging to these excerpts in problem 14, problem 17 and problem 19, about half of the students appropriately read and interpreted the data including graphs, tables, prose texts presented in the media channels which correspond to the second-tier and third-tier of Watson's (1997) hierarchy for statistical literacy. More specifically, these students provided statistical justifications and appropriate claims within the problem context. The rest of the students focused on the irrelevant points and therefore could not critically evaluate the statistics given in the media excerpts. This finding verifies the findings of Yolcu (2012) reporting that the students had difficulty in critically evaluating the statistical data. This situation might be related with the objectives in the elementary mathematics curriculum. In the curriculum, similar to other statistical content domains, objectives concerning making inferences about the statistical data and questioning of statistical data do not exist. For this reason, teachers do not focus on critical evaluation and making inferences from the data sufficiently.

When the students' answers in problem 23, problem 24 and problem 25 were examined, it was seen that the majority of the students were not aware that the use of numerical data is more appropriate to make inferences about mode, median and range of given data with context. More specifically, when the students were asked about the range of the data, the students determined the statement "the range of the data set collected by Selim teacher is 1 to 7" as true. However, the range was not "1 to 7"; it was " $7-1=6$ ". The majority of the students could not differentiate between

two statements. Similarly, the students chose the statement “the median of the data set collected by Selim teacher is industry” as correct. However, the median of the data set should be a numerical value as the middle value of the data set. These findings were in agreement with the findings of the study conducted by Sorto and White (2004) with pre-service teachers using a similar problem. Sorto and White (2004) gave some statistics regarding the frequencies of pets belonging to the students and provided statements similar to the present study. The researchers concluded that the pre-service teachers could not specify the errors about the mean, mode and range for categorical data.

5.2 Comparison between Problems with and without Context

In addition to examining the students’ achievement in the SLT and SLT-CP separately, the students’ achievement in both tests was compared by utilizing statistical analyses.

Initially, according to the comparison in terms of overall achievement in problems with context and problems without context, the mean scores were close to each other. More specifically, the students did not reveal outstanding achievement neither in problems with context nor problems without context. It could be inferred that the students’ achievement was not affected by the existence of the problem context.

When the students’ achievement in measures of central tendency problems were compared based on descriptive analyses, although the students showed better achievement in problems with context with a slight difference in mean for measures of central tendency problems, a statistically significant mean difference could not be found between the students’ achievement in measures of central tendency problems with and without context. This finding corresponds to the results of the research studies pointing to the ineffectiveness of the problem context in the literature (Boaler, 1993; Copper & Dunne, 1999; Gravemeijer, 1994; Monteiro & Ainley,

2003; Nisbet, Langrall & Mooney, 2007; Roth, 1996). It could be inferred that the utilization of context did not affect the students' achievement in reading and interpreting the measures of central tendency problems. It could be due to the fact that the use of environmental pollution context might have caused students to lose their attention towards the measures of central tendency in the problem by directing students' attention to environmental issues rather than to the measures of central tendency in the problems. On the other hand, the existence of context in the measures of central tendency problems might have increased some students' motivation about the problem and helped students to benefit from the real life experiences while solving the problem (Carragher et al., 1985; Chapman, 2006; Lappan & Phillips, 2009; Mitchell & Carbone, 2011; Nunes et al., 1993; Van-den Heuvel-Panhuizen, 2005, Widjaja, 2013).

Moreover, when the students' achievement in graph interpretation problems were compared, it was seen that although the students showed better achievement in problems without context with a slight difference in mean for graph interpretation, a statistically significant mean difference could not be found between the students' achievement in statistical literacy problems with context and problems without context. This means that the use of environmental issues in graph interpretation problems did not provide an advantage to students to recognize and critically evaluate the data in the graphs. This finding of the present study confirms the findings of research studies in the literature (Boaler, 1993; Copper & Dunne, 1999; Gravemeijer, 1994; Monteiro & Ainley, 2003; Nisbet, Langrall & Mooney, 2007; Roth, 1996). As a result, it could be stated that the use of environmental pollution context did not provide an advantage to students to indulge in the graph interpretation problems. This might have stemmed from the fact that the environmental pollution context itself might not provide convenience for students to engage in the graph interpretation problems. The students might need extra information about the problem situation.

Furthermore, when the students' achievement in sampling problems were compared, a statistically significant difference was found between the students' achievement in problems with context and problems without context in favour of problems without context. More specifically, the students demonstrated better achievement in sampling problems without context. This finding of the study confirms the findings of research studies in the literature (Boaler, 1993; Copper & Dunne, 1999; Gravemeijer, 1994; Monteiro & Ainley, 2003; Nisbet, Langrall & Mooney, 2007; Roth, 1996). Parallel with the results of these studies, the finding of the present study also confirms that the utilization of context in sampling problems might lead to a disadvantage by bringing about complicated cognitive processes besides arithmetic operations. In addition, extra factors could be existing which affect students' performance on sampling problems. These factors could be investigated through further studies. Moreover, the difference in students' achievement in sampling problems with context and sampling problems without context might have stemmed from the fact that the problems with context included a part concerning the problem context which extends the sampling problems. Therefore, because of the long introductory part, the students might have digressed from the aim of the sampling problems.

Lastly, as a result of the comparison between the students' achievement in questioning of statistical data problems with context and problems without context, it can be said that although the students showed better achievement in problems with context with a slight difference in mean for questioning of statistical data problems, a statistically significant mean difference could not be found between the students' achievement in questioning of statistical data problems with context and questioning of statistical data problems without context. This finding of the current study corresponds to the findings of research studies in the literature (Boaler, 1993; Copper & Dunne, 1999; Gravemeijer, 1994; Monteiro & Ainley, 2003; Nisbet, Langrall & Mooney, 2007; Roth, 1996). In line with the results of these research studies, the current study also confirms that the students' achievement in questioning of

statistical data problems with and without context cannot be directly related with the existence of the context in problems. More specifically, the problem context is not an incentive for students' achievement in questioning of statistical data problems. On the other hand, the existence of problem context might motivate some students to solve the questioning of statistical data problems by taking advantage of daily life experiences or experiences through media channels they encounter in daily life.

5.3 Implications for Educational Practices

In the present study, the 8th grade students' achievement and statistical literacy in problems with and without context regarding statistical content domains included in elementary mathematics curriculum were investigated. Furthermore, the mean difference in students' achievement in problems with and without context was examined. In this section, some inferences for teachers, textbook writers, curriculum developers and teacher educators are made in line with the findings of the current study.

The findings of the current study put forth that the students revealed procedural understanding rather than conceptual understanding in statistical literacy problems. More specifically, although the students conducted arithmetic procedures successfully, they could not interpret and critically evaluate statistical data given in the problems. In fact, the students should be able to develop a conceptual understanding about the statistical constructs (Friel, 1998). At this point, the teacher's manner plays a decisive role in developing students' conceptual understanding of statistical concepts. To develop students' statistical literacy, teachers could spare time on classroom discussion and in-class activities on the conceptualization of statistical constructs by enhancing students' relational thinking between these constructs.

In addition, it was observed that the students could not differentiate between the terms *mean* and *average*. It could be inferred that most of the teachers are not aware of the difference between average and mean. Thus, in teacher education programs, teacher educators should emphasize the difference between the use of average and mean as statistical constructs. Additionally, textbook writers should pay extra attention to distinguishing the concept of average from the concept of mean.

Moreover, as informed citizens, students need to be statistically literate about the excerpts from newspapers, journals, news, brochures and other media channels. It was reported that students need to be developed in terms of reading and interpreting the statistics they see in media channels. In this context, students should be more exposed to real-life data by bringing excerpts from media channels they encounter in daily life to class (Watson, 1997).

Furthermore, it was seen that the use of environmental context raised some students' awareness to the environmental issues. In that regard, the teachers and textbook writers could use this kind of context in problems more frequently to inform students about the environmental problems.

Additionally, it was observed that in the MoNE (2018), the objectives regarding the questioning of statistical data content domain is restricted. As a requirement of being a data-driven society, the curriculum should focus more on developing students' skill of questioning of statistical data. Thus, the objectives regarding the questioning of statistical data could be increased.

Lastly, the existence of context in statistical literacy problems affected students' achievement in different ways according to statistical content domains. For some problems, existence of context affected students' achievement positively. This means that the teachers and textbook writers could involve real-life data and real-life context for enhanced conception of statistical concepts. For some problems, the utilization of context affected the students' achievement negatively. At this point, it

could be inferred that the teachers, textbook writers and curriculum developers should be careful about involving context in statistical literacy problems. There exist critical points to involve context in statistical literacy problems.

5.4 Recommendations for Further Research

The present study investigated 8th grade students' achievement and statistical literacy in problems with context and problems without context. Future studies could be conducted with a larger number of students as the findings of the present study are limited with the accessible population from different districts of Ankara. The research studies could be conducted including different districts of Turkey to make more comprehensive generalizations. Additionally, research studies could be conducted to observe the difference in students' statistical literacy across different grade levels. Furthermore, longitudinal survey studies could be carried out to investigate the students' development of statistical literacy in future grade levels.

The instruments used in the current study were prepared by the researcher. The comprehensiveness of these instruments could be extended by adding extra items to measure students' statistical literacy in-depth. Furthermore, an instrument measuring students' affective domain could be added. The present study did not focus on the affective domain regarding students' statistical literacy. More precisely, the students' perception, attitude and behavior towards the statistics problems were not taken into consideration in the course of the present study. Future research studies could consider these affective domains of statistical literacy.

Additionally, the research studies could be conducted to observe the development of students' statistical literacy across the grade levels. This could be achieved in two ways. Firstly, the same students' development could be observed with longitudinal studies from grade 5 to grade 8 by testing the students' statistical literacy in certain time intervals. Secondly, 5th, 6th, 7th and 8th grade students' statistical literacy could

be observed at the same time by studying with the students from grade 5 to grade 8 simultaneously.

Besides, research studies could be conducted to observe the effect of the existence of the context in statistical literacy problems. In the present study, although the students were matched according to their academic achievement levels while distributing the instruments including problems with context and problems without context, the instruments can be distributed to the same students. More precisely, each student in the study may solve both the statistical literacy problems with context and statistical literacy problems without context.





REFERENCES

- Aoyama, K., & Stephens, M. (2003). Graph Interpretation Aspect of Statistical Literacy: A Japanese Perspective. *Mathematics Education Research Journal*, 15(3), 207-225.
- Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Carlton, Victoria: Curriculum Corporation.
- Begg, A. (1997). Some emerging influences underpinning assessment in statistics. *The assessment challenge in statistics education*, 17-26.
- Ben-Zvi, D., & Garfield, J. B. (Eds.). (2004). *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 3-16). Dordrecht, The Netherlands: Kluwer academic publishers.
- Biggs, J., & Collis, K. (1989). Towards a model of school-based curriculum development and assessment using the SOLO taxonomy. *Australian Journal of Education*, 33(2), 151-163.
- Boaler, J. (1993). Encouraging the transfer of “school” mathematics to the “real world” through the integration of process and content, context and culture. *Educational Studies in Mathematics*, 25(4), 341-373.
- Borasi, R. (1986). On the nature of problems. *Educational Studies in Mathematics*, 17(2), 125-141.
- Cai, J., Moyer, J. C., & Grochowski, N. J. (1999). Making the mean meaningful: An instructional study. *Research in Middle Level Education Quarterly*, 22(4), 1-24.

- Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology*, 3(1), 21–29.
- Chapman, O. (2006). Classroom practices for context of mathematics word problems. *Educational Studies in Mathematics*, 62(2), 211–230.
- Chick, H., Pfannkuch, M., & Watson, J. (2005). Transnumerative thinking: Finding and telling stories within data. *Curriculum Matters*, 1, 86 –107.
- Cimpoeru, S., & Roman, M. (2018). Statistical Literacy and Attitudes Towards Statistics of Romanian Undergraduate Students.
- Cobb, G.W. & Moore, D.S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*, 104(9), 801-823.
- Cooper, B., & Dunne, M. (1999). *Assessing Children's Mathematical Knowledge: Social Class, Sex, and Problem-Solving*. McGraw-Hill Education (UK).
- Cooper, L (2018) Assessing Students' Understanding of Variability in Graphical Representations that Share the Common Attribute of Bars. *Journal of Statistics Education*, 26(2), 110-124.
- Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. In *Educational Research*.
- Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18(5), 382-393.
- Curcio, F. R. (1989). *Developing Graph Comprehension. Elementary and Middle School Activities*. National Council of Teachers of Mathematics, Inc., 1906 Association Drive, Reston, VA 22091.

- Çatman Aksoy, E. (2018). *An Investigation of 7th grade students' statistical literacy about the concepts of average and variation on bar and line graphs* (Unpublished Master's thesis). Middle East Technical University, Ankara, Turkey.
- De Lange, J. (2003). Mathematics for literacy. *Quantitative literacy: Why numeracy matters for schools and colleges*, 80, 75–89.
- delMas, R., Garfield, J., & Ooms, A. (2005). Using assessment items to study students' difficulty reading and interpreting graphical representations of distributions. *Proceedings of the Fourth International Research Forum on Statistical Reasoning, Literacy, and Reasoning*.
- delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28-58.
- Doyle, P. (2008). *Developing statistical literacy with students and teachers in the secondary mathematics classroom* (Unpublished Masters Thesis). The University of Waikato, Waikato.
- English, L.D., & Watson, J. (2016). Making Decisions with Data: Are We Environmentally Friendly? *Australian Primary Mathematics Classroom*, 21(2), 3-7.
- Enisoğlu, D. (2014). *Seventh grade students' possible solution strategies, errors and misinterpretations regarding the concepts of mean, median and mode given in bar graph representations* (Master's Thesis). Middle East Technical University, Ankara, Turkey.
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education*. New York, the USA: McGraw-Hill.

- Fraenkel, J., Wallen, N., & Hyun, H. (2011). *How to Design and Evaluate Research in Education*. New York: The McGraw-Hill.
- Friel, S. N. (1998). Teaching statistics: What's average. *The Teaching and Learning of Algorithms in School Mathematics*, (pp. 208-217) Reston, VA: National Council of Teachers of Mathematics.
- Friel, S., Curcio, F., Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32 (2), 124-158.
- Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98(1), 29.
- Gal, I. Rothschild, K., Wagner, D. (1990). Statistical concepts and statistical reasoning in elementary school children: Convergence or divergence? Paper presented at the American Educational Research Association, Boston, MA.
- Gal, I. (1995). Statistical tools and statistical literacy: The case of the average. *Teaching Statistics*, 17(3), 97-99.
- Gal, I. (2000), *Adult Numeracy Development: Theory, Research, Practice*, Cresskill, NJ: Hampton Press.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Garfield, J., Delmas, R., & Chance, B. (2003). The web-based artist: assessment resource tools for improving statistical thinking. *Assessment Resources for Statistics*.

Gravemeijer, K. (1994), 'Educational development and educational research in mathematics education', *Journal for Research in Mathematics Education* 25(5), 443–471.

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics*, 39(1-3), 111-129.

Hardiman, P. T., Well, A. D., & Pollatsek, A. (1984). Usefulness of a balance model in understanding the mean. *Journal of Educational Psychology*, 76(5), 792–801.

Henriques, A., & Cruz, A. M. (2010). Developing Statistical Literacy in Primary Level: Results of A Teaching Unit 1.

Hoogland, K., Pepin, B., Bakker, A., de Koning, J., & Gravemeijer, K. (2016). Representing contextual mathematical problems in descriptive or depictive form: Design of an instrument and validation of its uses. *Studies in Educational Evaluation*, 50, 22–32. Retrieved from May 8 2020, <https://doi.org/10.1016/j.stueduc.2016.06.005>.

Jolliffe, F. (1997). *Issues in Constructing Assessment Instruments for the Classroom*. In I. Gal and J. Garfield, J. (Eds.), *The Assessment Challenge in Statistics Education*. (pp. 191-204). Netherlands: The International Statistical Institute

Kaynar Y., Halat E. (2011) “Sekizinci Sınıf Öğrencilerinin Sıklık Tablosu Okuma ve Yorumlama Becerilerinin İncelenmesi.” *Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi*

Kelly, A. E., Sloane, F., Whittaker, A. (1997). Simple approaches to assessing underlying understanding of statistical concepts. In I. Gal and J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*, (pp. 85-90). Amsterdam: IOS Press and the International Statistical Institute.

- Kirsch, I., & Mosenthal, P. (1990). Understanding the news. *Reading Research Quarterly*, 22(2), 83–99.
- Kliebard, H. M., & Franklin, B. M. (2003). The ascendance of practical and vocational mathematics, 1893-1945: Academic mathematics under siege. *A History of School Mathematics*, 1, 399-440. Reston, VA: NCTM.
- Koleza, E., & Kontogianni, A. (2013). Assessing statistical literacy: What do freshmen know. In *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education (CERME8)*. Ankara: Middle East Technical University.
- Konold, C & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics*, Reston, VA: National Council of Teachers of Mathematics, pp.193-215.
- Kranda, S. (2018). 7. Sınıf Öğrencilerinin Sosyal Bilgiler Dersindeki Grafik Okuryazarlık Düzeylerinin İncelenmesi (Doctoral dissertation, Lisansüstü Eğitim Enstitüsü (İlköğretim).
- Lappan, G. & Phillips, E. (2009). A Designer Speaks: Glenda Lappan and Elizabeth Phillips. *Educational Designer*, 1(3), 1-19.
- Lesh, R., Amit, M., & Schorr, R. (1997). Using “real-life” problems to prompt students to construct conceptual models for statistical reasoning. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 65–83). Amsterdam: IOS Press.
- Luke, A., & Freebody, P. (1997). Critical literacy and the question of normativity: An introduction. In S. Muspratt, A. Luke, & P. Freebody (Eds.), *Constructing Critical Literacies: Teaching and Learning Textual Practice* (pp. 1–18). St. Leonards: Allen and Unwin.

- Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, 13(1-2), 152-173.
- Memnun, S. (2013). Ortaokul yedinci sınıf öğrencilerinin çizgi grafik okuma ve çizme becerilerinin incelenmesi. *Electronic Turkish Studies*, 8(12), 1153-1167.
- Mevarech, Z. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics*, 4, 415-429.
- Meyer, M., Dekker, T., & Querelle, N. (2001). Innovations in Curriculum: Context in Mathematics Curricula. *Mathematics Teaching in the Middle School*, 6(9), 522-527. Retrieved April 17, 2020, from www.jstor.org/stable/41181019
- Millroy, W. (1994). Exploring the Nature of Street Mathematics. *Journal for Research in Mathematics Education*, 25(3), 304-309. doi:10.2307/749340
- Ministry of Education of China. (2001b). The standards of mathematics curriculum for compulsory education (Trial Draft). Beijing Normal University, Chinese.
- Ministry of Education. (2007). The New Zealand curriculum. Wellington: Learning Media
- Ministry of National Education (MoNE). (2018). Ortaokul Matematik Dersi Öğretim Programı Ankara: Milli Eğitim Basımevi.
- Mitchell, E., & Carbone, A. (2011). A typology of task characteristics and their effects on student engagement. *International Journal of Educational Research*, 50, 257-270.
- Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, 20-39.

- Monteiro, C., Ainley, J. (2003). Interpretation of Graphs: Reading Through the Data. *British Society for Research into Learning Mathematics* 23(3), 31-36
- Moore, D. (1990). Uncertainty. In L. Steen (Ed.), *On the Shoulders of Giants: New Approaches to Numeracy* (pp. 95–137). Washington, D.C.: National Academy Press.
- National Council of Teachers of Mathematics. (2000). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.
- Nisbet, S., Langrall, C., & Mooney, E. (2007). The role of context in students' analysis of data. *Australian Primary Mathematics Classroom*, 12(1), 16-22.
- Nunes, T., Carraher, T. N., Schliemann, A. D., & Carraher, D. W. (1993). *Street Mathematics and School Mathematics*. Cambridge University Press.
- OECD. (2010). PISA 2009 results: What students know and can do. Student performance in reading, mathematics, and science, 1, Paris: OECD.
- OECD (2013). PISA 2013: Assessment and Analytical Framework. Mathematics, Reading, Science, Problem Solving and Financial Literacy. Paris: OECD Publishing.
- Oruç, Ş. ve Akgün, İ. H. (2010). İlköğretim sosyal bilgiler 7. sınıf öğrencilerinin grafik okuma becerisini kazanma düzeyleri. *International Journal Of Eurasia Social Sciences*, 1(1), 51-58.
- Palm, T. (2009). Theory of authentic task situations. In L. Verschaffel, B. Greer, W. V. Dooren, & S. Mukhopadhyay (Eds.), *Words and Worlds - Modelling Verbal Descriptions of Situations* (pp. 3–20). Rotterdam, the Netherlands: Sense

- Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 17–47). Dordrecht: Kluwer.
- Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12, 191-204.
- Putnam, R. T., Lampert, M., & Peterson, P. L. (1990). Alternative perspectives on knowing mathematics in elementary schools. In Cazden, C. (Ed.), *Review of Research in Education*, 16, 57-150. Washington, DC: American Educational Research Association.
- Randall, G. (2006). An Exploration of Students' Statistical Thinking. *Teaching Statistics*, 28(1), 17-21.
- Roth, W. (1996). Where Is the Context in Contextual Word Problems? Mathematical Practices and Products in Grade 8 Students' Answers to Story Problems. *Cognition and Instruction*, 14(4), 487-527. Retrieved April 19, 2020, from www.jstor.org/stable/3233784
- Rumsey D.J. (2002) Statistical Literacy as a Goal for Introductory Statistics Courses. *Journal of Statistics Education*, 10(3)
- Schau, C., & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. *The assessment challenge in statistics education*, 51, 91-104.
- Schild, M. (2006). Statistical literacy survey analysis: Reading graphs and tables of rates and percentages. *Proceedings of the Sixth International Conference on Teaching Statistics*.

- Schiold, M. (2010). Assessing statistical literacy: Take care. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 133–152). Milton: John Wiley & Sons. Chapter 11
- Schnotz, W. (2002). Commentary: Towards an Integrated View of Learning from Text and Visual Displays. *Educational Psychology Review* 14, 101–120. <https://doi.org/10.1023/A:1013136727916>
- Schnotz, W., Baadte, C., & Müller, A. (2010). Creative thinking and problem solving with depictive and descriptive representations. In *Use of Representations in Reasoning and Problem Solving* (pp. 21-45). Routledge.
- Sharma, S. (2017) Definitions and models of statistical literacy: a literature review, *Open Review of Educational Research*, 4(1), 118-133, DOI: 10.1080/23265507.2017.1354313
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan Publishing Company.
- Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International Handbook of Mathematics Education* (pp. 205-237). Boston, MA: Kluwer.
- Skemp, R. R. (1979). *Intelligence, Learning, and Action a New Model for Theory and Practice in Education*. Chichester [Eng.] New York Wiley.
- Snell, L. (1999), “Using Chance media to Promote Statistical Literacy,” Paper presented at the 1999 Joint Statistical Meetings, Dallas, TX.
- Sorto, M. A., & White, A. (2004). Statistical knowledge for teaching. In *Tenth International Congress on Mathematical Education, Copenhagen*.

Steen, L.A. (Ed.). (1997). Why numbers count: *Quantitative literacy for tomorrow's America*. New York: College Entrance Examination Board

Steen, Lynn Arthur, ed. (2001). Mathematics and Democracy: The Case for Quantitative Literacy. Princeton, NJ: National Council on Education and the Disciplines.

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 18(1), 64-80.

Sullivan, P., Zevenbergen, R., & Mousley, J. (2003). The Contexts of mathematics tasks and the context of the classroom: Are we including all students? *Mathematics Education Research Journal*, 15(2), 107-121.

Şeker, V. (2018). Teaching Average Through Implementing Cognitively Demanding Statistical Tasks.

Tosun, T., Özen Ünal, D. (2019). Veri ve Olasılık Öğrenme Alanlarında Yapılmış Çalışmaların İçerik Analizi. *Ege Eğitim Dergisi*, 20 (1), 244-261. DOI: 10.12984/eegefd.489275

Thomas, R., & Gerofsky, S. (1997). An exchange about word problems. *For the Learning of Mathematics*, 17 (2), 21-23.

Thompson, A. G., Philipp, R. A., Thompson, P. W., & Boyd, B. A. (1994). Computational and conceptual orientations in teaching mathematics. In A. Oxford (Ed.), 1994 yearbook of the NCTM (pp. 79-92). Reston, VA: NCTM

Treffers, A. (1987). Three Dimensions. A Model of Goal and Theory Description in Mathematics Education: The Wiskobas Project, D. Reidel, Dordrecht

- Trewin, D. (2005). Improving statistical literacy: The respective roles of schools and the National Statistical Offices. In M. Coupland, J. Anderson, & T. Spencer (Eds.), *Making mathematics vital: Proceedings of the twentieth Biennial Conference of the Australian Association of Mathematics Teachers*, (pp. 11–19). Adelaide: Australian Association of Mathematics Teachers.
- Uçar, Z. & Akdoğan, E.N. (2009). 6.-8. Sınıf öğrencilerinin ortalama kavramına yüklediği anlamlar, *İlköğretim-Online*, 8(2), 391-400.
- Van Den Heuvel-Panhuizen, M. (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics*, 25 (2), 2-10.
- Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. (pp. 17-203). Lisse: Swets & Zeitlinger.
- Walkerdine, V.: 1990, 'Difference, Cognition and Mathematics Education', *For the Learning Mathematics* 10(3), 51-56.
- Wallman, K.K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, 88(421), 1-8.
- Watson, J. (1997). Assessing Statistical Thinking Using the Media. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107-121). Amsterdam: IOS Press and The International Statistical Institute.
- Watson, J. M. (2013). *Statistical Literacy at School: Growth and Goals*. Routledge.
- Watson, J. M., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3-46.
- Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, 2(1-2), 11-50.

Watson, J., & Callingham, R. (2004). Statistical literacy: from idiosyncratic to critical thinking. Curricular Development in Statistics Education: International Association for Statistical Education 2004 Roundtable. Voorburg, the Netherlands: International Statistical Institute, (pp. 116-162). Sweden. Retrieved from May 25, 2020 http://www.stat.auckland.ac.nz/~iase/publications/rt04/4.1_Watson&Callingham.pdf

Watson, J. M., Kelly, B.A. (2008). Sample, Random and Variation: The Vocabulary of Statistical Literacy. *International Journal of Science and Mathematics Education* 6, 741–767. Retrieved from June 14, 2020 <https://doi.org/10.1007/s10763-007-9083-x>

Watson, J. M., Kelly, B.A., Callingham, R.A., & Shaughnessy, J.M. (2003) The measurement of school students' understanding of statistical variation, *International Journal of Mathematical Education in Science and Technology*, 34(1), 1-29, DOI: 10.1080/0020739021000018791

Watson, J. M (1997). Assessing Statistical Thinking Using the Media. In *The Assessment Challenge in Statistics Education*

Widjaja, W. (2013). The Use of Contextual Problems to Support Mathematical Learning. *Indonesian Mathematical Society Journal on Mathematics Education*, 4(2), 157-168.

Wild, C., Triggs, C. & Pfannkuch, M. (1997). Assessment on a budget: using traditional methods imaginatively. In *The assessment challenge in statistics education*, Eds. I. Gal and J.B. Garfield, pp. 205-220. Amsterdam, The Netherlands: International Statistical Institution OS Press.

Yıldırım, H. H. (2006). *The Differential Item Functioning (DIF) Analysis of Mathematics Items in the International Assessment Programs* (Unpublished Doctorate Thesis). Middle East Technical University, Ankara, Turkey.

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, Ankara, Turkey.

Yotongyos, M., Traiwichitkhun, D., & Kaemkate, W. (2015). Undergraduate Students' Statistical Literacy: A Survey Study. *Procedia - Social and Behavioral Sciences*, 191, 2731–2734.



APPENDICES

A. APPROVAL FROM METU HUMAN SUBJECTS ETHICS COMMITTEE

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
ueam@metu.edu.tr
Sayı: 28620816/444



ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

21 KASIM 2019

Konu: Değerlendirme Sonucu

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi: İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Dr.Öğretim Üyesi Didem AKYÜZ

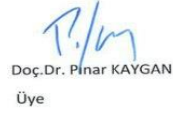
Danışmanlığını yaptığınız Tuğçe VURAL'ın "İlköğretim 8. Sınıf Öğrencilerinin İstatistiksel Okuryazarlığı ve İstatistiksel Düşünme Becerilerinin Bağlam Temelli ve Bağlam Temelli Olmayan Problemler ile İncelenmesi" başlıklı araştırması İnsan Araştırmaları Etik Kurulu tarafından uygun görülmüş ve 424 ODTU 2019 protokol numarası ile onaylanmıştır.

Saygılarımızla bilgilerinize sunarız.


Prof. Dr. Tülin GENÇÖZ

Başkan


Prof. Dr. Tolga CAN
Üye


Doç.Dr. Pinar KAYGAN
Üye


Dr. Öğr. Üyesi Ali Emre TURGUT
Üye


Dr. Öğr. Üyesi Şerife SEVİNÇ
Üye


Dr. Öğr. Üyesi Müge GÜNDÜZ
Üye


Dr. Öğr. Üyesi Süreyya Özcan KABASAKAL
Üye

B. PERMISSON FROM MINISTRY OF NATIONAL EDUCATION



T.C.
ANKARA VALİLİĞİ
Milli Eğitim Müdürlüğü

Sayı : 14588481-605.99-E.314526
Konu : Araştırma İzni

06.01.2020

ORTA DOĞU TEKNİK ÜNİVERSİTESİNE

İlgi : a)MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 2017/25 nolu Genelgesi.
b)27.12.2019 tarihli ve 171 sayılı yazınız.

Üniversiteniz Matematik ve Fen Bilimleri Eğitimi Anabilim Dalı Matematik Eğitimi Programı Yüksek Lisans Öğrencisi Tuğçe Varol'un "**İlköğretim 8. Sınıf Öğrencilerinin İstatistiksel Okuryazarlığı Ve İstatistiksel Düşünme Becerilerinin Bağlam Temelli Olmayan Problemler İle İncelenmesi**" konulu çalışması kapsamında İlimiz Çankaya, Mamak ve Yenimahalle İlçelerine bağlı, ekli listede belirtilen okullarda uygulama talebi ilgi (b) Genelge çerçevesinde incelenmiştir.

Yapılan inceleme sonucunda, söz konusu araştırmanın Müdürlüğümüzde muhafaza edilen ölçme araçlarının; Türkiye Cumhuriyeti Anayasası, Milli Eğitim Temel Kanunu ile Türk Milli Eğitiminin genel amaçlarına uygun olarak, ilgili yasal düzenlemelerde belirtilen ilke, esas ve amaçlara aykırılık teşkil etmeyecek, eğitim-öğretim faaliyetlerini aksatmayacak şekilde okul ve kurum yöneticilerinin sorumluluğunda gönüllülük esasına göre uygulanması Müdürlüğümüzce uygun görülmüştür.

Bilgilerinizi ve gereğini rica ederim.

Turan AKPINAR
Vali a.
Milli Eğitim Müdürü

Dağıtım:
Gereği:
Orta Doğu Teknik Üniversitesi

Bilgi:
Çankaya, Mamak, Yenimahalle
İlçe MEM

Adres:
Elektronik Ağ:
e-posta:

Bilgi için:
Tel: 0 () ____
Faks: 0 () ____

Bu evrak güvenli elektronik imza ile imzalanmıştır. <https://evraksorgu.meb.gov.tr> adresinden 4ac2-a836-3b0f-92d1-7d2e kodu ile teyit edilebilir.

C. STATISTICAL LITERACY TEST WITH CONTEXT PROBLEMS

İSTATİSTİK BAŞARI TESTİ:

İsim:

Şube:



Merhaba! Ben Sünger Bob ve arkadaşım Patrick. Çizgi filmlerden bizi tanımayan yoktur.

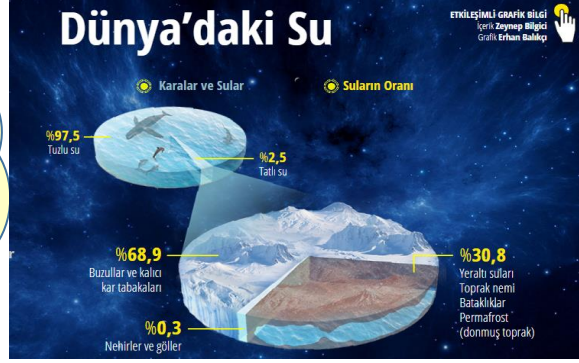
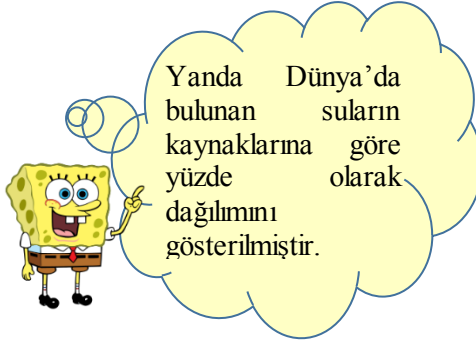
Okyanusların derinliklerinde yaşayan deniz canlılarıyız. Bugün sizlerle birlikte biz deniz canlılarının yaşam alanı olan ve insan yaşamının vazgeçilmezi olan su ilgili bir yolculuğa çıkacağız.

Aşağıda NASA tarafından 7 Aralık 1972'de Apollo 17 uzay aracının 29 bin kilometre mesafeden çektiği ve Dünya'yı bütün olarak gösteren 'Blue Marble' (Mavi Bilye) isimli fotoğrafı görmektesiniz. Fotoğraf ile ilgili açıklamayı okuyup soruları cevaplandırınız.



Mavi Bilye fotoğrafı dünyanın bu güne kadarki en detaylı gerçek renkli fotoğrafıdır. Bilim adamları uydu ile gözlemler yaparak dünya üzerindeki okyanusları, denizi, karaları, buz kütlelerini ve bulutları her bir kilometresine kadar hatasız bir şekilde birbirlerine bağladılar ve bu kusursuz fotoğrafı ortaya çıkardılar. Fotoğrafta da görüldüğü gibi dünyanın büyük bir kısmı sudan oluşmaktadır.

- 1) Sizce yeryüzünde hangi tür yerüstü suları bulunur? Bu kaynakların yeryüzündeki yaklaşık miktarları tahmin edebilir misiniz? Nasıl tahmin ettiğinizi açıklayınız.



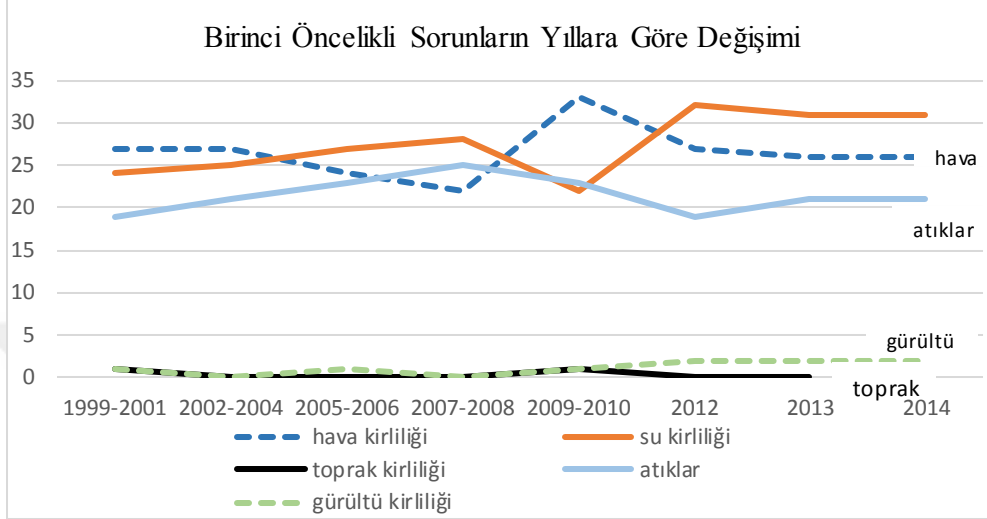
- 2) Yukarıda verilen gösterimdeki bilgilerden yararlanarak Dünya'da bulunan suların kaynaklarına göre dağılımını en uygun biçimde temsil eden grafik türünü hangisidir? Neden? Cevabınızı açıklayınız.
- 3) Dünya'da bulunan suların kaynaklarına göre dağılımını seçtiğiniz grafik türü ile ifade ediniz

İstatistiklerle de görüldüğü gibi su, yeryüzünün büyük çoğunluğunu oluşturmaktadır ve biz canlıların yaşamı suya bağlıdır. Yerüstü suları güneşin ısısını emerler ve atmosferin her yanına yayarlar. Böylece adeta bir klima gibi atmosfer ısısının düzenlenmesini sağlarlar.

Fakat, başta insanların bilinçsiz tutumları olmak üzere birçok etken su kirliliğine ve böylece iklim değişikliklerine sebep olmakta ve temel ihtiyacımız olan su için tehdit oluşturmaktadır. Gelin hep birlikte ülkemizdeki su kirliliği ile ilgili istatistiklere göz atalım.



Gelin hep birlikte çevre sorunları içerisinde bulunan su kirliliğine ve diğer çevre sorunlarına göz atalım. Aşağıdaki verilen grafikte ve tabloda Türkiye'deki çevre sorunlarının yıllara göre değişimlerini görmekteyiz.



	1999-2001	2002-2004	2005-2006	2007-2008	2009-2010	2012	2013	su
Hava Kirliliği	27	27	24	22	33	27	26	26
Su Kirliliği	24	25	27	28	22	32	31	31
Toprak Kirliliği	1				1			
Atıklar	19	21	23	25	23	19	21	21
Gürültü Kirliliği	1		1		1	2	2	2

Yukarıda verilen grafiğe ve tabloya bakarak aşağıdaki ifadelerin doğruluklarını tespit ediniz. Doğru ifadelerin başına D, yanlış ifadelerin başına Y yazınız.

Neden D veya Y yazdığınızı aşağıdaki boşluklara açıklayınız.

4) () 2014 yılında **su kirliliğinin** 1.öncelikli çevre sorunu olduğu il sayısı en fazladır.

Nedeni:.....

5) () **Gürültü kirliliği** yıllar ilerledikçe sürekli artış göstermiştir.

Nedeni:.....

6) () **Hava kirliliği** ve **su kirliliği** genel olarak ters orantılı ilişki göstermiştir.

Nedeni:.....

7) () **Hava kirliliğinin** 1.öncelikli çevre sorunu olduğu il sayısı sürekli olarak artış göstermiştir.

Nedeni:.....



“Türkiye’de günde **ortalama** 70 bin ton çöp atıldığını, atıkların yüzde 20’sini pet şişe, karton, teneke gibi ambalaj atıklarının oluşturduğunu söyledi.”

8) Yukarıda verilen haberde geçen ‘ortalama’ ifadesinden ne anlıyorsunuz? Sizce bu açıklamada neden ortalama kelimesi tercih edilmiştir? Açıklayınız.

9)



Bir hafta boyunca Türkiye’de her gün atılan ortalama çöp miktarını yaklaşık olarak hesaplamak istiyorum. Bir hafta boyunca her gün atılan çöp miktarını en yakın şekilde aşağıdaki yöntemlerin hangisini kullanarak **bulamayız**?

- A) Bir hafta boyunca üretilen en çok çöp miktarından en az çöp miktarını çıkararak
- B) Bir hafta boyu üretilen toplam çöp miktarını küçükten büyüğe doğru sıralayıp ortadaki sayıyı seçerek
- C) Bir hafta boyu günlük üretilen toplam çöp miktarını 7 ye bölerek
- D) Bir hafta boyu günlük üretilen toplam çöp miktarını yazıp en çok tekrar eden miktarı belirleyerek

10)

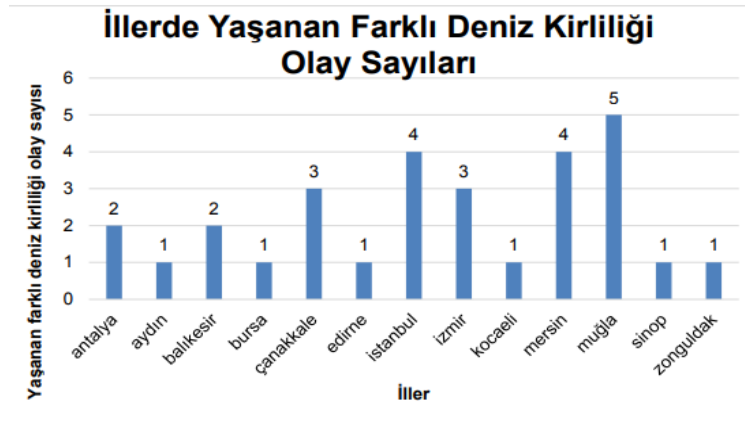


Bir arařtırmacı lkemizdeki deniz kirlilięini il bazında incelemek iin rastgele bir il seimi yapmak o ilin deniz kirlilięine sebep olan etmenleri incelemek istiyor. Buna gre yukarıdaki seimlerden hangisi veya hangilerini yaparsam **rastgele seim** yapmıř olur?

- I) Ulařımı kolay ve ucuz olduęu iin evresindeki en yakın denize kıyısı olan il gitmesi
- II) Marmara Denizinde sanayi faaliyetleri yoęun olduęundan dolayı o blgenin deniz kirlilięinin daha yoęun olacaęını dřunerek Marmara Blgesinden bir il semesi
- III) Bir torbaya denize kıyısı olan tm illerin plakalarını yazıp torbadan rastgele bir il plakası semesi
- A) Yalnız I B) I ve II C) Yalnız III D) II ve III



Trkiye’de illerde yařanan farklı deniz kirlilięi olayları sayılarını belirleyerek bazı istatistikler elde ettim ve topladıęım verileri bir grafik ile ifade ettim. Ařaęıda grafikte bazı illerde yařanan farklı deniz kirlilięi olay sayılarını verdim.



Buna grafięe, deniz kirlilięi olay sayılarınının;

11) Modu (tepe deęeri) nedir? Cevabınızın nedenini ve nasıl bulduęunuzu aıklayınız?

12) Medyanı (ortanca değeri) nedir? Cevabınızın nedenini ve nasıl bulduğunuzu açıklayınız?

13) Aritmetik ortalaması nedir? Cevabınızın nedenini ve nasıl bulduğunuzu açıklayınız?

14)

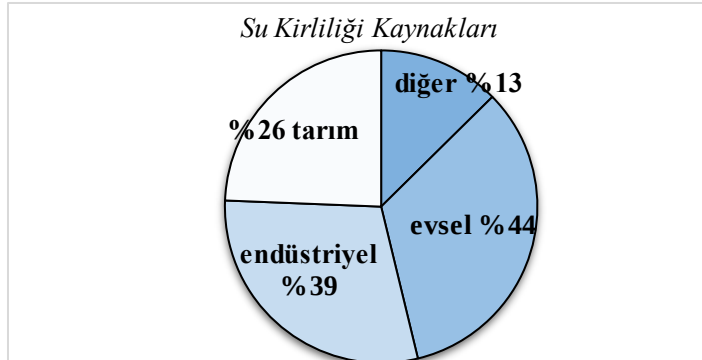


Aşağıda Çevre Mühendisleri Odası tarafından hazırlanan tabloda Türkiye'deki bölgelere 2016 yılında düşen yağış miktarları verilmiştir.

BÖLGELERİMİZİN YAĞIŞ DURUMLARI (01 Ocak 2016-31 Aralık 2016)							
BÖLGE	YAĞIŞ (mm)	NORMAL (mm)	GEÇEN YIL (mm)	DEĞİŞİM ORANI			
				NORMALE GÖRE (%)		GEÇEN YILA GÖRE (%)	
Türkiye Geneli	597,6	574,0	585,7	4,1	ARTMA	2,0	ARTMA
Marmara	646,6	659,2	638,3	-1,9	AZALMA	1,3	ARTMA
Ege	555,5	595,3	652,0	-6,7	AZALMA	-14,8	AZALMA
Akdeniz	559,3	663,7	647,5	-15,7	AZALMA	-13,6	AZALMA
İç Anadolu	436,5	407,8	441,9	7,0	ARTMA	-1,2	AZALMA
Karadeniz	882,6	698,0	727,2	26,4	ARTMA	21,4	ARTMA
Doğu Anadolu	583,6	565,2	533,3	3,3	ARTMA	9,4	ARTMA
Güneydoğu Anadolu	509,9	549,1	532,3	-7,1	AZALMA	-4,2	AZALMA

Bu tabloya göre 2016 yılında hangi bölgeye veya bölgelere normal yağış miktarından daha az miktarda yağış düşmüştür? Nasıl karar verdiğinizizi ayrıntılı olarak açıklayınız?

Aşağıda bir ülkede su kirliliği kaynakları daire grafiği ile ifade edilmiştir.



15) Bu grafiği nasıl yorumlarsınız? Grafikten ne anlıyorsunuz? Sizce su kirliliği kaynaklarının gösterildiği grafikte olağandışı bir durum var mı? Varsa ayrıntılı olarak açıklayınız.

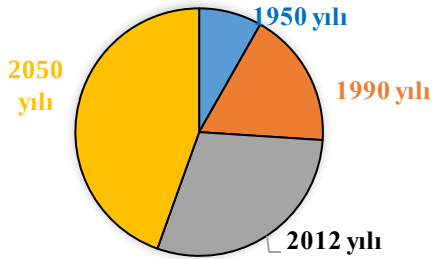
16)



2017 yılında UNICEF tüm Dünya’da kuraklığa dair farkındalık yaratmak amacıyla ‘Sular Ateş Altında ‘ diyerek bir rapor yayınlamış. Bu raporda kuraklığa dair acı gerçekleri istatistiklerle ve aşağıda gördüğünüz fotoğraflar ile gözler önüne sermiştir.



Yıllara Göre Su Kıtlığı Çeken Ülke Sayısı



‘Kuraklık Değerlendirme Raporu’ndan alınmış su kıtlığı çeken ülke sayılarına ait daire grafiği verilmiştir. Buna göre aşağıdaki sütun grafiklerinden hangisi daire grafiği ile aynı bilgiyi verir?



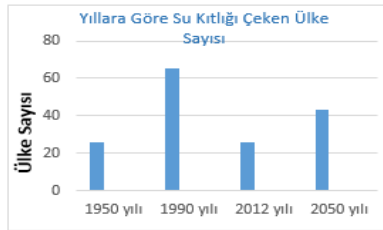
A)



B)



C)



D)





Kuraklığın sebeplerinden biri olan küresel ısınma, iklim değişikliği ve hava sıcaklıklarındaki dengesizlikleri de beraberinde getirmiştir. Aşağıdaki tabloda bir ile ait iki farklı aralıkta ölçülmüş dokuz günlük sıcaklıklar verilmiştir. (17. ve 18. soruları tabloya göre yanıtlayınız)

	1. gün	2. gün	3. gün	4. gün	5. gün	6. gün	7. gün	8. gün	9. gün
1.aralık	15	16	12	10	13	10	13	15	9
2.aralık	18	13	16	15	19	20	13	11	15

17)



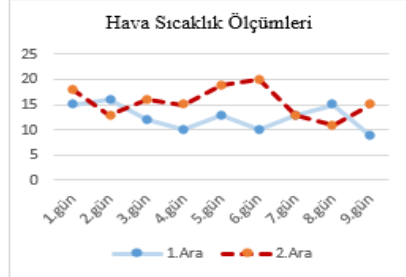
Bu tabloya göre bu il hangi aralıkta **daha sıcaktır**? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

18)

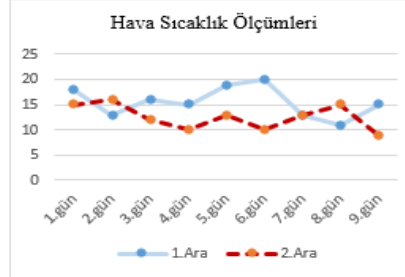


Aşağıdakilerden hangisinde tabloda verilen 1. ve 2. Aralıklardaki sıcaklıkların çizgi grafiği ile gösterilmiş biçimi doğru verilmiştir?

A)



B)



C)



D)

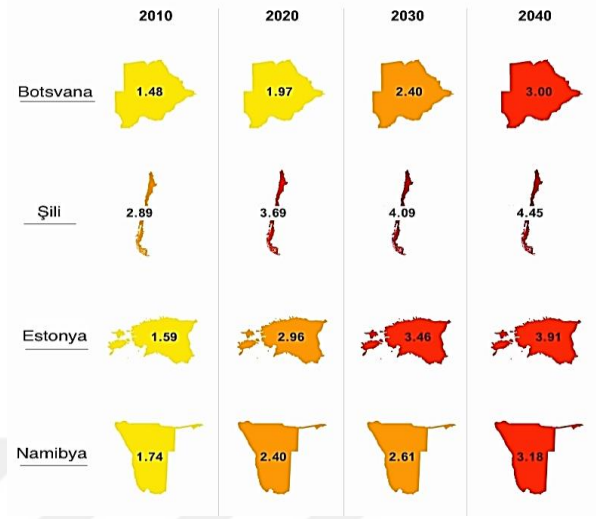


19)

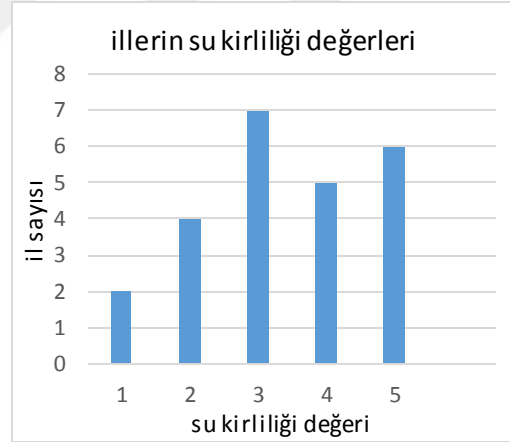
Yanda bir gazeteden alınan tabloda Dünya Kaynak Enstitüsü tarafından rapor edilen bazı ülkelere ait 2010-2040 yılları arası ‘Su Sıkıntısı’ değerleri verilmiştir. Bu ülkelerden 2010-2040 yılları arasında **su sıkıntısının en çok yaşandığı** ülkede inceleme yapmak isteyen bir araştırmacı hangi ülkeyi seçerse doğru karar vermiş olur? Cevabınızın sebebini açıklayınız.



2010'dan 2040'a Su Sıkıntısı



Artan su sıkıntılarının bir etkeni de suyun kalitesinin düşmesidir. Yeryüzündeki su kirliliği her geçen gün artmaktadır. Bir grup araştırmacı farklı illerdeki su kirliliği değerleri hakkında ölçümler yapmış ve yandaki sütun grafiğini elde etmişlerdir. Araştırmacıların kullandığı ölçek su kirliliğini 1 den 5 e kadar numaralandırmıştır. (1= az kirli ; 5= çok kirli)



20) Yapılan ölçümlere göre kaç ilin su kirliliği değeri 3'ten fazladır?

A)7

B) 18

C) 11

D) 22

21) Bu arařtırmacılar toplam kaç adet ilin su kirlilięi deęerlerini ölçmüşlerdir?

A) 6

B) 7

C) 15

D) 24

Çocuklar Çevre Dostu Deęil

Türkiye’de okul çağındaki çocukların sandığımız kadar da çevre dostu olmadığı ortaya çıktı. Bir çevre bilimci repörtajında “ Ankara’da bir grup 5.sınıf öğrencisi ile yapılan çalışmanın sonucu bizleri şok etti” ifadelerine yer verdi. Yürütölen çalışma öğrencilerin doğayı korumak adına çaba sarfetmedikleri, uzun süre duřta kalıp su israfı yaptıkları, çöpleri çevreye attıkları elektrik tasarrufu için çaba sarfetmedikleri sonuçlarını gözler önüne serdi.

Çevre bilimci, öğrencilerin çevre konusunda duyarlı olmadıklarının bu konuda daha dikkatli olmaları gerektiğinin altını çizdi.

22) Çevre bilimcinin iddiaları Türkiye’deki bütün çocuklara genellenebilir mi? Türkiye’deki her çocuk için bu iddialar doğru olabilir mi? Neden? Cevabınızın nedenini istatistiksel ifadeler ile ayrıntılı olarak açıklayınız.

Öğrencilerine enerji kaynaklarının tasarruflu kullanılması konusunda farkındalık yaratmak isteyen Selim Öğretmen sınıfında bir anket uygulamıştır. Türkiye İstatistik Kurumu'ndan elde ettiği araştırma sorusunu öğrencilerine sorarak sınıfından veri toplamıştır.

Araştırma sorusu şöyledir: “ Sizce ülkemizde elektrik tüketiminin en fazla olduğu alan hangisidir?

- a) ev b) ticaret c) resmi daire d) sanayi
e) aydınlatma f) tarımsal sulama g) diğer

Selim öğretmen, anketten elde ettiği verileri yandaki tabloda ifade etmiştir.

Tabloda sunulan verilere göre aşağıda verilen ifadelerin doğruluğunu tespit ediniz. Doğru ifadelerin başına D yanlış ifadelerin başına Y yazınız. Cevabınızın nedenini açıklayınız.

Elektrik Tüketilen Alan	Tercih Eden Kişi Sayısı
Ev	3
Ticaret	4
Resmi Daire	3
Sanayi	7
Aydınlatma	1
Tarımsal Sulama	3
Diğer	1

23) () Selim öğretmenin topladığı verilerin modu sanayidir.

Neden:.....

24) () Selim öğretmenin topladığı verilerin medyanı sanayidir.

Neden:.....

25) () Selim öğretmenin topladığı verilerin açıklığı 1 den 7 ye kadardır.

Neden:.....

D. STATISTICAL LITERACY TEST WITHOUT CONTEXT PROBLEMS

İSTATİSTİKSEL OKURYAZARLIK TESTİ

İsim:

Şube:

- 1) Sizce yeryüzünde hangi tür canlılar bulunur? Yeryüzünde yaşayan canlı türü sayısını tahmin edebilir misiniz? Nasıl tahmin ettiğinizi açıklayınız.

Aşağıda, bir bütünün A, B ve C parçalarının yüzdeleri verilmiştir.

A: %68,9

B: %0,3

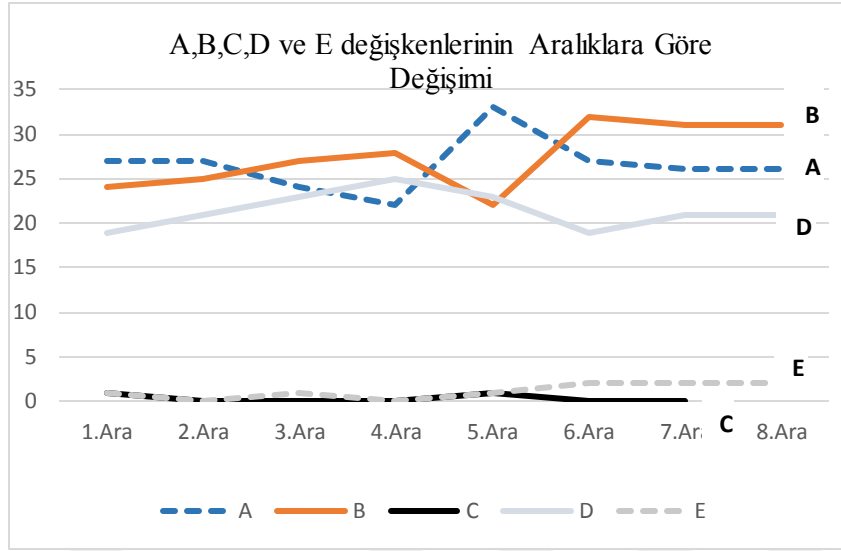
C: %30,8

- 2) Sizce bu bütünün parçalarının dağılımını en uygun biçimde temsil eden grafik türü hangisidir? Neden ? Cevabınızı açıklayınız.

- 3) A, B ve C parçalarının dağılımını seçtiğiniz grafik türü ile çizerek ifade ediniz.

Aşağıda A,B,C,D ve E değişkenlerinin birbirini takip eden belirli aralıklarda değişimini gösteren tablo ve grafik verilmiştir.(5,6,7 ve 8.soruları yanıtlayınız)

	1. Aralık	2. Aralık	3. Aralık	4. Aralık	5. Aralık	6. Aralık	7. Aralık	8. Aralık
A	27	27	24	22	33	27	26	26
B	24	25	27	28	22	32	31	31
C	1				1			
D	19	21	23	25	23	19	21	21
E	1		1		1	2	2	2



Verilen grafiğe ve tabloya bakarak aşağıda verilen ifadelerin doğruluklarını tespit ediniz.

Doğru ifadelerin başına D, Yanlış ifadelerin başına Y yazınız.

4) () 8.Aralıkta **B** değişkenler arasında en yüksek değere sahiptir.

Nedeni:.....

5) () E aralıklar ilerledikçe sürekli artış göstermiştir.

Nedeni:.....

6) () A ve B değişkenleri genelde ters orantılı ilişki göstermiştir.

Nedeni:.....

7) () A değişkeni aralıklara göre sürekli artış göstermektedir.

Nedeni:.....

8) “Ortalama” kelimesinden ne anlıyorsunuz? Sizce ortalama kelimesini neden kullanırsınız? Açıklayınız.

9) 7 adet sayının ortalamasını tahmin etmek için aşağıdaki yöntemlerden hangisini kullanamayız?

- A) En büyük sayıdan en küçük sayıyı çıkararak
- B) Sayıları küçükten büyüğe sıralayıp ortadaki sayıyı seçerek
- C) Sayıları toplayıp 7 ye bölerek
- D) Sayılar arasından en çok tekrar eden sayıyı seçerek

10) Aşağıdaki öncüllerden hangisi veya hangilerini yapılırsa veri toplanacak bölge rastgele seçilmiş olur?

- I) Kolay ve yakın bir bölgeye giderek
 - II) Araştırmacının araştırma sorularına pozitif yönde cevaplar verileceğini düşündüğü bölgede
 - III) Tüm bölgelerin isimlerini bir kağıda yazıp torbaya atarak rastgele bir kağıt seçerek
- A) Yalnız I B) I ve II C) Yalnız III D) II ve III

1, 4, 1, 2, 1, 5, 1, 2, 3, 3, 4, 1, 1 sayılarının;

11) Modu (tepe değeri) nedir? Cevabınızın nedenini açıklayınız?

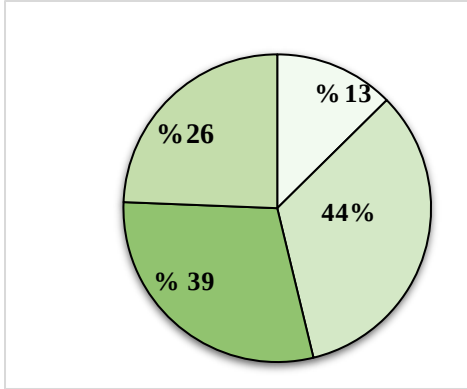
12) Medyanı (ortanca değeri) nedir? Cevabınızın nedenini açıklayınız?

13) Aritmetik ortalaması nedir? Cevabınızın nedenini açıklayınız?

14) Aşağıda elde edilen bazı ölçüm sonuçları verilmiştir.

	Ölçüm Sonucu	Beklenen Değer	Önceki Ölçüm Sonucu	Beklenene Göre Ölçüm Sonucu	Önceki Ölçüme Göre Ölçüm Sonucu
1. Ölçüm	597,6	574,0	585,7	4,1	2,0
2. Ölçüm	646,6	659,2	638,3	-1,9	1,3
3. Ölçüm	555,5	595,3	652,0	-6,7	-14,8
4. Ölçüm	559,3	663,7	647,5	-15,7	-13,6
5. Ölçüm	436,5	407,8	441,9	7,0	-1,2
6. Ölçüm	882,6	698,0	727,2	26,4	21,4
7. Ölçüm	583,6	565,2	533,3	3,3	9,4
8. Ölçüm	509,9	549,1	532,3	-7,1	-4,2

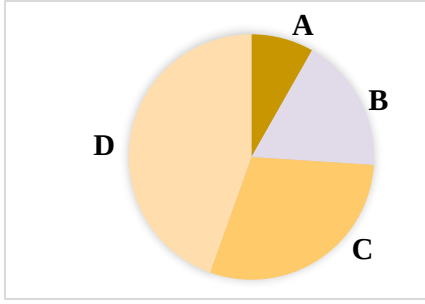
Buna göre hangi değer veya değerler beklenen değer in altında kalmıştır? Nasıl karar verdiniz?



Yanda bir bütünü oluşturan parçaların dağılımı daire grafiği ile ifade edilmiştir.

15) Bu grafiği nasıl yorumlarsınız? Sizce grafikte olağandışı bir durum var mı? Varsa nedir? Açıklayınız.

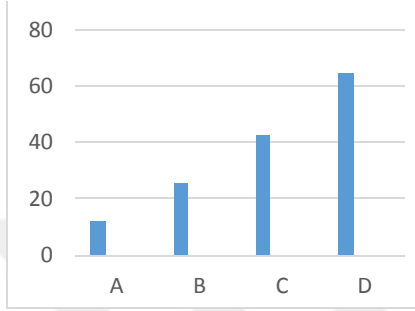
16)



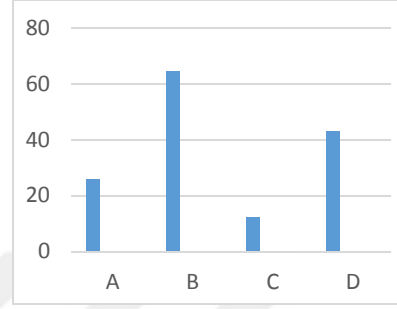
Bir bütünü oluşturan A,B,C ve D parçalarının dağılımı daire grafiğinde gösterilmiştir.

Aşağıdaki sütun grafiklerinden hangisi daire grafiği ile aynı bilgiyi verir?

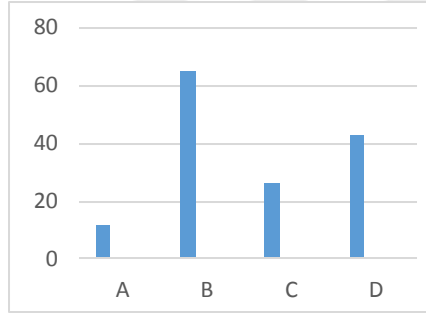
A)



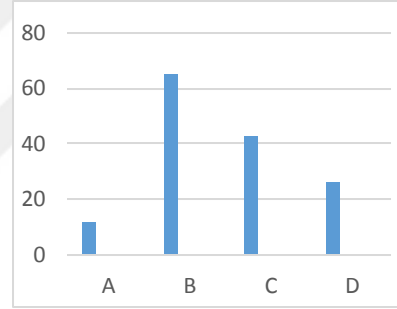
B)



C)



D)



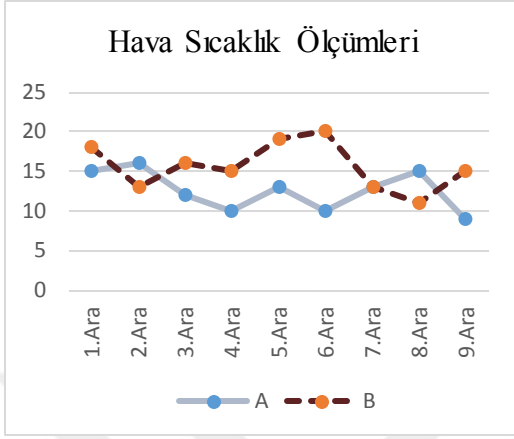
17) Aşağıda A ve B değişkenlerinin 9 farklı aralıkta aldığı değerler verilmiştir.

	1	2	3	4	5	6	7	8	9
A	15	16	12	10	13	10	13	15	9
B	18	13	16	15	19	20	13	11	15

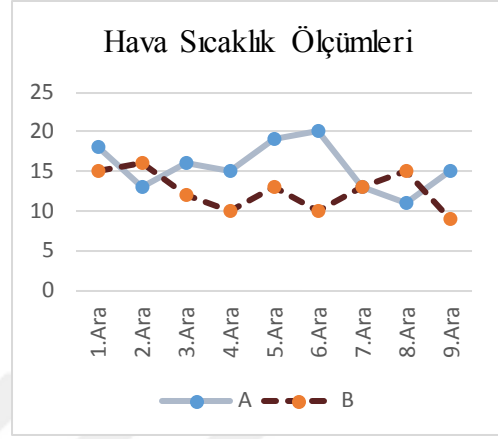
Sizce hangi değişken **daha yüksek değer** almıştır? Nasıl karar verdiniz? Cevabınızın nedenini açıklayınız.

18) Aşağıdakilerden hangisinde tabloda verilen A ve B değişkenlerinin aldığı değerler çizgi grafiği ile gösterilmiş biçimi doğru verilmiştir?

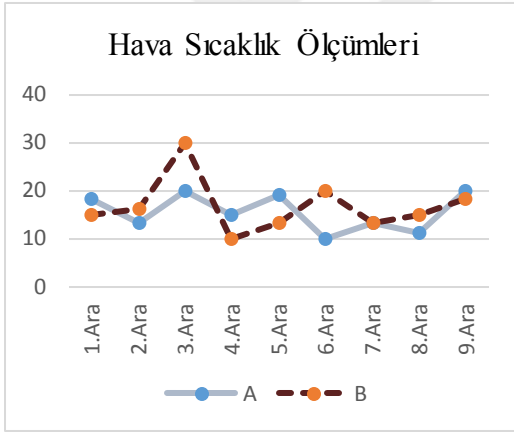
A)



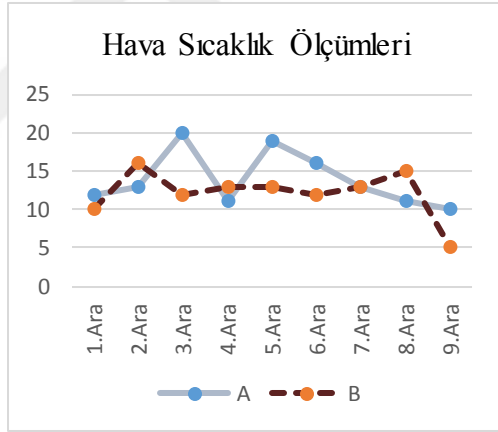
B)



C)



D)

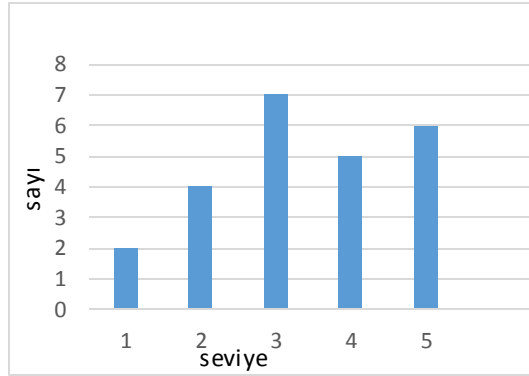


19)

Satır				
1.	1.48	1.97	2.40	3.00
2.	2.89	3.69	4.09	4.45
3.	1.59	2.96	3.46	3.91
4.	1.74	2.40	2.61	3.18

Yukarıdaki tabloda satırlara göre bazı değerler verilmiştir. Örneğin; 1. Satırda 1,48 – 1,97- 2,40 ve 3,00 değerleri vardır. Sizce bu dört satırdan hangisini seçersek en büyük değere sahip olan satırı seçmiş oluruz? Neden? Cevabınızın nedenini açıklayınız.

Aşağıda bir grubun seviyelerine ait yapılan ölçüm sonuçları sütun grafiği ile ifade edilmiştir. (20 ve 21. Soruları grafiğe göre yanıtlayınız.)



20) Kaç kişinin seviyesi 3. Seviyenin üstündedir?

- A) 7 B) 18 C) 11 D) 22

21) Bu ölçüm toplam kaç kişiye yapılmıştır?

- A) 6 B) 7 C) 15 D) 24

“Ankara’da bir grup öğrenci ile yapılan çalışma öğrencilerin doğayı korumak adına çaba sarf etmedikleri, tasarruf için çaba sarfetmedikleri, öğrencilerin çevre konusunda duyarlı olmadıkları ve bu konuda daha dikkatli olmaları gerektiği sonucunu verdi.”(22 ve 23. soruları cevaplayınız)

- 22) Yapılan çalışmanın iddiaları Türkiye’deki bütün çocuklara genellenebilir mi? Türkiye’deki her çocuk için bu iddialar doğru olabilir mi? Neden? Cevabınızın nedenini istatistiksel ifadeler ile ayrıntılı olarak açıklayınız.

	Tekrarlanma Sayısı
A	3
B	4
C	3
E	7
F	1
G	3
H	1

Aşağıdaki tabloda A,B,C,D,E,F ve G’nin bir aralıkta tekrarlanma sayıları ile ilgili veriler verilmiştir.

Tabloya göre aşağıda verilen ifadelerin başına doğru ise D yanlış ise Y yazınız.Cevabınızın nedenini açıklayınız.

- 23) () Verilerin modu E dir.

Neden:.....

- 24) () Verilerin medyanı E’dir.

Neden:.....

- 25) () Verilerin açıklığı 1’den 7’ ye kadardır.

Neden:.....

E. TABLE OF SPECIFICATION

Grade level	Objectives: Students should be able to.... MoNe (2018)	Related Problems
5	Solve problems for interpreting data represented with frequency table or bar graph.	P20, P21
6	Represent the data regarding two groups with frequency table and bar graph	P2, P3
6	Calculate and interpret the arithmetic mean of a data set	P8, P13
6	Use mean and range in comparing and interpreting data regarding two groups.	P17, P19
7	Create and interpret the line graph for the given data.	P4, P5, P6, P7
7	Find and interpret mean, median and mode of a data group.	P9, P11, P12, P13
7	Create and interpret the pie chart for a data group.	P2, P3, P15
7	Represent data by bar graph, circle graph, or line graph and conduct the appropriate conversions between these impressions.	P16, P18