AN INVESTIGATION OF EIGHTH GRADE STUDENTS' STATISTICAL LITERACY, ATTITUDES TOWARDS STATISTICS AND THEIR RELATIONSHIP

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

BY

AYŞE YOLCU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF ELEMENTARY EDUCATION

JUNE 2012

Approval of the Graduate School of Social Sciences

Prof. Dr. Meliha ALTUNIŞIK
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science

Prof. Dr. Jale ÇAKIROĞLU
Head of Department

This is to certify that we have read this thesis and that in our opinion
It is fully adequate, in scope and quality, as a thesis for the degree of Master Science.

Assist. Prof. Dr. Çiğdem HASER
Supervisor

Examining Committee Members

Assoc. Prof. Dr. Erdinç ÇAKIROĞLU (METU, ELE) ________________

Assist. Prof. Dr. Çiğdem HASER (METU, ELE) ________________

Assoc. Prof. Dr. Mine IŞIKSAL (METU, ELE) ________________

Assist. Prof. Dr. Elif YETKİN ÖZDEMİR (Hacettepe Uni., ELE) ________________

Assist. Prof. Dr. Sibel KAZAK (Pamukkale University, ELE) ________________

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 : Ayşe YOLCU

Signature :

ABSTRACT

AN INVESTIGATION OF EIGHTH GRADE STUDENTS' STATISTICAL LITERACY, ATTITUDES TOWARDS STATISTICS AND THEIR RELATIONSHIP

YOLCU, Ayşe
M.S., Department of Elementary Education
Supervisor: Assist. Prof. Dr. Çiğdem HASER
June 2012, 124 pages

The purpose of this study was to investigate the statistical literacy of 8^{th} grade students and their attitudes towards statistics. Moreover, the relationship between their statistical literacy and attitudes towards statistics was examined.

The study was conducted in Yenimahalle district of Ankara in the Spring semester of 2011-2012 academic year. The sample of this study was obtained through cluster random sampling. Nine schools were randomly selected for the study. A total of 1074 eighth grade students in these schools participated. The scales used in the data collection were Statistical Literacy Test (SLT) adapted from Probability Attitudes Scale previously developed for Turkish students (Bulut, 1994) and Attitude towards Statistics Questionnaire (ATSQ) developed by the researcher based on Watson's (1997) three tier framework.

The analysis of the mean scores of statistical literacy in terms of content domains revealed that although sample, graphs, and chance contents had closer mean scores to each other which was around moderate value; average, inference, and variation content domains had lower mean scores.

A one-way within subjects ANOVA indicated that there were significant differences between Tier 1, Tier 2 and Tier 3 aspects of statistical literacy. The pairwise comparisons indicated that students performed lowest in third tier of statistical literacy where students were required to evaluate inappropriate statistical claims. Although, students performed slightly higher in the first tier where they showed their ability in understanding statistical terminology; their performance was the highest in the second tier which was interpreting statistical claims in context.

Eighth grade students' attitudes towards statistics were positive with a mean score of 3.52 in five point scale. The correlation analysis indicated that there were positive and significant relationship between students' attitudes towards statistics and statistical literacy scores.

Key words: Statistical Literacy, Attitudes towards Statistics, Elementary students

ÖZ

İLKÖĞRETİM SEKİZİNCİ SINIF ÖĞRENCİLERİNİN İSTATİSTİKSEL OKURYAZARLIKLARININ, İSTATİSTİĞE YÖNELİK TUTUMLARININ VE BUNLAR ARASINDAKİ İLİŞKİNİN İNCELENMESİ

YOLCU, Ayşe
Yüksek Lisans, İlköğretim Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Çiğdem HASER
Haziran 2012, 124 sayfa

Bu çalışmanın amacı ilköğretim 8. sınıf öğrencilerinin istatistiksel okuryazarlıklarını ve istatistiğe yönelik tutumlarını araştırmaktır. Ayrıca 8. sınıf öğrencilerinin istatistiksel okuryazarlıkları ile istatistiğe yönelik tutumları arasındaki ilişki incelenmiştir.

Bu çalışma Ankara'nın Yenimahalle ilçesinde 2011-2012 öğretim yılının bahar döneminde gerçekleştirilmiş olup, çalışma grubu, küme rasgele örnekleme yoluyla elde edilmiştir. Dokuz okul çalışma için seçilmiş olup toplamda 1074 sekizinci sınıf öğrencisi katılmıştır. Çalışmanın verileri daha önce Türk öğrenciler için geliştirilen Olasılığa Yönelik Tutumlar Ölçeği'nden (Bulut, 1994) adapte edilen İstatistiğe Yönelik Tutumlar Ölçeği (İYTÖ) ile araştırmacı tarafından Watson'un (1997) üç aşamalı modeli temel alınarak geliştirilen İstatistiksel Okuryazarlık Testi (İOT) ile toplanmıştır.

İstatistiksel okuryazarlıkların konu alanları açısından ortalama değerlerinin analizi sonucunda bu konu alanlarının ortalama değerleri arasında farklılıklar vardır. Örneklem, grafikler ve olasılık konu alanlarında birbirine ve orta değere yakın ortalama değerlere sahip olmasına rağmen, ortalama, çıkarım ve yayılma konularında

bu değer daha düşük bulunmuştur.

Öğrencilerin istatistiksel okuryazarlıklarının üç aşaması arasındaki farkı incelemek üzere yapılan çıkarımsal analizin sonuçları bu üç aşama arasında anlamlı bir fark olduğunu göstermektedir. İkili karşılaştırmalara göre öğrenciler uygun olmayan istatistiksel iddiaları değerlendirmeleri gereken üçüncü aşamada en düşük performans göstermişlerdir. Birinci aşama özelliği olan istatistiksel terminolojiyi anlamada biraz daha yüksek performans göstermelerine rağmen, en yüksek performansı istatistiksel bilgileri bir bağlamda yorumlama olan ikinci aşamada göstermişlerdir.

Sekizinci sınıf öğrencilerinin istatistiğe yönelik tutumları 3.52 ortalama değeri ile olumlu bulunmuştur. Son olarak, yapılan korelasyon analizi öğrencilerin istatistiksel okuryazarlıkları ile istatistiğe yönelik tutumları arasında pozitif ve anlamlı bir ilişki olduğunu göstermiştir.

Anahtar Kelimeler: İstatistiksel Okuryazarlık, İstatistiğe Yönelik Tutumlar, İlköğretim Öğrencileri

To My Family

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Assist. Prof. Dr. Çiğdem Haser for her scientific guidance, wisdom and encouragement throughout this research. She always challenged me to do my best and helped me to focus on my research when I was lost in several ideas. Thanks for creating valuable footsteps for me to follow.

I am very thankful to my dearest family for their encouragement and belief in me throughout my undergraduate and graduate studies. My sister Işıl Yolcu, my mother Sevgi Yolcu, my father Hüseyin Yolcu and my uncle Bekir Güçbey, thank you all.

I sincerely thank to my committee members Assoc. Prof. Dr. Erdinç Çakıroğlu, Assoc. Prof. Mine Işıksal, Assist. Prof. Dr. Elif Yetkin Özdemir for their invaluable contributions for my study. A special gratitude to Assist. Prof. Dr. Sibel Kazak for her suggestions, feedbacks and comments for my study from the very beginning to willingness to serve as a committee member.

I am also grateful to Assoc. Prof. Dr. Semra Sungur for her suggestions and contributions for the instruments in this study. A special thanks to my colleague Bilge Gök who dealt with all my questions during the test development process.

I want to express my gratitude to my colleague in the department Bahadır Yıldız who helped me very much while finalizing this thesis and his contributions for the study.

I would like to express my thanks to my officemates Şenay Özen Altınkaynak, Ayşegül Deniz and Seda Ata for their sincere support, help and insights during this research even in my desperate times at work. My special thanks are for my friends Rezzan Doktoroğlu, Halime Samur, Nagehan Orhan for their help and support for this study.

I would like to thank TUBITAK for their scholarship which helped me pursue my master study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATSQ: Attitudes towards Statistics Questionnaire

ANOVA: Analysis of Variance

H0: Null Hypothesis

HEC: Higher Education Council

KMO: Kaiser-Meyer-Olkin

M: Mean

MoNE: Ministry of National Education

N: Number of Participants

NCTM: National Council of Teachers of Mathematics

PAS: Probability Attitude Scale

PASW: Predictive Analytics SoftWare

SATS: Student Attitudes toward Statistics

SD: Standard Deviation

SE: Standard Error

SLT: Statistical Literacy Test

SOLO: Structure of Observed Learning Outcomes

CHAPTER 1

INTRODUCTION

Statistics has become positioned in our daily lives at present times from health care to politics. Put it differently, individuals should have some knowledge of statistics in order to contribute in an informed way to the debates which include both statistical messages and wider social context such as politics, health and education related issues (Carmichael, 2010). In the context of this study, such knowledge refer the term “statistical literacy” which can be defined as the capability to understand, interpret, and critically evaluate statistical messages (Watson, 1997). Statistical literacy provides emancipations of citizens through enabling them not only to read data but also evaluate and communicate with statistical messages (François & van Bendegem, 2010), which have crucial roles in both personal and public decision making where daily lives of individuals and society are full of statistics (Wallman, 1993). In addition, statistical literacy is regarded as a way for participatory democracy and equity due to the fact that students are not simple consumers of data, rather they are becoming more engaged in public policy discourse which involves quantitative or statistical information (Cobb, 1999). However, statistical descriptions appear in everyday life or media often involve bias or lack of objectivity in a considerable number of situations (Frankenstein, 1998; Gal, 2004). The need for understanding and evaluating critically those one-sided, misleading, or biased statistical claims have been acknowledged by several researchers (e.g. Frankenstein, 1998; Gal, 2004; Watson, 2006). Therefore, school systems and mathematics curricula do have an important role in the development of statistical literacy which helps building active and critical citizenship (François & van Bendegem, 2010; Frankenstein, 1998; Gal, 2004; Watson, 2006).

Statistical literacy, which consists of statistics and probability concepts, might be considered as the “meeting point of the statistics and probability curriculum and everyday world” (Watson, p.11, 2006). Being statistically literate is very important for every individual as informed citizens who can interpret the statistical messages in various contexts. Thus, statistical literacy, which is becoming a part of the school mathematics curriculum, plays a key role in preparing students to encounter the needs of society when they complete their elementary education (Watson & Callingham, 2004).

Correspondingly, “Statistics and Probability”, as one of the five content areas of elementary mathematics curriculum in Turkey (MoNE, 2005), is closely related to individuals’ daily lives and contributing them to develop critical thinking about data and ability to make judgment based on data which is essential for becoming informed citizens (NCTM, 2000). For this reason, it is aimed in many curricula that students should not only know and apply required statistical knowledge and skills in their daily lives and other subject areas, but also they should develop awareness of the importance of statistics and probability (MoNE, 2005; NCTM, 2000; Watson & Callingham, 2003).

Having known the importance of statistical literacy in daily lives of individuals as the ability of making decisions in social issues, elementary mathematics curriculum revisions regarding those concepts can be seen both in Turkey and throughout the world. For example, National Council of Teachers of Mathematics (NCTM) in the United States (2000) indicates developing critical thinking and judgment based on data as one of the fundamental goals of the mathematics curriculum. In the same way, Turkey has undergone through some revisions in the elementary mathematics curriculum considering the inclusion of statistics and probability as one of the five content areas of school mathematics. Statistics and probability domain of the school mathematics curriculum consists of concepts such as sampling, measure of central tendency, graphs and tables, measure of spread, probability, and beginning inferences (MoNE, 2005; 2009). Objectives regarding statistics and probability in elementary mathematics curriculum in Turkey generally aim at developing informed citizens who possess knowledge of statistical concepts and who have appreciation of the importance of statistics in society. In addition, appropriate interpretation, conjecturing, and predicting based on data are emphasized. Additionally, the dispositional issues such as building positive attitudes towards statistics and probability are also emphasized (MoNE, 2005). Yet, in Turkish context, evidence whether elementary school students are statistically literate or not is scarce. Therefore, this study aims investigating statistical literacy of 8^{th} grade students.

In the context of study, statistical literacy is defined as the ability of understanding, interpreting, and evaluating statistical messages in various contexts and it is represented in a three tiered framework as for meeting the needs of cognitive aspects in the middle school context (Watson, 1997). More specifically, the first tier refers to the familiarity with terminology used in statistical messages in media. To illustrate, understanding the term "average" in context or defining "average" is a feature of tier 1. On the other hand, the second tier includes the interpretations of these terms where they are contextualized in statistical claims. For example, interpreting or applying ideas of average in a variety of context is a characteristic of tier 2. The last tier is the ability to question others' statistical reports critically; in other words, the critical evaluation of biased statistical information and posing possible critical questions to this information constitute the third tier of statistical literacy. For instance, examining whether mean or median is an appropriate average in a given statistical report is a characteristic of tier 3 (Watson, 1997, 2006; Watson & Moritz, 2000).

In the literature, there were several studies investigating statistical literacy in the context of middle school students with respect to age or grade level. The results of these studies consistently reveal that as grade level increases, students' performance in statistical literacy also increases (e.g. Aoyama & Stephens, 2003; Watson & Kelly, 2008). In addition, effect of utilization of media reports on understanding statistics in a positive way is documented (Doyle, 2008; Merriman, 2006). Besides, although inconsistent, there was a considerable research regarding statistical literacy in tertiary level. Yet, the research on statistical literacy considering Turkish context is rather scarce. Therefore, this research aims to investigate statistical literacy of 8^{th} grade Turkish students.

The theoretical framework of the statistical literacy concerns not only the cognitive aspects, but also appreciates the importance of affective or dispositional issues

(Watson, 2006; Gal, 2004). Gal (2004) states that certain beliefs and attitudes are required in order to critically evaluate statistical messages as a part of the statistical literacy. They also have a central role in the teaching and learning process during class time, for the statistical behavior out of the class, and enrollment in further statistics related courses (Gal, Ginsburg & Schau, 1997). Students' attitudes can help or hinder statistical thinking and they do have influence on the application of knowledge and skills in variety of contexts (Gal, Ginsburg & Schau, 1997). In addition to these, Watson (2006) indicates that low motivated students about statistics or students holding procedural beliefs about statistics tend to perform less in statistical literacy.

Although attitudes towards statistics are considered very important to investigate, there is no exact definition of these constructs. Indeed, researchers often define attitudes in relation to what their assessment instruments measure (Gal, Ginsburg & Schau, 1997). According to Gal et.al (1997), where they apply McLeod's (1992) terminology, the emotions and feelings including positive and negative responses experienced by individuals during learning statistics constitute attitudes towards statistics (Gal, Ginsburg & Schau, 1997). There is a significant body of research concerning attitudes in mathematics education (Mcleod, 1992); however, there is a lack of research on affective issues in statistics education and particularly statistical literacy (Carmichael, 2010; Gal & Ginsburg, 1994). The studies have been done in statistics education investigating affective domain are limited to undergraduate and graduate levels of university education (e.g. Duisburg, & Brisbane, 2005; Evans, 2007; Schau, 2003; Tempelaar, Loeff, & Gijselaers, 2007).In Turkish context, although some research investigating affective issues in the context of statistics education or statistics related courses have been done with undergraduate and graduate students (Aksu, & Bikos, 2002; Doğan, 2009), there is a dearth of research studies in statistics education related to affective variables including attitudes in middle school context except a few experimental studies which examine the change in attitudes towards statistics (Yılmaz, 2006). Therefore, this study also aims at investigating attitudes towards statistics and the relationship between statistical literacy of 8^{th} grade students and their attitudes and beliefs about statistics. For the purposes of the current study, Gal, Ginsburg, and Schau's (1997) definition is extended with the inclusion of opinions and thoughts regarding statistics. Hence, opinions and thoughts about statistics together with emotions and feelings experienced by students while learning statistics refer attitudes towards statistics in the context of this study.

1.1. Purpose of the Study

Having emphasized the importance of being statistically literate in everyday data driven world, this study has several purposes. The first purpose is to investigate 8th grade students' statistical literacy with respect to three tiers. The other purpose is to investigate attitudes towards statistics that contribute the statistical literacy of 8th grade students and the relationship between these attitudes and the statistical literacy.

1.2. Research Questions and Hypothesis

This study intends to investigate these research questions and related hypothesis;

1. What is the statistical literacy of 8th grade students in terms of content domains (sample, average, graph, chance, inference and variation)?

2. What is the statistical literacy of 8th grade students in terms three tiers?

3. Is there a significant difference between the mean scores of first, second and third tier of statistical literacy?

   H_0: There is no significant mean difference between the mean scores of first, second and third tier of statistical literacy.

4. What are attitudes of 8th grade students' towards statistics?

5. Is there a significant relationship between statistical literacy of 8th grade students and their attitudes towards statistics?

   H_0: There is no significant relationship between statistical literacy of 8th grade students and their attitudes towards statistics.

1.3. Significance of the Study

We live in a world that is full of quantitative information. Statistical messages take place in media including various kinds of arguments, advertisements, or suggestions (Ben-Zvi & Garfield, 2004). The ability to understand, interpret, and critically evaluate the statistical messages in daily lives of individuals which refers to statistical literacy (Watson, 1997) becomes very important in information societies. Being statistically literate holds a crucial role in economic, social and political participation of every individual as active and critical citizen in the data driven world. Thus, statistical literacy, which is a part of the school mathematics curriculum, has an important place while preparing students to encounter the needs of society when they complete their compulsory education (Watson & Callingham, 2004). Elementary school curriculum in Turkey aims at developing informed citizens who possess knowledge of statistics with an appreciation of the importance regarding the position of statistics in society (MoNE, 2005). Since the importance of statistics, particularly statistical literacy, is acknowledged by several researchers, educators and curriculum documents both international and national context, it should be questioned that whether students near the end of their elementary education are statistically literate or not. Although there has been studies in the mathematics education literature examining statistical literacy from different aspects such as sampling or graphing in terms of grade level (e.g. Aoyama & Stephens, 2003; Watson & Kelly, 2008), the research is scarce in middle school context in Turkey in the accessible literature. Therefore, one of the purposes of this study is to investigate the statistical literacy of 8th grade students in terms of Watson's (1997) three tiers and content domains.

Additionally, since dispositional elements are highly recognized in statistics education (e.g. Shaughnessy, 2007; Wild & Pfannkuch, 1999) and particularly in statistical literacy (Carmichael, 2010; Gal, 2004; Watson, 2006), the other aim of this research is to investigate the attitudes towards statistics of 8th grade students and the relationship between attitudes towards statistics and statistical literacy. In order to interpret, evaluate, and question statistical messages in various contexts critically, some set of attitudes are required. Due to the fact that the research investigating affective issues in statistics education is limited with undergraduate or graduate students (e.g. Evans, 2007; Schau, 2003; Tempelaar et.al 2007) and number of studies examining attitudes towards statistics in middle school context is few (e.g. Calderia, 2010; Yılmaz, 2006). Hence, there is a need in researching the middle school students' attitudes towards statistics and its relationship with statistical literacy to fill a gap in the statistics education literature (Gal, Ginsburg & Schau, 1997; Shaughnessy, 2007). Therefore, this research provides an opportunity to examine 8^{th} grade students' attitudes towards statistics and the relationship between attitudes towards statistics and statistical literacy.

It is possible to understand statistical literacy of 8^{th} grade students and its relationship with attitudes towards statistics through this study, which does not only contribute to the mathematics/statistics education literature but also provides teaching implications for teachers, curriculum developers, and educational policy makers. Students' responses to the items which are contextualized with the first, second, and third tier of statistical literacy are categorized through what students are able to do or not able to do; therefore, this study provides a brief reflection of the current elementary mathematics curriculum in Turkey with respect to students' capabilities in statistical literacy. In addition, these categorizations inform both pre-service and in-service teachers regarding students' abilities in statistical literacy in terms of three tiers and specific content domains of statistics. The results may point out further research in the context of both mathematics curriculum development and teacher education. Since this research also includes the dispositional aspects of statistics education, the implication of results could not only to be extended to the classroom culture including mathematics teaching and communication but also teachers' affect regarding statistics. In the same way, the results considering the relationship between attitudes towards statistics and statistical literacy will contribute to the literature which does not have sufficient number of studies related to these issues.

1.4. Definitions of Important Terms

Statistical Literacy: The present study employs the definition of statistical literacy given by Watson (1997) as the ability of understanding, interpreting, and evaluating statistical messages in various contexts and it is represented in a three tiered framework as measured by Statistical Literacy Test (SLT). The descriptions of these tiers as follows:

> Tier 1: Familiarity with terminology used in statistical messages
>
> Tier 2: Interpretations of these statistical terms where they are contextualized in statistical claims which appears in the media or elsewhere
>
> Tier 3: The ability to question others' statistical reports critically; in other words, the critical evaluation of biased statistical information and posing possible critical questions to this statistical information

Attitudes towards Statistics: For the purposes of this study, attitudes towards statistics are defined as opinions and thoughts about statistics together with emotions and feelings experienced by students while learning statistics as measured by Attitudes towards Statistics Questionnaire.

CHAPTER 2

REVIEW OF RELATED LITERATURE

The purpose of this study is to investigate the statistical literacy of 8th grade students in terms of content domains and three tiers and to investigate the relationship between students' statistical literacy and attitudes toward statistics.

In this section, the studies in the literature and the theoretical perspective this study employs are introduced. Since the purpose of this study is twofold, literature review is separated into two main parts: In the first part, the review of related literature about statistical literacy is presented and the second part of the literature review of this study is devoted for the studies which examined attitudes toward statistics and the relationship between statistical literacy and attitude toward statistics. In each part, the definitions or theoretical perspectives from different researchers and related studies are presented.

2.1. Statistical Literacy

2.1.1. Definitions of Statistical Literacy

Although the importance of statistical literacy and statistical understanding is acknowledged by many teachers, educational researchers, and curriculum documents both in Turkey and in the international arena (Ben-Zvi & Garfield, 2004; MoNE, 2005; NCTM, 2000), it is still a new area of research in the field of mathematics education and there is not a consensus on what the statistical literacy construct is (Shaughnessy, 2007). Therefore, definitions of statistical literacy construct are discussed below.

Carmichael addressed the dictionary definition of statistical literacy might be "an ability to interpret statistical messages and where necessary communicate such messages using the written or spoken word" (2010, p.9). However, this kind of definition is regarded as too narrow or incomplete since it lacks an indication of evaluating statistical messages with a critical eye (Carmichael, 2010). In parallel with this idea, Wallman (1993) provides the following definition for statistical literacy:

> "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 (p.1).

In this brief definition, Wallman (1993) includes not only the ability to understand data related arguments in context but also critically evaluate them with an appreciation of statistical thinking. In addition to the cognitive demands of statistical literacy in this definition, the personal and societal needs of individuals are emphasized. In other words, statistical literacy is essential for either private or public decision making where daily lives of individuals or collective groups are permeated by statistical results. Consistently, the highlight of context is important in terms of statistical literacy, since statistical literacy basically deals with data in context (Watson, 2006).

Schield (1999) defines statistical literacy as one's ability of addressing a critical thinking about opinions by considering statistics as evidence in the context of business college students. According to Schield, to be statistically literate addresses to be able to distinguish between casual and not causal relations, sample statistic, and population parameter, and some other characteristics related to college level statistical knowledge such as power test and inferential statistics. Although such kind of knowledge is presented usually in introductory statistics courses, statistical literacy is considered as a daily life skill (Schield, 1999). It is also reported that students are not willing to take further statistics course unless they have to do (Schield, 2004), which can be related to the attitudes or the affect of those students.

Gal (2004) offers a statistical literacy conceptualization and its elements in a model for adults or "future adults", in his term. In this model, communication with statistics, interpretation and judging of statistical claims are treated as the possible skills of statistically literate individuals. These skills are essential for active citizenship in societies where information pervades.

Ben-Zvi and Garfield (2004) make a distinction between statistical literacy, reasoning, and thinking, and define statistical literacy as a collection of basic and important skills which are essential for comprehension of statistical information or research findings, quite different from the previous definitions given. Organizing data, constructing graphs and tables, representing data, and understanding basic statistical terminology constitute those basic skills. Although critical evaluation of statistical claims is involved in statistical thinking, which addresses a higher cognitive level than statistical literacy, they state that the fundamental goal of statistics instruction is to make students statistically literate (Ben-Zvi & Garfield, 2004).

The definition of statistical literacy which this study employs is the capability to understand, interpret, and critically evaluate the statistical messages in daily lives of individuals (Watson, 1997). Statistical literacy is described along a three-tiered continuum where the first tier refers to the familiarity with terminology used in statistical messages in media, the second tier includes the interpretations of these terms in a variety of context, and the last tier is the ability to question others' reports critically (Watson, 1997). Statistical literacy, which is "the meeting point of data and chance curriculum and everyday world" (Watson, 2006, p.11), plays a key role in training students to meet the needs of society when they complete their compulsory education and it is regarded as a survival skill for out of school contexts (Watson & Callingham, 2003).

Most of the definitions of statistical literacy include understanding statistical messages in various contexts and evaluate the appropriateness of those messages. Hence, statistical literacy does not only address knowing formulas and definitions in the curriculum, but also to integrate such kind of knowledge with various contexts where statistical messages, claims, or statements appear (Watson & Callingham, 2003). For example, statistical literacy deals with whether mean is the best representation of data or not and whether the sample is suitable to generalize findings to a particular population or not. Critical judgment of statistical claims which appear in the media (Gal, 2004; Watson, 1997) is also common in many of the definitions. Although there is a difference between statistical literacy, reasoning, and thinking (Ben-Zvi & Garfield, 2004), statistical thinking is related with statistical reasoning and literacy (Chance, 2002). Therefore, most of the definitions emphasize critical evaluation of statistical claims through statistical thinking as an important part of statistical literacy (Gal, 2004; Schield, 1999; Wallman, 1993; Watson, 1997; Watson & Callingham, 2003). Hence, together with understanding and interpreting statistical messages, this study also focuses on students' questioning of statistical claims in daily life. This aspect of statistical literacy includes asking critical questions about claims without proper statistical foundation.

2.1.2. Models /Theoretical Frameworks for Statistical Literacy

Although definitions for statistical literacy are given by numerous authors, few of them conceptualize statistical literacy in depth (e.g. Gal, 2004; Watson, 1997). Therefore, in this section of literature review, the existing statistical literacy models and frameworks given by scholars are discussed.

2.1.2.1. Gal's Model of Statistical Literacy

Gal's (2004) model of statistical literacy consists of knowledge and dispositional elements presented in Table 2.1. Specifically, knowledge elements consist of literacy skills, mathematical knowledge, statistical knowledge, critical questions, and context knowledge; dispositional elements consist of critical stance and attitudes and beliefs. Each knowledge and dispositional element is described in the following table.

Table 2.1 Model of Statistical Literacy

Source: (Gal, 2004, p.51)

Knowledge Elements	Dispositional Elements
Literacy skills Statistical knowledge Mathematical knowledge Context knowledge Critical Questions	Beliefs and attitudes Critical stance
Statistical Literacy	

2.1.2.1.1. Knowledge Elements

Understanding or interpreting statistics necessitated not only knowledge of statistics but also other knowledge types as mentioned in Table 2.1. It is proposed that each knowledge base is contributed for statistical literacy of individuals (Gal, 2004) and these are described below in detail.

Since most of the statistical messages are represented through oral or written text form, activation of specific literacy skills such as understanding or making inference is required in order to deal with those messages. Additionally, awareness of certain statistical terminology such as random, average, and representativeness is counted as literacy skills as a component of knowledge elements.

"Knowing why data are needed and how it is produced, familiarity with basic terms and ideas related to descriptive statistics and graphical and tabular displays, understanding basic notions of probability and knowing how statistical conclusions or inferences are reached" (Gal, 2004, p.58) constitutes the statistical knowledge element for statistical literacy. For example, knowledge of big ideas in statistics, such as variation, is fundamental for understanding data sets or distributions. Additionally, knowledge of how data is represented in graphical form is one of the basics of statistical knowledge element since graphical or tabular displays are the most frequent forms of statistics that are confronted in daily life.

Although extra emphasis on mathematical computations while doing statistics is not good for the development statistical ideas and concepts, certain mathematical knowledge such as calculating mean and percent or proportional reasoning is required in order to understand the statistical ideas in the context of statistical literacy. Therefore, *mathematical knowledge* constitutes a knowledge element for the statistical literacy.

Certain context or world knowledge is important in order to make sense of data. While looking for sources of variation and error, *context knowledge* is the main source for such a familiarity, and therefore an important knowledge element.

Sources of statistical messages in the media are not always objective and are usually biased. Therefore, a critical evaluation is needed for those messages. Gal (2004) lists some "worry questions" such as whether sample size is large enough, whether a given graph is properly drawn, or does it alter certain tendencies in the data, whether any unintended variable explains the findings other than the variables included in statistical information. Hence, in order to evaluate objectivity or credibility of statistical reports, asking *critical questions* is crucial.

Each of the knowledge elements might seem separate; however, statistically literate individuals use these elements interdependently in a dynamic relationship (Gal, 2004). For example, while dealing with statistical reports in the newspapers, one should have literacy skills together with context knowledge to grasp the meaning. Also, statistical and mathematical knowledge is needed to understand how data are produced and why they are presented in a certain way. For critical evaluation of those messages, asking critical questions is required.

2.1.2.1.2. Dispositional Elements

Critical evaluation of statistical reports has been emphasized in many definitions; therefore, statistical literacy involves a specific type of action, not just knowing terminology and passive interpretation of them (Gal, 2004). In order to activate the knowledge elements, certain dispositions such as a critical stance and attitudes and beliefs are required. Gal (2004) refers them in his model as "dispositional elements" of statistical literacy. Statistically literate persons should have a questioning attitude toward quantitative messages through asking "worry questions." This questioning attitude is a part of *critical stance* and it is an essential part of statistical literacy since those messages could be misleading, biased, or one-sided. In order to have a critical stance toward data and have motivation for taking action, certain *beliefs and attitudes* are required (Gal, 2004). For the purpose of defining them, Gal (2004) utilizes McLeod's (1992) conceptualization of affective domain in mathematics education in which attitude is described as an affective construct related to positive and negative responses about statistics while belief is considered as a cognitive construct related to opinions or ideas.

2.1.2.2. Watson's Three-Tiered Framework for Statistical Literacy

Watson (1997) develops a three tiered statistical literacy framework for addressing the cognitive characteristics of statistical literacy. The level of complexity of each tier is consistent with learning models in developmental psychology such as Biggs and Collis's (1982) SOLO (Structure of Observed Learning Outcomes) model (as cited in Watson, 1997). These tiers are the foci of this study and they are described below in detail.

The first tier of this framework is basic understanding of terminology. The skills in this tier include the understanding the terminology of statistics without considering the context (Watson, 1997). This terminology includes specific concepts in the curriculum such as sampling, average, graphing, random, and variation. This tier also involves calculation of measures of central tendency or measure of spread without any reference to the social issues in the daily lives of students.

The second tier of the framework, which is contextualization of terminology including statistical language and concepts, requires students to read reports of statistics and interpret them, rather than only performing the statistical computations (Watson, 1997). The level of statistical literacy in this tier involves embedding the statistics terminology in the context of wider social discussions such as a comprehension of risk in situations where decision making is needed or drawing conclusions and inferences from graphs and charts.

Questioning of statistical claims is the third tier of this framework. Students confidently challenge the statistical claims in the media in the third tier. The skills in this tier include constant questioning attitude towards statistical conclusions without proper statistical foundation. The important thing for statistical literacy in this tier is that students develop intelligent questions for data and related claims instead of believing everything they read in the media (Watson, 1997).This framework is mainly meeting with cognitive part of statistical literacy and affective issues are not directly indicated.

2.1.2.3. Watson's Components of Statistical Literacy

In addition to previous conceptualizations and frameworks of statistical literacy, Watson (2006) constructed a conceptual map for statistical literacy which shows the contribution of each element and the relationship between them. This framework for statistical literacy consists of mathematical and statistical skills, context, task motivation, task format, literacy skills, and variation which is presented in Figure 2.1 and explained below.

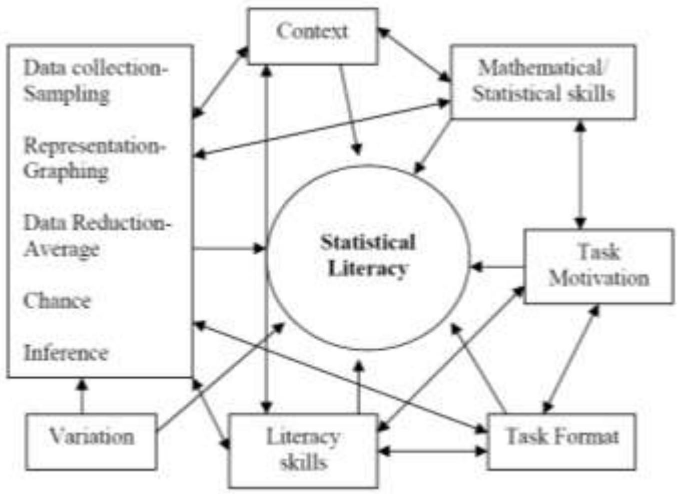


Figure 1 Links among the Components of Statistical Literacy (Watson, 2006, p.248)

Mathematical skills related to statistical literacy do include understanding proportions, percents, and part-whole relationships which is also an expectation of the middle school curriculum and should be achievable at the end of the middle school. Likewise, statistical skills consist of understanding and calculation of average and probability. Definition of basic terms is also counted as statistical skills (Watson, 2006).

Role of context in statistics education is appreciated by many researchers (Gal, 2004; Langrall, 2010; Pfannkuch, 2011). In the same way, Watson (2006) considered context a significant part of statistical literacy and mentioned that context exists in three forms: isolated context (e.g. flipping coins), familiar context (e.g. school surveys), and unfamiliar context (e.g. extracts from media).

General literacy skills as described by Luke and Freebody (1997) are important parts of statistical literacy since statistics is presented in some form of text (as cited in Watson, 2006). Watson (2006) portrayed literacy skills in four elements where element one performance includes code-breaking such as understanding a graph presenting information, element two pertains making the meaning of data in context, element three refers to creating alternative meanings from information presented in statistical messages, and critical aspects such as questioning statistical claims constitute element four. These four elements are closely in line with three-tiered statistical literacy framework which includes understanding, interpreting, and questioning information.

Multiple-choice and open-ended tasks can be used while assessing statistical literacy. Multiple-choice questions can be more helpful because they allow students to agree with an answer instead of constructing it, which is demanded by most of the statistical literacy levels. On the other hand, open-ended items can allow students to create answers and perform the questioning aspect of statistical literacy (Watson, 2006).

For any task format, students have some dispositions towards it. Therefore, task motivation is also labeled as a component of statistical literacy. Similar to researchers who have mentioned about attitudes toward statistics (Gal, 2004) or about dispositions as a part of statistical investigations (Wild & Pfannkuch, 1999); task motivation can assist in drawing conclusions or making informal inferences (Watson, 2006).

Watson's (1997) three tiered framework is at the core of present study and used as a main analysis of students' statistical literacy. As other models enhance this framework through mentioning affective dimension during statistical investigations, still three tiered framework pervades cognitive domain of statistical literacy. To illustrate the similarity between other models, understanding statistical terminology which is the first tier can be found in them as literacy skills (Gal, 2004) and as a part of statistical skills (Watson, 2006). In addition, the last two frameworks for statistical literacy have several issues in common. For example, both Gal (2004) and Watson (2006) argue that certain mathematics, but not advanced, is a requirement for statistical literacy. Although mathematical knowledge is not mentioned in Watson's (1997) three tiered framework, understanding and interpreting statistical messages are also a similarity between these three models. Moreover, it is declared that the role of affect, attitudes, or motivation as a disposition of individuals is important while critically evaluating or questioning statistical claims. Besides, context is appreciated as an important part of statistics or statistical literacy; hence, these three frameworks take context into account.

In this study, the two aspects of statistical literacy are examined where cognitive aspect of statistical literacy is conceptualized through Watson's (1997) three tiered framework while dispositional aspect is attitudes towards statistics. The tiers include understanding, interpreting, and critical evaluation of statistical claims. Understanding terminology which refers first tier can be counted as a literacy skill as stated by Gal (2004) and Watson (2006). Although context is not mentioned directly in the definition, three tiers are associated with diverse contexts since all of the frameworks take context into consideration. In the light of these two aspects, statistical literacy of eighth grade students were examined through these three tiers and attitudes towards statistics.

2.1.3. Research on Statistical Literacy

In the literature, studies investigating statistical literacy are very limited in number. One of the pioneering researchers is Watson and her colleagues, especially in survey studies conducted with middle school students. For instance, Watson and Moritz (2000) investigated the understanding of sampling concept related to statistical literacy with middle school students in grades 3, 6 and 9. The research instrument was consisting of 11 items in relation to three tiered framework for statistical literacy (Watson, 1997) and students' answers were evaluated as pre-structural, uni-structural, multi-structural, and relational levels where individuals possess non-statistical, single, and multiple and interrelated statistical ideas respectively in those levels. The results revealed that a developmental sequence existed in conceptualizing sampling as in the first and second tier of the statistical literacy. That is, the performances in questions related to first and second tier showed a correspondence with pre-structural, uni-structural, multi-structural, and relational levels in terms of grade level of participants. However, performances of students in the third tier tasks showed that questioning claims in sampling context were suitable for students who were not in the relational level yet. In addition to these results, to examine the longitudinal development of sampling concept related to statistical literacy, they conducted the same study after two years and after four years with the same students. The results indicated that 15% of the students performed at a lower level, 48% performed at a higher level and 37% did not change their level compared to previous assessment. However, within four years, only 7% of the students performed at a lower level, 24% of them did not change their level and 60% performed at a higher level. Performing at lower level might be related to motivation of students whereas performing higher level within four years indicated that understanding sample concept might develop gradually.

In addition to this progressive three tiered framework, Watson and Callingham (2003) conducted a study to verify the hierarchical nature of statistical literacy across grades from 3 to 9 with approximately 3000 students using Rasch analysis techniques. The results showed parallel structure with Watson's three tiered framework. More specifically, they identified six hierarchical levels from idiosyncratic thinking to critical- mathematical with respect to statistical literacy. The skills that students exhibit at these levels are given in the following table:

Table 2.2 Levels of statistical literacy construct (Watson & 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 and 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 associates with 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 two levels, which were idiosyncratic and informal, were related to intuitive understanding of terminology and concepts with no context engagement, whereas inconsistent and consistent non-critical levels showed an ability to perform calculations and little engagement with context. In the last two levels, critical and critical-mathematical, students could both show appropriate calculations and appreciation with context. The only difference between these last two levels was the proportional reasoning that was included in the last level. Although this study showed a consistency with the previous research since the first two, the next two and the last two levels of the hierarchy were related to tier 1, tier 2 and tier 3 respectively, the results suggested that appropriate use of terminology and critical evaluation of statistics appeared in critical mathematical level. In this study majority of students were placed in informal, inconsistent and consistent non-critical levels whereas a small percentage of students' responses were categorized in critical and critical-mathematical levels (Watson & Callingham, 2003).

In addition to putting statistical literacy in a framework which shows a hierarchical nature, in this study, the importance of dispositional elements are not ignored through indicating certain attitudes and beliefs which are required for questioning engagement with the tasks in critical level (Watson & Callingham, 2003).

Graph interpretation aspect of statistical literacy was investigated to document the differences regarding grade levels by Aoyama and Stephens (2003). Their study revealed that there is an identification of levels of graph interpretation from A-basic reading tables and graphs- to F -creating new dimensional information-, which is the highest level. The participants of this study were 55 students from grades 5 and 8 who answered the basic graph reading and Level F tasks. The results revealed that 95% of eighth grade students and likewise 82% of fifth grade students could read the beyond data in lower level tasks. However, there was no appropriate response from both of the groups for Level F task which is consistent with the sixth level –critical mathematical- of Watson and Callingham (2003) statistical literacy levels. The reason for this result was that students lacked sufficient experience with evaluating statistical information in graphs both in and out of school settings.

Watson and Kelly (2008) investigated the literacy aspect of statistical literacy where they asked vocabulary of statistical literacy across grades. More explicitly, they wanted students who are in grades 3 and 5 (N = 359) to answer the definition of sample while students who are in grades 7 and 9 (N = 379) to answer additionally definitions of random and variation. There is a significant difference between grades on sample related tasks. Although they did not find significant difference between grades 7 and 9, there was a significant difference between grades 3 and 5 with a medium effect size and there was a significant difference between grades 5 and 7 with a small effect size. Additionally, they did not find any significant difference between grades on random and variation related tasks. The researchers also investigated the effect of a specialized instruction related to chance and data with a smaller group of students. The change after instruction was evaluated with paired sample t-tests which indicated that there were significant differences in students grade 3, 5 and 7 from pre-test to post-test in sample related tasks. On the other hand, grade 7 students improved significantly on both for the terms random and variation whereas grade 9 students performed significantly better only on definition of random.

Using media reports and real world data has an important place in the development of statistical literacy in school mathematics (Ben-Zvi & Garfield, 2004; Shaughnessy, 2007). For instance, Merriman (2006) investigated the effect of a specifically designed course unit using media reports on statistical literacy of students aged 14. Students' responses were evaluated through hierarchical coding based on SOLO (Structure of Observed Learning Outcomes) framework. The results revealed that students' responses showed a consistency with previous research conducted by Watson and Callingham (2003) which revealed a hierarchical level. In addition, there was a significant improvement in students' statistical literacy scores after the course unit which was utilizing media reports. Although there was not a significant correlation between mathematical ability and statistical literacy scores, English language and statistical literacy had positive linear correlation since reading media reports required effort and time. Both teachers and students claimed that using media reports for teaching statistics was interesting and worthwhile.

In an action research conducted by Doyle (2008), teaching statistical concepts through adopting critical statistical literacy approach by highlighting language and critical thinking skills in the classroom was investigated. In this study, three participating teachers integrated statistical literacy in their mathematics curriculum and adopted critical statistical literacy approach in secondary mathematics classroom in New Zealand where students were exposed to critical thinking skills and questioning statistical reports. The activities in this research included media reports and real world data where the language learning principles was the focus for the development of statistical literacy. The results revealed that there was an improvement in students' understanding and conceptions in statistical literacy tasks where language and literacy skills were highlighted. Both Merriman's (2006) and Doyle's (2008) studies indicated that language did have an important place in development of statistical literacy with school students.

Several studies could be found in the literature which examined statistical literacy in tertiary level since statistical literacy was also labeled as a goal for introductory statistics courses (Rumsey, 2002). One of the descriptive studies, apart from middle school students, was the study of Schield (2006) in which he investigated the statistical literacy of college students, professional data analyst and college teachers (n=169) where survey instrument focused on informal statistics such as reading tables and graphs. In general, 44%, 65% and 81% of participants misread row table, pie chart and X-Y plots respectively. Among 49% of college students, 44% of data analysts and 28% of college teachers made error on average in this survey instrument. Moreover, the effect of introductory statistics course together with different teaching methods had been a concern for research in statistics education in tertiary level. However, the results of those studies regarding introductory statistics courses were conflictive rather than consistent. For example, students misunderstood confidence intervals which was a part of statistical literacy after introductory statistics course since these courses mainly focused on computation and memorization skills (McAlevey & Sullivan, 2010). In addition, using media reports to promote statistical literacy for non-quantitative and quantitative majors made little difference in their understandings regarding statistical concepts in those reports based on the interviews conducted (Budgett & Pfannkuch, 2010) as opposed to studies conducted with middle school students (e.g. Merriman, 2006). Similarly, effect of stand-alone online introductory statistics course compared to traditional teaching methods on statistical literacy in the context of tertiary students was not found to be significant (Meyer & Thille, 2010). On the other hand, introductory statistics course, before research methods course, in social science undergraduate curriculum (Wade & Goodfellow, 2009), statistics course utilizing daily life examples (Martinez-Dowson, 2010) and exposure to instructional program (Wilson, 1994) made a significant difference in statistical literacy post test scores of tertiary students.

The review of literature indicated that the studies investigating statistical literacy in middle school students have examined the age or class level differences through different aspects of statistical literacy where a developmental sequence was revealed (e.g. Aoyama & Stephens, 2003; Watson & Kelly, 2008) or positive effects of utilization of media reports on understanding statistics were reported (Doyle, 2008; Merriman, 2006). Besides, although inconsistent, there were a considerable number of studies regarding statistical literacy in the tertiary level. However, the research on statistical literacy considering Turkish context is rather scarce. Therefore, this research aims to investigate statistical literacy of 8^{th} grade Turkish students.

2.2. Attitudes toward Statistics

2.2.1. Importance of Attitudes toward Statistics and Definitions

Attitude toward statistics is an important part of statistical literacy models and attitudes are regarded as a factor effecting statistical literacy (Gal, 2004; Watson, 2006). Especially for critical evaluation of statistical claims, attitudes are taken into consideration. For example, Gal (2004) states that certain beliefs and attitudes are required in order to critically evaluate statistical messages as a part of the statistical literacy. Attitudes also have an important role in the teaching and learning process during class time, for the statistical behavior out of the class, and enrollment in further statistics related courses (Gal, Ginsburg, & Schau, 1997). Students' attitudes towards statistics can help or hinder statistical thinking and they do have influence of the application of knowledge and skills in variety of contexts (Gal, Ginsburg, & Schau, 1997). In addition to these, Watson (2006) indicates that students with low motivation or negative attitudes about statistics perform less in statistical literacy. Therefore, this study also aims at investigating attitudes toward statistics and its relationship to statistical literacy.

Attitudes toward statistics are considered very important to investigate; however, there is no exact definition of this construct. Indeed, researchers often define attitudes in relation to what their assessment instruments measure (Gal, Ginsburg, & Schau, 1997). For example, according to Schau (2003), who conceptualized attitudes toward statistics based on expectancy value model for achievement (Eccles, Adler, Futterman, Goff, Kaczala, Meece, &Midgley, 1983 as cited in Schau, 2003), students' attitudes toward capability of doing statistics, students' opinions and thoughts about the difficulty of statistics and about the value of doing statistics successfully are components of attitudes toward statistics. Therefore, Student Attitudes toward Statistics (SATS) questionnaire developed by Schau and her colleagues (1995) consists of affect, difficulty, value, and cognitive competence subscales.

Gal, Ginsburg, and Schau (1997) employed McLeod's (1992) definition for attitudes toward mathematics and considered attitudes towards statistics as the emotions and feelings including positive and negative responses experienced by individuals during learning statistics. For the purposes of the current study, this definition is employed which describes general attitudes toward statistics.

2.2.2. Research Related to Attitudes toward Statistics

In the literature, there were numerous studies investigating attitudes towards statistics, its relationship between statistical outcomes such as statistical achievement, literacy or reasoning and change in attitudes after a teaching method with pre-college students (Calderia, 2010; Carmichael, 2010; Leong, 2007; Yılmaz, 2006; Yingkang & Yoong, 2007) and college students (Aksu & Bikos, 2002;

Alajaaski, 2006; Allredge, Johnson & Sanchez, 2006; Biajone, 2006; Bingham, 2010; Carlson & Winquist, 2011; Carnell, 2008; Chiesi & Primi, 2010; Dempster, 2009; Diri, 2007; Doğan, 2009; Emmioğlu, 2011; Evans, 2007; Mocko & Jacobe, 2010; Mvududu, 2003; Nasser, 2004; Posner, 2011; Ragasa, 2008; Schau, 2003; Tempelaar, Loeff & Gijsealers, 2007;Vanhoof, 2006). As seen, majority of literature regarding attitudes towards statistics consisted of studies conducted with college students. In this part of the literature review, these studies were examined and the relations between them were provided.

One of the descriptive studies conducted with pre-college students who were 1128 secondary students from 7th, 8th, 9th, 10th and 12th grade in Lisbon where they responded the 11, 10 and 14 item questionnaires composed of close ended statements related to views about statistics (Calderia & Mouriño, 2010). For each grade, more than 70% of students found statistics interesting. Similarly 10th and 12th grade students found probability as an interesting subject. However, they were not sure whether statistics was a difficult topic or not within mathematics curriculum. These results might an indication that students are more interested in applied topics rather than abstract concepts (Calderia & Mouriño, 2010).

In another study conducted with middle school student, Yingkang and Yoong (2007) investigated attitudes towards statistical graphs with respect to gender and grade level. The questionnaire they used consisted of five subscales which were enjoyment, confidence, usefulness, learning preferences and critical views. The descriptive results revealed that their attitudes with respect to enjoyment, confidence, and usefulness were neutral to positive. In addition, students preferred traditional methods of statistics teaching which included clear explanation and practice and they were unlikely to have a questioning or critical attitude towards statistical graphs. As students' grade level increased, their attitudes became more positive whereas there were no gender differences.

These two descriptive studies indicated that school students tended to have attitudes ranged between neutral to positive. The studies investigating the relationship between statistical literacy and attitudes towards statistics in pre-college context were very limited. A research which investigated the possible relationship between statistical literacy and affective domain could be found in Carmichael's (2010) study where he conducted a study with 204 students in order to examine the possible relationship between interest in statistical literacy and self-efficacy in statistical literacy and students' statistical literacy. In order to analyze data, structural equation modeling techniques were utilized and the results revealed that interest had weak and insignificant relationship (r=.14) with statistical literacy while self-efficacy had an effect on both (r=.24) interest and (r=.16) statistical literacy achievement scores which indicated that students' competency values strongly predicted both of these constructs.

Statistics and probability had become a strand of school mathematics recently (MoNE, 2005; NCTM, 2000). Therefore, intervention studies which utilized different teaching methods conducted to examine the change in attitudes towards statistics of school students. For instance, Leong (2007) investigated the attitudes and beliefs of high school students in a service course, where students were required to provide real life or person based scenarios related to statistics, through obtaining data from journals, narratives, and an open-ended survey. The qualitative results indicated that, in terms of attitudes, students liked statistics which was real life related and that did not involve difficult and complex mathematics. Moreover, Cobb and Hodge (2002) conducted a study through utilizing computer tools as they formed a constructivist climate in the classroom where they focused on students' construction of identities including their interest and motivation for studying statistics. The analysis of teaching experiment revealed that social climate of classroom where every individual became as doers of statistics rather than passive learners promoted development of positive attitudes among 12-year old students (Cobb & Hodge, 2002). Another research, which was conducted in Turkish context, investigated effects of calculator supported and real data used statistics instruction on statistics achievement and attitudes toward statistics of 7^{th} grade students within school mathematics (Yılmaz, 2006). According to the results, students were not sure about whether they were interested in the graphs used in media after intervention. Additionally, students tended to be neutral about whether they enjoyed preparing tables or drawing graphs for statistical data. These studies indicated that literature was not only limited in number but also unclear in terms students' attitudes towards statistics after different teaching methods since either their attitudes stayed in neutral position or changed in a positive way.

Research in tertiary education provided wealthier information for attitudes towards statistics compared to pre-college context. These studies investigated attitudes of students in tertiary level either through the examining the change after a teaching method (Alajaaski, 2006; Allredge, Johnson & Sanchez, 2006; Biajone, 2006; Bingham, 2010; Carlson & Winquist, 2011; Carnell, 2008; Doğan, 2009; Evans, 2007; Mocko & Jacobe, 2010; Posner, 2011; Ragasa, 2008) or examining the relationship between attitudes and statistical outcomes (Chiesi & Primi, 2010; Dempster, 2009; Nasser, 2004; Tempelaar, Loeff & Gijsealers, 2007; Vanhoof, 2006) or consisted of description of attitudes (Aksu & Bikos, 2002; Diri, 2007; Mvududu, 2003).

Mvududu (2003) described attitudes towards statistics through constructivist learning environment which included personal relevance, student negation, critical voice, and uncertainty dimensions which revealed a positive relationship between these variables and attitudes. Another study conducted to investigate the predictors of attitudes toward statistics with 88 and 140 graduate students from educational departments (Aksu & Bikos, 2002). The significant predictor of statistics attitudes was departmental affiliation whereas previous experience in statistics and gender difference were not significant. Furthermore, Diri (2007) conducted a study which described the attitudes toward statistics lesson of 135 college students in a vocational school with a scale consisting of six different components. For the components of anxiety and importance, students reported positive attitudes whereas for the components of interest and confidence they reported negative attitudes. Also, for the components profession and enjoyment, students were not sure whether they agreed with positive or negative attitude statements. These studies indicated that predictors of attitudes towards statistics might be constructivist learning environment

(Mvududu, 2003) and departmental affiliation (Aksu & Bikos, 2002) and other affective constructs such as confidence or anxiety (Diri, 2007).

In correlational studies, one of the most frequent questionnaire tools used was SATS (Schau, 2003) which was constructed based on expectancy model of achievement. Accordingly, for the purpose of summarizing these studies, Emmioğlu (2011) conducted a meta-analysis with the studies which examined the relationship between statistics outcomes and attitudes toward statistics through using SATS. The effect sizes of each dimension were .34, .36, .23, .20 and statistically significant for affect, cognitive competence, value, and difficulty dimensions in order. In addition, there are many studies conducted with university students which revealed a positive relationship between attitudes towards statistics and statistics achievement (Chiesi & Primi, 2010; Dempster, 2009; Diri, 2007; Nasser, 2004; Vanhoof, 2006). On the other hand, there was an example of a research which documented weak or insignificant relationships between dimensions of attitudes towards statistics (affect, value, cognitive competence, and difficulty) and statistical reasoning of undergraduate students before an introductory statistics course, which might be related that study focused on statistical reasoning which was a naïve knowledge when formal statistical knowledge had been forgotten by students (Tempelaar, Loeff, & Gijsealers, 2007).

Related literature about attitudes towards statistics at the university level indicated that there was a considerable number of research which examined the change in these attitudes after a teaching method. However, these studies revealed inconsistent results. To illustrate, research investigated the effect of project work on attitudes towards statistics revealed significant increase in attitudes (Biajone, 2006; Bingham, 2010) whereas Carnell's (2008) study reported that there were no significant change in attitudes after utilizing project-work which might be related to course conditions, though unreported (Carmichael et.al., 2009).

Technology or computer assisted learning had an important place in statistics education (Ben-Zvi & Garfield, 2004; Shaughnessy, 2007). Related to this, there are several studies in the literature which examined the effect of technological tools on attitudes towards statistics in the context of university students. Some studies indicated that technology promoted to develop positive attitudes towards statistics such as real life related video viewing (Allredge, Johnson & Sanchez, 2006), audio response system (Mocko & Jacobe, 2010) and utilizing statistics software (Doğan, 2009). On the other hand, there are studies which documented that students' attitudes remain steady after exposure to technology such as computer assisted statistics instruction (Ragasa, 2008) or web-based technology (Alajaaski, 2006) which might be related to how these experiments were conducted and subject characteristics.

There were other teaching methods such as proficiency based assessment (Posner, 2011) and workbook approach (Carlson & Winquist, 2011) which improved attitudes towards statistics. On the other hand, reform oriented teaching did not likely to have an influence on attitudes although statistical self-efficacy and statistical reasoning abilities were positively affected most probably due to teacher encouragement (Olani et.al, 2011). College students' relatively positive attitudes also might be related to the fact that these attitudes were already positive since those students had already been taught statistics under reformed curriculum of NCTM (1989) before attending college.

The review of literature related to attitudes towards statistics pointed out that research with middle school context was scarce compared to the research conducted with university students. Although middle school students' attitudes towards statistics varied between neutral and positive (e.g. Calderia & Mouriño, 2010) in descriptive studies, its relationship with statistical literacy was not clear since those studies were limited in number and focused only specific dimension of attitudes such as interest (e.g. Carmichael, 2010). The results of intervention studies which investigated the effect of a teaching method on attitudes towards statistics were ambiguous since either attitudes remained the same around neutral (e.g. Yılmaz, 2006) or changed in a positive way (e.g. Cobb & Hodge, 2002). The change in the positive direction after providing constructivist climate in the classroom regarding middle school students (Cobb & Hodge, 2002) was validated in tertiary education where there existed positive relationship between constructivist learning environment and attitudes towards statistics (Mvududu, 2003). However, in college or university education context, the research was inconsistent where the results revealed that either positive relationship between attitudes and statistical outcomes (e.g. Nasser, 2004; Vanhoof, 2006) or weak and insignificant relationships (e.g. Tempelaar, Loeff & Gijsealers, 2007). This inconsistency could be observed also in intervention studies where the results indicated that either these interventions promoted the development of positive attitudes (e.g. Biajone, 2006, Doğan, 2009) or attitudes towards statistics remained the same (e.g. Carnell, 2008; Alajaaski, 2006). As seen, a considerable body of research existed in the literature. However, the studies investigating the relationship between attitudes towards statistics and statistical literacy with middle school students were very limited. Therefore, this research could provide an opportunity to examine this relationship in Turkish context.

2.3. Summary of Literature Review

The review of literature indicated that statistical literacy was a new area of research in mathematics education and there was a concern for researching statistical literacy within the scope of school mathematics (Shaughnessy, 2007; Watson & Callingham, 2003) since the ultimate goal of statistics instruction was labeled as statistical literacy (Ben-Zvi & Garfield, 2004). Since statistical literacy had entered into the literature relatively new, there was no common definition for this construct. However, in this study statistical literacy defined as understanding, interpreting, and evaluating statistical claims together with attitudes where a combined framework was constructed. In this framework, cognitive aspect of statistical literacy was conceptualized through Watson's (1997) three tiered framework while dispositional aspect was attitudes towards statistics taking Gal's (2004) model into account. The tiers included understanding, interpreting, and critical evaluation of statistical claims. Understanding terminology which was referred by the first tier could be counted as the literacy skill as stated by Gal (2004) and Watson (2006). In addition, three tiers were associated with diverse contexts since all of the frameworks took context into consideration. In the light of this combined framework, statistical literacy of eighth grade students were examined through these three tiers and attitudes towards statistics.

The research considering middle school students investigated different aspects of statistical literacy in terms of grade level and results of these researches revealed that statistical literacy was developed as grade level increased (e.g. Aoyama & Stephens, 2003; Watson & Kelly, 2008). In addition, there were a few remarkable intervention studies where the effect of project work (Merriman, 2006) and critical thinking skills (Doyle, 2008) improved conceptions and understanding regarding statistical literacy. There existed a considerable research in the context of university students; however, results of those studies were not consistent as stated in previous sections.

Attitudes towards statistics were considered very important to investigate since they had role in the teaching and learning process during class time, in the statistical behavior out of the class, and enrollment in further statistics related courses (Gal, Ginsburg & Schau, 1997). Despite of its importance, there were several definitions of attitudes towards statistics in the literature. For the purpose of this study, attitudes towards statistics could be described as the emotions and feelings experienced by individuals during learning statistics which addressed general attitudes toward statistics.

The review of literature related to attitudes towards statistics revealed that there were a considerable number of studies in tertiary education context while research in pre-college context was limited in number. Though explanatory studies conducted with middle school students indicated that students' attitudes differentiate between neutral and positive (e.g. Calderia & Mouriño, 2010), relationship between statistical literacy and effects of various interventions still remained ambiguous either suggesting weak relationships (e.g. Carmichael, 2010) or stayed in the same position after a teaching method (e.g. Yılmaz, 2006). It was possible to see this inconsistency in the tertiary education literature considering attitudes towards statistics as mentioned.

CHAPTER 3

METHODOLOGY

The purpose of this study is to investigate the statistical literacy of 8^{th} grade students and to investigate the relationship between statistical literacy and attitudes toward statistics. The following research questions have guided the present study:

1. What is the statistical literacy of 8^{th} grade students in terms of content domains (sample, average, graph, chance, inference and variation)?

2. What is the statistical literacy of 8^{th} grade students in terms three tiers?

3. Is there a significant difference between the mean scores of first, second and third tier of statistical literacy?

4. What are attitudes of 8^{th} grade students' towards statistics?

5. Is there a significant relationship between statistical literacy of 8^{th} grade students and their attitudes towards statistics?

The focus of this chapter is to describe methodology used to conduct this study. This chapter provides information about the research design, sample and its major characteristics, reliability and validity of instruments, data collection procedures, and statistical methods used for data analysis. Lastly, internal and external validity of the study is presented.

3.1. Design of the Study

The main purposes of this research were to investigate the statistical literacy of 8^{th} grade students and their attitudes towards statistics. The other purpose was to investigate the relationship between statistical literacy and attitudes toward statistics. In order to examine the related research questions, quantitative methods were utilized. Due to the fact that the first research question aimed at describing some aspects and characteristics such as knowledge, according to Fraenkel and Wallen

(2006), survey research design was employed. Particularly, this study was designed as a cross-sectional survey with the aim of collecting data at one point of time from a sample selected to describe a population from the inferences what is found (Fraenkel & Wallen, 2006). However, the second research question of this study intended to describe a possible relationship between variables which were attitudes towards statistics and statistical literacy, correlational research design was preferred (Fraenkel & Wallen, 2006). Data were analyzed through obtaining frequencies, percentages and calculating Pearson Product-Moment Correlation Coefficient between these two variables. In addition, in order to examine whether there were significant differences between tiers one way within subjects of ANOVA was conducted.

3.2. Population and Sample

The target population of the current study is defined as all 8th grade students attending public schools in Ankara. Yet, the accessible population of this study is defined as all 8th grade students enrolled in public schools in Yenimahalle district since it was not feasible to reach the target population. According to the information gathered from Yenimahalle Directorate of National Education, there are approximately 8700 students in 8th grade in 2011-2012 school year. The number of students participated in this study was 1074 which constituted at least 10% of accessible population.

Cluster random sampling method was used in order to select the sample of the study. In cluster random sampling method, schools are randomly selected rather than students (Fraenkel & Wallen, 2006). The determination of these elementary schools participated for the study as follows: First of all, a random list of public elementary schools in Yenimahalle district was generated. The contact information of these schools was gathered from the website of Yenimahalle Directorate of National Education. According to this information, these schools were contacted, the aim of research and the other details such as time duration for the tests were explained to the principles, and they were asked whether they could participate for this study. In addition, information regarding the number of students in these schools was obtained through these contacts. Though most of the schools accepted to participate for this study, some of them did not want to allocate their time. Therefore, this study was conducted with these schools in the random list who agreed to participate. The findings of the present study could be generalized for this accessible population. The rationale for selecting 8^{th} grade students was to make an analysis of statistical literacy of elementary students near the end of their elementary education. In addition to this, since elementary mathematics curriculum is spiral in nature, most of the statistical topics such as standard deviation are covered in the 8^{th} grade curriculum (MoNE, 2005). Therefore, 8^{th} grade students were the participants of this study. At the end of the study, the number of sampled schools was 9 and the number of classes where this study was conducted was 48. The more detailed characteristics of sample are displayed in Table 3.1.

Table 3.1 Major Characteristics of Sample

	Frequency	Percentage
Gender		
Male	499	46.5
Female	574	53.4
Age		
13	3	0.3
14	997	92.8
15	74	6.9
Fall Semester Mathematics Grade		
1	151	14.1
2	140	13.0
3	185	17.2
4	248	23.1
5	324	30.2

As shown in Table 3.1, the sample composed of 46.5% of males and 53.4% of females. The mean age of students was 14.06 while the range lies between 13 and 15.

The mean of participants' final mathematics grade was 3.43 at the end of the 2011-2012 fall semester. In addition, 53% of the students' grade was 4 or above.

Several indicators of socioeconomic level of students were also gathered during data collection. Two of these indicators were parents' employment status and educational level. The percentages of these indicators are presented in Table 3.2.

Table 3.2 Employment Status and Educational Level of Parents

Indicator	Percentages	
Employment	Father	Mother
Employed	87.3	31
Unemployed	1.6	61.8
Irregular	1.8	1.5
Retired	8.1	5.5
Education Level		
Illiterate	0.1	0.4
Primary School	9.4	21.9
Middle School	14.8	16.6
High School	35.8	38.1
University	34.7	20.7
Graduate School	3.3	2.2

Table 3.2 indicated that majority of students' father is employed whereas most of the mothers are unemployed. In addition, most of the parents are graduated either from high school or university, yet fathers' educational level (36% high school and 35% university) is higher than mothers (38% high school and 20% university).

Number of books at home, presence of study room, frequency of buying newspaper and number of siblings might be regarded as other indicators of socioeconomic status. These characteristics of sample are presented in Table 3.3.

Table 3.3 Other Indicators of Socioeconomic Status of the Sample

Indicator	Percentages
Number of books at home	
0-10 books	4.3
11-25 books	18.8
26- 100 books	43.0
101-200 books	20.4
More than 200 books	13.4
Newspaper	
Never	5.4
Sometimes	61.2
Always	33.4
Study Room	
Exist	86.4
Non-exist	13.6
Number of Siblings	
0	11.3
1	56.1
2	24.7
3	6.2
4 and more	1.8

As can be seen in Table 3.3, majority of students indicated that there were their own study rooms at home and most of them had one sibling (56%). It was also found that newspaper was bought sometimes to their home. In addition, many students had books in their homes.

Briefly, it can be claimed that although there were variations between students, majority of the sample of this study were coming from middle socioeconomic families living in urban district.

3.3. Data Collection Instruments

The data for this study were collected through statistical literacy test and attitude toward statistics questionnaire. These instruments and validity and reliability of them are explained below in detail.

3.3.1. Statistical Literacy Test (SLT)

3.3.1.1. Development and Validity of SLT

Statistical Literacy Test (SLT) was developed by the researcher of this study (See Appendix B). The items used in the study were devised to measure various aspects of the statistics and probability curriculum which had been implemented since 2004-2005 academic year in Turkey. In this test, content domains were sampling, average, probability, variation, tables and graphs, inference. The items used in the test also included wide variety of social context such as media, health and school related issues. By means of inclusion of such content domains and context related to them, it was intended to measure the statistical literacy of 8th grade students. While constructing the items, three tiered statistical literacy framework (Watson, 1997) was taken into account for determining the objectives. These tiers refer understanding, interpreting and evaluating statistics respectively.

A question pool was constructed including 25 questions and sub-questions contextualized according to the first, second and third tier of statistical literacy framework. The test had 17 main question and its sub-questions. Questions 1, 2, 3, 4, 5 and their sub-questions were related to understanding terminology, 6, 7, 8, 9, 10, 11 were related to interpreting statistical messages and questions 12, 13, 14, 15, 16, 17 were related to evaluating or questioning the appropriateness of those statistical messages. Table of specifications of statistical literacy test is presented below:

Table 3.4 Table of Specification for Statistical Literacy Test

	Sample	Average	Graphs and Tables	Probability	Inference	Variation
Tier 1 (understand)	1a, 1b	2a, 2b	3	4a, 4b	-	5a, 5b,5c
Tier 2 (interpret)	6	7	8	9	10a, 10b	11
Tier 3 (evaluate)	12a, 12b	13a, 13b, 13c	14a, 14b	15a, 15b	16a, 16b	17a, 17b

As can be seen from the specification table, there were questions in each of the content domains placed in each of tiers. The sub-questions labeled as “b” or “c” required examinees to give explanation or exemplification of the answer in the first part of the question.

In the selection of questions the researcher consulted the textbooks, curriculum documents, a variety of statistics books, and the existing literature. This test consisted of multiple-choice, yes/no, and open-ended questions. Multiple-choice questions were preferred for most of the items because they allowed students to “show recognition rather than creation of an appropriate answer” (Watson, 2006, p. 251) which was demanded by statistical literacy. In addition, the evaluation of statistical claims was formatted in yes/no questions. However, since it was not possible to demonstrate questioning ability and statistical thinking in multiple-choice format (Watson, 1997) or yes/no questions; open-ended items were also utilized in questions in the third tier which were sub-questions asking for the explanation of students’ responses.

Some of the questions were taken or adapted from existing literature. For example, 7^{th} question was an adaptation of a checklist item in statistical literacy assessment scale developed by Watson and Callingham (2003). Similarly 9^{th} question was taken from statistical reasoning assessment scale (Garfield, 2003). Questions fifteen and sixteen were adapted from items in statistical literacy assessment scale (Watson & Callingham, 2003). In addition, the graphical representation in question eight was taken from graph interpretation test (Aoyama & Stephens, 2003) and adapted into multiple choice question format.

Although the test was prepared according to the three tiered statistical literacy framework, the objectives of national mathematics education curriculum were taken into account. Therefore, with respect to those objectives the following table of specification was prepared.

Table 3.5 Table of Specification for SLT with respect to MoNE (2005)

Objectives of MoNE	Contents					
	Sample	Average	Graphs and Tables	Probability	Inference	Variation
To show data in appropriate statistical representations			Q3			
To explain random in context				Q4a		
To identify sample for a situation	Q6					
To interpret arithmetic mean of data		Q7				
To conjecture through interpreting graphs and tables			Q8			
To explain probability of an event				Q9		
To predict based on data.					Q10	
To conjecture through interpreting measure of central tendency and variation.						Q11
To explain situations of misinterpretation of bar graphs			Q14			

This table indicated that the objectives of national curriculum did not meet the purposes of statistical literacy defined by Watson (1997). That is, there was no correspondence between the some questions in the first tier and in the third tier and objectives of national elementary mathematics curriculum. To illustrate, although there was an objective regarding misinterpretation of bar graphs, there were not any objectives for other content domains of statistics such as average or sampling. In other words, objectives did not focus on the subject of sampling biases, misuse of measures of central tendency. Therefore, while obtaining the expert opinions, the table of specification which was constructed according to the three-tiered framework of statistical literacy was used. However, each of the questions in the test, in terms of their content, is instructed within the scope of regular mathematics instruction. The sample items in the tiers are presented in Table 3.6.

Table 3.6 Sample Items of First, Second, and Third Tier

Tiers	Sample Items
Tier 1	"Last year, an average of 20 people had died due to traffic accident." What do you understand the word **"average"** in this sentence?
Tier 2	A researcher who lives in a town consisting of 50 families has found the mean of children per family as 2.2. Which one of the followings is absolutely true? a. Half of the families in this town has two children b. The number of families with 3 children more than families with 2 children. c. There are 110 children in this town. d. The mean of children per adult is 2.2.

Table 3.6 (continued)

Tier 3 The number of problems solved in a math class is counted and represented in the following table

Student	Number of problem
A	2
B	6
C	2
D	22
E	3
F	2
G	1
H	2

In order to summarize these data the mean is calculated and found 5.

a. Do you agree with this claim?

b. Explain your answers with reasons.

A holistic rubric was prepared in order to classify students' responses in open-ended items and eliminate subjectivity (See Appendix E). Literature was reviewed before developing this rubric. Accordingly, students' responses were coded as non-statistical/incorrect, pre-statistical, and statistical. However, some items were not suitable for partial credit (item 10b and item 14b). These were coded as dichotomously where the score of 1 for true explanations and the score 0 for false and empty responses were given. In addition, explanations of terminology with arithmetic procedures were coded separately in item 2a and item 5b.

Validity is the "appropriateness, correctness, meaningfulness and usefulness of the inferences" (Fraenkel & Wallen, 2006, p. 151) claimed based on the data collected for the study. Validation of an instrument is a process of collecting evidence such as appropriateness of content, comprehensiveness, structure of the construct. Validity evidences for Statistical Literacy Test (SLT) are described below in detail.

In order to check content validity of SLT; firstly, table of specification was constructed based on Watson's (1997) three tiered framework as mentioned in Table 3.4. Afterwards, the form of constructed items was given to three mathematics education researchers and two elementary mathematics teachers. The SLT was submitted to experts with a summary description of statistical literacy and its three tiers. They were also provided a checklist including following categories:

- whether the purpose of each question is consistent with relevant statistical literacy tier,
- whether each question is appropriate for 8thgrade level in terms of content and format,
- whether question's wording is understandable, and
- whether the question's content and context is appropriate to the statistics curriculum of MoNE.

After the experts filled the checklists for items, the comments and evaluations of them were examined by the researcher. Through the suggestion of one of the teachers, instead of asking the meaning of sample directly, it was changed to ask with an example. The other question was identified as memorization type of item; hence, it was removed from the test. Moreover, some of the questions were revised or reworded since there were comprehension problems as indicated by the experts. For question fifteen, a table which explains the root of the item, was added to make the question more understandable. In addition to these, with advises of the experts, the questions in the negative form was highlighted and underlined in order to gain students' attention. Necessary revisions such as wording of statistical concepts were also done to make items appropriate for national elementary mathematics curriculum according to the refinements of experts. Hence, quality of items and content validity of SLT were assured.

3.3.1.3. Pilot Study of SLT

The pilot study was conducted after ensuring the content validity of items. The answers of students were entered into computer by using PASW (Predictive

Analytics SoftWare) 18 statistics program.

According to the responses of students, some adjustments were made in SLT. The analysis of the responses indicated that questions 13a and 13b did have very low proportion of correct answers and they were regarded as difficult items. Since there should be a question related to average concepts in third tier for the purpose of statistical literacy, these questions were revised rather than removed. In more detail, the previous form of this item included a context related to teachers. Since students might respond accordingly, the statement was revised into passive form and teacher context was removed. Before pilot study, the initial question of tier three items were constructed as multiple choice formats where the alternatives were "Yes" and "No". Since it was observed that the multiple choice format lead students to choose alternatives by chance factor rather than providing appropriate answer, the alternatives were removed and they remained as an open ended format where the second part still asking for explanation of initial answers.

3.3.1.4. Reliability of SLT

Reliability, basically, is the stability of the scores obtained, that indicates how consistent they are for each individual from one administration to another (Fraenkel & Wallen, 2006). Methods for reliability are presented in the following sections.

Internal consistency methods were utilized to examine reliability of the instrument. For the statistical analyses of the internal consistency, Cronbach's alpha coefficient was obtained through PASW 18 program. The reliability estimate for scores on Statistical Literacy Test was found as .72 according the data gathered from pilot study. Following to the adjustments made on SLT, the reliability estimate increased to .75 which was obtained in the actual study. Since reliability coefficient of a scale should be at least .70 (Fraenkel & Wallen, 2006) to have a reliable instrument, it can be said that SLT is a reliable instrument.

In addition to internal consistency between items, there should be scoring agreement which refers inter-rater reliability since there were open ended items in the test and they was scored according to a rubric (Crocker & Algina, 1986). The open-ended items codes according to the rubric were item 1, 2, 5b, 12b, 13b, 15b, 16b and 17b. In order to check inter-rater reliability 10% of cases were randomly selected to be scored by a different rater who was a mathematics teacher. Interclass Correlation Coefficient (ICC) for two raters was used to measure inter-rater reliability where this coefficient should be .70 and higher (Shrout & Fleis, 1979). The value of .87 indicated a quite high reliability between scorers. In addition, the Pearson Product Moment Coefficient was calculated between two scorers. This correlation coefficient was found as .89 which indicated the high consistency between scorers.

The all findings in this section indicated that SLT is a valid and reliable instrument which measures eighth grade students' statistical literacy within Turkish context.

3.3.2. Attitudes towards Statistics Questionnaire (ATSQ)

3.3.2.1. Development and Validity of ATSQ

Attitudes toward Statistics Questionnaire (ATSQ) was modified by the researcher of this study from Probability Attitude Scale developed by Bulut (1994). This instrument was utilized in this study since statistics is one of the domains of school mathematics and probability was the closest one to this domain. The other rationale for modifying this instrument was that it was constructed for 8^{th} grade students. To be precise, the items in the questionnaire were appropriate for 8^{th} grade students, who were at the age of fourteen, in terms of language, feelings and emotions. Additionally, since this scale was developed, piloted, and implemented in Turkish context, the contextual factors were less of a concern.

Probability Attitude Scale (PAS) was originally composed of 28 items on a six point Likert scale to be responded, Strongly Disagree, Tend to Disagree, Tend to Agree, Agree, and Strongly Agree. Obtaining high scores means that students have more positive attitudes. There were 15 positive and 13 negative items constituting the PAS. The statements were selected from an item bank consisting of 80 items. According to the factor solution of the scale, although there were 5 sub-dimensions, they were not named as sub-dimension since all positive and all negative items loaded in the same dimensions. Therefore, the scale was unidimensional, which was describing general attitudes toward probability. The reliability coefficient Cronbach's alpha was calculated 0.94 for PAS by Bulut (1994).

Modification of PAS was made through replacing the word probability by the word statistics. Whilst 28 of all the items were remained, five-point scale was utilized since there could be also neutral responses.

Since attitudes towards statistics questionnaire was a modification of another attitude scale (Bulut, 1994) whose content validity in terms of representativeness of students' attitudes had already ensured, there was a need to find evidence of comprehensiveness of the questionnaire. For this purpose, the modified questionnaire was given to three mathematics education researchers and two elementary mathematics teachers. They were asked for whether the items are understandable or not and whether there were any grammatical mistakes or not. According to their comments, one necessary revision was done to make the wording of the item more appropriate.

Followed by the final form of attitudes toward statistics questionnaire, the pilot study was carried in those schools stated in the following sections in order to find evidence related to validity and reliability of ATSQ.

In order to reduce the number of observed variables exploratory factor analysis was conducted with the aim of grouping the variables in constructs. Factor analysis with the data of the attitudes toward statistics questionnaire was run to determine which sets of observed variables sharing common variance characteristics define constructs. Before running factor analysis, the assumptions which were sample size, factorability of correlation matrix, linearity and outliers among cases were checked.

Principal components analysis with varimax rotation method was carried out for ATSQ by PASW 18. The exploratory factor analysis, including 28 observed variables, was considered in order to know the underlying structure of this questionnaire.

Descriptive statistics for the items in the questionnaire indicated that there were maximum 9 missing values for each of the item which constituted less than 5%. Hence, replacing missing scores with mean method was utilized in order to deal with missing data (Tabachnick & Fidell, 2007). At the end, there were 272 cases for pilot study of ATSQ which indicated that the sample size assumption were assured since there were approximately at least 10 cases for each variable (Tabachnick & Fidell, 2007).

Pallant (2007) stated that the correlation matrix should show at least some correlations of r = .30 or greater in order to be considered suitable for factor analysis of variables. Since most of the r values are higher than .30 in the correlation matrix of the variables, this rule was guaranteed. In addition to this, the values of Kaiser-Meyer-Olkin and Bartlett's test of Sphericity are presented in Table 3.7.

Table 3.7 KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,947
Bartlett's Test of Sphericity	Approx. Chi-Square	3851,120
	df	378
	Sig.	,000

The KMO measure revealed a value greater than 0.60 which indicated that evaluating the distribution of values is adequate for conducting factor analysis. In addition to this, the Bartlett's Test of Sphericity was statistically significant (χ^2= 3851.120 and p=.000). Thus, the multivariate distribution was normal and the hypothesis that variables constitute an identity matrix was rejected, and acceptable for factor analysis.Furthermore, since factor analysis was based on correlation, it was assumed that the relationship between the variables was linear since examining all scatterplots were not feasible and sample size was adequate (Pallant, 2007). Last assumption, outliers among cases, was checked and it was found that there were no outliers in the data set. All of the findings indicated that it was safe to conduct a factor analysis without violating assumptions.

In order to determine the smallest number of factors that represents the interrelations between variables, Principal components analysis method was utilized. According to Pallant (2007), there are different methods to determine the number of factors. Kaiser's criterion is one of them. In this method, only factors whose eigenvalue is greater than 1 are retained for further investigation.

Table 3.8 Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	11.67	41.66	41.66
2	2.41	8.60	50.26
3	1.08	3.87	54.13
4	1.07	3.83	57.96

According to eigenvalues, there were 4 underlying factors among variables whose eigenvalues were greater than 1. These four components explained a total of 57.96 percent of the variance. However, Büyüköztürk (2011) stated that the number of factors can be determined through identifying two third of total variance explained initially. In this case two third of total variance was explained in the first component. Therefore, the number of factors was retained as one. In addition, 30 percent or higher are regarded as adequate variance explained when the scales have one factor (Büyüköztürk, 2011). Since 42 percent of variance was accounted for by the first component, this scale has single factor.

Another method for determining number of factors is analyzing scree plot which becomes important when there are many components (Pallant, 2007). Scree plot of factor analysis is given below:

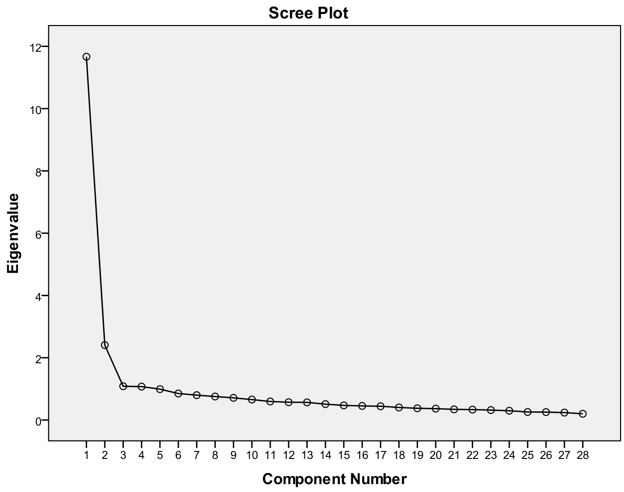


Figure 2 Scree Plot for ATSQ

The change in the shape of the plot suggests information regarding the number of the factors. The examination of the plot showed that there was a sharp decrease between first and second factors and the other changes were almost flatted. Therefore, number of factors could be retained as one (Büyüköztürk, 2011).

In order to analyze the factor structure of the scale more precisely, the component matrix table was examined with unrotated loadings to determine the number of factors. Table 3.8 illustrates the component matrix loads.

Table 3.9 Component Matrix

	Component
att10	,774
att13	,773
att28	,736
att11	,708
att16	,706
att1	,701
att20	,699
att23	,699
att5	,698
att14	,688

Table 3.9 (continued)

att6	,683
att18	,662
att25	,661
att2	,660
att24	,658
att12	,640
att19	,622
att8	,622
att22	,614
att21	,608
att3	,606
att7	,605
att4	,598
att9	,594
att26	,585
att17	,582
att27	,250
att15	,430

Pallant (2007) stated that an item load quite strongly, if the loading exceeds .40. The examination of the loading of all items revealed that these loadings were higher than .40 except item 27. Since this item loaded .25 in the first component which is lower than criterion value, it was removed from the scale. Hence, the analysis of component matrix yielded that there was one factor solution of scale.

3.3.2.5. Reliability of ATSQ

To establish reliability of the instrument, internal consistency methods were used. For the statistical analyses of the internal consistency, Cronbach alpha coefficient was computed. Cronbach alpha coefficient of a scale should be at least 0.70 (Crocker &Algina, 1986). In the pilot study, Cronbach alpha coefficient was calculated as .94 which showed that there were satisfactory reliability and internal consistency between items. In addition to checking internal consistency values of ATSQ, item inter-correlation values were examined. These values ranged from .40 to .74 which indicated that there were quite strong relationships between items (Pallant, 2007).

The findings presented above indicated that ATSQ was a valid and reliable instrument which would assess middle school students' attitudes toward statistics. For the final form of this scale see Appendix C.

3.4. Pilot Study

Following to the assurance the content validity of both of the instruments, they were pre-piloted with one 8th grade student. The aim of this implementation was to control comprehensiveness of items and convenience of time duration given for the implementation. The student completed both of the tests within the given time. After that, the clarity and difficulty of items were asked. She stated there were no items that she could not understand. Also, she declared that graphs, tables, and figures in the items were comprehensible.

The pilot study was carried out in order to check construct validity and reliability of the instruments. Another aim of pilot study was to determine possible difficulties that might occur in the actual study. Students were also asked to indicate whether the items were comprehensible and there was anything that they did not understand in the items in the pilot study. Thus, pilot study was conducted with 8th grade students who would not be in the sample of the study.

The pilot study was conducted in conveniently selected public schools in İstanbul, Uşak, and other districts of Ankara. The participants of pilot study were students whose teachers addressed as having average achievement level in mathematics. The demographic information sheet, attitudes towards statistics questionnaire and statistical literacy test were administrated to 272 eighth grade students by their mathematics teachers within one class hour in their classrooms. Nearly half of the participants in pilot study was female (52%) and the other half was male (48%). In addition, the mean of their latest mathematics grade was 3.6.

After pilot study, necessary analyses of items and scales, stated in the validity and reliability sections in detail, was conducted with the data of pilot study in order to ensure construct validity and reliability of instruments.

In addition to this, no observed difficulties during administration were reported by the teachers and one class hour was adequate to complete the tests.

3.5. Data Collection Procedures

The data were collected during spring semester of 2011-2012 academic year. In the fall semester of the same academic year, Statistical Literacy Test and Attitudes towards Statistics Questionnaire were developed based on related literature. Then, expert opinions were taken, where they asked to indicate the consistency of the items with relevant statistical literacy tier, appropriateness of the questions to 8th grade, the clearness of the questions, and appropriateness of the content and context of the questions to the statistics curriculum of MoNE. After necessary revisions of items were done according to the expert comments on both of the instruments, pilot study was conducted in order to examine validity and reliability of these instruments. The necessary official permission was obtained from Middle East Technical University Human Subjects Ethics Committee before the data collection process. Subsequent to the approval of ethics committee, required permissions were taken from Ministry of National Education (see Appendix A) in order to conduct the main study.

The data of actual study was collected during the spring semester of 2011-2012 academic year. In the beginning of the spring semester, the schools were visited and explained the purpose of this study to both the school principals and mathematics teachers. The mathematics teachers of 8th grades were asked about their available class periods and appointments were made with them. Afterwards, instruments were administrated to 8th grade students during their regular class periods by the researcher of this study. First of all, researcher introduced herself and the reason to conduct this study to the students before answering the questions. The students were explained how to respond the items in the questionnaire and the test. In addition, it was declared that all their responses would be kept completely confidential and would only be used for the study. Each administration took approximately 40 minutes and the mathematics teachers were present at the class during the administration, yet they were not obtrusive. However, in one of the schools, the data was not collected by the researcher since mathematics teachers did not allow to collect data in their own mathematics classes. Instead, instruments were administrated in students' guidance and counseling sessions by their class teachers. These teachers were informed about the importance about research and data collection procedures and were kindly asked to implement the scales in the same way the researcher implemented. The whole data collection process lasted for four weeks.

3.6. Data Analysis

In the present study, quantitative research methodologies were used to analyze data through a number of descriptive and inferential statistics by using PASW 18 software. First of all, a codebook was prepared and a number of data screening procedures were carried out such as controlling data entry errors and controlling missing data prior to running primary data analysis.

In terms of defining the items that would be analyzed, the missing percentages of the questionnaire and statistical literacy test were essential criteria. In this sense, each of the items in the questionnaire and test was analyzed in detail. The general criterion for the missing data was 5%. The missing data did not exceed 5% of total cases. Since there were less than 5% missing data, pair wise deletion method was used in order to handle missing data.

First, demographic information and attitude towards statistics questionnaire was evaluated. The demographic information and each questionnaire item were analyzed by using frequencies, percentages, means, and standard deviations. The responses to questionnaire items were assigned a numeric value from 1 to 5 with 1 the least favorable response and 5 with the most favorable. For the items whose wording indicated a negative affective domain, the scale was reversed. Following to recoding of negative items, the numeric value 1 showed an indication of negative attitude (or strongly disagrees) and 5 indicated positive attitude.

Second, the responses of Statistical Literacy Test were classified according to the codes in the rubric. These codes were summarized as frequencies and percentages. Total literacy scores of individuals were calculated through statistical literacy survey which was analyzed according to the codes in the rubric. Then, for descriptive statistics; mean, standard deviation, percentages, and frequencies were calculated.

In addition to these, Pearson product- moment correlation analyses were run to examine the relationship between attitudes towards statistics and statistical literacy. Yet, prior to running this analysis, the assumptions were checked. Eta square was calculated to investigate the practical significance of the results.

3.7. Assumptions and Limitations

In this section the assumptions and limitations of the current study are discussed. First of all, since the variables in this study were based on self-report questionnaire and test, it was assumed that the participating 8th grade students gave a careful attention on each item both in statistical literacy test and attitudes towards statistics questionnaire, and their responses were honest and based on their personal attitudes and beliefs and feelings. It was also assumed that their statistical literacy and attitudes towards statistics could be measured through both of the instrument.

The study was conducted only in Yenimahalle district of Ankara and therefore, the findings of this study might be limited in its application to a more generalized population of 8th grade students. Yet, the results can be generalized to students whose context is the same as in this study. Another limitation was that the results of the present study were based on quantitative data collected from participants through a questionnaire. Therefore, the study is limited by the representation of the items on the test and questionnaire.

3.8. Internal Validity

Internal validity of a study means that any relationship observed between variables should be clear and not be caused or related by any unintended variable (Fraenkel & Wallen, 2006). Implementation, history, maturation, attitudes of subjects and regression threats are the internal validity threats which were not applicable since there was no intervention (Fraenkel & Wallen, 2006). Therefore, the following threats were considered during the study.

When characteristics of individuals are correlated, there can be a possibility that an unintended individual variable can explain the relationship. In this study, the sampled group was eighth grade students who were at the same age and from public schools. In addition, due to the fact that elementary mathematics education curriculum is national and standardized across the nation; students were assumed to have similar mathematical experiences in their own classes. Therefore, the subject characteristics threat was assumed to be reduced.

The mortality threat, which is loss of subject for internal validity, was not an issue in this correlational study, since lost subjects were excluded from the study as suggested (Fraenkel & Wallen, 2006).The study was conducted with 8^{th} grade students in their own mathematics lessons. Since absenteeism through the end of semester due to national examination is very common in Turkish schools, data collection period was arranged at the beginning of spring semester. Hence, the maximum participation was ensured and loss of subject was not a problem for this study.

Location is a threat when different individuals are tested in different locations (Fraenkel & Wallen, 2006). However, since instruments were administrated in their own classroom environment which was very similar in public schools, this threat was controlled.

There can be testing threat if there is there is an influence of first instrument on second instrument (Fraenkel & Wallen, 2006). However, in this study, attitudes towards statistics questionnaire did not likely have an effect on the performance of statistical literacy test since they measured different things. Testing threat did not constitute an internal validity problem.

Another threat for internal validity of the study was instrumentation consisting of instrument decay, data collector characteristics, and data collector bias (Fraenkel &

Wallen, 2006). Instrument decay was not a problem for this study, since the instruments were not implemented to students many times. However, there could be data collector bias since there were open-ended items in Statistical Literacy Test. Yet, in this study, two scorers scored the responses of students and the interclass correlation coefficient was calculated. As this coefficient lied between sufficient values, this threat was controlled. Data collector characteristics threat occurs when different persons administer both instruments (Fraenkel & Wallen, 2006). However, this issue did not serve as a threat since researcher collected most of the data. In addition, the teachers who collected data on their own were carefully informed about the data collection procedures. Hence, this threat was handled.

3.9. External Validity

External validity is the degree of generalization the results to a population (Fraenkel & Wallen, 2006). The target population of the current study was defined as all 8th grade students in Ankara, attending public schools. However, accessible population was the 8th grade students enrolled in public schools in Yenimahalle district. In this research sample was obtained through cluster random sampling method where the schools were selected randomly and students in those schools were surveyed. Since selected students were the representative of accessible population, the results were generalized to this accessible population. Fraenkel and Wallen (2006) describe ecological generalizability as "…the degree to which the results of a study can be extended to other settings and conditions" (p. 106). This research was conducted with 8th grade students in urban public schools where elementary mathematics education curriculum was the same and most of the schools were using the mathematics textbook provided by MoNE. Therefore, the results of this study were considered to be generalizable to similar settings.

CHAPTER 4

RESULTS

The purpose of this study is to investigate the statistical literacy of 8^{th} grade students in terms of content domains and three tiers and to investigate the relationship between students' statistical literacy and attitudes toward statistics.

Data of this study were collected through demographic information sheet (see Appendix D), Attitudes towards Statistics Questionnaire (ASTQ) and Statistical Literacy Test (SLT). The demographic information of sample was already presented in the methodology section. This chapter consists of the data analyses that were conducted to answer the research questions of this study. Firstly, statistical literacy of 8^{th} grade students in terms of content domains and three tiers respectively were described through descriptive statistics. Then, the difference between mean scores of tiers was explored through inferential statistics utilizing one-way within subjects analysis of variance. In addition, descriptive statistics, including minimum and maximum values, mean and standard deviation, related to attitudes towards statistics were presented. In order to examine the relationship between attitudes towards statistics and statistical literacy of 8^{th} grade students, Pearson-product moment correlation analysis was conducted. In addition, in order to examine whether there were significant differences between tiers one way within subjects of ANOVA was conducted.

4.1. 8^{th} Grade Students' Statistical Literacy in terms of Content Domains

The content domains of statistical literacy were sample, average, graphs, chance, inference, and variation. Students' statistical literacy scores in terms of content domains were divided by maximum score that they could get for each domain similar to the calculation of total statistical literacy scores. The descriptive statistics for each content domain was represented in Table 4.1.

Table 4.1 Descriptive Statistics for SLT Scores with respect to Content Domains

	N	Min.	Max.	Mean	SD	Skewness		Kurtosis	
						Stat	SE	Stat	SE
Sample	1063	,00	1,00	,53	,28	-,29	,08	-1,1	,15
Average	1063	,00	1,00	,30	,17	,64	,08	,97	,15
Graph	1060	,00	1,00	,52	,31	-,08	,08	-1,03	,15
Chance	1060	,00	1,00	,51	,24	,31	,08	-,37	,15
Inference	1058	,00	1,00	,39	,22	,59	,08	,65	,15
Variation	1056	,00	1,00	,35	,23	,48	,08	-,48	,15

As Table 4.1 indicated the mean scores for each content domain were varied. Although sample (*M*=.53, *SD*=.28), graph (*M*=.52, *SD*=.31) and chance (*M*=.51, *SD*=.24) content domains had closer mean scores to each other which was around moderate value, average (*M*=.30, *SD*=.17), inference (*M*=.39, *SD*=.22) and variation (*M*=.35, *SD*=.23) content domains had lower mean scores than other contents. These values indicated that students performed differently for each content domain. Skewness and kurtosis values were exceeded -1.00 and +1.00. However, according to Kunnan (1998), these values could be regarded as in the acceptable range which placed -2.00 and +2.00; hence these values did not violate normality assumptions (Kunnan, 1998).

The next sections of this chapter are devoted for item based descriptive results for each content domain. Since open ended items were classified as wrong, pre-statistical and statistical, the descriptive statistics including frequencies and percentages were represented through the classifications given to the students' responses based on the rubric.

4.1.1 Sample

The items related to sampling content were items 1, 6 and 12a/b. These items intended to measure tier one, two, and three in the respective order. Item 1 and item 12b were open-ended items whilst item 6 and 12a were multiple choice and true/false type respectively.

4.1.1.1. Tier 1: Understanding Sample Terminology

This category of statistical literacy was measured through item 1 which intended to measure students' understanding of sample in context. Item 1 was presented below.

> "A study is conducted where the sample is mathematics teachers worked in Ankara". What do you understand of the word "**sample**" in this sentence?

The mean score for this item out of 1 was .31 with standard deviation of .31. The responses of participants for this item were categorized in three themes which were incorrect or unrelated responses, pre-statistical responses including intuitive conceptions and correct and statistical responses. The frequencies and percentages regarding the categorizations related to first item were represented in Table 4.2.

Table 4.2 Descriptive Statistics for Item 1

Classification of Responses	Students' Responses	f	p
Blank/Wrong or unrelated responses		493	45.9%
	- Topic of a research	87	8.1%
	- Other blank/wrong responses	406	37.8%
Pre-Statistical		495	46.1%
	- Example	326	30.4%
	- Subjects of a research	169	15.7%
Statistical		80	7.4%
	- Part of a whole population	58	5.4%
	- Representatives of a population	22	2%

Table 4.2 showed that majority of students either left this item blank and gave

incorrect responses (45.9%) or provided pre-statistical answers (46.1%). Regarding the classifications in the rubric some of the students referred sample as “topic of a research” (8.1%) which was an example for wrong responses. Students who gave pre-statistical responses understood sample either as “example” (30.4%) or as “subjects of a research” (15.7%). 5.4 % percent of students referred sample as “part of a population” and 2% percent of students as “representativeness of a population” where these statistical responses constituted 7.4% of total percentage.

4.1.1.2. Tier 2: Interpreting Sample in Context

The next item related to sample in the second tier group was item 6. This item intended to measure students’ application of ideas related to sample in context. Item 2 was presented below.

> Ali is a member of library club in an elementary school and he wants to search the number of books at students’ home. Which one of the followings identified the representative sample of the school for this research?
>
> a) Randomly chosen 30 students from the library club
>
> b) Randomly chosen 30 students from the school
>
> c) Randomly chosen 30 students from Ali’s class
>
> d) Randomly chosen 30 female students from the school

The mean score for this item out of 1 was .74 with standard deviation of .44. The responses of students are given in Table 4.3.

Table 4.3 Descriptive Statistics for Item 6

	f	p
Incorrect responses	283	26.4%
Correct response	784	73.5%

Table 4.3 indicated that although some of the students (26.4%) chose the distracters,

majority of them (73.5%) provided correct response for this question. It could be inferred that for sampling content in the second tier, majority of students had performed more than moderate.

4.1.1.3. Tier 3: Evaluating Sample Claims

The last two items related to sample content were 12a and 12b. These items try to measure students' evaluation of sample in statistical claims which involve bias or inappropriate generalizing. The former one was a true/false type item addressing the evaluation of a statistical report which involved generalization of a sample to a population. The latter asked for an explanation of responses in the previous item. These items are presented below.

> The sample of a study investigating how many hours do children watch TV was 5th grade students in School A. As a result, students who were participated in this study watched TV for 3 hours in a day. The results of this study is announced as follows:
>
> "Every elementary school students in Turkey watch TV for 3 hours a day."
>
> a) Do you find this sentence as an acceptable statistical claim?
>
> b) Provide a statistical explanation for your answer.

The mean score for this item out of 1 was .69 with standard deviation of .46. The frequencies and percentages regarding the responses of students for the former item were presented in Table 4.4.

Table 4.4 Descriptive Statistics for Item 12a

	f	p
Incorrect responses (True/Yes/Blank)	334	31.1%
Correct response (False/No)	729	67.9%

As seen, majority of students gave correct responses (67.9%) which indicated that they could critique a statistical claim in sampling context. The rest of them (31.1%) either accepted this statistical claim or left this item blank. The explanations for this item were requested in item 12b where the mean score for this item out of 1 was .58 with standard deviation of .47. These explanations were categorized as incorrect or unrelated responses, pre-statistical responses including intuitive conceptions and correct and statistical responses. The frequencies and percentages regarding these explanations were represented in Table 4.5.

Table 4.5 Descriptive Statistics for Item 12b

Classification of Responses		Students' Responses	f	p
Blank/Wrong or unrelated responses			394	36.7%
Pre-Statistical			111	10.3%
	-	Explanations related to daily life	71	6.6%
	-	Contextual beliefs	40	3.7%
Statistical			559	52%
	-	Cannot generalize with one school	314	29.2%
	-	Cannot generalize with one class	166	15.5%
	-	Cannot generalize with very few students	79	7.4%

The students rejected this sample claim which involved generalization to a population either asserting that "there was one school" (29.2%), "there was one class" which was fifth graders (15.5%), or "there were very few students" (7.4%). However, some of the students justified their incorrect responses either by giving examples related to their daily lives (6.6%) such as "I do not watch TV for 3 hours" or their beliefs (3.7%) such as "All children should not watch TV for 3 hours". These explanations were counted as pre-statistical and constituted 10.3% of total participants. The rest of students (36.7%) provided inadequate explanations or did not attempt to justify their responses.

The detailed analysis of items related to sampling content revealed that students performed better while applying ideas of sample which was a second tier behavior and evaluating statistical reports which was a third tier characteristics in sampling context. In addition, almost half of the participants were able to explain correctly their evaluations of statistical report whereas they performed poorly in defining sample as a statistical terminology which was a characteristic of the first tier.

4.1.2. Average

The items related to average content were items 2a/b, 7 and 13a/b. These items intended to measure tier one, two, and three respectively. Items 2b and 7 were multiple choice type items whereas items 2a and 13b were open-ended items.

4.1.2.1. Tier 1: Understanding Average Terminology

Item 2a intended to measure students' understanding of average in context. This item was presented below

> "Last year, an average of 20 people had died due to traffic accidents." What do you understand of the word **"average"** in this sentence?

The mean score for this item out of 1 was .42 with standard deviation of .26. The answers of students were classified through four categories which were blank or incorrect responses, pre-statistical responses, responses through measures of central tendency and statistical responses. The frequencies and percentages regarding this classification related to item 2a are represented in Table 4.6.

Table 4.6 Descriptive Statistics for Item 2a

Classification of Responses		Students' Responses	f	p
Blank/Wrong or unrelated responses			170	15.8%
Pre-Statistical			523	48.7%
	-	Almost	374	34.8%
	-	Approximately	56	5.2%
	-	More or less	94	8.8%
Descriptions via Measures of Central Tendency			318	29.6%
	-	Arithmetic Mean	281	26.2%
	-	Median	14	1.3%
	-	Mod	23	2.1%
Statistical			56	5.2%
	-	Balance point	23	2.1%
	-	Representative value of data set	33	3.1%

Table 4.6 indicated that majority of students either explained the term "average" through pre-statistical words (48.7%) or described through measures of central tendency (29.6%). The most notable response in pre-statistical responses was "almost" (34.8%) while "arithmetic mean" or "add them up and divide" algorithm were the most frequent descriptions (26.2%) for those who explained average through measures of central tendency. However, statistically correct responses constituted only 5.2% percent of total responses.

Item 2b intended to measure students' familiarity with methods for finding average or central tendency. This item was presented below.

Which one of the followings was not a method for finding average 20 people had died due to traffic accidentlast year?

a) Add the number of people died in a year and divide with 12.
b) Put the number of people died each year in order and choose the middle one.
c) Find the most frequent number in the data set involve number of people died each year.
d) Subtract the largest number of people died each year from the smallest number of people died each year.

For this item, 44% of students labeled “range” which was not a method for finding average. Yet, 36% of them labeled either “median” (18.1%) or “mod” (18.5%) as if these were not a method for finding average. This finding indicated that almost one third of the participants did not count median and mod as a measure of central tendency.

4.1.2.2. Tier 2: Interpreting Average in Context

The next item related to average which was in the second tier group was item 7. This item intended to measure students’ application of ideas related to average in context. This item was presented below.

A researcher who lives in a town consisting of 50 families has found the mean of children per family as 2.2. Which one of the followings is absolutely true?

a) Half of the families in this town has two children.
b) The number of families with 3 children more than families with 2 children.
c) There are 110 children in this town.
d) The mean of children per adult is 2.2.

The mean score for this item out of 1 was .41 with standard deviation of .46. The responses of students are given in Table 4.7.

Table 4.7 Descriptive Statistics for Item 7

	f	p
Incorrect responses	635	59.1%
Correct response	432	40.2%

Table 4.7 showed that 40.2% of students correctly interpreted average in context whereas others (59.1%) chose the incorrect interpretations. It could be inferred that for average content in the second tier, only less than half of the participants had performed properly.

4.1.2.3. Tier 3: Evaluating Average Claims

The last two items related to average content were 13a and 13b which were placed in third tier group. The first one was an evaluation of a statistical claim which involved calculating arithmetic mean with an outlier in true/false format where mean was .17 and standard deviation was .37. The second item required students to explain their responses in the previous item. These items are presented below.

The number of problems solved in a math class is counted and represented in the following table.

Student	**Number of problem**
A	2
B	6
C	2
D	22
E	3
F	2
G	1
H	2

In order to summarize these data, the mean is calculated and found 5.

a) Do you agree with this?
b) Explain your answers with reasons.

The mean for this item was .05 and standard deviation was .20. Since this item was open-ended, responses of students were classified as incorrect, pre-statistical and statistical. The frequencies and percentages of the students' responses are presented in Table 4.8 and Table 4.10.

Table 4.8 Descriptive Statistics for Item 13a

	f	p
Incorrect responses (True/Yes/Blank)	888	82.7%
Correct response (False/No)	176	16.4%

As seen, majority of students gave incorrect responses or left this item blank (82.7%) which indicated that they were not able to critique a statistical claim in average context. The rest of them (16.4%) could correctly evaluate the appropriateness of this claim. The explanations regarding this item including classifications of these explanations are presented in Table 4.9.

Table 4.9 Descriptive Statistics for Item 13b

Classification of Responses	Students' Responses	f	p
Blank/Wrong or unrelated responses		1000	93.1%
	- Justification with arithmetic mean	563	52.4%
	- Wrong explanations related to context	26	2.4%
	- Other blank/wrong explanations	411	38.3%
Pre-Statistical		24	2.2%
	- Notice the difference between numbers	14	1.3%
	- Notice the outlier/extreme value	10	0.9%
Statistical		40	3.7%

Table 4.9 indicated that most of the participants provided wrong or unrelated

responses (93.1%). These students accepted the statistical claim in average context without criticizing either providing wrong explanations related to context (2.4%) such as “Five questions can be solved in a class period” or justifying the results with arithmetic mean (52.4%). The rest of the participants gave pre-statistical (2.2%) or statistical (3.7%) responses. The statistical responses included either recognizing outlier in the data set or stating that getting average with median or mode is more appropriate. The difference between these statistical and pre-statistical responses was the appreciation of variability in the data set occurred in statistical explanations whereas recognizing outlier appeared in pre-statistical responses.

The detailed analysis of items revealed that majority of students had inadequate knowledge regarding average content. The most notable finding was that most of the students understood average which was a characteristics of the first tier behavior as “add them up and divide” algorithm which referred to the arithmetic mean and they did not consider median and mode as a way of finding average. In addition, only less than half of the participants were able to interpret average in context as a characteristic of second tier of statistical literacy. The majority of participants had failed to evaluate a statistical claim which was contextualized as third tier where they could not recognize outlier or justified this claim by providing evidence through arithmetic mean.

4.1.3. Graphs

The items related to graphs content were items 3, 8 and 14a/b measuring tier one, two, and three respectively. Items 3 and 8 were multiple choice type items whereas item 14b was an open-ended item.

4.1.3.1. Tier 1: Understanding Graph Terminology

Item 3, which was a tier 1 question, required respondents to choose appropriate graphical representation among others for a given data set in context. This item is presented below.

The data below represents what a 5 TL lunch includes and price of each item.

- 2 TL main meal
- 0.5 TL soup
- 1.5 TL desert
- 1 TL salad

Which one of the following graphs type represents best this data?

a) Pie Chart b) Histogram c) Line Graph d) Bar Graph

The mean score for this item out of 1 was .38 with standard deviation of .49. The frequencies and percentages regarding the alternatives related to this item are represented in Table 4.10.

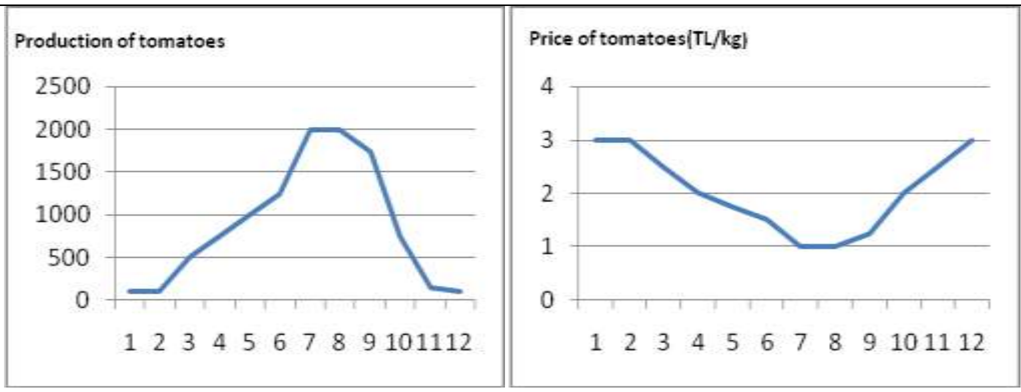
Table 4.10 Descriptive Statistics for Item 3

	Distracters	f	p
Correct responses (Pie Chart)		401	37.3%
Incorrect response		667	62.1%
	- Histogram	197	18.3%
	- Line Graph	114	10.6%
	- Bar Graph	356	33.1%

As it is seen in Table 4.10, majority of students responded for this question incorrectly. Almost one third of the participants (33.1%) labeled “bar graph” as an appropriate representation for the given data set which involved part-whole relationship. Yet, 37.3% of students provided “pie chart” answer, which was the most suitable representation. Although these two responses were acceptable where pie chart was the more appropriate since it represents part whole relationships, still, the rest of the participants labeled histogram (18.3%) and line graph (10.6%) which could represent only continuous data sets.

4.1.3.2. Tier 2: Interpreting Graph in Context

The next item related to graphs concepts in the second tier group was item 8. This item intended to measure students' interpretations of graphical representations in context. This item is presented in below.



The above graphs represent production of tomatoes and price per kg. According to the graphs, which can be inferred?

I. The price of tomatoes is less in summer, more in winter times.

II. The production of tomatoes is more in summer, less in winter.

III. Since production of tomatoes is less in winter, the price is high.

a) Only I b) I and II c) II and III d) I, II and III

The mean score for this item out of 1 was .69 with standard deviation of .46. The responses of students are given in Table 4.11.

Table 4.11 Descriptive Statistics for Item 8

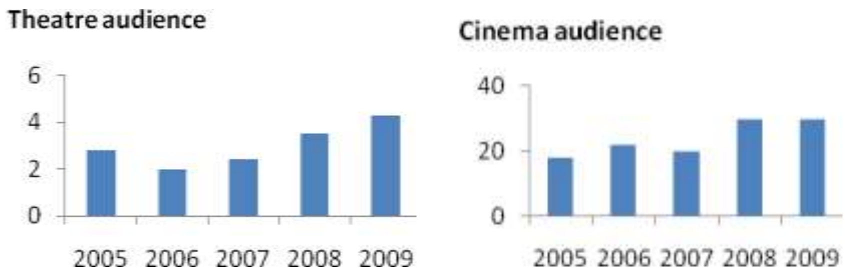
	f	p
Incorrect responses	330	30.7%
Correct response	737	68.6%

Table 4.11 indicated that majority of students interpreted the collection of two line graphs correctly. This finding provided the inference that most of the students could

perform more than moderate in graphs concepts in the second tier.

4.1.3.3. Tier 3: Evaluating Graphs

The last two items associated with graphs concepts were 14a and 14b which belonged to tier three group. These items, which are presented below, intended to explore the critical evaluation of graphical representations involving misleading.



An announcer showed these graphs and said that "Although the number of audience in theatre and cinema differs before 2009, the number of audience reaches to the equal number in 2009".

a) Do you think that the claim announced is acceptable?

b) Provide a statistical explanation for your answer.

The mean score for the former item out of 1 was .56 with standard deviation of .49 while the mean score for the latter item out of 1 was .45 with standard deviation of .50. The frequencies and percentages regarding these two items are presented in Table 4.12.

Table 4.12 Descriptive Statistics for Item 14a and 14b

Item #		f	p
Item 14a	Incorrect responses (True/Yes/Blank)	463	43.1%
	Correct response (False/No)	597	55.6%
Item14b	Incorrect explanations	583	54.3%
	Correct explanations	478	44.5%

As seen, although more than half of the students (55.6%) could critically evaluate graphical representations, less than half of them (44.5%) could correctly explain why these graphs were misleading where they indicated that the scales of these graphs were different or the number of audience for cinema and theatre were different. It could be inferred that almost 10% of students who could critically evaluate these graphs had failed to provide appropriate explanations for their answers.

The detailed analysis of items showed that almost half of the participants performed better while interpreting graphical representations. In addition, almost half of the participants were able to evaluate misleading bar graphs and explain correctly their evaluations of statistical report whereas they performed poorly in choosing appropriate graphical representations which was demanded by the first tier compared to other tiers.

4.1.4. Chance

The items related to chance concept were items 4, 9, and 15a/b. These items intended to measure tier one, two, and three in order. Item 4 and 9 were multiple choice type items. Item 15a was formatted in true/false type whereas item 15b was an open-ended item.

4.1.4.1. Tier 1: Understanding Chance Terminology

Item 4, which was a tier 1 question, required respondents to choose appropriate random selection among others for an isolated context which was choosing marbles from a bag. This item is presented below.

Which one of the followings are random selections?

I. Selection of red marbles after putting them in a bag and mixed
II. Selection of any two marbles after putting them in a bag and mixed
III. Selection of every 5^{th} marble without putting in a bag

a) Only I b) Only II c) I and II d) I and III

The mean score for this item out of 1 was .63 with standard deviation of .48. The frequencies and percentages regarding these two items are presented in Table 4.13.

Table 4.13 Descriptive Statistics for Item 4

	f	p
Incorrect responses	386	35.9%
Correct response	681	63.8%

Table 4.13 indicated that almost two third of the students chose the correct alternative for the item contextualized in understanding chance. More precisely, 63.8 percent of the participants for the first tier question in chance context.

4.1.4.2. Tier 2: Interpreting Chance in Context

Item 9 was one of the tier 2 questions which asked for interpretation of a risk situation in health context. This item is presented below.

The following message is printed in a bottle of skin cream:
“WARNING: For application to skin areas there is a 15% chance of developing rash. If a rash develops, consult your doctor.
Which of the following is the best interpretation of this warning?

a) About 15 of 100 people who use this medication develop a rash.
b) If a rash develops, it involves only 15% of the skin.
c) There is hardly any chance of getting a rash using this medication.
d) If you use this cream, apply only 15% of your skin.

The mean score for this item out of 1 was .69 with standard deviation of .46. The frequencies and percentages regarding these two items are presented in Table 4.14.

Table 4.14 Descriptive Statistics for Item 9

	f	p
Incorrect responses	330	30.7%
Correct response	737	68.6%

Table 4.14 indicated that the second tier question had similar results with the first tier question where almost two third of the students chose the correct alternative for these item, too. More precisely, 68.6 percent of them for the second tier in chance context performed correctly.

4.1.4.3. Tier 3: Evaluating Chance Claims

Item 15a and 15b were the last two items related to chance concepts contextualized in the third tier group. The former item required students to critically evaluate a statistical report in chance context. However, the possible statistical questions that could be asked regarding the appropriateness of this report was asked in the latter one. These items are presented below.

A study found that those who smoked a pack of cigarette a day for less than 49 years doubled the risk of premature wrinkling while for more than 50 years, the risk was 4.7 times greater compared to those who do not smoke. The table below summarizes this information.

	Less than 49 years	**More than 50 years**
Risk of non smokers	A	B
Risk of smokers	2A	4,7B

a) Is the result of this report acceptable?

b) What kind of do you ask to examine the validity of this report?

The mean score for this item out of 1 was .76 with standard deviation of .43 for the former item while the mean score for this item out of 1 was .22 with standard deviation of .37 for the latter one. The frequencies and percentages related to items 15a and 15b are displayed in Table 4.15.

Table 4.15 Descriptive Statistics for Item 15a and 15b

Item #		f	p
Item 15a	Incorrect responses (False/No/Blank)	250	23.3%
	Correct response (True/Yes)	811	75.5%
Item15b	Unrelated questions/Incorrect responses	747	69.6%
	Questioning the report but not statistical	153	14.2%
	Questioning the report and statistical	160	14.9%

The first part of Table 4.15 indicated that majority of students (75.5%) correctly evaluated this statistical report which was related to chance concepts. However, 69.6% of the students did not question the claim where they did not provide anything or asked unrelated questions such as “Do you smoke?” The rest of the participants were able to ask questions related to the results of the report. They either could question the report but not in a statistical way (14.2%) through asking questions such as “How did you conduct this study?” and “Why did you conduct this study?” or could question the report statistically (14.9%) through asking questions such as “How many people were asked about smoking?”, or “Does one year make a difference?”.

The analysis of items related to chance concepts indicated that although 8th graders could perform better for questions in the first and second tier, they failed to ask questions for a statistical report in chance context. More precisely, almost two third of the participants were able to understand “random” as a terminology; likewise, two third of them understood risk situations and interpreted correctly. Similarly, 75% of them could evaluate the appropriateness of a statistical report in probability context. However, majority of the participants (70%) had failed to question this report either statistically or in other ways.

4.1.5. Inference

The items related to inference were items 10a/b and 16a/b. These items intended to measure tier two and three respectively. There was not any item for the first tier since this content did not have a specific statistical terminology as opposed to other content domains.

4.1.5.2. Tier 2: Making Inferences in Context

There were two items which were 10a and 10b required to make students inferences in context. These two items were open-ended items while the former ones could be considered as true/false type. Item 10a intended to measure students' predictions based on data where 10b required students to explain their responses in the "a" part. These items are presented below.

The weight of a baby for each month from the birth is provided in below table.

Age (month)	**Weight (kg)**
0	3,5
1 month	4,5
2 month	5
3 month	6
4 month	7
5 month	7,5
6 month	8

a) According to this, predict that how much kg baby will weighted at the end of 7^{th} month.

b) Explain how you predict your answer.

The mean score for the former item out of 1 was .89 with standard deviation of .30 while the mean score for the latter item out of 1 was .66 with standard deviation of .47. The frequencies and percentages related to this item are represented in Table 4.17.

Table 4.16 Descriptive Statistics for Item 10a and 10b

Item #		f	p
Item 10a	Incorrect responses	108	10.1%
	Correct response	908	89.2%
Item10b	Incorrect explanations	360	33.5%
	Correct explanations	706	65.7%

As seen from Table 4.16, very high percentage of participants (89.2%) predicted appropriate values for the given data set. Yet, nearly two third of participants (65.7%) were able to explain their predictions correctly.

4.1.5.3. Tier 3: Evaluating Inferences

The next questions placed in inference domain were items 16a and 16b. The former one required respondents to critically evaluate an inference related to a biased two-way chart. More precisely, the chart in the question was interpreted as if there were cause-effect relationship, though there was not. These items are presented below.

The following information is from a survey about smoking and lung disease among 250 people.

	Lung Disease	**No lung disease**	**Total**
Smokers	90	60	150
Nonsmokers	60	40	100
Total	150	100	250

a) Using this information a researcher states that “The reason for lung disease is smoking.” Do you think that is this claim acceptable?
b) Explain your answer statistically.

The mean score for the former item out of 1 was .20 with standard deviation of .40 while the mean score for the latter item out of 1 was .08 with standard deviation of

.22. The frequencies and percentages related to the first part of the item are displayed in Table 4.17.

Table 4.17 Descriptive Statistics for Item 16a

	f	p
Incorrect responses (True/Yes/Blank)	847	78.9%
Correct response (False/No)	213	19.8%

As table 4.17 indicated, a considerable percentage of students (78.9%) had failed to critically evaluate the incorrect inference of a two way chart. That is, they accepted the improper statistical claim without questioning. The classification of explanations regarding their responses for this item is represented in Table 4.18.

Table 4.18 Descriptive Statistics for Item 16b

Classification of Responses	Codes based on Students' Responses	f	p
Blank/Wrong or unrelated responses		912	84.9%
	- Contextual Beliefs	96	8.9%
	- Looking for numbers not examining relationships	321	29.9%
	- Other blank/wrong explanations	495	46.1%
Pre-Statistical		117	10.9%
Statistical		31	2.9%
	- Equal ratios	24	2.2%
	- Equal probabilities	7	0.7%

Table 4.18 indicated that majority of students could not provide either pre-statistical or statistical explanations for their evaluations where they gave either incorrect or unrelated responses (84.9%). These students accepted the statistical inference

without criticizing either by providing wrong explanations related to context beliefs (8.9%) such as “Smoking is harmful” and “My grandparent got cancer due to smoking” or looking for numbers without examining the relationships between them (29.9%) such as “The number/ratio of people who smoke is higher”. 10.9% of participants gave pre-statistical explanations such as “There are smokers who are not cancer”. The rest of the participants, which was a small percentage (2.9%), statistically explained their responses through indicating either ratios (2.2%) or probabilities (0.7%) were equal for those who were lung cancer or not.

The detailed analysis of these items revealed that students performed differently in the second and third tiers. In other words, majority of students were able to make predictions based on data and explain their predictions, whereas most of them had failed to evaluate critically an inference without appropriate statistical foundation. Yet, very small percentage of participants explained their responses statistically.

4.1.6. Variation

The items related to variation content were items 5a/b, 11 and 17a/b. These items intended to measure tier one, two, and three respectively. Item 5a and 11 were multiple choice type items whereas items 5b and 17a/b were open-ended items.

4.1.6.1. Tier 1: Understanding Variation Terminology

Item 5a required students to select the data set which had more variability among others without context. This item is presented below.

Which of the data sets involve more variability? Provide your answer without calculation. a) 10, 11, 12, 13, 14, 15 b) 13, 13, 13, 13, 13, 13 c) 11, 12, 12, 13, 13, 14 d) 10, 12.5, 12.5, 12.5, 12.5

The mean score for this item out of 1 was .61 with standard deviation of .49. The frequencies and percentages of correct responses for this item are represented in Table 4.19.

Table 4.19 Descriptive Statistics for Item 5a

	Distracters	f	p
Correct responses (a)		653	60.8%
Incorrect response		318	29.8%
	- b	171	16.1%
	- c	39	3.6%
	- d	108	10.1%

As seen, majority of students (60.8%) were able to choose the data set with more variability. Since this item was in multiple-choice format, the explanations regarding this item were asked in item 5b. More specifically, students were required to provide explanations for their selections. The mean score for this item out of 1 was .24 with standard deviation of .31. The frequencies and percentages regarding the classification of responses provided by participants for item 5b are presented in Table 4.20.

Table 4.20 Descriptive Statistics for Item 5b

Classification of Responses	Codes based on Students' Responses	f	p
Blank/Wrong or unrelated responses		625	58.2%
	- All numbers are same	70	6.5%
	- Other blank/wrong explanations	555	51.8%
Pre-Statistical		130	12.1%
	- Numbers are increasing	51	4.7%
	- Numbers are different	79	7.4%

Table 4.20 (continued)

Descriptions via Measures of Spread		292	27.2%
	- Range	248	23.1%
	- Inter quartile range	1	0.1%
	- Standard deviation	43	4%
Statistical		19	1.8%
	- Larger variability	13	1.2%
	- Away from average	6	0.6%

Table 4.20 indicated that majority of students (58.2%) either gave wrong responses or did not explain anything related to their answers in the first part. Of these, those who selected the data set which had the same numbers explained their responses through stating “all numbers are the same” (6.5%). Some of the participants (12.1%) provided pre-statistical explanations either stating that “numbers are increasing” (4.7%) or “numbers are different” (7.4%). A considerable percentage of students (27.2%) explained their responses through measures of spread. The most notable response in this category was “range” (23.1%) while “standard deviation” response was quite frequent (4%). Yet, very small percentage of participants (0.1%) explained their response through “inter quartile range”. Statistically correct responses constituted only 1.8% percent of total responses where they either indicated the large variability in data set (1.2%) or distance from the average value (0.6%).

4.1.6.2. Tier 2: Interpreting Variation in Context

The next item related to variation concept was 11th item which was contextualized with the second tier. It was a multiple choice item and required students to interpret statistical claims involving variability. This item is presented below.

Some statistics regarding the grades of mathematics for 8A and 8B sections in an elementary school is presented in the table below.

	Arithmetic Mean	Standard deviation
Section A	80	5,2
Section B	76	3,5

Which one of the followings is true?

I. If arithmetic mean is examined, the grades in section A higher than section B.

II. If standard deviation is examined, the variation in section B is smaller.

III. If standard deviation is examined, the variation in section A is smaller.

a) Only I b) Only II c) I and II d) I and III

The mean score for this item out of 1 was .75 with standard deviation of .44. The frequency and percentage of correct responses are presented in Table 4.21.

Table 4.21 Descriptive Statistics for Item 11

	f	p
Incorrect responses	272	25.3%
Correct response	795	74.0%

As seen, majority of students (74%) were able to interpret statistical claims involving variability. It could be inferred that variability in the second tier was accomplished by most of the participants.

4.1.6.3. Tier 3: Evaluating Variation Claims

The next items which were in the third tier group were items 17a and 17b. Students did evaluate the data sets and chose the one had more appropriate variability among others in the former item, while they were asked for explanations for their responses in the latter one. These items are presented below.

A group of students noted the highest temperatures in Ankara during one year. They find the highest average temperature in Ankara as 16^{o}. Different from this group, three students predicted possible highest temperature for six different days in a year.

Students	**Predicted Temperature**
Seda	16, 35, 1, 5, 29, 10
Zeynep	16, 16, 16, 16, 16, 16
Umut	16, 15, 14, 26, 8, 17

a) Which students provide the data set regarding average temperatures with the most appropriate variability?

a) Seda
b) Zeynep
c) Umut

b) Explain your answer.

The mean score for the former item out of 1 was .20 with standard deviation of .40 while the mean score for the latter item out of 1 was .08 with standard deviation of .22. The descriptive statistics involving frequencies and percentages of item17a is presented in Table 4.22.

Table 4.22 Descriptive Statistics for Item 17a

Distracters	f	p
Correct responses (Umut)	253	23.6%
Incorrect response	807	75.1%
- Seda	222	20.7%
- Zeynep	329	30.6%
- Other blank/wrong responses	256	23.8%

The analysis of Table 4.22 revealed that majority of students gave incorrect response where only 23.6% of the participants did choose the data set with more appropriate variability. Of the incorrect responses, 20.7% of students did choose “Seda” which had greater variability, whereas almost one third of the students labeled “Zeynep” which consisted of the same numbers. The classification of the explanations regarding their answers is given in Table 4.23 below.

Table 4.23 Descriptive Statistics for Item 17b

Classification of Responses	Codes based on Students’ Responses	f	p
Blank/Wrong or unrelated responses		785	73.1%
	- Same numbers in the data set	151	14.1%
	- Equal to the average	96	8.9%
	- Other blank/wrong responses	539	50.2%
Pre-Statistical		125	11.6%
	- More difference between numbers	111	10.3%
	- Different numbers	14	1.3%
Statistical		149	13.9%
	- Appropriate variation	38	3.5%
	- Different numbers but closer	38	3.5%
	- Around average value	73	6.8%

Table 4.23 indicated that a high percentage of students (73.1%) either gave wrong and unrelated responses or left the explanation part blank. Those who picked “Zeynep” as data set which had the most appropriate variation explained their answers either as “the numbers were equal to the average” (8.9%) or “numbers were the same” (14.1%). The pre-statistical explanations included either more difference between numbers (10.3%) or different numbers (1.3%). Still, there were statistical explanations (13.9%) which consisted of responses such as “appropriate variation”

(3.5%), “different but closer numbers” (3.5%), and “around average value” (6.8%).

The detailed analysis of items related to variation concept indicated that students obviously performed differently in different tiers. For instance, although it was possible to say that there were inadequate knowledge in understanding and evaluating variability, almost 75% of participants correctly interpreted variation in context. One of the interesting findings regarding the evaluation of responses in item 5b and 17b was that 6.5% and 14.1% of students, respectively, indicated that more variation was involved where the data set consisted of same numbers. In addition, very small percentage of students (1.8%) gave statistically correct explanation regarding understanding of variation whereas most of them (27.2%) described variation through measures of spread.

In this section of result chapter, frequencies and percentages of each item in relation to content domains were presented. The next section consisted of results regarding three tiers.

4.2. 8th Grade Students’ Statistical Literacy in terms of Three-Tiers

In this study, statistical literacy was composed of three tiers as mentioned in detail before. Students’ statistical literacy scores in terms of tiers were obtained through dividing by the maximum score that they could get for each tier similar to the calculation of total statistical literacy scores. The descriptive statistics for each tier were represented in Table 4.24.

Table 4.24 Descriptive Statistics of SLT Scores in terms of Tiers

	N	Min.	Max.	Mean	SD	Skewness		Kurtosis	
						Stat.	SE	Stat.	SE
Tier 1	1063	,00	,83	,35	,17	,31	,08	-,55	,15
Tier 2	1065	,00	1,00	,69	,23	-,53	,08	-,33	,15
Tier 3	1056	,00	,94	,31	,19	,37	,08	-,36	,15

Table 4.25 indicated that mean scores for each tiers were different. That is, although mean scores for tier 1 (*M*=.35, *SD*=.17) and tier 3 (*M*=.31, *SD*= 19) were closer to each other, tier 2 had relatively higher mean score (*M*=.69, *SD*=23). Skewness and kurtosis values showed that the distribution of mean scores for each tier was normal where these values did not surpass -1.00 and +1.00. The next section of this chapter was allocated for inferential statistics which explored the difference between mean scores of tiers through utilizing one-way within subjects analysis of variance.

4.3. The Difference between Tiers of Statistical Literacy

Another research question of this study was "Is there a significant difference between the mean scores of first, second and third tier of statistical literacy?" Pallant (2007) stated that in order to compare differences between two or more conditions that had been undertaken by the same participants or each participant measured on three different questions or items, one-way within subjects analysis of variance (or repeated measures ANOVA) should be used provided that the measures were in the same response scale. Therefore, in order to examine the difference between mean scores of tiers in this study one-way within subjects analysis of variance was conducted. Prior to running the analysis, the statistical assumptions associated with one-way within subjects ANOVA were checked.

4.3.1. Assumptions

The assumptions to be assured before conducting one-way within subjects ANOVA were level of measurement, random sampling, and independence of observations, normality, and sphericity (Pallant, 2007).

The variables for one-way within subjects ANOVA were mean scores of tier 1 , tier 2 and tier 3 for statistical literacy which were continuous variables. Hence, level of measurement assumption was assured.

The cluster sampling method was utilized for this study which indicated that the schools in this study were chosen randomly. Hence, scores of individuals were obtained using a random sample of population.

Each measurement of tiers was not influenced by others, since each tier consisted of different items or questions. That is, the measurements were independent from each other; therefore, there was no violation of this assumption.

For parametric techniques, mean scores for each variable should be normally distributed (Pallant, 2007). In order to check normality of each statistical literacy scores in terms of tiers, histograms, normal Q-Q plots, and skewness and kurtosis values were examined. The shape of these graphs indicated that the distributions were normal. In addition, skewness and kurtosis values of each variable (tier) were represented in Table 4.26.

Table 4.25 Skewness and Kurtosis Values for Tier 1, Tier 2 and Tier 3

	Skewness	Kurtosis	N
Tier 1	,31	-,55	1063
Tier 2	-,53	-,33	1065
Tier 3	,37	-,36	1056

Table 4.25 indicated that skewness and kurtosis values for each variable were placed in the acceptable range. In addition, the sample sizes for each variable were quite large. Therefore, normality of distribution assumption was assured for Tier 1, Tier 2 and Tier 3 scores.

Another assumption for one-way within subjects ANOVA was sphericity which meant that the variance of population difference scores for any two conditions were the same as the variance of the population difference scores for any other two conditions (Pallant, 2007). In order to check sphericity assumption, Mauchly’s Test of Sphericity was examined. This test was significant which rejected the hypothesis of equal variances in difference scores. However, in Pallant’s (2007) point of view, this assumption was commonly violated similar to this case. Therefore, results of multivariate statistics, which did not require sphericity assumption, were interpreted for this analysis.

4.3.2. Results

A one-way within subjects ANOVA was conducted to compare mean scores on Tier 1, Tier 2, and Tier 3 aspects of statistical literacy. The means and standard deviations are presented in Table 4.26. Preliminary analyses were performed to ensure that there was no problem related to assumptions. There was a significant difference between three tiers of statistical literacy, Wilks' Lambda = .25, F (2, 1048) = 1579.56, p<.0005 with multivariate partial eta squared = .75. The effect size was interpreted as large using Cohen's (1988) guidelines. As a follow up test, paired samples t-tests were conducted and results were evaluated using the Holm's sequential Bonferonni procedure. There were statistically significant differences between mean scores of Tier 1 and Tier 2, Tier 1 and Tier 3, and Tier 2 and Tier 3.

Table 4.26 Descriptive Statistics for Tier 1, Tier 2 and Tier 3

	N	Mean	SD
Tier 1	1063	,35	,17
Tier 2	1065	,69	,23
Tier 3	1056	,31	,19

4.4. 8th Grade Students' Attitudes towards Statistics

In this study, measure of students' attitudes towards statistics was obtained through the final version of ATSQ. This one-dimensional questionnaire consisted of a total 27 items which were five point Likert type items to be responded as Strongly Disagree, Undecided, Agree, and Strongly Agree. Since there were 12 items in negative form, before running the analyses, these were re-coded. A mean attitude score for each student was calculated by taking average of students' attitude scores for each item. Therefore, the maximum attitude value for each participant was 5 while minimum was 1 where obtaining high scores from this questionnaire meant holding positive attitudes towards statistics.

The descriptive statistics for ATSQ including mean and standard deviation are presented in Table 4.28.

Table 4.27 Descriptive Statistics for ATSQ Scores

	N	Min.	Max.	Mean	SD	Skewness		Kurtosis	
						Stat.	SE	Stat.	SE
ATSQ	1034	1,14	5,00	3,52	,74	-,38	,08	-,11	,15

Table 4.27 indicated that there were 1034 participants whose ATSQ scores were examined; therefore, there were 40 missing cases (4%) which constituted less than 5% of the sample. The mean value of attitudes towards statistics scores of participants was 3.52 out of 5 where standard deviation was .74. In addition, the scores ranged between 1.14 and 5.00. From the descriptive statistics regarding ATSQ, it could be inferred that students had slightly positive attitudes towards statistics while the mean value was close to neutral. Skewness and kurtosis values showed that the distribution of ATSQ mean scores was normal where these values did not surpass -1.00 and +1.00.

Item mean distribution for each item is displayed in Table 4.28 in order to provide an in depth idea for students' attitudes towards statistics. "R" addresses reversed items.

Table 4.28 Item Mean Distribution for ATSQ

Items		Mean
1.	I like statistics.	3,57
2.	Statistics is unlikeable. (R)	3,62
3.	I enjoy discuss about statistics.	3,25
4.	Information related to statistics is annoying. (R)	3,54
5.	Statistics helps for mental development.	3,67
6.	Information related to statistics makes me anxious. (R)	3,76
7.	I want more class hours related to statistics.	2,88
8.	Statistics can be learned easily.	3,62
9.	I am scared of exams related to statistics. (R)	3,67
10.	Statistics draws interest on me.	3,37
11.	Statistics has an important role for decision making.	3,59
12.	Statistics makes me confused. (R)	3,54
13.	I study statistics lovingly.	3,36
14.	If I could I do not learn statistics. (R)	3,59
15.	Statistics is not an interesting subject. (R)	3,13
16.	I want to learn statistics in advance level.	3,25
17.	Statistics is used almost every occupation.	3,67
18.	I get bored when I study statistics. (R)	3,48
19.	Statistics teaches individuals to think.	3,65
20.	I get frustrated when I heard statistics. (R)	4,00
21.	I am scared of statistics. (R)	3,78
22.	Everybody needs to learn statistics.	3,42
23.	I do not like statistics. (R)	3,58
24.	Statistics enhances one's estimation ability.	3,80
25.	I get bored while statistics is taught. (R)	3,44
26.	Statistics has an important place in daily lives.	3,56
27.	Statistics is enjoyable.	3,37

The analysis of mean distribution of items indicated that 8th grade students generally tended to agree with the attitude statements, yet their responses were around a neutral stance. The most notable finding based on mean distribution was that statements related to opinion had slightly higher mean scores compared to interest related statements. Put it differently, participants tended to agree with opinion statements such as “Statistics enhances one’s estimation ability” or “Statistics helps for mental development” where they were not sure about interest statements such as “Statistics is not an interesting subject” or “I want more class hours related to statistics”. In addition, before recoding, students tended to disagree with the attitude statement incorporating anxiety or frustration such as “I get frustrated when I heard statistics” or “I am scared of statistics”.

4.4. The Relationship between Eighth Grade Students’ Statistical Literacy and Their Attitudes towards Statistics

In order to examine the possible relationship between attitudes towards statistics and statistical literacy of eighth grade students, Pearson-product moment correlation analysis was conducted. The variables for this correlational analysis were mean scores for attitudes towards statistics and mean scores for statistical literacy. The mean scores of attitudes towards statistics have already been presented in the previous section in Table 4.27 and Table 4.28. The total statistical literacy scores of participants are presented below in detail.

4.4.1. 8th Grade Students’ Statistical Literacy Scores

Statistical literacy of 8th grade students was measured through Statistical Literacy Test. In this test, there were both multiple choice and open-ended items. For convenience, students’ total statistical literacy scores were divided by the maximum score that they could get. Therefore, the maximum statistical literacy score that students could gain was 1 whereas minimum was 0 for each student. The descriptive statistics for total statistical literacy scores of participants is displayed in Table 4.29.

Table 4.29 Descriptive Statistics for SLT Scores

	N	Min.	Max.	Mean	SD	Skewness		Kurtosis	
						Stat.	SE	Stat.	SE
SLT	1050	,06	,97	,46	,16	,21	,08	-,28	,15

Table 4.29 indicated that there were 1050 participants whose SLT scores were examined; therefore, there were 24 missing cases (2%) which constituted less than 5% of the total sample. The mean value of statistical literacy scores of participants was .46 out of 1 where standard deviation was .16. In addition, the scores ranged between .06 and .97. From the descriptive statistics regarding SLT, it could be inferred that students have lower mean scores than the middle point of the test. Skewness and kurtosis values showed that the distribution of SLT mean scores was normal where these values did not surpass -1.00 and +1.00.

Prior to running the analysis, the statistical assumptions associated with correlation analysis were checked.

4.4.1. Assumptions of Pearson Product Moment Correlation Analysis

The assumptions to be assured before conducting analysis were level of measurement, related pairs, independence of observations, normality, linearity, and homoscedasticity (Pallant, 2007).

The variables for correlational analysis were mean scores for attitudes towards statistics and mean scores for statistical literacy which were continuous variables. Hence, level of measurement assumption was assured.

Pallant (2007) stated that providing a score on both variables was another assumption of correlational analysis. Since pairwise deletion method was used while dealing with missing data, participants, who were included in the correlation analysis, had both scores which were attitudes towards statistics and statistical literacy. Therefore, this assumption was ensured.

The issue that each measurement was not influenced by other was already mentioned. That is, the measurements were independent from each other; therefore, there was no violation of this assumption.

In correlational analysis, mean scores for each variables should be normally distributed (Pallant, 2007). In order to check normality of attitude towards statistics and statistical literacy scores, histograms, normal Q-Q plots, and skewness and kurtosis values were examined. The shape of these graphs indicated that the distributions were normal. In addition, skewness and kurtosis values of each variable are represented in Table 4.30.

Table 4.30 Skewness and Kurtosis Values of ATSQ and SLT

	Skewness	Kurtosis	N
ATSQ	-.38	-.11	1034
SLT	.21	-.28	1050

Table 4.30 indicated that skewness and kurtosis values for both variables were placed in the acceptable range. In addition, the sample sizes for each variable were quite large. Therefore, normality of distribution assumption was assured for ATSQ and SLT scores.

Another assumption in correlational studies was linearity which referred that the relationship between variables should be linear (Pallant, 2007). In order to examine the linearity between variables scatterplot for ATSQ mean scores and SLT mean scores were constructed.

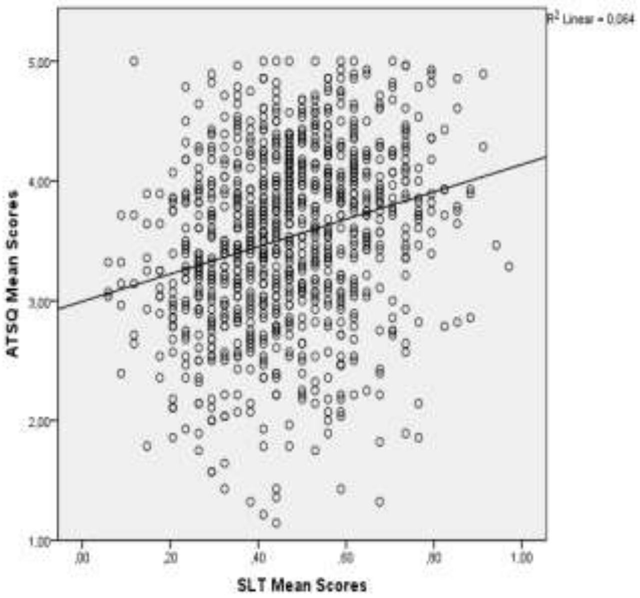


Figure 3 Scatterplot of ATSQ Mean Scores and SLT Mean Scores

As seen, there was a linear drawn in the scatterplot which indicated the linear relationship between these two variables. Hence, linearity assumption was ensured. The direction of this relationship was positive. Specifically, those who had higher mean scores on SLT had more positive attitude towards statistics. Regarding the slope of the line, it could be inferred that the relationship between ATSQ and SLT scores were almost moderate.

Homoscedasticity, which referred that variability of ATSQ mean scores should be similar to SLT mean scores, was another assumption of correlation analysis (Pallant, 2007). In order to examine homoscedasticity assumption, scatterplot in Figure 4.1 was checked. There was a fairly cigar shape in this figure which indicated that homoscedasticity assumption was met.

4.4.2. Pearson Product Moment Correlation Analysis Results

The relationship between attitudes towards statistics (as measured by the ATSQ) and statistical literacy (as measured by SLT) was investigated through the Pearson product-moment correlation coefficient. Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity, and homoscedasticity.

Results indicated that there was a significant positive relationship between the two variables, r = .25, p< .01. This meant that as higher level of 8th grade students' attitudes towards statistics associated with higher level of their statistical literacy scores. The strength of the relationship was considered as small according to Cohen's (1988) categorization. The coefficient of determination was calculated as .06 which indicated that there were 6 percent shared variance between ATSQ scores and SLT scores.

4.5. Summary of Results

The purpose of this study was to investigate the statistical literacy of 8th grade students with respect to statistics domains in the elementary mathematics curriculum and tiers, and to investigate the attitudes towards statistics and the relationship between statistical literacy and attitudes toward statistics.

Based on descriptive results, it could be inferred that students in this study performed lower than moderate in statistical literacy. However, their attitudes towards statistics were placed between neutral to agree in five point scale. In other words, 8th grade students in this study had positive attitudes towards statistics. The analysis of the mean scores of statistical literacy in terms of content domains revealed that there were variations between these content domains. Although sample, graphs, and chance content domains had closer mean scores to each other which was around moderate value, average, inference, and variation content domains had lower mean scores than other contents. Likewise, mean scores for each tier were different from each other. That is, although mean scores for Tier 1 and Tier 3 were closer to each other, Tier 2 had relatively higher mean score than others.

Furthermore, as for inferential statistics, there were significant differences between Tier 1, Tier 2, and Tier 3 aspects of statistical literacy with large effect size. The pair wise comparisons indicated that students performed lowest in third tier of statistical literacy where students were required to question or evaluate inappropriate statistical claims. In other words, there were significant differences between the mean scores of Tier 1 and Tier 2, Tier 1 and Tier 3, and Tier 2 and Tier 3. Although, students performed slightly higher in the first tier where they showed their ability in understanding statistical terminology; their performance was the highest in the second tier which was interpreting statistical claims in context.

The correlation analysis indicated that there were positive and significant relationship between students' attitudes towards statistics and statistical literacy scores. This meant that higher levels of 8^{th} grade students' attitudes towards statistics were associated with higher levels of their statistical literacy scores.

CHAPTER 5

DISCUSSION AND CONCLUSION

The purpose of this study was to investigate statistical literacy of 8th grade students with respect to content domains in the statistics in Elementary Mathematics and Watson's (1997) three tiers. In addition, this study aimed at investigating students' attitudes towards statistics and the relationship between students' statistical literacy and attitudes towards statistics. In this chapter, findings of this study is summarized and then discussed, implications for educational practices are addressed, and further research is recommended.

5.1. Discussion of Findings

In this study, students performed less than moderate in statistical literacy. In addition, there were significant mean score differences between Tier 1 and Tier 2, Tier 1 and Tier 3, Tier 2 and Tier 3 where Tier 2 had the highest mean score and tier 3 had the lowest mean score. Similarly, there were differences between content domains such as average and sample. In addition, 8th grade students in this study had positive attitudes towards statistics. The correlation coefficient between their attitudes towards statistics and statistical literacy was found as .25 which indicated that 8th grade students' higher level of attitudes towards statistics was associated with higher level of their statistical literacy scores. The details of these results and related discussion of these findings are presented in the following sections.

5.1.1. 8th Grade Students' Statistical Literacy

The results of this study clarified that 8th grade Turkish students' statistical literacy near the end of elementary school could be regarded as low. Although Elementary Mathematics Curriculum in Turkey indicated that developing informed citizens who possess knowledge of statistical concepts to deal with existing statistical resources in society is vital (MoNE, 2005), students who were taught by this curriculum performed poorly in statistical literacy. This result could be regarded as critical since statistical literacy does play an important role in society and building active and critical citizenship through informed contribution to the debates ranged from politics to environment. This study revealed that students' statistical literacy was not adequate when they were about to graduate from elementary schools. Since cognitive aspect of statistical literacy is composed of three tiers (Watson, 1997) and students' responses were analyzed through these tiers, statistical literacy performance of students is discussed accordingly.

5.1.2. 8th Grade Students' Statistical Literacy in Terms of Tiers

There were significant differences between Tier 1 and Tier 2, Tier 1 and Tier 3, and Tier 2 and Tier 3 aspects of statistical literacy with a large effect size. The pairwise comparisons indicated that students performed lowest in the third tier of statistical literacy where students were required to question or evaluate inappropriate statistical claims. Although students performed slightly higher in the first tier where they showed their ability in understanding statistical terminology, their performance was the highest in the second tier which was interpreting statistical claims in context. This result was consistent with previous research on statistical literacy which indicated that appropriate usage of statistical terminology and critical evaluation of statistical claims appeared in the same level in statistical literacy hierarchy (Watson & Callingham, 2003; Watson, 2006).

Relevant to the present study, Watson and Kelly (2008) investigated students' understanding and related definitions of the terminology of statistical literacy across grades which indicated that the majority of students performed either in the lowest level or one level higher which involved idiosyncratic responses and one single related idea respectively. At this point, Watson and Kelly's (2008) results supported the findings related to Tier 1 performance of students in this study where students mean score for the first tier is found to be low.

The low performance in the definition of statistical terminology, which is emphasized in Tier 1, could also be the consequence of the nature of statistical knowledge that mathematics teachers had. Teachers might not use appropriate statistical language including clear indication of these terms in the classroom while teaching statistics and probability. Indeed, a study conducted with Turkish pre-service mathematics teachers to examine their competencies in using appropriate mathematical language indicated that pre-service teachers' use of mathematical language were not adequate (Yeşildere, 2010). Miller (1993) also pointed that mathematics teachers avoided using correct terminology while teaching and preferred using everyday language instead (as cited in Watson & Kelly, 2008). Therefore, while interpreting these results, it is important to note the teachers' lack of correct usage of terminology in the instructional language might affect students' usage of statistical terminology.

Eighth grade students in this study have performed relatively higher in the second tier of statistical literacy compared to other tiers which was in line with the previous research conducted with statistical literacy (Watson & Callingham, 2003). Students' higher performance in the second tier of statistical literacy was an expected result due to the fact that majority of objectives in the Elementary Mathematics Curriculum in Turkey regarding statistical concepts were contextualized around the ability of application or interpretation of statistics which refers Tier 2 in this study. Considering that greater part of upper elementary mathematics instruction in Turkey is devoted for the application of mathematical content (Doğan, 2006), the results considering relatively higher performance on the Tier 2 of statistical literacy.

In this study, 8^{th} grade students performed poorly in the third tier questions where they were required to critically evaluate statistical claims. This result is consistent with previous research conducted with middle school students on statistical literacy where majority of students were placed between informal and consistent-noncritical levels while very small percentage of students were in the critical and critical-mathematical level (Watson & Callingham, 2003). The low performance in Tier 3 might be derived from existing Elementary Mathematics Curriculum in Turkey. As stated before, the objectives considering evaluation of statistical claims, which refers third tier, are limited in number in Elementary Mathematics Curriculum whereas objectives regarding understanding and interpreting statistics, which are first and second tier, respectively are relatively more. For example, there is only one objective targeting students' evaluation of statistical claims, messages, or representations in graphs concept. Therefore, the result of this study which indicated different performances in each tier and insufficient performance on questioning statistics might be a reflection of the Elementary Mathematics Curriculum in Turkey.

The inadequacy of evaluation of statistical claims documented in this study could also be attributed to the regular mathematics instruction in schools. The examination of existed pattern in mathematics classroom while teaching of a unit indicated that the sequence of instruction was consisted of introducing new content, practicing this content and assigning or doing homework (Doğan, 2006). Although there has been a revision of Elementary Mathematics Curriculum in Turkey, it is evident that statistics concepts are likely to be instructed in Turkish middle schools with rather traditional teaching methods despite the constructivist approach emphasized in the curriculum, which was observed in teaching of graphs concepts through the revised curriculum (Tortop, 2011). As seen, majority of time in mathematics classrooms in Turkey is devoted for application of the content in a traditional way; hence, there remains a small amount of time for reflection, discussion and evaluation of those contents including statistical messages which is a supportive idea for the results of this study.

Teachers' knowledge and experience of taking critical positions towards data may have an influence on students' poor performance in the third tier of statistical literacy (Chick & Pierce, 2011; Watson, 2006). Since mathematics teachers' knowledge for teaching statistics has a strong influence on students' achievement, the low performance on statistical literacy of 8th grade students might be derived from the knowledge required for teaching statistics concepts. The relatively lower performance on third tier statistical literacy might also be originated from teachers' affect including beliefs and attitudes regarding both statistics and statistical literacy. Since teachers' beliefs do have an influence on the practices of teachers during teaching statistics (Pierce & Chick, 2011) and their attitudes towards statistics play an important role on development of statistical outcomes (Estrada, Batanero & Lancaster, 2011), it is crucial to remark that teachers' affect might have an influence on students' ability of evaluating statistical descriptions appear in everyday life or media often involve bias or lack of objectivity.

The reason for relatively lower performance in the first and third tier of statistical literacy compared to the second tier might be originated from the item formats. Third tier requires students to explain the reasons of their answers; hence, these questions were in open-ended form while majority of questions in the second tier were in multiple choice formats. The outperformance of 8^{th} grade students in the second tier questions which were in multiple choice format and failure to provide explanations and justifications required in the first and third tier could be attributed to 8^{th} graders' familiarity with multiple choice items due to national examinations for placement in more competitive secondary schools.

5.1.2. 8^{th} Grade Students' Statistical Literacy in Terms of Content Domains

The results of this study were also examined through content domains of statistical literacy (which are sample, average, graphs, chance, inference and variation) which revealed that students performed differently in different contents. Therefore related discussion of results of each content domain is presented separately below.

5.1.2.1. Sample

Students performed better while implementing ideas of sample and evaluating statistical reports in sampling context. Although students have the ability to evaluate statistical claims related to sampling, they performed poorly in defining sample as a statistical terminology. This result is consistent with previous research done by Watson and Moritz (2000) where they indicated that students, who could not give related statistical ideas for definition of sample, were able to question claims in the sampling context. In addition, appropriate selection of sample is aimed in Turkish elementary mathematics curriculum. Hence, it could be inferred that the current curriculum contributes applying and evaluating ideas related to sampling more than it contributes to defining statistical terminology.

Almost one third of 8^{th} grade students referred sample as "example". This response might be due to the confusion due to the confusion of Turkish version of sample

(Turkish: örneklem) and example (Turkish: örnek). A similar situation was also observed in studies conducted in English language where students referred "sample" as "simple" (Watson & Kelly, 2008). Since Turkish version of sample and example resemble each other students might have responded as such. Durkin and Shire (1991) have declared that students who did not know the traditional meaning of mathematical words used an unfamiliar one which had a similar sound (as cited in Watson & Kelly, 2008).

5.1.2.2. Average

Understanding average as "add them up and divide" algorithm was the most frequent response provided by students which is a consistent result with the previous research conducted with Turkish students where students have understood average as the arithmetic mean (Toluk-Uçar & Akdoğan, 2009). Also, most of them did not take median and mode into account as other ways of finding average of a given data set. In addition, only less than half of the participants were able to interpret average in context, which might be derived from students' procedural understanding of average concept. Similarly, their performance in evaluation of a statistical claim involving average as a representative value were poor as they could not recognize extreme values or explained this claim by providing evidence through arithmetic mean. In a relevant study which examined students' conceptions of average, it was found that students did not consider average as a representative value for the given data set (Mokros & Russell, 1995), which is similar to the findings for this study.

The lack of understanding average as a summarizing or representative value for students in this study might be related to the elementary mathematics curriculum. In Turkish curriculum average concept is represented through measures of central tendency which are mean, median and mode. Therefore, students might conceptualize average concepts through mean. Additionally, although average concept is instructed each year in line with the spiral nature of curriculum, students begin to learn average through arithmetic mean which may result in understanding average as "add them up and divide" algorithm. In addition, while teaching average concept, teachers may not focus on its characteristics of representative value of a data set; instead they may devote majority of instructional time for computational skills.

5.1.2.3. Graphs

Another content domain of statistical literacy was graphs concept. The analysis of items related to graphs concept indicated that majority of participants performed better while interpreting graphical representations compared to their ability to evaluate misleading bar graphs in which only half of the students correctly explained their evaluations of statistical report. Aoyama and Stephens (2003) indicated that students did not have sufficient knowledge and experience for evaluating graphs whereas they correctly could read graphs which was not considered as critical. As seen, students performed relatively higher in the graphs concept than other content domains which might be derived from Turkish elementary mathematics curriculum where the graph concept appears in the curriculum from pre-school to the 8th grade (MoNE, 2005). Therefore, students' capability of reading graphs was an expected result. However, although there is an objective considering critical evaluation of misleading graphs, half of the students failed to critically evaluate bar graphs in Statistical Literacy Test and they performed poorly in choosing appropriate graphical representations which was demanded by the first tier of statistical literacy. The reason for this result could be explained by findings in the Turkish context which indicated that teachers did not cover all of the objectives regarding graph concepts and caused errors and misconceptions about graphs (Tortop, 2011).

5.1.2.4. Chance

Majority of participants performed well in items related to chance content; though they failed to ask questions for a statistical report in chance context. Chance or probability has been one of the oldest topics in Elementary Mathematics Curriculum and accordingly it is expected that teachers have the required knowledge and experience with understanding and application of these concepts. In addition, there are objectives regarding understanding randomness and interpreting chance in context in the national curriculum. However, since objectives regarding the critical evaluation of statistical claims in chance context do not exist in the curriculum and teachers may not allocate time for evaluation of chance related claims during instruction, students might perform poorly in third tier of chance content.

5.1.2.5. Inference

The different performances of 8th grade students were also observed while they were doing inference based on statistics. Majority of students were able to make predictions based on data and explain their predictions, whereas most of them had failed to evaluate critically an inference without appropriate statistical foundation. This result is closely connected to the Elementary Mathematics Curriculum where there is an objective considering predictions based on data set while critical evaluation of inferences are not placed in the curriculum like in other content domains.

5.1.2.6. Variation

Similar to the other content domains, the performances of students in the second tier of variation are relatively higher than the first and third tier of statistical literacy, which could be attributed to objectives in the curriculum and statistics instruction in schools. Almost one third of the 8th grade students explained the meaning of variation through the measure of spread range, which was the easiest to calculate. Similar to the explanations of average, these responses might be originated from the procedural understanding of statistics and particularly the variation concept. Likewise, in Turkish curriculum variation concept is represented through measures of spread which are standard deviation, range and interquartile range. Therefore, students might conceptualize variation concepts through range. Although majority of participants interpreted variation in context, their responses to items contextualized in the first and third tier addressed more variation where the data set consisted of the same numbers. This response might be regarded as a sign of possible misconception about variation concept of 8th grade students.

5.1.2. 8th Grade Students' Attitudes towards Statistics

One of the purposes of this study was to investigate 8th grade students' attitudes towards statistics. Accordingly, their attitudes towards statistics were measured through Attitude towards Statistics Questionnaire (ATSQ) and the descriptive analysis of this questionnaire indicated that 8th grade students had slightly positive attitudes towards statistics while the mean value was close to neutral. Several studies have documented that the attitudes towards statistics ranged between neutral to positive in the context of pre-college students (e.g. Calderia & Mourino, 2010; Yingkang and Yoong, 2007) which is consistent with the result of this study. The positive attitudes of 8th grade students in this study might be connected to curriculum revision conducted in 2005. The revised curriculum, in the context of statistics and probability, aims at developing positive orientations towards statistics and probability so that students would understand the importance of statistics. Therefore, the slightly high attitudes of 8th grade students towards statistics could be the result of the emphasis on such attitudes. In addition to this, it is documented that statistics is a methodological discipline distinct than mathematics (delMas, 2004). Moreover, statistical subjects consist of applied topics rather than abstract concepts (Calderia & Mourino, 2010). Therefore, students' attitudes did not position through the negative responses. In other words, slightly positive attitudes of 8th grade students might be derived from distinct nature of statistics. However, it is important to note that students' attitudes towards statistics are not very high; instead their attitudes are placed closed to neutral degree. Due to the fact that previous research revealed that positive attitudes towards statistics are closely related to constructivist learning environment during statistics instruction (Cobb & Hodge, 2002; Mvududu, 2003), it could be inferred that statistics topics might still be instructed in Turkish middle schools with rather traditional teaching methods based on computation and procedural skills, which was documented for graphs concepts in the 7th grade mathematics classrooms (Tortop, 2011). Therefore, despite teachers might indicate the importance of statistics during regular mathematics instruction as suggested by national curriculum, students' attitudes stayed close to neutral due to the rather traditional nature of instruction in their mathematics classrooms.

5.1.3. The Relationship between 8th Grade Students' Statistical Literacy and Attitudes towards Statistics

The results considering the relationship between 8th grade students' statistical literacy and their attitudes towards statistics revealed that there was a significant positive relationship between 8th grade students' statistical literacy and their attitudes towards statistics. Several perspectives could be found in the literature indicating that dispositional aspects of statistics instruction, such as attitudes and beliefs or task motivation, do play an important role in statistical literacy and these perspectives included dispositions into their statistical literacy models or frameworks (Gal 2004; Watson, 2006).

The two aspects of statistical literacy models and frameworks for statistical literacy was examined in this study. According to the results, in Turkish context, it could be said that statistical literacy is composed of both cognitive aspects which are represented in Tier 1, Tier 2, and Tier 3 and affective aspects which are attitudes towards statistics. Also, the positive relationship between these two aspects of statistical literacy has confirmed this combined framework.

Several studies have investigated the relationship between the attitudes towards statistics and statistical outcomes in the context of university students and they indicated a positive relationship (Chiesi & Primi, 2010; Dempster, 2009; Diri, 2007; Emmioğlu, 2011; Nasser, 2004; Vanhoof, 2006) which is in line with the current study conducted with 8th grade students. The findings of this study revealed that students with relatively higher attitudes towards statistics tended to perform higher on statistical literacy test. Nonetheless, it is essential to emphasize that although there is a significant positive relationship between 8th grade students' attitudes towards statistics and their statistical literacy, the relationship is not strong. This result might be related to the nature of statistical knowledge needed for statistical literacy. To be more precise, due to the fact that statistical literacy is regarded as a bridge between everyday life and statistical concepts (Watson, 2006), students might use informal statistical knowledge for their performance on statistical literacy depending on context. In addition, another study with undergraduate students revealed that weak or insignificant relationships between dimensions of attitudes towards statistics and statistical reasoning are derived from the nature of knowledge for statistical reasoning which was naïve knowledge when formal statistical knowledge had been forgotten by students (Tempelaar, Loeff, & Gijsealers, 2007). Furthermore, the items in the attitude questionnaire are stated through the word “statistics”. These items might be confused students since they did not have a full understanding of what statistics was. Indeed, statistics in school mathematics might be instructed through statistical concepts such as sample or average without indicating these topics are within the scope of statistics and what statistics referred to as a concept. Therefore, in the present study, 8th grade students’ attitudes towards statistics built in mathematics classroom do play small but still significant role on their statistical literacy since students’ performance of statistical literacy is both related to their informal knowledge of statistics and their image regarding statistics in their minds.

5.2. Implications for Educational Practices

In this section, some implications for mathematics teachers, mathematics curriculum developers and mathematics teacher educators are presented in relation to findings and discussion of findings of this study.

The results of this study regarding three tiers of statistical literacy could be attributed to statistics instruction in schools. It seems that the majority of class time was devoted for the application of statistical ideas while there remains small amount of time for opportunity to develop conceptual understanding and critical evaluation. Instead, a teaching approach including project work together with real life data and media reports could be employed in statistics concepts in order to increase statistical literacy of elementary students (Merriman, 2006). Teachers might incorporate daily news including statistical reports appearing in the media to the statistics lessons. In addition, there needs to be integration of more contexts into statistics teaching since statistical literacy basically deals with data in context and plays role of a bridge between statistics and everyday life. Technological tools should be integrated into statistics instruction in schools so that students might handle the procedural aspects of application of statistical ideas easily and devote their time for conceptual understanding. The results of this study indicated that students could not evaluate critically and could not question statistical claims. A relevant study where critical thinking skills were emphasized during statistics instruction indicated that students' statistical conception enhanced (Doyle, 2008). Therefore, similar critical thinking approach might be employed in classrooms. Besides, teachers should be aware that statistical claims appeared in media may be one-sided, biased, or misleading. This addresses that teachers should use the kind of pedagogy where students reflect, discuss, and evaluate statistics rather than accepting without questioning.

The findings of this study revealed that 8th grade students had performed moderately low in statistical literacy. Since statistical literacy is an important feature for building active and critical citizens, elementary mathematics curriculum should aim at developing statistical literacy within statistics and probability content area in each grade level. Furthermore, objectives might be modified in relation to support for statistical literacy. There was only one objective regarding evaluation statistical messages in the context of graph content domain. Therefore, curriculum makers or planners should identify and include objectives regarding critical evaluation and questioning of statistical claims to promote the development of statistical literacy within elementary school students. For instance, evaluation of arithmetic mean as a representative value or evaluation of a given sample in terms of generalization to a particular population should take place as objectives in the curriculum so that there would be the possibility of instruction those objectives.

The inadequate performance considering critical evaluation of statistical claims might be derived from teachers' knowledge, affect and practices during statistical instruction. Since most of the topics were recently added to the elementary mathematics curriculum such as variability or central tendency measures in addition to the concept of arithmetic mean (MoNE, 2005), in-service teachers may not have sufficient knowledge, background, and experience regarding teaching of those concepts. Therefore, similar to the enhancement of understanding of statistical terminology through in-service trainings, critical evaluation and questioning statistical claims should be another focus of those trainings so that teachers would become capable of planning instructional time in accordance with reflection, discussion, and evaluation of statistical messages. Since these results could be connected to teachers' affect including their attitudes and beliefs, another focus of in-service teacher education in the context of statistics should be building positive dispositions towards statistics and probability. In addition, most of the statistics related concepts have recently entered to the elementary mathematics curriculum in 2005; therefore, in-service training of teachers regarding adequate usage of these statistical concepts should be increased.

The implications regarding in-service teacher education could be extended to pre-service teacher education. The reformed pre-service teacher education offered by Higher Education Council (HEC) includes two obligatory courses regarding statistics and probability. Although these courses might have a positive influence on the development of required content knowledge, the overall program lacks specific courses on teaching statistics and probability. Since statistics is regarded as a methodological discipline distinct from mathematics (delMas, 2004), teacher education programs should include courses related to specific teaching methods of statistics. The inclusion of such kind of courses would not only develop statistics instruction in terms of teachers' content and pedagogical content knowledge for teaching statistics, but also might have an influence on pre-service teachers' attitudes and beliefs about statistics and statistics teaching and learning. In other words, both in-service and pre-service teacher education should focus on development of statistical literacy and learning to teach statistical literacy within elementary schools.

The most notable finding regarding statistical literacy of 8th grade students in terms of content domains was explanations of average and variation contents through measures of central tendency and spread. This indicated that students' conceptions were mainly procedural in these domains. In Elementary Mathematics Curriculum, the meaning of these concepts should be clearly presented in addition to measures of central tendency and spread. Likewise, while teaching these subjects, the meaning of these subjects should be emphasized at first before introducing measurements. After students understand the meaning of statistical meaning of these concepts and reasoning behind them, the measures such as mean or standard deviation should be instructed. In addition, incorporating more contextual examples would probably provide students with a clear understanding of the meanings of these terms.

In this study, results considering 8th grade students' attitudes towards statistics revealed that their attitudes ranged between neutral to positive. Since constructivist learning environments do play role on building positive orientations towards statistics (Cobb & Hodge, 2002), the classroom activities, regarding statistics concepts, should be arranged accordingly. In other words, classroom activities should not merely focus on procedural skills based on memorization; rather the focus of those activities should make students as doers of statistics. In addition, since technological tools such as statistics software (e.g. Doğan, 2009) or video recordings (e.g. Allredge, Johnson & Sanchez, 2006) had positive impact on attitudes towards statistics in the context of university students, these technological tools might be utilized in the instruction of statistics in middle school as well.

Another result of this study was the positive relationship between attitudes towards statistics and statistical literacy which meant higher attitudes towards statistics implied higher statistical literacy scores for 8th grade students. In terms of educational implications, teachers should be aware this issue and focus also on development of positive attitudes during statistics instruction so that students' statistical literacy would increase. Furthermore, the strength of this relationship was categorized as small which could be explained through the needed for statistical literacy. Since statistical literacy is a kind of bridge between everyday life and statistics, teachers should incorporate daily life issues to the statistics activities. Therefore, the gap between everyday and formal knowledge becomes small and students develop positive attitudes towards statistics. The same implications regarding statistics activities for teachers might be extended to curriculum developers and textbook writers. They should also include daily life issues to their proposed activities so that students might both develop positive attitudes towards statistics and increase their performance in statistical literacy.

5.3. Implications for Further Research

In this study, statistical literacy, attitudes towards statistics, and relationship between statistical literacy and attitudes towards statistics of 8th grade students were examined. The generalization of results were limited with accessible population, therefore, the same research might be replicated nationwide with broadened sample which is a representative of all Turkish 8th grade students. In addition, cross sectional surveys can be done where these constructs are examined with respect to grade level and gender with minor modifications in the instruments so that how statistical literacy and attitudes towards statistics alter with respect to those variables become clearer. The changes in students' statistical literacy and attitudes towards statistics might also be investigated in a longitudinal study since same students may give a better idea about the changes in conceptions of statistical literacy and attitudes towards statistics. Findings of such longitudinal studies provide detailed information in relation to objectives in the elementary mathematics curriculum.

The results of this study considering the tiers of statistical literacy were attributed to the regular mathematics instruction. However, the analysis of statistics instruction and teacher practices while teaching statistics in each grade level would also be beneficial since there might be differences in the instruction of different contents in elementary mathematics in different grades. In other words, the instructional approaches employed for teaching average concept may differ from the instructional approaches for teaching sample. Therefore, specific examination of instruction regarding these specific contents would give wealthier information regarding the implementation of intended elementary mathematics curriculum.

Though this research provided an examination of each content domain of statistical literacy, it is still limited with the instrument used. More research is necessary to be conducted on these specific contents such as sample or variation so that detailed examination of them would be possible. For example, typical errors and misconceptions might be investigated and statistics instruction in elementary schools and activities in curricula and textbooks could be modified according to the results.

Several intervention studies might be conducted to provide a cause-effect relationship with statistical literacy. The research considering effect of technological tools on both statistical literacy and attitudes towards statistics was scarce in the middle school context. The findings of experimental studies in the context of tertiary students might not always be informative for the middle school context. Therefore, investigating the effect of technological tools such as statistics software or calculators on statistical literacy and on attitudes towards statistics could provide a substantial contribution to the field of statistics education in the middle school context. Similarly, examining the effects of several teaching approaches on these constructs will provide wealthier information regarding how to develop statistical literacy and positive orientations toward statistics.

Future research seems to require an examination of student and teacher related variables in the context of statistical literacy through statistical modeling approaches. Therefore, it might be possible to understand which teacher and student related variables and to what extent these variables contribute to the statistical literacy of 8^{th} grade students.

In this study, the results revealed that there was a significant positive relationship between 8^{th} grade students' statistical literacy and attitudes towards statistics; however, this relationship was not so strong. Since the attitude questionnaire used in this study described general attitudes towards statistics, affect dimension is limited to the attitudes represented with this instrument. Therefore, further research might be conducted with other affect constructs such as beliefs, self-efficacy, or anxiety in the context of statistics.

In addition to this, the affective domain that is related to statistical literacy is limited with this questionnaire. The reason for the small size of relationship between attitudes towards statistics and statistical literacy was explained that students might not have an idea of what statistics is in their minds. As a suggestion, more attitude instruments should be developed that have specific contextualization. For example, items might be “I like analyzing data”, “I enjoy doing surveys” or “I hate calculating mean”. Hence, the relationship between statistical literacy and affective domain would become clearer.

The instruments used for this study is specifically developed for this study for 8^{th} grade students who are near the end of their elementary education. Although there was not a problem in the understandability of items in both of the instruments in the pilot and actual study, these instruments still needs to be revised in terms of the wordings of items in the middle school context for further usage. Students’ understandings of the terms in the items were unknown in the study context. Additionally, the same instruments might be utilized for university students from non-quantitative majors or adults to examine their statistical literacy which is required for active and critical citizenship. The possible results of such kind of research conducted might have an implication for undergraduate programs or community centers for adult education.

Lastly, further qualitative research might be conducted for an in-depth examination of both statistical literacy and attitudes towards statistics through observations during statistics instruction and interviews. Therefore, it would be possible to understand how students evaluate statistical claims in a critical way and how their responses differ in relation to their attitudes towards statistics. Additionally, the two approaches including quantitative and qualitative methods might be utilized as a research design for the purpose of validation of students’ responses.

REFERENCES

Aksu, M. &Bikos, L. (2002). Measuring and predicting graduate students' attitudes toward statistics. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 23, 22-31.

Alajaaski, J. (2006). How does web technology affect students' attitudes towards the discipline and study of mathematics/statistics? *International Journal of Mathematical Education in Science and Technology*, *37*(1), 71-79.

Allredge, J. R., Johnson, H. D., & Sanchez, J. J. (2006). Does viewing video of statistics in action affect student attitudes? In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/17/C335.pdf

Aoyama, K., & Stephens, M. (2003). Graph Interpretation Aspect of Statistical Literacy: A Japanese Perspective. *Mathematics Education Reseacrh Journal* , *15* (3), 207-225.

Ben-Zvi, D., & Garfield, J. (2004). Statistical Literacy, Reasoning and Thinking: Goals, Definitions and Challenges. in D. Ben-Zvi, & J. Garfield, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 3-15). Dordrecht: Kluwer Academic Publishers.

Biajone, 2006 Biajone, J. (2006). Promoting positive attitudes towards statistics in pedagogy students through project work. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/17/C408.pdf

Bingham, A.D. (2010). Student attitudes to real-world projects in an introductory statistics course. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International*

Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_C203_BINGHAM.pdf

Budgett, S. &Pfannkuch, M. (2010). Using media reports to promote statistical literacy for non quantitative majors. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*.Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_7G1_BUDGETT.pdf

Bulut, S. (1994).*The effects of different teaching methods and gender on probability achievement and attitudes toward probability.* Unpublished Doctoral Dissertation, Middle East Technical University, Ankara.

Büyüköztürk, Ş. (2011).*Sosyal Bilimler için Veri Analizi El Kitabı, İstatistik, Araştırma Deseni, SPSS Uygulamaları*. Ankara: Pegem Akademi Yayıncılık

Calderia, S. &Mouriño, H. (2010). Students' opinion on the subjects of statistics and probability in secondary schools of lisbon, Portugal. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_C116_CALDEIRA.pdf

Carlson, K. A. & Winquist, J.R. (2011).Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach. *Journal of Statistics Education, 19* (1). Retrieved from www.amstat.org/publications/jse/v19n1/carlson.pdf

Carnell, 2008 Carnell, L.J. (2008). The effect of a student-designed data collection project on attitudes toward statistics. *Journal of Statistics Education*, *16*(1). Retrieved from www.amstat.org/publications/jse/v16n1/carnell.html

Carmichael, C. S. (2010). *The Development of Middle School Children's Interest in Statistical Literacy.* Unpublished Doctoral Dissertation, University of Tasmania, Tasmania.

Chance, B. (2002).Components of statistical thinking and implications for instructionand assessment. *Journal of Statistics Education, 10* (3).Retrieved from http://www.amstat.org/publications/jse/

Chick, H. L. & Pierce, R. (2011). Teaching for statistical literacy: Utilising affordances in real-world data. *International Journal of Science and Mathematics Education*, *10*(2), 339-362.

Chiesi, F. &Primi, C.(2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal, 9*(1), 6-26.

Cobb, P. (1999). Individual and collective mathematical development: Thee case of statistical data analysis. *Mathematical Thinking and Learning, 1*(1), 5-43.

Cobb, P., & Hodge, L. (2002).Learning, identity, and statistical data analysis. In B. Philiphs (Ed.), *Developing a statistically literate society: Proceedings of the Sixth International Conference on Teaching Statistics, Cape Town, South Africa*. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/1/2e1_cobb.pdf

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Crocker, L., &Algina, J. (1986).*Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.

delMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79-95). Dordrecht: Kluwer Academic Publishers.

Dempster, M. &McCorry, N.K. (2009).The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students.*Journal of Statistics Education*, 17(2). Retrieved from www.amstat.org/publications/jse/v17n2/dempsterpdf.html

Diri, F. Ü. (2007). *İstatistik dersine yönelik tutumların araştırılması meslek yüksek okul örneği.* Gazi University, Ankara.

Doğan, N. (2009). Bilgisayar destekli istatistik öğretiminin başarıya ve istatistiğe karşı tutuma etkisi. *Eğitim ve Bilim*, *154* (34), 3-15.

Doğan, O. (2006). *A study on pattern of 6th grade elementary mathematics lesson.*Unpublished Masters Theses, Middle East Technical University.Ankara.

Doyle, P. (2008). *Developing statistical literacy with students and teachers in the secondary mathematics classroom*. Unpublished Masters Theses, The University of Waikato , Waikato.

Emmioğlu, E. & ÇapaAydın, Y. (2011).A Meta-Analysis on Students' Attitudes toward Statistics. *Paper presented 58th World Statistics Congress of the International Statistical Institute,* Dublin.

Estrada, A., Batanero , C. & Lancaster, S. (2011). Teachers’ attitudes towards statistics. In C. Batanero, G. Burril & C. Reading (Eds.), *Teaching statistics in school mathematics challenges for teaching and teacher education*. (pp. 163-174). New York: Springer Publishers.

Evans, B. (2007). Student Attitudes, Conceptions, and Achievement in Introductory Undergraduate College Statistics. *The Mathematics Educator*. *17* (2), 24-30.

Fraenkel, J., & Wallen, N. (2006). *How to Design and Evaluate Research in*

Education. New York: McGraw-Hill.

François , K. & van Bendegem, J.P. 2010. Ethical-political aspects of statistical literacy. . In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_C258_FRANCOIS.pdf

Frankenstein, M. (1998). Reading the world with maths: Goals for a critical mathematical literacy curriculum. In *Proceeding of the 1st International Mathematics Education and Society Conference*. Nottingham, UK. Retrieved from http://www.nottingham.ac.uk/csme/meas/papers/frankenstein.html

Gal, I. (2004). Statistical Literacy: Meanings, Components, Responsibilities. In D. Ben-Zvi, & J. Garfield, *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Dordrecht: Kluwer Academic Publishers.

Gal, I. & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: towards an assessment framework, *Journal of Statistics Education. 2*(2). Retrieved from http://www.amstat.org/publications/jse/v2n2/gal.html

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring Attitudes and Beliefs in Statistics Education. In I. Gal, & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37-51). Amsterdam: IOS Press.

Garfield, J. B. (2003). Assessing statistical reasoning.*Statistics Education Research Journal*, *2*(1), 22-38.

Kunnan, A. J. (1998). An introduction to structural equation modeling for language assessment research. *Language Testing, 15*(3), 295-332.

Langrall, C.W. (2010). Does context expertise make a difference when dealing with data? In C. Reading (Ed.), *Data and context in statistics education: Towards* *an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_2A2_LANGRALL.pdf

Leong, J. (2007). *High school students' attitudes and beliefs regarding statistics in a service-learning-based statistics course.* Unpublished doctoral dissertation, Georgia State University, Atlanta.

Mokros, J. & Russell, S. (1995). Children's Concepts of Average and Representativeness, *Journal for Research in Mathematics Education*, *26*, 20-39.

Martinez-Dowson, R. (2010). *The effects of a course on statistical literacy upon students' challenges to statistical claims made in the media*. Unpublished Doctoral Dissertation, Clemson University, Clemson.

McAleyev, L. & Sullivan, C. (2010). Statistical literacy and sample survey results. *International Journal of Mathematical Education in Science and Technology, 41*(7), 911-920.

McLeod, D. B. (1992), Research on Affect in Mathematics Education: A Reconceptualization, in *Handbook of research on mathematics teaching and learning*, ed. D. A. Grouws, NY: Macmillan, pp. 575-596.

Merriman, L. (2006). Using media reports to develop statistical literacy in year 10 students. *Proceedings of the seventh international conference on teaching statistics*. Voorburg, The Netherlands: International Statistical Institute. Retrieved from www.stat.auckland.ac.nz/~iase/publications/17/8A3_MERR.pdf

Meyer, O., &Thille, C. (2006).Developing statistical literacy across social, economic and geographical barriers using a "stand-alone" online course. *Proceedings of the seventh international conference on teaching statistics*. Voorburg, The

Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/17/6C3_MEYE.pdf

Ministry of National Education [MoNE]. (2005), *İlköğretim Okulu Ders Programları: Matematik Programı 6-7-8 (Elementary Curricula Programs: Mathematics Curricula Program for Middle Grades)*. Ankara: MEB.

Mocko M. & Jacobe, T.(2010**), T**he effects of an audience response system (ars) on achievement and attitudes towards statistics in an introductory statistics class. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute.

Mvududu, N. (2003). A cross-cultural study of the connection between students' attitudes towards statistics and the use of constructivist strategies in the course. *Journal of Statistics Education 11*(3). Retrieved from http://www.amstat.org/publications/jse/v11n3/mvududu.html

Nasser, F. (2004).Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking pre-service teachers in introductory statistics. *Journal of Statistics Education*, *12*(1). Retrieved from www.amstat.org/publications/jse/v12n1/nasser.html

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics* .Reston, VA

Olani, A., Hoekstra, R., Harskamp, E., van der Werf, G. (2010). Fostering Students' statistical reasoning, self-efficacy, and attitudes: Findings from a comprehensively reformed undergraduate statistics course. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_C142_OLANI.pdf

Pallant, J. (2007). *SPSS Survival Manual.* New York: Open University Press.

Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: a classroom study. *Mathematical Thinking and Learning, 13*(1-2), 27-46.

Pierce, R. & Chick, H. (2011).Teachers' beliefs about statistics. In C. Batanero, G. Burril & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education*. (pp. 151-162). New York: Springer Publishers.

Posner, M. (2011).The impact of a proficiency-based assessment and reassessment of learning outcomes system on student achievement and attitudes. *Statistics Education Research Journal, 10*(1), 3-14.

Ragasa, C. Y., (2008). A comparison of computer-assisted instruction and the traditional method of teaching basic statistics. *Journal of Statistics Education 16*(1) Retrieved from http://www.amstat.org/publications/jse/v16n1/ragasa.pdf

Rumsey, D.J. (2002) Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10*(3) Retrieved from www.amstat.org/publications/jse/v10n3/rumsey2.html

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp 957–1009). Reston, VA: The National Council of Teachers of Mathematics.

Schau, C. (2003).Students' attitudes: the "other" important outcome in statistics education. *Paper presented at the Joint Statistical Meetings - Section on Statistical Education,* San Francisco.

Schau, C., Stevens, J., Dauphinee, T.L., & Ann, D.V. (1995). The development and validation of the survey of attitudes towards statistics. *Educational and Psychological Measurement, 55*(5), 868-875

Schield, M. (1999). Statistical literacy: Thinking critically about statistics. *Of Significance Journal. The Association of Public Data Users, 1*(1). Retrieved from , http://www.statlit.org/pdf/1999SchieldAPDU.pdf

Schield, M. (2004).Statistical literacy curriculum design.In Burrill, G., & Camden, M. (Eds.). (2005). *Curricular Development in Statistics Education: International Association for Statistical Education 2004 Roundtable.* Voorburg, the Netherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/rt04/2.4_Schield.pdf

Schield, M. (2006). Statistical literacy survey results: Reading graphs and tables of rates andpercentages. *Conference of the International Association for Social Science Information Service and Technology,* Retrieved from, www.StatLit.org/pdf/2006SchieldIASSIST.pdf

Shrout, P. E., & Fleiss, J. L. (1979).Intraclass Correlation: Uses in Assessing Rater Reliability. *Psychological Bulletin*, *86*, 420-428.

Tempelaar,D., Loeff, S., &Gijselaers, W. (2007). A structural equation model analyzing the relationship of students' attitudes toward statistics, prior reasoning abilities and course performance. *Statistics Research Education Journal.6*(2), 78-102.

Tabachnick, B. G., &Fidell, L. S. (2007).*Using Multivariate Statistics*. Boston, MA: Pearson Education Inc.

Uçar, Z. &Akdoğan, E.N. (2009). 6.-8. Sınıf öğrencilerinin ortalama kavramına yüklediği anlamlar, *İlköğretim-Online, 8*(2), 391-400.

Tortop, T. (2011).*7th grade students' typical errors and possible misconceptions in the graphs concept before and after the regular mathematics instruction*. Unpublished Masters Theses, Middle East Technical University, Ankara.

Vanhoof, S., Sotos, A. E. C., Onghena, P., Verschaffel, L. (2006).Attitudes toward statistics and their relationship with short- and long-term exam results. In A. Rossman& B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics,* Salvador, Brazil. Voorburg, TheNetherlands: International Statistical Institute. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/17/C133.pdf

Wade, B. &Goodfellow, M. (2009). Confronting statistical literacy in the undergraduate social science curriculum.*Sociological Viewpoints, 25*, 75-90.

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 Media. In I. Gal, & J. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (s. 107-122). Amsterdam: IOS Press.

Watson, J., & Moritz, J. (2000). Development of Understanding of Sampling for Statistical Literacy. *Journal of Mathematical Behavior, 19*, 109-136.

Watson, J., & Callingham, R. (2003). Statistical Literacy: A Complex Hierarchical Construct. *Statistics Education Research Journal, 2* (2), 3-46.

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). Sweeden. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/rt04/4.1_Watson&Callingham.pdf

Watson, J. (2006) *Statistical Literacy at School: Growth and Goals*, Mahwah, NJ: Lawrence Erlbaum Associates.

Watson, J. & Kelly, B. A. (2008).The Vocabulary of Statistical Literacy.

International Journal of Science and Mathematics Education, *6*, 741-767

Wild, C. J. & Pfannkuch, M. (1999).Statistical thinking in empirical enquiry (with discussion).*International Statistical Review*, *67*(3), 223–265.

Wilson, B. L. (1994). *The development and evaluation of an instructional program in statistical literacy for use in post-secondary education*, Unpublished Doctoral Dissertation, Illinois State University, Illinois.

Yeşildere, S. (2010). İlköğretim matematik öğretmen adaylarının matematiksel alan dilini kullanma yeterlikleri. *Boğaziçi Üniversitesi Eğitim Dergisi*, *24*(2), 61-70.

Yılmaz, S. (2006). *The effects of real data based and calculator supported statistics activities on 7th grade of students' statistics performance and attitude toward statistics.* Unpublished Master Thesis, Middle East Technical University, Ankara.

Ying Kang, W., & KhoonYoong, W. (2007).Exploring attitude toward statistical graphs among Singapore secondary school students. *The Mathematics Educator*, *10*(1), 39-58.

APPENDICES

APPENDIX A

PERMISSION OBTAINED FROM MINISTRY OF EDUCATION

T.C.
ANKARA VALİLİĞİ
Milli Eğitim Müdürlüğü

BÖLÜM : İstatistik Bölümü
SAYI : B.08.4.MEM.0.06.20.01-60599/3522
KONU : Araştırma İzni
Ayşe YOLCU

13/01/2012

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
(Öğrenci İşleri Daire Başkanlığı)

İlgi : a) MEB Bağlı Okul ve Kurumlarda Yapılacak Araştırma ve Araştırma Desteğine Yönelik İzin ve Uygulama Yönergesi.
b) Üniversiteniz 27/12/2011 tarih ve 15886 sayılı yazısı.

Üniversiteniz İlköğretim Fen ve Matematik Eğitimi Anabilim Dalı Yüksek Lisans öğrencisi Ayşe YOLCU'nun **"İlköğretim 8. sınıf öğrencilerinin istatistiksel okuryazarlıklarının ve istatistiğe yönelik tutumları ve ilişkisinin araştırılması"** konulu tezi ile ilgili çalışma yapma isteği Müdürlüğümüzce uygun görülmüş ve araştırmanın yapılacağı İlçe Milli Eğitim Müdürlüğüne bilgi verilmiştir.

Mühürlü anket örnekleri (9 sayfadan oluşan) ekte gönderilmiş olup, uygulama yapılacak sayıda çoğaltılması ve çalışmanın bitiminde iki örneğinin (CD/disket) Müdürlüğümüz İstatistik Bölümüne gönderilmesini rica ederim.

Gülçin UYSAL
Müdür a.
Müdür Yardımcısı

EKLER : Anket (9 Sayfa)

20.01.12*001377

İl Milli Eğitim Müdürlüğü-Beşevler
İstatistik Bölümü
Bilgi İçin: Nermin ÇELENK

Tel : 223 75 22
Fax: 223 75 22
istatistik06@meb.gov.tr

APPENDIX B

STATISTICAL LITERACY TEST

İstatistiksel Okuryazarlık Testi

Bu test istatistik ve olasılık konuları ile ilgili 17 sorudan oluşmaktadır. Bazı sorular bir ya da birkaç alt soru içermektedir. Bazılarında ise açıklama yapmanız istenmektedir. **Lütfen tüm soruları cevaplamaya çalışınız.**

1. “Örneklemi Ankara’da çalışan matematik öğretmenleri olan bir araştırma yürütülmektedir.”
Yukarıdaki cümlede geçen “örneklem” kelimesinden ne anlıyorsunuz?

“Geçtiğimiz bir yıl boyunca her ay ortalama 20 kişi trafik kazasında hayatını kaybetti.”
2. a) Yukarıdaki gazete haberinde yer alan “ortalama” kelimesi sizce ne ifade etmektedir?

b) Bir yıl boyunca her ay trafik kazalarında ortalama kaç kişinin hayatını kaybettiği hangi yöntemle <u>bulunamaz</u>?

a) Bir yıl boyunca her ay hayatını kaybeden kişi sayısını toplayıp 12 ye bölmek
b) Bir yıl boyunca her ay hayatını kaybeden kişilerin sayısını küçükten büyüğe doğru sıralayıp ortadaki sayıyı seçmek
c) Bir yıl boyunca her ay hayatını kaybeden kişilerin sayılarında en çok tekrar eden sayıyı bulmak
d) **Bir yıl boyunca her ay hayatını kaybeden kişilerin sayılarından en büyüğünden en küçüğünü çıkararak**

3) Aşağıdaki veriler 5 TL'lik bir öğle yemeğinin içinde neler olduğunu ve fiyatlarını göstermektedir.

- 2 TL ana yemek
- 0.5 TL çorba
- 1.5 TL tatlı
- 1 TL salata

Buna göre veriler aşağıdaki grafik türlerinden hangisiyle en uygun biçimde temsil edilir?

a) Daire Grafiği b) Histogram c) Çizgi Grafiği d) Sütun Grafiği

4. Aşağıdakilerden hangisi ya da hangileri rastgele seçimdir?

I. Bilyeleri torbaya koyup iyice karıştırdıktan sonra kırmızı bilyeleri seçmek

II. Bilyeleri bir torbaya koyup iyice karıştırıp herhangi ikisini seçmek

III. Bilyeleri torbaya koymadan her beşinci bilyeyi seçmek

b) Yalnız I **b) Yalnız II** c) I ve II d) I ve III

5. a) Aşağıdaki veri gruplarından hangisinde yayılma <u>en fazladır</u>? İşlem yapmadan cevaplayınız.

a) 10, 11, 12, 13, 14, 15
b) 13, 13, 13, 13, 13, 13
c) 11, 12, 12, 13, 13, 14
d) 10, 12.5, 12.5, 12.5, 12.5,

b) Cevabınızın nedenini açıklayınız

6. Bir ilköğretim okulunun kütüphanecilik kolundaki Ali okuldaki öğrencilerin evlerindeki kitap sayısını araştırmak istiyor. Aşağıdakilerden hangisi bu araştırma için tüm okulu temsil edebilecek uygun örneklemi belirtir?

a) Kütüphanecilik kolundan rastgele seçilen 30 öğrenci
b) Okuldan rastgele seçilen 30 öğrenci
c) Ali'nin sınıfından rastgele seçilen 30 öğrenci
d) Okuldan rastgele seçilen 30 kız öğrenci

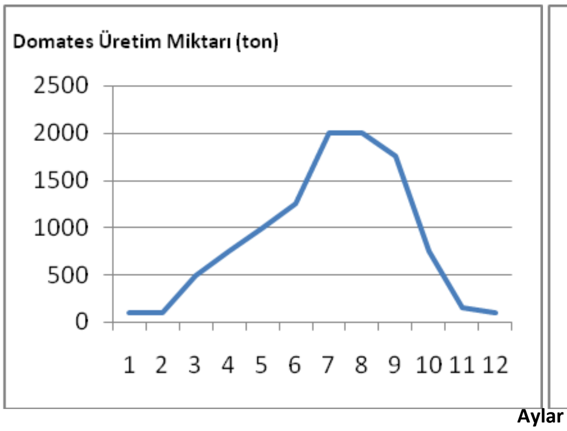
7. Bir kasabada 50 aile yaşamaktadır. Bu kasabada yaşayan bir araştırmacı aile başına düşen ortalama çocuk sayısını 2,2 olarak buluyor. Buna göre aşağıdakilerden hangisi kesinlikle doğrudur?

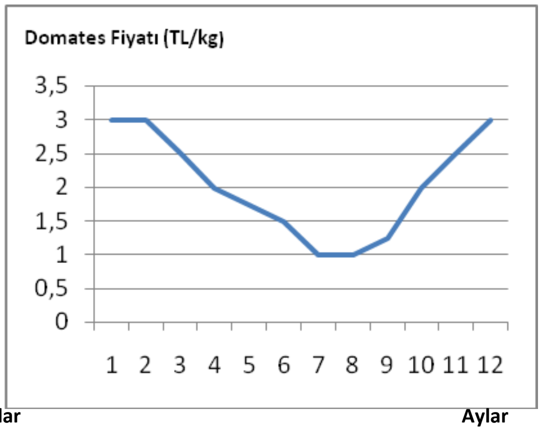
a) Bu kasabadaki ailelerin yarısı iki çocukludur.

b) 3 çocuklu aile sayısı 2 çocuklu aile sayısından fazladır.

c) **Bu kasabada 110 tane çocuk vardır**.

d) Her yetişkin başına 2,2 adet çocuk düşmektedir.





8. Yukarıdaki grafiklerde aylara göre domates üretim miktarı ve kilogram (kg) fiyatı verilmiştir. Bu grafiklere göre aşağıdaki sonuçlardan hangileri çıkarılabilir?

I. Domatesin fiyatı yaz aylarında düşük kış aylarında yüksektir.

II. Domatesin üretim miktarı yaz aylarında yüksek kış aylarında düşüktür.

III. Domatesin üretim miktarı kış aylarında düşük olduğu için fiyatı da yüksektir.

a) Yalnız I b) I ve II c) II ve III **d) I, II ve III**

9. Bir cilt kreminin üzerinde şöyle yazıyor:

“UYARI: Deri bölgesindeki uygulamalarda %15 cilt kızarıklığı riski vardır. Eğer cilt kızarıklığıyla karşılaşırsanız doktorunuza başvurun.”

Buna göre aşağıdakilerden hangisi doğrudur?

a) **Bu cilt kremini kullanan 100 kişiden yaklaşık 15’i cilt kızarıklığıyla karşılaşmıştır**.

b) Eğer kremi kullanıp cilt kızarıklığı ile karşılaşırsanız, bu kızarıklık cildinizin %15 ini kapsar.

c) Eğer bu kremi kullanırsanız kızarıklıkla karşılaşma ihtimali neredeyse yoktur.

d) Eğer bu kremi kullanacaksanız, cildinizin %15 inden fazlasına uygulamayın

10) Aşağıda bir bebeğin doğumundan itibaren her ay ölçülen ağırlığı verilmiştir.

Yaş (ay)	**Ağırlık (kg)**
0	3,5
1 ay	4,5
2 ay	5
3 ay	6
4 ay	7
5 ay	7,5
6 ay	8

a) Buna göre bebeğin 7. ayın sonundaki ağırlığının yaklaşık olarak kaç kilogram (kg) olabileceğini tahmin ediniz.

b) Bu sonuca nasıl ulaştığınızı açıklayınız.

11) Aşağıdaki tabloda bir ilköğretim okulunun 8A ve 8B şubelerindeki öğrencilerin matematik sınavından aldığı puanların bazı istatistik değerleri verilmiştir.

	Aritmetik Ortalama	Standart Sapma
A Şubesi	80	5,2
B Şubesi	76	3,5

Buna göre aşağıdaki yorumlardan hangisi ya da hangileri doğru olabilir?

I. Aritmetik ortalamaya bakarsak bu iki sınıftan A şubesindeki öğrencilerin notları daha yüksektir.

II. Standart sapmaya bakarsak B şubesindeki öğrencilerin notlarının yayılması daha küçüktür.

III. Standart sapmaya bakarsak A şubesindeki öğrencilerin notlarının yayılması daha küçüktür.

a) Yalnız I b) Yalnız II **c) I ve II** d) I ve III

12) Çocukların günde kaç saat televizyon izlediklerini araştıran bir çalışmanın örneklemi A İlköğretim Okulundaki 5. sınıf öğrencileridir. Bu araştırmaya katılan öğrencilerin günde ortalama 3 saat televizyon izledikleri ortaya çıkıyor. Bu araştırmanın sonuçlarını haber sunucusu şöyle anons ediyor:

"Türkiye'deki bütün ilköğretim öğrencileri günde ortalama 3 saat TV izlemektedir."

a) Sunucunun bu sözlerinin kabul edilebilir bir yorum olduğunu düşünüyor musunuz?

b)Yanıtınızı desteklemek için istatistiksel bir açıklama yapınız.

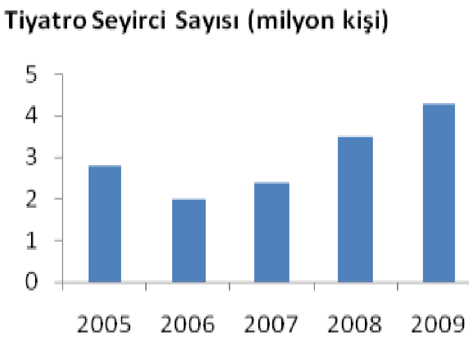
13) Bir ders boyunca 8 öğrencisinin çözdüğü soru sayıları aşağıdaki gibidir.

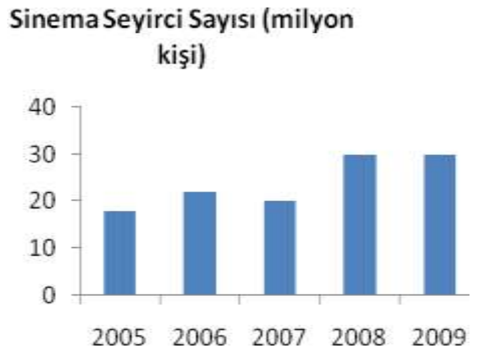
Öğrenci	**Soru Sayısı**
A	2
B	6
C	2
D	22
E	3
F	2
G	1
H	2

Yukarıdaki verileri özetleyen değeri bulmak için verilerin aritmetik ortalaması hesaplanarak öğrencilerin bir ders saati boyunca ortalama 5 soru çözdüğü sonucuna varılıyor.

a) Bu sonuca katılıyor musunuz?

c) Cevabınızı nedenleriyle açıklayınız.





Yıllar Yıllar

14. Bir kültür programının sunucusu yukarıdaki iki grafiği gösterip;
"2009 yılından önce sinema seyirci sayısı ile tiyatro seyirci sayısı değişiklik gösterse de, 2009 yılında sinema ve tiyatro hemen hemen eşit seyirci sayısına ulaşmıştır." dedi.

a) Sunucunun yorumunun kabul edilebilir olduğunu düşünüyor musunuz?

b)Yanıtınızı desteklemek için istatistiksel bir açıklama yapınız.

15) Sigara içenler ve içmeyenlerle yapılan bir araştırmaya göre 49 yıldan az bir süredir günde bir paket sigara içenlerde sigara içmeyenlere göre erken yaşlanma riski 2 katına, 50 yıldan fazla bir süredir günde bir paket sigara içenlerde içmeyenlere göre erken yaşlanma riski 4,7 katına çıkmaktadır. Aşağıdaki tablo bu bilgiyi özetlemektedir.

	49 yıldan az	**50 yıldan fazla**
Sigara içmeyenlerin riski	A	B
Sigara içenlerin riski	2A	4,7B

a) Yukarıdaki raporun sonucu kabul edilebilir midir?

b) Yukarıda verilen raporun **<u>geçerliliğini</u>** sorgulamak için araştırmacılara hangi soru ya da soruları yöneltirsiniz?

16) Aşağıdaki tabloda akciğer kanseri ve sigara kullanımı ile ilgili 250 kişiye yapılan bir araştırmanın sonuçları gösterilmiştir.

	Akciğer Kanseri Olan	**Akciğer Kanseri Olmayan**	**Toplam**
Sigara İçen	90	60	150
Sigara İçmeyen	60	40	100
Toplam	150	100	250

a) Verileri değerlendiren bir araştırmacı "Sigara içmek akciğer kanserine neden olmaktadır." diyor. Araştırmacının bu sözlerinin kabul edilebilir bir yorum olduğunu düşünüyor musunuz?

b) Cevabınızı istatistiksel olarak açıklayın.

17) Bir grup öğrenci bir yıl boyunca her gün hava durumu haberlerini izleyip, günlük olarak Ankara'daki en yüksek sıcaklıkları not alıyorlar. Ankara'nın yıllık ortalama en yüksek sıcaklığını 16^{o} olarak buluyorlar.

Bu gruptan farklı 3 öğrenci yılın herhangi altı farklı gününde olabilecek en yüksek sıcaklıkları aşağıdaki gibi tahmin ediyorlar.

Öğrenciler	**Tahmin Ettikleri Hava Sıcaklıkları**
Seda	16, 35, 1, 5, 29, 10
Zeynep	16, 16, 16, 16, 16, 16
Umut	16, 15, 14, 26, 8, 17

a) Sizce hangi öğrenci yıllık ortalama sıcaklıklar hakkında en uygun yayılmayı gösteren veriyi oluşturmuştur?

a) Seda

b) Zeynep

c) Umut

b) Neden böyle düşündüğünüzü açıklayınız.

APPENDIX C

ATTITUDES TOWARDS STATISTICS QUESTIONNAIRE

İstatistiğe Yönelik Tutum Anketi

Aşağıda istatistik hakkındaki duygu ve düşüncelerden oluşan ifadeler bulunmaktadır. Her ifade ile ilgili görüş, kişiden kişiye değişebilir, hiçbirisinin kesin cevabı yoktur. Bunun için vereceğiniz yanıtlar kendi görüşünüzü yansıtmalıdır. Her cümleyle ilgili görüş belirtirken önce cümleyi dikkatle okuyunuz, sonra cümlede belirtilen düşüncenin, sizin düşünce ve duygunuza ne derecede uygun olduğuna karar veriniz. Bu cümleler için ifade edilen düşüncelere sizin ne derece katılıp katılmadığınızı belirtmeniz için "*Kesinlikle Katılmıyorum*", "*Katılmıyorum*", "*Kararsızım*", "*Katılıyorum*", "*Kesinlikle Katılıyorum*" seçenekleri verilmiştir. Lütfen **tüm** soruları dikkatlice okuyup **boş bırakmadan,** sizin için en uygun seçeneği işaretleyiniz.

	Kesinlikle Katılmıyorum	**Katılmıyorum**	**Kararsızım**	**Katılıyorum**	**Kesinlikle Katılıyorum**
1. İstatistik konularını severim.	☐	☐	☐	☐	☐
2. İstatistik konuları sevimsizdir.	☐	☐	☐	☐	☐
3. İstatistik konularını tartışmaktan hoşlanırım.	☐	☐	☐	☐	☐
4. İstatistikle ilgili bilgiler can sıkıcıdır.	☐	☐	☐	☐	☐
5. İstatistikle ilgili bilgiler zihin gelişimine yardımcı olur.	☐	☐	☐	☐	☐
6. İstatistik ile ilgili konular beni huzursuz eder.	☐	☐	☐	☐	☐
7. İstatistikle ilgili ders saatlerinin daha çok olmasını isterim.	☐	☐	☐	☐	☐
8. İstatistik konuları rahatlıkla/kolaylıkla öğrenilebilir	☐	☐	☐	☐	☐
9. İstatistikle ilgili sınavlardan korkarım.	☐	☐	☐	☐	☐
10. İstatistik konuları ilgimi çeker.	☐	☐	☐	☐	☐
11. İstatistiğin doğru karar vermemizde önemli rolü vardır.	☐	☐	☐	☐	☐
12. İstatistik konuları aklımı karıştırır.	☐	☐	☐	☐	☐
13. İstatistik konularını severek çalışırım	☐	☐	☐	☐	☐
14. İstatistik konularını, elimde olsa öğrenmek istemezdim.	☐	☐	☐	☐	☐
15. İstatistik, ilginç bir konu değildir.	☐	☐	☐	☐	☐
16. İstatistikle ilgili ileri düzeyde bilgi edinmek isterim.	☐	☐	☐	☐	☐
17. İstatistik hemen hemen her iş alanında kullanılmaktadır.	☐	☐	☐	☐	☐
18. İstatistik konularını çalışırken canım sıkılır.	☐	☐	☐	☐	☐
19. İstatistik, kişiye düşünmesini öğretir.	☐	☐	☐	☐	☐
20. İstatistiğin adını bile duymak sinirlerimi bozuyor.	☐	☐	☐	☐	☐
21. İstatistik konularından korkarım.	☐	☐	☐	☐	☐
22. İstatistik konularını herkesin öğrenmesi gerekir.	☐	☐	☐	☐	☐
23. İstatistik konularından hoşlanmam.	☐	☐	☐	☐	☐

24. İstatistikle ilgili bilgiler, kişinin tahmin (etme) yeteneğini artırır.					
25. İstatistik konuları anlatılırken sıkılırım.					
26. İstatistikle ilgili bilgilerin, günlük yaşamda önemli bir yeri vardır.					
27. İstatistik konuları eğlencelidir.					

APPENDIX D

DEMOGRAPHIC INFORMATION SHEET

1. Okulunuzun adı: ………………………… 2. Sınıfınız:…………………….

3. Doğum tarihiniz (yıl): ……………. 4. Cinsiyetiniz: □Kız □Erkek

5. Geçen dönemki Matematik karne notunuz: ………….

6. Siz hariç kaç kardeşiniz var?

□ Kardeşim yok □ 1 □ 2 □ 3 □ 4 ve üstü

7. Anneniz çalışıyor mu?

□Çalışıyor □Çalışmıyor □Düzenli bir işi yok □Emekli

8. Babanız çalışıyor mu?

□Çalışıyor □Çalışmıyor □Düzenli bir işi yok □Emekli

9. Annenizin Eğitim Durumu

□Hiç okula gitmemiş
□İlkokul
□Ortaokul
□ Lise
□ Üniversite
□ Yüksek lisans / Doktora

10. Babanızın Eğitim Durumu

□Hiç okula gitmemiş
□ İlkokul
□Ortaokul
□ Lise
□Üniversite
□ Yüksek lisans / Doktora

11. Magazin dergileri, gazete ve okul kitapları dışında evinizde kaç tane kitap bulunuyor?

□ Hiç yok ya da çok az (0 – 10)
□ Bir rafı doldurmaya yetecek kadar (11 – 25 tane)
□ Bir kitaplığı doldurmaya yetecek kadar (26 – 100 tane)
□ İki kitaplığı doldurmaya yetecek kadar (101- 200 tane)
□ Üç veya daha fazla kitaplığı doldurmaya yetecek kadar (200 taneden fazla)

12. Ne kadar sıklıkla eve gazete alıyorsunuz?

☐Hiçbir zaman ☐Bazen ☐Her zaman

13. Evinizde aşağıdakilerden hangileri vardır (Her sırada sadece bir kutuyu işaretleyiniz):

	Var	**Yok**
Bilgisayar	☐	☐
İnternet erişimi	☐	☐
Çalışmak için ayrı oda	☐	☐
Çalışma masası	☐	☐
Sözlük	☐	☐
Hesap makinesi	☐	☐
Bulaşık makinesi	☐	☐

APPENDIX E

RUBRIC FOR OPEN ENDED ITEMS

İstatistiksel Okuryazarlık Testi Açık Uçlu Sorular İçin Dereceli Puanlama Anahtarı

1.Madde

0. Yanlış cevaplar

   Örneğin: Hiçbir şey anlamıyorum

   Araştırmanın konusu

1. İstatistiksel olmayan cevaplar

   Örneğin: Örnek

   Araştırma yapılan kişiler

2. En üst seviyede açıklama yapanlar

   Örneğin: Bütünün (bir) parçası

   Popülasyonu (tamamını) temsil eden kişiler

2. Madde

0. Yanlış cevaplar

   Örneğin: Bağlamla ilişkili yanlış açıklamalar

   Diğer yanlış açıklamalar

1. İstatistiksel olmayan cevaplar

   Örneğin: Yaklaşık

   Tahminen

   Aşağı yukarı

2. Merkezi eğilim ölçüleri ile açıklayanlar

   Örneğin: Aritmetik ortalama anlamı: Toplayıp bölme

   Medyan anlamı: Ortadaki sayı

   Mod anlamı: En çok tekrar eden sayı

3. En üst seviyede açıklama yapanlar

   Örneğin: Herhangi bir ay için ölen kişileri (temsil eden) değer

   Denge noktası

5. Madde

0. Yanlış cevaplar

   Örneğin: Açıklama yapılmamış

   Bütün sayılar aynı

1. İstatistiksel olmayan cevaplar

   Örneğin: Sayılar 1'er 1'er artmış.

   Sayılar daha çok farklılık gösteriyor.

2. Yayılma ölçüleri ile açıklayanlar

   Örneğin: Açıklığı hesaplayanlar

   Çeyrekler açıklığını hesaplayanlar

   Standart sapmayı hesaplayanlar

3. En üst seviyede açıklama yapanlar

   Örneğin: Değişkenlik daha fazla

   Ortalamaya en uzak sayılar

11. Madde (b)

0. Yanlış cevaplar

   Örneğin: Tahmin ettim.

1. Doğru cevaplar

   Örneğin: Tablodaki verilerden ulaştım, Örüntüyü takip ettim.

12. Madde (b)

0. Yanlış cevaplar
1. İstatistiksel olmayan yanlış cevaplar

   Örneğin: Kendi yaşantısından örnek verenler: Ben 3 saat izliyorum. Sonuç doğrudur

   Contextualbeliefs: Çocuklar 3 saat TV izlememeli.

2. İstatistiksel/ En üst seviyede açıklama yapanlar

   Örneğin: Tek okul ile genelleme yapılamaz

   Sadece 5. Sınıflar ile genelleme yapılamaz.

   Genelleme yapılamaz örneklem sayısı çok az

13. Madde (b)

0. Yanlış cevaplar / açıklamalar

Örneğin: Bağlamla ilişkili yanlış açıklamalar

(aritmetik ortalama) kontrol edip sonucu doğrulayanlar

1. İstatistiksel olmayan açıklamalar

Örneğin: Çözülen soru sayılarını yadırgayanlar

22'nin farklılığını fark edenler: bir derste 22 soru çözülemez

2.Uygun istatistiksel açıklamalar

Örneğin: Uç değer (22) olduğundan ortalamaya katılmamalıydı

Uç değer olduğundan medyan

Yayılmayı göz önünde bulunduranlar (aralığı geniş, st.sapması büyük)

14. Madde (b)

0. Eşit, aynı ya da bağlamla ilgili yanlış açıklamalar
1. Sayılar farklı, eksenlerin aralığı farklı

15. Madde (b)

0. İlgisiz sorular

Örneğin: Sigara içiyor musunuz?

1. İstatistiksel olmayan ancak raporu sorgulayan sorular

Örneğin: Bu araştırmayı neden yaptınız?

Verileri nasıl topladınız?

2. İstatistiksel ve aynı zamanda raporu sorgulayan sorular

Örneğin: Neden 49 yıl?

Yaşlanmanın tek sebebi sigara içmek midir?

Kaç kişi?

Örneklemi nedir?

16. Madde (b)

0. Yanlış açıklamalar

Örneğin: Bağlamla ilişkilendirenler:

Sigara içmek zararlıdır.

1. Sadece sayılara bakanlar ilişkileri incelemeyenler

Örneğin: Sigara içenler(in oranı) fazla

2. İstatistiksel /doğru açıklamalar

Örneğin: Sigara içmeyenler de kanser oluyor

Oranları eşit

Olasılıkları eşit

17. Madde (b)

0. Zeynep

Örnek açıklama: Hepsi aynı

Ortalama ile eşit

1. Seda

Örnek açıklama: Farklılık çok fazla

Değişik sayılar

2. Umut

Örnek açıklama: Farklı sayılar

Bir nokta (ortalama =16) etrafında toplanmış

APPENDIX F

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü [ ]

Sosyal Bilimler Enstitüsü [X]

Uygulamalı Matematik Enstitüsü [ ]

Enformatik Enstitüsü [ ]

Deniz Bilimleri Enstitüsü [ ]

YAZARIN

Soyadı : YOLCU
Adı : Ayşe
Bölümü : İlköğretim Bölümü

TEZİN ADI (İngilizce) : An Investigation of Eighth Grade Students' Statistical Literacy, Attitudes Towards Statistics and Their Relationship

TEZİN TÜRÜ : Yüksek Lisans [X] Doktora [ ]

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın. [X]

2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullanıcılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) [ ]

3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) [ ]

Yazarın imzası .... Tarih ....