INVESTIGATION OF PRE-SERVICE MATHEMATICS TEACHERS' CRITICAL THINKING PROCESSES THROUGH STATISTICAL AND PROBABILISTIC KNOWLEDGE IN THE CONTEXT OF POPULAR MEDIA TEXTS

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

BY

MEHTAP ÖZEN

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

JANUARY 2013

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 of Science.

_____

Doç. Dr. Erdinç ÇAKIROĞLU
Supervisor

Examining Committee Members

Doç. Dr. A. Kürşat ERBAŞ (METU, SSME) _____

Doç. Dr. Erdinç ÇAKIROĞLU (METU, ELE) _____

Yrd. Doç. Dr. Didem AKYÜZ (METU, 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 : Mehtap ÖZEN

Signature :

ABSTRACT

INVESTIGATION OF PRE-SERVICE MATHEMATICS TEACHERS' CRITICAL THINKING PROCESSES THROUGH STATISTICAL AND PROBABILISTIC KNOWLEDGE IN THE CONTEXT OF POPULAR MEDIA TEXTS

Özen, Mehtap

M.S., Department of Elementary Science and Mathematics Education

Supervisor : Doç. Dr. Erdinç Çakıroğlu

January 2013, 197 pages

The aim of the study is to investigate pre-service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts. The study was conducted with a qualitative case study method. Participants of the study consisted of four senior pre-service middle school mathematics teachers enrolled in a public university. Data were collected through in-depth interviews with the participants. Analysis of the data was conducted on the basis of two dimentions; critical thinking skills, and statistical and probabilistic knowledge.

The results of the study indicated that pre-service middle school mathematics teachers reflected different critical thinking skills and made use of different statistical and probabilistic knowledge in different contexts. They mostly reflected interpretation skill on the basis of their statistical and probabilistic knowledge. Moreover, to what extent they made use of critical thinking skills was differentiated

on the basis of their statistical and probabilistic knowledge. They reflected complicated critical thinking process dealing with conditional probability statements. They had difficulty with probabilistic statements underlying conditional probability especially in this process.

Keywords: Critical thinking, Statistical literacy, Probabilistic literacy, Mathematics education

ÖZ

MATEMATİK ÖĞRETMENİ ADAYLARININ POPÜLER MEDYA METİNLERİNDE İSTATİSTİKSEL VE OLASILIKSAL BİLGİLERİ BAĞLAMINDA ELEŞTİREL DÜŞÜNME SÜREÇLERİNİN İNCELENMESİ

Özen, Mehtap
Yüksek Lisans, İlköğretim Fen ve Matematik Alanları Eğitimi Bölümü
Tez Yöneticisi : Doç. Dr. Erdinç Çakıroğlu

Ocak 2013, 197 sayfa

Bu çalışmanın amacı, ortaokul matematik öğretmeni adaylarının popüler medya metinlerinde istatistiksel ve olasılıksal bilgilerini kullanarak eleştirel düşünme süreçlerini incelemektir. Araştırma nitel durum çalışması ile gerçekleştirilmiştir. Çalışmanın katılımcılarını, bir devlet üniversitesinde öğretmen yetiştirme programına kayıtlı olan 4 son sınıf ortaokul matematik öğretmeni adayı oluşturmaktadır. Veriler, katılımcılardan derinlemesine görüşmeler aracılığıyla toplanmıştır. Verilerin analizi eleştirel düşünme becerileri ve istatistiksel ve olasılıksal bilgi olmak üzere iki boyut altında yürütülmüştür.

Çalışmanın sonuçları katılımcıların farklı içeriklerde farklı eleştirel düşünme becerileri yansıttıklarını ve farklı istatistiksel ve olasılıksal bilgi kullandıklarını ortaya koymuştur. İstatistiksel ve olasılıksal bilgilerini kullanarak çoğunlukla yorumlama becerisi yansıtmışlardır. Eleştirel düşünme becerilerinden ne derecede

yararlandıkları istatistiksel ve olasılıksal bilgilerinin yapısına göre farklılaşmaktadır. Koşullu olasılıkla ilgili ifadeler üzerinde çalışırken karmaşık bir eleştirel düşünme süreci yansıtmışlardır. Katılımcıların özellikle bu süreçte zorluk yaşadığı gözlenmiştir.

Anahtar Kelimeler: Eleştirel düşünme, İstatistiksel okuryazarlık, Olasılıksal okuryazarlık, Matematik eğitimi

To My Lovely Mother

ACKNOWLEDGMENTS

I would like to express my deepest appreciation and gratitude to my supervisor Doç. Dr. Erdinç Çakıroğlu for his guidance, encouragement, patient and greatest support throughout my M.S. study. He encouraged me to pursue my goal whenever I was in the complicated processes of the research. His advices and comments have broadened my view point towards academical working.

I also extend my sincere thanks to my thesis examining committee members; Doç. Dr. A. Kürşat Erbaş and Yrd. Doç. Dr. Didem Akyüz for their inavaluable comments, suggestions and insightful contributions to my thesis.

I wish to express my deep gratitude to Doç. Dr. Mine Işıksal and Yrd. Doç. Dr. Çiğdem Haser for their suggestions, feedbacks, and invaluable views during my thesis work.

I would like to express my special thanks to Dr. Nihal Akdere and Özgür Asar for their support and invaluable views, feedbacks, and contributions to my thesis.

I would also like to express my special thanks to my kind-hearted friend Seçil Yemen Karpuzcu. I feel lucky that I have a friend like Seçil. She was always with me from the beginning to the end of the thesis. I am very grateful for her valuable views and contributions to my study. Her greatest support, encouragements in every time are not unforgettable. She always listened to me and encouraged me to do my best when I have lost in complicated processes of the thesis. And, another special thanks to Hüseyin Karpuzcu for his patience and support, providing us working conditions.

I would like to thank my office mates; Savaş Pamuk, Nursel Yılmaz, Burak Çaylak, and Erdinç İşbilir. I am very lucky to have such friendly office mates. Their supports, encouragements are invaluable, providing positive environment in the office.

Another special thanks to my collegues; Seçil Tısoğlu, Celal İler, Berna Aygün, Ayşenur Kubar, Sibel Akın, Nilay Öztürk, Gözde Didiş, Dilek Girit, Mustafa Cansız, Esra Balgalmış, Yasemin Taş for their support and contributions. I apologize from the people whose names I might have forgotten to add and thank all of them for their contributions.

I would like to thank participants of the my study. Their willigness and greatest support is not unforgettable. They were always helpful to make easy procedures of data collection. This study could not be conducted in a best way if I had not worked with such helpful and kindhearted participants.

I would like thank the Scientific and Technological Research Institute of Turkey (TUBITAK) for its support with master of science studies scholarship.

Finally, I would like to express my special thanks to my lovely mother and dear father. They were always with me with unconditional support and love. They always believed me to accomplish my work. And, another special thanks for my lovely sisters; İrem Özen and Meltem Özen for their patience and insight. They helped me feel positive through the thesis.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AEC	Australian Education Council
GAISE	National Statement on Mathematics for Australian Schools, and Guideline for Assessment and Instruction in Statistics Education Report
NCTM	National Council of Teachers of Mathematics
PCK	Pedagogical Content Knowledge

CHAPTER 1

INTRODUCTION

Statistics and probability are an integral part of everyday life. Social and scientific texts such as newspapers, journals, and advertisements mostly include information, arguments, or claims based on statistical studies. Forecasts, economic trends over the years, risks of having an accident, political decisions, the effects of drug use, for example, are presented in multiple and various representations such as words, tables, or in graphs. To what extent citizens understand and evaluate these appropriately and critically reported results based on statistical studies and to what extent journalists present valid information or arguments to the public is a negotiable issue. It is, however, clear that there is a need to prepare individuals as statistically thinking citizens in the society (Wallman, 1993).

In recent years, there have been attempts in mathematics education to integrate statistical thinking needed for efficient citizenship into mathematics curricula. Principles and Standards proposed by the National Council of Mathematics Teachers (NCTM, 2000) in the U.S.A., Australian Education Council (AEC, 1991, National Statement on Mathematics for Australian Schools), and Guidelines for Assessment and Instruction in Statistics Education Report (GAISE, 2005) are among predominant documents emphasizing understanding statistics and probability in real life, which can be seen as indicators of the increasing number of attempts in relation to integrating statistics into mathematics education. For example, the GAISE Report as a pre-K-12 curriculum framework put emphasis on statistical literacy as a prominent goal in education and stated;

> Every high school graduate should be able to use sound statistical reasoning to intelligently cope with the requirements of citizenship, employment, and family and to be prepared for a healthy, happy, and productive life (GAISE, 2005, p., 1).

In a similar way, the Australian Educational Council (1991) reflected the need for statistically literate students in society with the standards of "understand and explain social uses of chance" (p.175) and "understand the impact of statistics on daily life" (p.178).

In Turkey, elementary mathematics curriculum also included a new domain called statistics and probability, which includes basic concepts such as measures of central tendency, data representation, measures of spread, probabilistic events, and basic inference. Similar to the other curriculum efforts, the Turkish curriculum has pointed out the importance of statistics and probability in real life and the fact that students should be able to interpret statistics in real life contexts and make decisions on the basis of their statistical and probabilistic knowledge (MNE, 2005).

In addition to these curriculum efforts, various studies underscored the importance of statistical literacy for efficient citizenship, indicating that citizens need to think statistically to participate actively in the society (Gal, 2002; Utts, 2003; Wallman, 1993; Watson, 1997; Watson, 1998). Wallmann (1993) described statistical literacy as;

> […]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)

Watson (1997; 1998; 2006), one of the pioneer studies regarding statistical literacy in mathematics education, addressed the need for students to develop understanding and evaluation of claims in news media based on knowledge of statistical and probabilistic concepts (sampling, average, chance, inference, data representation, and variation). Supportively, Gal (2004; 2005) expressed the need for statistical

literacy for adults, which requires the abilities of interpretation and critical evaluation of claims or arguments in the statistical context in order to make explanations based on statistical and probabilistic knowledge, and to discuss with other people.

These studies emphasized the importance of having basic statistical and probabilistic knowledge to engage with media texts including bias, misleading information, one-sided arguments, and ambiguous language, which could lead people to make inappropriate decisions regarding their life. In this regard, Ridgway, Nicholson, and McCuskes (2011) have suggested more recently that both students and teachers be active in society by interpreting and thinking critically while reading reports in the media.

Critical Thinking and Statistical and Probabilistic Knowledge

The need for citizens to think critically and actively participate in the society raises the questions of "What critical thinking refers to" and "How we integrate critical thinking into a specific subject domain" in statistics education. Critical thinking, which dates back to early years, has been discussed among many researchers. The importance of critical thinking as an educational goal has been acknowledged in various research areas such as philosophy, psychology, education, and cognitive sciences (Ennis; 1985; Facione, 1990; Halpern, 1997; Kennedy, Fisher, & Ennis, 1991; Kuhn, 1999; Paul, 1984; Siegel, 1988). One of the reasons behind this interest is that there is need for qualified and critically thinking citizens to actively participate in the society (Kennedy et. al., 1991). Another reason is that people should keep pace with rapid changes in the world. Yet another is that every person is required to make decisions regarding economy, political election, health or education etc., which have possible effects on people's future life (Halpern, 1997). The last but not least is that critical thinking plays a crucial role in making educational decisions as regards "…what we should teach, how we should teach,

how we should organize educational activities, what the points of many of those activities are, how we should treat students and others in the educational setting…" (Siegel, 1988; p. 46). Moreover, characteristics of an ideal person that educators try to nurture in their students are conceptualized on the basis of critical thinking In other words, critical thinking constitutes a guideline for educational decisions (Siegel, 1988). These reasons provoked researchers to make a movement toward teaching critical thinking as an educational goal (Kennedy et. al., 1991).

In this regard, researchers attempted to conceptualize the term of critical thinking, identify its constituents by proposing models or frameworks, determine the factors that have possible affect in the improvement of critical thinking, propose instructional strategies or methods to promote both students and teachers' critical thinking, and design instructional instruments to assess students and teachers' critical thinking.

Moreover, many researchers proposed suggestions to improve both students' and teachers' critical thinking. Facione (1990) reported a long-term project, supported by the American Philosophical Association, to develop a consensus on the conceptualization of critical thinking. In the report, it is indicated that the learning and teaching of critical thinking foster students' both cognitive skills and dispositions in critical thinking, which aids students in dealing with educational, personal, and social concerns. For this reason, the report recommends that all grade levels in K-12 curriculum be taught critical thinking immersed into specific subject matters. Supportively, Brown (1997) advocated that students' critical thinking could be developed through serious matters that they are engaged with within the context of a specific subject matter. In a similar way, Halpern (1998) addressed the issue of transferability of critical thinking across different domains and advocated that students can learn critical thinking with an appropriate educational design. To achieve this goal, critical thinking should be transferable to real-world or out of school settings in various subject matters.

Moreover, in their overview of studies regarding critical thinking, Kennedy et. al. (1991) suggested that there has been an agreement regarding the usage of real-world contexts, problems or popular media texts such as newspaper articles, magazine articles, advertisements or television programs in the classroom environments in teaching critical thinking.

Another agreement among researchers is that statistical and probabilistic knowledge have a crucial role in critical thinking in the scientific and social contexts (Facione, 2011a; Halpern, 2003; Osana & Seymour, 2004). For example, they draw attention to questioning sample size, quality or reliability of research studies, generalizability of the reported findings, statistical and probabilistic concepts such as conditional probability to be effective critical thinkers in the society.

In conclusion, studies about statistical literacy and critical thinking intersect with each other and meet on a common ground indicating the need for efficient citizenship in the society, which requires thinking critically in the statistical and probabilistic sense and having the dispositions to do so.

Utts (2003), however, argues that many citizens do not have enough statistical and probabilistic knowledge to interpret and critically evaluate the statistical information in most of the newspapers. Many studies supported this claim that most of the students had difficulty especially in questioning and critically evaluating claims or arguments presented in the media (Schield; 2006; Watson, 1997; Watson; 1998; Yolcu; 2012). In this circumference, Shaughnessy (2007) briefly reviewed studies related to both students and teachers in statistics education and advocated that teachers have an important role in the education of statistics and probability. Supportively, Paul, Elder, and Bartell (1997) pointed out the role of teachers in the development of critical thinking. They suggested that teachers must learn to think critically in real life and their professional subject domain and appreciate the role of critical thinking in education because improving the critical thinking skills of

students as future citizens who can keep pace with the changes in the world of 21st century would be impossible if teachers themselves lack with the critical thinking. In summary, these two different and closely related research domains; statistical literacy and critical thinking point out the important role of teachers in educating students as active participants in a democratic society.

However, studies in Turkey, are limited regarding the investigation of teachers' statistical and probabilistic knowledge in context. Yolcu (2012) recommended in her study on students' statistical literacy that there is a need for investigation of teachers' statistical literacy and to what extent it has role in the development of students' statistical literacy. Similarly, from the point of critical thinking, there is not enough studies regarding the development of teachers' critical thinking skills, especially in mathematics education although Seferoğlu and Akbıyık (2006) suggested that teachers provide students with the opportunity to transfer critical thinking skills to different contexts.

In this regard, this study aimed to investigate pre-service middle school mathematics teachers' critical thinking skills through statistical and probabilistic knowledge when they engage in real-life contexts.

1.1. Purpose of the Study

The purpose of the study is to investigate pre-service middle school mathematics teachers' critical thinking processes in combination with their statistical and probabilistic knowledge they utilized in the thinking process while reading popular media texts involving statistical and probabilistic information. The following research questions guided the study:

Research Question: To what extent do senior pre-service middle school mathematics teachers make use of critical thinking skills through statistical and probabilistic knowledge in the context of popular media texts?

Sub-Questions:

1. What knowledge of statistics and probability do senior pre-service middle school mathematics teachers utilize while reading popular media texts including statistical and probabilistic information?

2. What critical thinking skills do senior pre-service middle school mathematics teachers make use of while reading popular media texts including statistical and probabilistic information?

1.2. Significance of the Study

In the last decades, there has been an increased attention among researchers and in curriculum documents regarding the need to nurture skills in students that enable them to become efficient citizens and participate actively in the democratic society (Facione, 2011a; Ten Dam & Volman, 2004; Utts, 2003). To keep up with the demands of the rapidly changing world, students should be able to think critically to make proper decisions on the basis of claims, arguments, or information in real-life contexts which mostly include statistical and probabilistic statements (Watson, 2006). The need for critical and democratic citizenship has been considered under the headings of critical thinking and statistical literacy. Studies conducted in both educational areas consistently indicated that teacher preparation programs should aim to teach prospective teachers how to design a learning environment supporting critical thinking because students have difficulty in questioning claims or arguments in the social or scientific contexts. In addition, studies mostly focused on students' development of critical thinking and statistical literacy. There is a need to develop

pre-service teachers' critical thinking on the basis of statistical and probabilistic sense in the light of the recommendations arising from the educational research community. In this context, this study, which aims to investigate pre-service middle school mathematics teachers' critical thinking processes by means of statistical and probabilistic knowledge in the context of popular media texts, can be considered as an attempt to satisfy this growing need in teachers' education in mathematics.

There are two main underlying premises of this study. The first premise is that the findings of the study would contribute to the literature in statistics education since there are limited studies in statistics with regard to teachers, especially in the issue of statistics and probability in context. Training prospective teachers in interpreting and critically evaluating statistical and probabilistic arguments, or claims reported in the real world would be a step in nurturing students from early grades to high grade levels toward thinking statistically and making rational decisions in their daily life. In this regard, statistics education would go beyond computational approaches like rolling a dice, or pulling out balls from an urn, enabling citizens in the society as consumers of data to be in accord with the shift in world demands.

The second premise is that the findings of the study would lead to the reorganization of learning and teaching environments. The possible developments in statistics education in terms of students, teachers, teacher educators, curriculum designers, or policy makers could be expected as follows: development of open-ended items in the assessment of both middle school students and pre-service middle school mathematics teachers, which requires critical thinking in context including statistical and probabilistic information, overview of the content of teaching methods of mathematics and statistics courses in teacher education programs, usage and development of efficient appropriate tasks or activities including critical questions from the media; revision of the current curriculums taking into consideration statistical literacy and critical thinking, development of textbooks or other sources in

mathematics education connecting statistical concepts with mathematical concepts, as wells as the real world.

1.3. Definition of Important Terms

Critical Thinking: Critical thinking refers to "…purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which judgment is based." (Facione, 1990; p. 3). Critical thinking is conceptualized in two dimensions: cognitive skills and affective dispositions. This study focuses only cognitive skill dimension of critical thinking.

Statistical and Probabilistic Knowledge: Gal (2004) defined statistical knowledge for adults on the basis of statistical literacy. Statistical knowledge includes five basic knowledge bases; "knowing why data are needed and how data can be produced, familiarity with descriptive statistics, familiarity with graphical and tabular displays, understanding basic notions of probability, and knowing how statistical conclusions or inferences are reached." (p. 58). In a further study, Gal (2005) also identified knowledge elements required for probabilistic literacy in five sub-dimensions: big ideas of probability, the terminology and language of probabilities, critical questions, and context knowledge. Moreover, Watson (2006) determined statistical and probabilistic knowledge for students at grades from 6 to 9 on the basis of five concepts; sampling, average, data representation, chance, inference, and variation. These knowledge elements for statistical literacy were combined in line with the purpose of the study as; bases of reported findings, reported findings as summary statistics, and generalizability of the reported findings, which are explained in the method section.

Popular Media Texts: Popular media texts refer to social or scientific written sources such as newspapers, advertisements, journals, magazines widely

encountered in the media or the Internet. This study is limited with the newspaper articles published in the websites of common newspaper associations in Turkey. The selection process of the newspaper articles used in this study is identified in the method section.

CHAPTER 2

LITERATURE REVIEW

The purpose of the study is to investigate pre-service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts. Theoretical background and related research studies were overviewed throughout the chapter. The review of literature is presented at two main topics; statistical literacy and critical thinking.

The first part of the chapter, statistical literacy, begins with definitions of statistical literacy and continues with theoretical models related to statistical and probabilistic literacy. Then, research studies were presented with the topics of statistics and probability in context and research in teacher education regarding statistics and probability in context. At the end of the first part, studies regarding learning and teaching statistics and probability in Turkey were reviewed.

The second part of the chapter, critical thinking, consisted of four topic; conceptualization of critical thinking, theoretical model of critical thinking, learning and teaching of critical thinking, and research regarding critical thinking in Turkey.

At the end of this chapter, a summary was stated to emphasize the underlying rationale of this research by combinations of all studies mentioned in the first and second part.

2.1. Statistical Literacy

In recent years, statistical literacy has gained a considerable attention among the educational research community for the purpose of educating students as statistical citizens in society. Although many studies have emphasized its crucial role on making decisions in social life and being citizens who question claims, arguments made in the society, there is still no consensus on the question of "what statistical literacy refers to" (Batanero, 2002).

The term of statistical literacy has been defined, or identified, by several researchers. Wallman (1993) defined statistical literacy as "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). She also stated that there is a need for further clarification of statistical literacy construct in order to nurture people in the society as statistically thinking. In this regard, many researchers attempted to define or describe the construct of statistical literacy. For example, Watson (1997), one of the pioneer studies in statistical literacy, defined statistical literacy for students in the developmental process. This developmental process includes three tiers; understanding basic terminology of statistical and probabilistic concepts in a mathematical concept, understanding and interpretation of statistical and probabilistic concepts in social or scientific context, and critical questioning of the claims or arguments in the real life. In a further study, Gal (2004) defined statistical literacy for adults. He defined statistical literacy with two main components. The first ability required for statistical literacy is to interpret and critically evaluate claims or arguments in the statistical context. The second ability is to make discussion of these claims of arguments and transfer own ideas, knowledge or concern to the other people. In this regard, he also identified two main factors that have potential role in the development of statistical literacy; knowledge and dispositional aspects.

In a similar way, Rumsey (2002) discussed the definition of statistical literacy construct and how it can be developed in introductory statistics courses. He defined statistical literacy at four main points; appreciation of role of data in real life, understanding basic statistical concepts, interpretation of these concepts in the context, and communication of statistical information to someone else. He preferred to use the terminology of statistical competence rather than statistical literacy since statistical literacy has broader meaning. Similar with Rumsey (2002), Schield (1999) defines statistical literacy in terms of competency, which requires critically thinking and interpreting claims or arguments based on statistical and probabilistic concepts. He also identified the characteristics of people with the ability of statistical literacy. For example, a statistically literate person can recognize the difference between correlational and experimental studies and identify possible confounding variables that have possible effect on the result of the study. Moreover, a person could ask critical questions themselves as how could be bias in measurement or could be controlled.

To conclude, these studies have stressed several common points of statistical literacy construct; knowledge base for statistical literacy and dispositional aspect and interpretation, critical evaluation, communication with other people, which are possible factors to affect statistically literacy.

There is another discussion among several researchers beside discussion of what fundamental structures of statistical literacy. Researchers (Ben-Zvi & Garfield, 2004; delMas, 2002; Shaugnessy; 2007) also argued what critical points statistical literacy differs from statistical thinking and statistical reasoning. delMas (2002) argues the differences among these three concepts. The distinction between those domains occurs when asking different questions regarding context. To illustrate, statistical literacy includes skills of giving examples of a statistical or probabilistic concept, expressing the meaning of the graphs or relationship between variables, restating and interpreting findings of a research. On the other hand, statistical

reasoning is mostly related to the questions of how results of the study could be reached or why random sampling has advantages or why median and mean is used as a descriptive statistics in a study. Statistical thinking includes application of school learning to the real life situations by criticizing, assessing results of the study given in context. In a further study, Ben-Zvi and Garfield (2004) stated that statistical literacy includes fundamental skills which are essential to understand statistical and probabilistic information. These basic skills are organizing, representing data, understanding different representation of the data, and language of statistics. Statistical reasoning includes skills of understanding, explaining and interpreting statistical findings. Similar with delMas(2002), Ben Zvi and Garfield (2004) defined statistical thinking as the reason and the way behind statistical investigations could be conducted and critical evaluation of the findings of these statistical investigations. From a different perspective, Shaugnessy (2007) made a distinction between those domains in a brief literature review about the statistical research, which is more compatible with the previous definitions mentioned at the beginning of this part. Statistical literacy mostly includes critical thinking skills about statistical and probabilistic information, which are crucial for both students and adults in order to make proper decisions in their life. On the other hand, statistical thinking includes thinking process of procedures of statistical information, statisticians and statistical reasoning is related to cognitive and developmental process in which students encounter difficulties, know or understand statistical concepts.

2.1.1. Theoretical Models for Statistical Literacy and Probabilistic Literacy

This part includes models related to statistical and probabilistic literacy. While two models (Watson' model and Gal's model) are presented with respect to statistical literacy, one model (Gal's model) is summarized in probability literacy. These models were summarized respectively.

2.1.1.1. Theoretical Models for Statistical Literacy

2.1.1.1.1. Watson's Statistical Literacy Model

In mathematics education, Watson's studies (Watson, 1997; Watson, 1998; Watson & Moritz, 2000a; 2000b) have been among preceding studies with regard to statistical literacy. Watson (1997) proposed a framework to assess students' statistical literacy skills. This framework suggests three tiers including hierarchy among each other. The first tier includes understanding of fundamental concepts in relation to statistics and probability. In this tier, students are expected to understand the basic statistical and probabilistic concepts such as data representation, percentage, measures of central tendency, measures of spread, probability in a mathematical context. The second tier is that students apply these statistical and probabilistic concepts beyond the mathematical context. In other words, students understand and interpret these concepts in social and scientific contexts. The last tier includes not understanding and interpreting statistical and probabilistic terminology in mathematical, social or scientific context, but also questioning claims in these contexts.

In further studies (Watson & Moritz, 2000a; 2000b), Watson and her colleagues used these hierarchical tiers to investigate students' understanding about the various statistical and probabilistic concepts such as chance, sampling, average, inference, and data representation embedded in media texts. Based on these previous studies, Watson and Callingham (2003) proposed a statistical literacy scale for students in grades 3 to 9. These scales consist of six hierarchical levels ranging from idiosyncratic to critical mathematical. The first stage, idiosyncratic, includes students' intuitive and idiosyncratic responses and their lack of engagement with the context. To illustrate, students were asked to interpret a probabilistic statement reported in a newspaper article, they gave idiosyncratic responses such as "good chance" and "hardly any chance". At the second stage, informal, students still have

lack engagement with context but they started to recognize single aspects of the statistical and probabilistic concepts. For example, they recognize the sample concept but not representative role of the sample. Third stage, inconsistent, includes more engagement with context by using qualitative language although they could not detect bias in the study reported in the media texts. At the fourth stage, consistent non-critical, students engage with the context; but, they could not still think critically. At the last two stages, critical and critical mathematical, students engage with the context critically by applying their mathematical skill to the context. However, quantitative reasoning and proportional reasoning is prevailing aspect of the last stage. For example, they could recognize bias in sample and misleading graphs. It is clear that the last two stages are highly related to third tier of Statistical Literacy Hierarchy while middle two stages are related to the second tier.

Combining the previous studies, Watson (2006) suggested a model indicating the relationships between statistical literacy's elements (see in Figure 1). The model includes major components of statistical literacy construct, which are closely related to each other. There are six main components grouped as context, mathematical/statistical skill, statistical and probabilistic concepts (sampling, average, data representation, chance, and variation), task format, task motivation, and literacy skills, explained respectively in below:

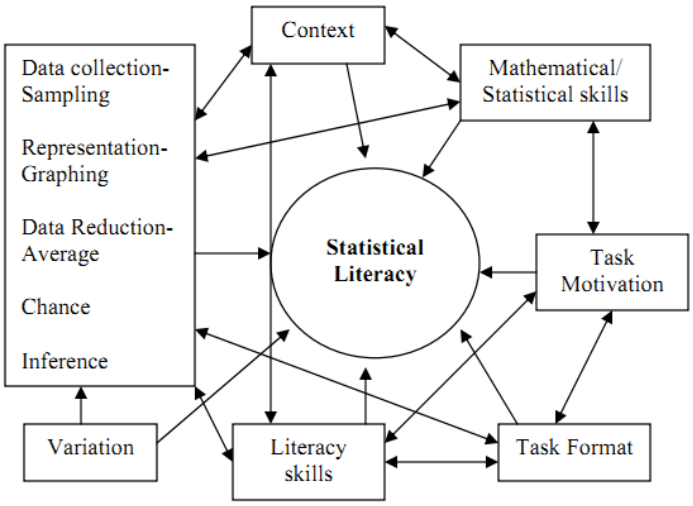


Figure 1. Relationship between elements of statistical literacy. Adapted from "Statistical Literacy- A Global Goal," by J. Watson, 2006, *Statistical Literacy at School: Growth and Goals*, p. 248. Copyright 2006 by Lawrence Erlbaum Associates, Inc.

The first component of the model is statistical and probabilistic concepts such as sampling, data representation, average, chance, and inference, which are intertwined each other. Students should be able to understand the meaning of sample concept and understand role of sample to make inference from sample to population, the need for representative samples, variability among the samples drawn from the same population, and bias in sampling. Regarding data representation, students should be able to interpret graphical displays and detect misleading graphs. Students also should be able to understand the meaning average concept in a given context, representative property of the measures of central tendency, and in which conditions they could represent the data the best. With regard to chance concept, it is important to use appropriate chance language, interpret risk and bias in chance situations, find several types of probabilities including single events, conditional events, and independent events. Regarding basic inference, students should be able to make inference from graphical representations such as prediction, hypothesizing, and comparison of two groups and critically evaluate the inferences reported in the media. Finally, regarding variation concept students should be able to understand

the meaning of variation, the relationship between sample and variation concepts, and recognize the variation in the graphs to compare two data sets.

The second component of the model is mathematical and statistical skill for middle grade students, which requires understanding the mathematical concepts such as proportion, rate, percentages, and part-whole relationship and statistical/probabilistic concepts such as average, probability of an event. Moreover, definitions of the terms are also crucial for mathematical and statistical skill.

The third component of the model is the context in which statistical and probabilistic concepts are embedded, which have critical role in statistical literacy. Watson (2006) identified three different contexts; namely, isolated contexts (rolling die), familiar contexts with school experience (survey in school), and unfamiliar contexts (media texts). To reach higher levels of statistical literacy, students should be able to engage with the unfamiliar contexts.

The fourth component of the model is literacy skills. Luke and Freebody (1997) proposed four elements for readers to be literate in the texts (as cited in Watson, 2006). Element 1 includes code-breaking such as appreciation of role of graphical displays in the text. Element 2 and Element 3 are mostly related to interpretation of the concepts in the text. For example, whereas Element 2 includes understanding different meanings of average concepts in the context, Element 3 includes creating possible meanings with the usage of statistical and probabilistic concepts such as data, sample, results and graphs in order to shape the meaning of the text. Element 4 includes critical questioning of the text.

The fifth component is that tasks could be presented in the forms of multiple choices and open-ended. Watson (2006) suggests that multiple choices tasks would be more useful for students since it allows students many alternatives rather than creating of an answer.

The last component is task motivation that involves students' dispositions such as attitudes, beliefs toward the context, or statistics and probability, which has potential effect in the development of statistical literacy.

2.1.1.1.2. Gal's Statistical Literacy Model

Gal (2004) proposed a model to assess adults' statistical literacy. The model involves two interrelated components of statistical literacy, namely knowledge and dispositional component. The knowledge component includes five knowledge bases: Literacy skills, statistical knowledge, mathematical knowledge, context knowledge, and knowledge of critical questions while interpreting the claims. The dispositional element involves a person's beliefs and attitudes toward statistics and probability and critical stance towards statistics and probability. These two main components are interrelated, not separate elements. Moreover, each of the knowledge bases interacts with each other. The knowledge and dispositional elements of statistical literacy are, respectively, described in detail.

Literacy skill is one of the knowledge bases for statistical literacy, which requires comprehension of prose and non-prose texts such as graphs, tables, and charts or oral texts. The second type of knowledge base is statistical knowledge, which includes "knowing why data are needed and how data can be produced, familiarity with descriptive statistics, familiarity with graphical and tabular displays, understanding basic notions of probability, and knowing how statistical conclusions or inferences are reached." (Gal, 2004, p.58). Gal (2004) mentioned about several key points for "knowing why data are needed and how data can be produced". First of the key points is to be aware of data reduction made by producer. Thus, it is essential to understand and appreciate the use of summary statistics such as mean, graphs despite of the existence of variation in the statistical investigation. One another key point is that adults should have knowledge of research design such as

experimental, survey, census behind the given data-related information. Moreover, the notions of representativeness and the method of sampling are significant factors to make inference from sample to population. Familiarity with descriptive statistics includes two key concepts, percent and central tendency measures since these concepts are commonly used in many texts. In this regard, it is essential that adults possess knowledge of measures of central tendency as summary statistics and in which conditions they would be the best representative of data set. Another statistical knowledge for statistical literacy is to be familiar with graphs and tables displayed in the media texts, which requires knowing the role of graphical and tabular displays in organization of data set and possible misleading errors or information created by authors. Understanding basic notions of probability includes to understand probability of events, risk estimation can be explained through different ways such as percentages, odds, ratios, or verbally and to be familiar with randomness and variation concepts in chance processes. The last statistical knowledge includes knowing how researchers reached the findings of the study or made inferences from sample to population and evaluating significance of the results. Further, it is useful to question whether obtained differences between the groups are large enough to come up with these conclusions or occurred by chance.

Mathematical knowledge is another important knowledge type to be statistically literate, which includes basic knowledge of mathematical procedures underlying statistical information and number sense. This basic knowledge includes calculating the basic terms such as the mean, percent, probability and being familiar with the derivation of these terms.

Third knowledge type, context or world knowledge, requires the world knowledge, background information, in the given context to interpret appropriately texts including statistical information since the statistical context makes sense to adults when combined with the numbers and context.

The last knowledge type for statistical literacy is knowledge of critical questions. Messages in the media may be misleading. Journalists, politicians, manufacturers, or advertisers may make biased reporting. Adults should question the reasonableness of claims, validity, and credibility of messages. They should ask themselves "sample worry questions" Gal (2004) identified in his study.

Another aspect of statistical literacy is dispositions including critical stance, beliefs and attitudes. Adults should have a tendency to hold a critical stance, which they adopt an interrogative attitude toward messages including misleading statistical and probabilistic information. Moreover, adults should have a positive belief and attitude in regard to being confident in exploring, interpreting and questioning statistical and probabilistic messages.

Gal (2004) emphasized that these interconnected dispositional elements are important factors that influence a person's statistical literacy beside five knowledge bases mentioned above. To maintain their critical stance, adults should have a belief that being critical about the claims, arguments involving data-based information is an essential personality in the society although they have lack formal knowledge in statistics or mathematics.

2.1.1.2. Theoretical Model for Probabilistic Literacy

2.1.1.2.1. Gal's Probabilistic Literacy Model

Gal (2005) proposed a model to develop adults' probabilistic literacy similar to the statistical literacy model. In other words, the model includes both knowledge and dispositional dimensions. Knowledge elements of the probabilistic literacy consist of five sub-dimensions: big ideas of probability, the terminology and language of probabilities, critical questions, context knowledge. Variation, randomness, and independence were accepted as big ideas of probability. Their connection with the

terms of stability, regularity, and co-occurance was also emphasized in the model. Figuring probabilities require understanding of the probabilistic statements and explanation of their thought and estimates about the probabilities of the events to the other people. For instance, familiarity with the concepts of conditional probability, and Bayes' theorem is one of the essence skills in order to understand the difference between P(A|B) and P(B|A) and so, detect errors in probabilistic statements . To be familiar with the language of chance is to realize that the likelihood of events can be explained in different forms such as verbal or numerical forms. Contextual knowledge requires understanding the role of the probability on different situations in the real life such as medicine, public policy, personal decisions, and technology. The last element of the knowledge base of probability is to be aware of critical question with the purpose of identification of errors, deficiencies, biases made by journalists, advertisements, or policy makers. To illustrate, it is important to be able to critically evaluate misleading graphs, correlation versus causation, practical significance versus statistical significance, or term of margin of error. In dispositional aspects of the probability literacy includes the elements of critical stance, beliefs and attitudes and these elements were described in a similar way with statistical literacy.

To conclude, while there are many differences among these models, all models strike a balance in view of which there is a need to educate students, teachers and the other people as statistically literate citizens keeping face with the chances in the world by the means of understanding and interpreting statistical and probabilistic statements and critically evaluating these statements, claims or arguments made in media. This study includes combination of those three theoretical models regarding statistical and probabilistic literacy. In the method section, it is explained in detail how these models were combined to investigate participants' statistical and probabilistic knowledge.

2.1.2. Statistics and Probability in Context

For the last decade, statistics and probability have an increased concern among the research community including, statisticians, cognitive psychologists, statistics, and mathematics educators. Studies on the development of the statistics through the years emphasized that statistics is a new domain of education and there is a need to raise generations thinking statistically to come up to the changes in the world, rather than focusing on computational procedures (Ben-Zvi & Garfield, 2008; Utts, 2003).

Many studies were conducted regarding the importance of statistics and probability in real life contexts (Gal, 2002; 2004; Scheaffer, 2003; Steen; 2001; Utts, 2005; Watson; 2004). Steen (2001) emphasized that statistical and probabilistic concepts such as chance, sampling, variation, inference, average, and graphical representation are situated in many social and scientific contexts. For example, headlines of the newspapers mostly includes numbers, charts and graphs describing risks, increases, changes in a specific area such as health, finance, or education on the basis of statistics. Steen (2001) also argued that statistics was introduced into school curriculum on the basis of quantitative literacy due to anxiety toward statistics in the society. In a further study, Scheaffer (2003) put more emphasize on the strong link between statistics and quantitative literacy in real life. Since statistics is prevailing on real life situations, it plays significant role on the development of quantitative literacy. For example, statistical thinking prevents people make sudden decisions with the appreciation of variation when a change occurred in any area. In addition, Scheaffer (2003) proposed several suggestions in order to develop quantitative literacy based on statistical thinking. School curriculums should be designed on the basis of quantitative literacy, which is closely related to statistics curriculum. Moreover, it is necessary that teachers evaluate students' understanding, interpretation of daily life examples with authentic assessments. Supportively, Watson (2004) proposed authentic problem solving tasks with integrating quantitative literacy with statistics for mathematics teachers. This problem solving

tasks are embedded in newspaper article, which requires to detect erroneous or unclear statements and read critically. These studies have common point with integrating numeracy with statistics in real life context, which indicates transition toward statistical literacy.

Gal (2002; 2004), however, argued that there is no consensus on the basic statistical knowledge required for adults to interpret and evaluate media texts. In one of the early studies, Joram, Resnick, and Gabriele (1995) investigated rational numbers in twenty-one magazines regarding children, teenagers, and adults. The analysis was made through the coding of frequency and type of rational numbers such as fractions, percentages, and averages. The findings of the study indicated that percentages are the most common concept in the magazines. The frequency of rational numbers, tables, or graphs is much more in adult magazines compared to children's and teenagers' magazines. They suggested that tasks encouraging students to interpret quantitative statements in newspapers be a goal for instruction rather than just perform computational procedures. Moreover, Utts (2003) identified several significant points to be considered by the citizens in order to understand and interpret real life contexts including statistical and probabilistic information. Understanding difference between cause-effect and correlation, statistical significance versus practical significance, and misleading language in the contexts, variability, and conditional probabilities are the most common points arising from the study. In a further study, Utts (2005) pointed out three noteworthy points that are misleading statistics in the media. The first, many articles may involve just relative risk, not baseline risk, which aims to influence people's decisions as if there has been a big difference with the treatment. The second important point is that some studies may not mention about how it was conducted. For example, it is essential to include the age of the participants in a study with the result of increase in woman's breast cancer risk since baseline risk for being breast cancer differs among the different age groups. The last significant point is that the risk reported in the media can be misleading because of confounding variables.

In line with the previous studies, Gigerenzer and Edwards (2003) mentioned possible confusions about the health risks reported in the articles patients and doctors encounter and remarked the importance of evaluation of probabilistic statements. He investigated the confusions in three topic; probabilities of single events, conditional probabilities, and relative risks. For example, the likelihood of particular events (e.g. “There is a 30% chance of rain tomorrow”, p.1) or conditional statements (e.g. “If a woman has breast cancer the probability that she will have a positive result on mammography is 90%.”, p.3) reported in the articles may include different meanings, or the ambiguous language.

In a more comprehensive study, Gal (2004; 2005) proposed, respectively, two models related to statistical literacy and probabilistic literacy. These models basically based on the phenomenon that adults should understand and have a sense of critical questions related to the sources and the meanings of the statistical and probabilistic statements in the media contexts. He proposed a comprehensive base for statistical knowledge by reviewing the studies regarding mathematics and statistics (see in the previous part) and identified statistical knowledge on the five basic knowledge; knowing why data are needed and how data can be produced, familiarity with basic terms and ideas related to descriptive statistics, familiarity with basic terms and ideas related to graphical and tabular displays, understanding basic notions of probability and knowing how statistical conclusions or inferences are reached (Gal, 2004, p.58) (see in the previous part). In addition to statistical knowledge, the framework includes mathematical knowledge, context knowledge, language skills, knowledge of critical questions, and dispositions, which are possible factors closely related with statistical literacy.

In conclusion, these studies pointed out four noteworthy points. The first noteworthy point is that statistical and probabilistic knowledge is crucial to understand, interpret, and evaluate media texts through critical thinking. The second

important point is that there is still no consensus on what knowledge are required to understand, interpret and evaluate statistical an probabilistic information reported in the media. Another noteworthy point is that media text could include ambiguous language and misleading information, which could lead people to make inappropriate decisions regarding their life. The last important point is that language skills, mathematical knowledge, context knowledge, and dispositional skills for critical investigation are possible factors required for understanding statistical and probabilistic information embedded in the media texts.

Utts (2003), however, argues that many citizens have not enough knowledge to interpret and critically evaluate the statistical information in most of the newspapers. In addition, more recently, Ridgway, Nicholson, and McCuster (2011) stressed the need for development of students' and teachers' statistical literacy and integration of statistical literacy with the school curriculums. In this circumference, many studies were conducted in order to examine students' and teachers' understanding statistics and probability in the context.

Studies related students' statistical literacy were focused on the investigation of students' statistically literacy levels (Aoyama &Stephen, 2003; Calligham; 2006; Watson, 1997; Watson; 1998; Watson & Callingham, 2003); interest in statistical literacy (Carmichael, Calligham, Watson, & Hay, 2009; Carmichael, Calligham, Hay, &Watson, 2010); longitudinal development of statistical literacy (Watson & Moritz, 2000b; Watson & Kelly, 2008) and effect of instruction based on statistical literacy (Doyle, 2008; Merriman, 2006). For example, Watson (1997), one of the early studies, assessed students' statistical thinking in the media contexts on the basis of graphical representation and sampling. The participants of the study consisted of students at Grades 6 and 9. The qualitative analysis of the data indicated that most of the Grade 6 students had difficulty in understanding the meaning of the pie chart. In the sampling task, most of the Grades 6 and 9 students could not make any criticism regarding the article. However, there were some Grade

9 students who understand and interpret the role of sample to make generalization in the given context. An interesting finding is that students who evaluate critically the question that requires higher level thinking could answer the questions which require the interpretation of the situation in the articles. However, this finding could not be generalized to the inverse relationship between two tiers. Supportively, in a different context (risk assessment), Watson (1998) conducted a study with Grade 9 students. Students were asked to interpret and critically evaluate four conditional statements reported in the newspaper article. Only 28% of the students appropriately interpreted the complex conditional statement. In a similar way, students had difficulty in Tier 2 and Tier 3 regarding the cause-effect relationship reported in the newspaper extract. Only 20% of the students could critically question this relationship.

From a different perspective with previous studies, Merriman (2006), furthermore, designed a lesson including media reports in order to investigate the effect of the Grade 10 students' development of statistical literacy. Pre-test and post-tests were implemented to the participants. The data were analyzed on the basis of SOLO taxonomy. The results of the study indicated that there is a statistically significant difference between the pre-test and post- test in terms of statistical literacy improvement. Although before and after the instruction students were at the multi-structural level, they were more advanced position in that level after instruction. Another finding was that there was no statistically significant relationship between mathematical ability and students' scores in statistically literacy test although there is statistically significant positive relationship between English ability and students' performances in statistical literacy. After the instruction, students had aware of misleading nature of the media reports and appreciated the importance such an lesson to be statistically literate although they thought that media reports require higher level thinking and are difficult to read. Teachers also expressed that they could prefer such a lesson based on media reports to the traditional method since it

allow students to communicate with each other, transfer their statistical knowledge to the real life situation, and be motivated in the instruction.

In general, Shaughnessy (2007) briefly reviewed the studies related to both students and teachers in statistics education and advocated that teachers have important role on learning of statistics and probability; thus, it is essential to examine teachers' knowledge and beliefs towards statistics and probability (Shaughnessy, 2007). Specifically, Stohl (2005) discussed problems arising from by teachers' beliefs, misconceptions, lack of knowledge and mathematics teachers' computation approach towards probability by summarizing studies related to probability in teacher education. Moreover, she addressed suggestions to prepare teachers to be effective in their educational efforts since teacher education programs have not still included professional learning and teaching of statistics and probability (Shaughnessy, 1992).

2.1.3. Research in Teacher Education Regarding Statistics and Probability in Context

Several studies about teachers regarding statistics and probability in context were conducted in line with suggestions in teacher education (Kvatinsky & Even, 2002; Roca & Batanero, 2006; Watson, 2001; Watson, Callingham, & Nathan, 2009). For instance, Watson (2001) developed an instrument, prepared based on knowledge types identified by Shulman (1987), to assess teachers' competence and confidence in teaching statistics and probability concepts**.** The study was conducted through a survey with 43 teachers in Australia (15 primary and 28 secondary teachers**).** There are three important points that arise from the article. The first main point is that sampling concept is less familiar to the teachers when compared to the concept of average. The second important point is that teachers have the lowest confidence in teaching odds and have the highest confidence in teaching data representation. However, high school teachers are more confident in teaching of statistical and

probabilistic concept than primary teachers. The third important point is that some teachers have difficulty in interpretation of the probabilistic statement in the context and most of the primary teachers reported that the concepts of sampling and odds could be difficult concepts for their students with supporting the findings of Watson (1997). The last significant point is that there is a statistically significance between high school teachers and primary school teachers about usefulness of media texts including statistical and probabilistic information.

Similarly, Kvatinsky and Even (2002), one of the studies related to teachers, proposed a framework for teachers' subject matter knowledge with respect to probability. It was asserted that teachers should understand the power of probability in real life, have knowledge of different forms of representations such as Venn diagram, tree diagram, and area model in order to understand probability, and understand that the nature of mathematics and probability have different aspects.

Contreras, Batanero, Diaz, and Fernandes (2011) conducted a research with 183 prospective primary school teachers in order to investigate their common and specialized knowledge. They were asked to find simple, compound, and conditional probability in a contextual problem represented in the two-way table and identify fundamental concepts behind problem. The findings of the study indicated that most of the participants have poor common and specialized knowledge in an open ended task including two-way table. They had difficulty in calculating simple, compound, and conditional probabilities of the events. Moreover, few of the participants had used mathematical symbols and applied computational procedures correctly. The researchers suggest that pre-service teacher education program need to be developed in probability education.

More recently, some studies also focused on the investigation of teachers' Pedagogical Content Knowledge (PCK) regarding statistics and probability. For example, Watson, Callingham and Nathan (2009) investigated middle school

teachers' pedagogical content knowledge in the pictograph task, which requires making beginning inference including uncertainty. Participants of the study consisted of 40 teachers taught in grade level 5 to 12 and involved a professional development project about statistics in Australia. The analysis of the interview protocols was on the basis of four non-hierarchical components; *Recognizing Big Ideas, Anticipates Student Answers, Employs Content-specific Strategies,* and *Constructs Shift to General* (Watson et. al, 2009, p.567). The analysis of the data revealed four main results. The first result was that most of the teachers could recognize the concept of uncertainty as a big idea behind student task. The second result was that 9 teachers revealed high level of PCK whereas 14 teachers displayed medium and 17 teachers revealed low level of PCK. The third significant result was that teachers in the medium group reflected inconstant responses in the different component and some teachers in the low level of PCK made immediate decisions regarding appropriateness of students' answers and did not questioned them. They also pointed out that content knowledge is fundamental building block for the other component even though they are not hierarchical.

In the further study, Watson and Nathan (2010a) conducted a study to examine 40 middle school teachers' PCK through media task regarding sampling. This media extract about legalization of marijuana had been previously asked to the students and three students' responses to the task were presented to the teachers. The analysis of the interview data elaborated by integrating sub-levels (Code 0, 1, and 2) into each component was done on the basis of four components mentioned in the previous research. In the first and second component of the PCK, more than half of the teachers were scored as Code 2, which was related with teacher confidence in the context knowledge regarding sampling concept. Another finding was that teachers have more confident in suggestion of appropriate strategies regarding students' responses compared to the construction a change to the general concepts. More specifically, they had difficulty in application of the sampling problem to wide statistical contexts and association it with the other concepts.

In a similar way, Watson and Nathan (2010b) carried out a research study to investigate teachers' PCK on the basis of contextual problem including two-way table and conditional probability concept. 29 middle school teachers were participated to the study and interviewed on the basis of three main points; big ideas behind the problem, suggestion of students' inappropriate and appropriate responses to the task, respond to the sample students' answers to the task. Analysis of the interviews was conducted by using previously hierarchical rubric previously developed. Most of the teachers recognized partially big idea behind the problem and stated as proportional reasoning and relationship between two variables. Moreover, only 9 of the teachers could formulate appropriate and inappropriate students' responses toward the task. When two students sample responses were presented to the students, few teachers' answers were in Code 3 and employed content-specific strategy for students' responses and appreciated the complex nature of the language of conditional statements and inverse relationship between them. In addition, almost half of the teachers proposed appropriate mathematical solution to the problem which requires understanding conditional probability, relationship between two variables, percentage, or fractions. They suggested that teacher have PCK to encourage students think critically and school curriculums include such authentic tasks to achieve this goal.

In a similar study, which makes a relationship between two way tables and conditional probability, Roca and Batanero (2006) carried out a descriptive study with 65 pre-service teachers to investigate their understanding of conditional probability. They were asked a contextual problem presented in two-way table which includes four questions about simple, compound and conditional probability. Frequency of the responses toward four questions indicated that although majority of the students could calculate simple probability, many errors were made in the other questions requiring understanding of compound and conditional probability. These errors as a semiotic conflict were grouped. Future teachers had confusion

mostly in inverse relationship between conditional statements, difference between conditional probability and joint probability. The study indicated teachers' difficulties in reading two-way table and conditional probability in the contextual problem. In this regard, researchers of the study underlined the necessity for teacher education in conditional probability and probabilistic language to nurture students statistically literate.

In line with these studies, Watson (2011) proposed teachers to use a media article including conditional statements in order to develop students' understanding of conditional probability by using two-way tables. In another research to develop teachers' quantitative and statistical literacy, Watson and Moritz (2002) set up a quantitative literacy project in order to educate pre service teachers who encourage students to be quantitatively literate members of the society with using quantitative literacy tasks. The participants of the study consist of 40 pre-service teachers. In this project, a website including newspaper articles with different contexts such as sports, transportation, health was created. In the website, pre service teachers chose a newspaper article related to chance and data, or numeracy and made comments, discussions related to the usefulness of the selected article in the classroom and submitted lesson plans developed through using the newspaper article. The findings indicated that pre-service teacher could be educated via these tasks and they believe usefulness of them during their teaching practices.

To summary, the focus of these studies was to investigate mainly teachers' content and pedagogical knowledge of statistics and probability. They have difficulty in recognizing the big idea behind the contextual problems. Moreover, they have low confidence regarding probability and difficulties in understanding contextual probabilistic statements, especially conditional probability statements. Finally, it is suggested that teachers could be trained through contextual tasks such as media reports to develop their knowledge in statistics and probability.

2.1.4. Research Regarding Statistics and Probability in Turkey

Ulutas and Ubuz (2009) investigated research trend in mathematics education in Turkey with content analyses of 129 articles from 2000-2006. Analysis of the articles indicated that most studies were about elementary students and pre-service teachers. Another finding was that number and geometry learning domain are one of the most common points of the articles. The researchers underlined the necessity for studies regarding statistic and probability, and assessment and evaluation.

In this regard, some studies about students and teachers in learning and teaching statistics and probability are presented. With respect to teachers, Bulut (2001) investigated 125 prospective secondary mathematic teachers' performances in the probability problems. The results of the quantitative study revealed that they could not recognize the fundamental probabilistic concept behind the problems and they especially have difficulty in solving non-mutually exclusive events. They concluded that prospective teachers have not enough computational skills in the probabilistic problems. In the further study, Bulut, Yetkin and Kazak (2002) examined senior prospective secondary mathematics teachers' achievement in probability, attitudes toward probability. Quantitative analysis of the data indicated that males' probability achievement were statistically significant compared to that of females while females' attitude toward probability was statistically higher than males. There was no statistical significant difference in attitudes toward probability in terms of gender. Another main finding was that there was statistically significant relationship between males' achievement in probability and attitudes toward probability whereas there is no statistically significant relationship for females.

With regard to students, more studies in statistics and probability domain were conducted in Turkey compared to teachers. In one of the studies regarding students, Ucak and Akdogan (2009) investigated Grades 6 to 8 students' understanding average. Participants of the study consisted of 18 student for each grade level. Most

of the students understood the average concepts as arithmetic mean rather than appreciate it as a representative value. In a similar way, Akkas (2009) conducted a study with middle school students. This mixed research consisted of thirty participants in order to examine their statistical thinking on the four levels; idiosyncratic, transitional, quantitative, and analytical proposed by Mooney (2002). The researcher concluded that majority of the students were in the levels of transitional and quantitative and there was no statistically significant change in students' statistical thinking levels across grade levels.

With regard to probabilistic concepts, Memnun (2008) conducted an experimental study to investigate the effect of active learning method in teaching permutation and probability concepts to the eight grade level students. The study included 197 (90 students for experimental group, 107 students for control group) eight grade students. In general, the results of the study indicated that active learning has statistically significant positive effect on students' achievement in the permutation and probability.

In a comprehensive study regarding learning in probability, Kazak (2008; 2009) reviewed the studies related to high school and elementary students' difficulties and misconceptions in learning of probability respectively and presented suggestions made by the previous researchers. Elementary students approach to the probabilistic situation intuitively on the basis of personal experiences, which could contradict with probability theories. One of the suggestions is that students be encouraged to discuss their ideas and communicate with each other by designing social environment. Another suggestion is to support the integration of technological software such as Probability Explorer and Tinkerplots so that students have ability for test and validate their intuitive strategies in the technological environment. Moreover, Kazak (2008; 2009) advocated the need for teachers' knowledge regarding the students' misconceptions or difficulties in probability.

However, it is clear that there is not enough study to investigate statistics and probability in context, especially regarding teachers to nurture students as a statically thinkers in the society.

2.2. Critical Thinking

2.2.1. Conceptualization of Critical Thinking

This part includes the historical sketch of the debate about critical thinking. In one of the early studies, Dewey (1933) as a pioneer in the research area of thinking defined good thinking as reflective thinking. Reflective thinking refers to "active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends" (Dewey, 1933, p.9). It involves in doubtful, ambiguous states that require mental difficulties and acts such as conscious and effortful testing and inquiring on the basis of evidence. For example, a person who makes decisions or conclusions immediately without considering the premises it builds upon might not be enough critical about his or her ideas. To think reflectively, willingness to survive for handling the difficulties is also essential. In other words, Dewey argues that reflective thinking as an educational aim requires not only skilled method but also attitudes.

Critical thinking was conceptualized by various researchers in terms of philosophical, psychological, cognitive perspectives and these conceptualizations mostly rest on the Dewey's works about the way human thinks. However, there is still no consensus on the conceptualization of critical thinking although they share common points regarding critical thinking, which is elaborated in the following.

From the perspective of philosophy of education, firstly, Ennis (1985), one of the pioneering studies from philosophical perspective, made an expanded definition of

critical thinking as “reflective and reasonable reflective thinking that is focused on deciding what to believe or do” (p. 45) and a combination of abilities (judge the credibility, analyse arguments, draw conclusions, define terms or concepts, etc.) and dispositions (be open-minded, be willing to examine reasons of the statements; to try to be informed and insightful, etc.).

Paul (1984) conceptualized critical thinking in terms of two different senses; a weak sense and a strong sense. Weak sense critical thinking involves a set of technical skills without extrinsic to the personal characteristics whereas strong sense critical thinking skills are “ultimately intrinsic to the character of the person” (Paul, 1984, p.5). Paul suggested that weak sense critical thinking be replaced with strong sense critical thinking because he argues that critical thinking involves not just pure skills’ mastery or utilization but also intellectual traits with avoidance of egocentrism, sociocentrism, and self-deception. In the extended conceptualization of critical thinking, Paul and Elder (2007) characterized critical thinking in terms of three aspects; intellectual standards, intellectual traits, and question of element of thoughts. Firstly, intellectual standards include relevance, accuracy, precision, clarity, depth, and breadth, which are implemented to thinking to reason quality of the problem or information. Secondly, intellectual traits consist of intellectual humility, intellectual autonomy, intellectual integrity, intellectual courage, intellectual perseverance, confidence in reason, intellectual empathy, and fair-mindedness (p. 21). Finally, purpose, questions, information, inferences or conclusions, concepts, assumptions, implications or consequences, and points of view constitute of elements of thought. People who think critically implement intellectual standards to the elements of their thoughts with purpose of development of intellectual traits.

From a different perspective, McPeck (1981) argued that critical thinking should regarded as a subject specific skills rather than generalizable because it is not possible to apply a single critical thinking skill to the diverse specific content areas,

which requires different subject matter knowledge. McPeck exemplified his argument as "the ability to identify assumptions" in mathematics require different skills from those of a political or scientific debate. To teach students what is assumption cannot be translated to the specific subject matters in order to explore the underlying assumptions of that subject matter. Moreover, he offered conceptualization of critical thinking as a combination of tendencies and skills to be involved in a specific subject matter with reflective skepticism (as cited in Siegel, 1988).

Siegel (1988) agrees with McPeck (1981) from a point of view that specific knowledge is necessary to engage in a subject matter. However, he also argued that both logical knowledge and specific subject matter knowledge is necessarily required for critical thinking. For example, knowing what is assumption or valid argument as logical knowledge helps people to engage in a specific context by means of subject specific skills. He conceptualized critical thinking as:

> "…Critical thinking is best conceived, consequently, as the *educational cognate* of rationality: critical thinking involves bringing to bear all matters relevant to the rationality of belief and action; and education aimed at the promulgation of critical thinking is nothing less than education aimed at the fostering of rationality and the development of rational persons"(Siegel, 1988, p.33)

Siegel (1988) identified two components of critical thinking, namely, reason assessment component and critical spirit. The first component is that reason assessment, which requires assessing reasons and one's ability to support beliefs, claims, and actions in an appropriate way. Thus, subject specific and subject neutral principles are required for reason assessment in specific contexts. Assessing reasons is not sufficient itself and it is also necessary that students have dispositions to do so, which brings out the second component of critical thinking. The second component is critical spirit or critical attitude, which involves habits of mind to seek evidence or reasons; disposed to seek evidence or make judgment on the basis of reasons.

From psychological point of view; Halpern (1998) conceptualized critical thinking as transferable skills on the basis of a developmental model, which aims to enhance teaching and learning critical thinking. The proposed model included four components; (a) dispositions required for conscious and effortful thinking; (b) skills required for critical thinking, (c) training for promotion of transferable skills through structures aspects of contexts, and (d) a metacognitive thinking process. The skills required for critical thinking consisted of five main hierarchical skills; namely verbal reasoning skills, argument analysis skills, scientific thinking skills such as hypothesis testing, thinking regarding likelihood and uncertainty, and decision making and problem-solving skills. Halpern also suggested that transferability of critical thinking skills to the real world should be aim of the instruction and stated that it is worth examining naturally what people thinks when reading a newspaper or making decision regarding their career.

Although there is a clear agreement among researchers regarding critical thinking components as a combination of skill and dispositions, there is dispute on whether critical thinking is conceived as a set of general skills and dispositions and skills (Ennis, 1985; Paul, 1984) or subject specific skills and dispositions (McPeck, 1981).

More recently, some authors (Facione, 1990; Jones, 1995; Gibson, 1995) favored a different view beyond the dispute about distinction between general and subject specific skills and dispositions. They hold view that general skills and dispositions of critical thinking have wide range area for application to the various subject matter domains with the appreciation of role of familiarity with specific context as well. This research was conducted on the basis of this view by combining the general skills of critical thinking and subject specific knowledge, statistical and probabilistic knowledge.

Facione (1990) reached a consensus on the characterization of critical thinking in a long term project with experts widely known and experienced in different domains

such as philosophy, education, social sciences, and physical sciences. Critical thinking conceptualized at two dimensions; cognitive skills and affective dispositions, which are explained in detail in the following part.

Supportively, in a national survey, employers, policy-makers, and educators reached a consensus in regard to constitutes of critical thinking for college graduates; namely, skill dimension and dispositional dimension (Jones, Hoffman, Moore, Ratcliff, Tibbetts & Click, 1995). In the dimension of critical thinking skills, they had consensus on five skills; namely, interpretation (categorizing, detecting indirect persuasion, and clarifying meaning), analysis (examining ideas and purpose, detecting and analyzing arguments), evaluation, inference (collecting and questioning evidence, develop alternatives and hypotheses, and drawing conclusions), presenting argument, and reflection, which correspond to self-regulation skill in the conceptualization of Facione (1990).

To conclude, various researchers defined the term of critical thinking from different perspectives. In this study, we favor of consensus definition proposed by American Philosophical Association (Facione, 1990) on the basis of integration of critical thinking skills into mathematics education as a specific subject matter domain.

2.2.2. Theoretical Model of Critical Thinking

Facione (1990) reported a long-term project, supported by American Philosophical Association, to develop a consensus on the conceptualization of critical thinking. The aims with the conceptualization of critical thinking were to develop critical thinking assessment tool, prepare critical thinking program for college level students, and facilitate integration of critical thinking into K-12 curriculum and instructional environments. Facione was pioneer of the study, which involved forty six experts widely known and experienced in critical thinking, instruction and theory. Until they reached a consensus on the characterization of critical thinking,

the study continued two year in which they reviewed the others' comments or classifications of critical thinking organized and presented them by Facione. At the end of the project, the experts conceptualized critical thinking in two dimensions: cognitive skills and affective dispositions. The experts identified five cognitive skills of critical thinking; interpretation, analysis, evaluation, inference, explanation, and self-regulation. However, they highlighted that listing of the skills is not intended to indicate any order or hierarchical. Whereas some of them may overlap to each other, some skills may be prerequisite for the others in different contexts. Consensus on the descriptions of each cognitive skill and sub-skill was briefly presented by associating with the further studies of the author in the following.

Interpretation refers to making sense of evidential (facts, experiences, statements), conceptual (ideas, theories, ways of seeing the world), methodological (strategies, techniques, approaches), criteriological (standards, benchmarks, expectations), or contextual (situations, conditions, circumstances) considerations and expression of the meaning and importance of them (Facione, 2011a, p.17). Interpretation includes sub-skills of *categorization*, *decoding significance*, and *clarifying meaning*.

Categorization means to classify the information, data or findings, to construct categories on the basis of criteria in order to describe the meaning in a articulable way. Decoding significance refers to notice the intent, purpose of what is being communicated such as authors' indirect intentions and purposes or to distinguish the main idea from the extraneous ideas (Facione, 2011b). Clarifying meaning is to restate or paraphrase the given statement; to make explicit ambiguous or vagueness language of the statement. Facione (2011a) elaborated this description and also mentioned about vagueness or ambiguity of the terms in the texts. Whereas vagueness refers to unclear meaning of the language in a given context, ambiguity refers to multiple meanings of the language in the context, which leads to misunderstanding or doubtful meanings of the statement. A person with ability of clarifiying meaning could resolve these problems by contextualizing, clarifyig

intent, negotiating, qualifying, stipulating (Facione, 2011a, p.46). Clarification of the meaning is important since people could easily believe to what is intended to express in a given context when they have lack knowledge of context (Facione, 2011a).

Analysis includes identification of implicit and explicit relationships among representations such as descriptions, statements, or concepts expressed to communicate with the people. Analysis includes sub-skills of *examining ideas*, *detecting arguments*, and *analyzing arguments.* Examining ideas refers to identify closely related statements, to compare and contrast the statements or concepts, to make complicated situation simpler (Facione, 1990). Detecting arguments means to determine if authors' claim includes the reason(s) supporting it, or to identify unstated premises of a claim (Facione, 2011a; Facione, 2011b). Analyzing arguments is to identify background information for the reasons supporting the claim in a given argument (Facione, 1990). Facione (2011a) also suggested mapping for analyzing arguments, which indicates one' reasoning process through reasons, intermedieate claims, and reached conclusions.

Evaluation refers to assess credibility, truthfulness, strengths of the statements or claims. Evaluation includes sub-skills of *assessing claims* and *assessing arguments*. Assessing claims includes to realize the factors affecting the credibility of the statements, to determine if the claim is true or false (Facione, 1990). To assess the claims, one should be able to intergorate the sources of information if it is trustworthy or not, identify if the claims can be confirmed with the other information in the context, or recognize statements in a given context which contradict with each other (Facione, 2011a). Assessing Arguments is, given an argument, to judge its logical strengths and confidence level or to determine if it was based on false, biased premises (Facione, 1990). In the further study, Facione (2011a) identified four hierarchical tests to evaluate arguments. The first test, truthfulness of premises, includes deciding if the reasons of the claims are true. The

second test, logical strength, is to recognize if the reason of the claims is true, claim should quite likely be true on the basis of true reasons. The third test, relevance, requires judgement of which truth of the claim rely on the truth of reasons. The last test, non-circularity, includes judgement of which one should not use truth of claims for support the truth of the reasons. Moreover, Facione (2011a) emphasized the need for evaluation of deductive and inductive reasoning on the basis of logical strength. In evaluating inductive arguments, one should assess generalization of the reported findings on the basis of sample characteristics, data collection method, sample size, sample representativeness. Moreover, one should be able to distinguish correlational relationships from causality and detect exaggerated generalizations and numbers or statistics.

Inference refers to seek information needed to be addressed, project alternatives and draw conclusions. Inference includes sub-skills of *querying evidence, conjecturing alternatives*, and *drawing conclusions*. Querying evidence is to determine the background information need to be addressed. Conjecturing alternatices includes to propose alternative(s) to resolve a problem. Drawing conclusions refers to educe new conclusions by using relevant information (Facione, 1990).

Explanation refers to examine one's research with a large perspective, to present or write one's research results, procedures persuasively and consistently, and offer arguments regarding a claim in the research (Facione, 1990; Facione, 2011b). Explanation includes sub-skills of *stating results*, *justifying procedures*, and *presenting arguments*. Stating results is to state one' research findings in a coherent way. Justifying procedures refers to " …present the evidential, conceptual, methodological, criteriological and contextual considerations which one used in forming one's interpretations, analyses, evaluation or inferences, so that one might accurately record, evaluate, describe or justify those processes to one's self or to others, or so as to remedy perceived deficiencies in the general way one executes those processes" (Facione, 1990, p.18). Presenting arguments is to explain reasons

for holding a view, to argue and write one' thoughts, objections on an issue (Facione, 1990).

Self-Regulation refers to check own thinking and decision-making process, which requires judging reflectively and self-consciously (Facione, 2011a). Self-regulation includes monitoring the all processes (interpretation, analysis, evaluation, inference, and explanation) of critical thinking as a recursive process. People with strong critical thinking evaluate own strenghts and weaknesses by asking themselves questions of "Do I have any unclear status regarding the issue", "Is my evidence is enough good", or "What is the missing point I overlooked?" (Facione, 2011b) . Self-regulation includes sub-skills of self-examination and self-correction. Self-Examination is to make objective self-assessment of process of thinking, to reread the sources to make sure that one has not overlooked important information. Self-Correction is to make corrections or revisions when self-examination uncovers mistakes or inadequacies (Facione, 1990).

The dispositions toward critical thinking were listed in two main topic; the *approaches to life and living in general* and *approaches to specific issues, questions or problems* (Facione, 1990, p. 25). Approaches to life and living in general are exemplified as having curiosity to be well-informed, readiness to make use of critical thinking skills, being trustful in own reasoning process, being open-minded to the different ideas, opinions, or being flexible in thinking process by considering a variety of alternatives. On the other hand, being clear in their statement, opinions, or questions, being systematic and decisive in working with difficult situations, being careful regarding the focus of the situation are among the examples of approaches to specific issues. In the further study, Facione (2011a) grouped these approaches under seven main dispositions; truth-seeking, open-minded, analytical, systematic, confident in reasoning, inquisitive, and judious (p. 31).

2.2.3. Learning and Teaching Critical Thinking

Critical thinking has been argued among the researchers since 1940s (Paul, 1984). There, however, has been considerable interest in critical thinking as an educational goal, especially in the 1980s (Ennis, 1993; Halpern, 1997; Kennedy et. al., 1991; Ten Dam & Volman, 2004). One of reasons behind such an interest was that there is need for qualified and critically thinking citizens for political election to actively participate in the society. Another was that there have been rapid changes in the world people need to keep pace with. The other reason was that in the 1980s, American students could not present sufficient thinking ability. These reasons yielded researchers to make a movement toward teaching of critical thinking as an educational goal (Kennedy et. al., 1991).

In one of the early studies, Siegel (1988) regarded critical thinking as an educational goal since it orients us to decide what to teach, how to teach, how to design educational activities, or how to treat students in the classroom environments. In other words, critical thinking constitutes a guideline for educational activities, assessment these activities, educational decisions. Moreover, characteristics of ideal person that educator try to nurture are conceptualized on the basis of critical thinking. In this circumference, Siegel (1988) conceptualize critical thinking as a regulative ideal in the education as stating "it [regulative ideal] aids us in evaluating and choosing between, alternative curricula, teaching methods, theories, policies, and practices" (p. 47).

Halpern (1997) also advocated that critical thinking should be regarded as a compulsory ability for people in 21st century even though its importance has always been emphasized among the researchers. The reason behind Halpern's claim is that every people is required to make decisions regarding economy, political election, usage of nuclear power or education etc., which have possible effect on the life of future generations.

Although there is a clear agreement among researchers regarding critical thinking as an educational aim, there is dispute on whether critical thinking is conceived as a set of general skills and dispositions and skills (Ennis, 1985; Paul, 1984) or subject specific skills and dispositions (McPeck, 1981). The debate between researchers regarding subject-specific versus subject-neutral arises the question of to what extent critical thinking skills could be transferable to the other domains (Ten Dam & Volman, 2004).

In this regard, Brown (1997) advocated that students' critical thinking could be developed through serious matters that they are engaged with in the context of a specific subject matter. In 1998, Halpern also addressed the issue of transferability of critical thinking across different domains and advocated that students can learn critical thinking with the appropriate educational design. To achieve this goal, critical thinking should be transferable to real-world or out of the school settings in various subject matters. Kennedy et. al. (1991) also remarked that teaching of critical thinking.

However, some authors (Facione, 1990; Jones et. al., 1995; Gibson, 1995) favored a different view beyond the dispute about distinction between general and subject specific skills and dispositions. They hold view that general skills and dispositions of critical thinking have wide range area for application to the various subject matter domains with the appreciation of role of familiarity with specific context as well. Regardless of learning and instruction of critical thinking, Facione (1990) made various recommendations in their report regarding conceptualization of critical thinking. There are three main recommendations arising from the report. The first recommendation was that learning and instruction of critical thinking foster both students' critical thinking cognitive skills and dispositions since critical thinking aids students to deal with educational, personal, and social concerns. The second recommendation was that critical thinking should also a fundamental goal of elementary education. Therefore, all grade levels in K-12 curriculum should be

taught for critical thinking by integrating it into specific subject matters. Another recommendation was that to integrate critical thinking into K-12 curriculum, teacher educators should model critical thinking and foster teacher candidates' confidence in teaching critical thinking.

With regard to development of critical thinking of students and teachers, Ten Dam and Volman (2004) reviewed about 55 studies in the literature of critical thinking by presenting different approaches toward learning and teaching of critical thinking. They argued that critical thinking should be regarded as a competence to be critical citizenships in the society beside as a higher-order skill. They conceptualized the critical thinking in terms of this perspective by linking the learning how to think critically with social constructivist approach, which supports the learning as an intrinsic social process. Students could acquire the competence of thinking critically by critical participation in the social practices. Moreover, learning context should make sense for students and encourage them to participate critically in that social practice. In a similar way, Williams (2005) proposed improvement of citizens' critical thinking in the society and stressed the important role of students and teachers to achieve this goal. He gave attention for the importance of critical thinking in all disciplines in education and the role as an inseparable part of teacher education. Development of students' critical thinking depends on future teachers' improvement in critical thinking. Thus, critical thinking should be emphasized in the teacher education programs by creating learning environment in which teachers have opportunity to question situations, claims, views or ideas in the news or societal contexts. These studies support Kennedy et. al. (1991)'s study in which they considered teaching critical thinking and transferability of the critical thinking to the other subject matters are curial issues. In addition, they suggested that experts in the area of critical thinking have an agreement regarding the usage of real-world contexts, problems or popular media texts such as newspaper articles, magazine articles, advertisements or television programs in the classroom environments in teaching of critical thinking.

Paul, Elder and Bartell (1997), one of the early studies regarding teacher education in critical thinking, reported how pre-service teachers in teacher education programs in California were trained regarding teaching critical thinking and problem-solving skills in elementary and secondary schools. The project was based on interviews with faculty members in private and public colleges and universities. The results of the study pointed out that most of the faculty members did not have a comprehensible understanding of critical thinking and how to integrate it with instruction. On the contrary, faculty members who had been trained through a professional development program related with critical thinking gave more detailed and appropriate responses regarding their approaches toward critical thinking in the classroom environment. Moreover, they made various recommendations with respect to development of critical thinking in teacher education programs. One of the recommendations was that subject matters such as mathematics, physics should focus on critical thinking so that prospective teachers have ability of critical thinking about learning domains in the school curricula. The second recommendation was that teacher education programs should aim to teach prospective teacher how to design a learning environment supporting critical thinking. However, before achieving this goal, teachers must learn to think critically in real life and their professional subject domain and appreciate the role critical thinking in education. It is impossible that teacher with lack of critical thinking could improve students' critical thinking as future citizens to keep pace with the changes in the world of 21st century. In this circumference another recommendation was that there is a need for assessment of prospective teachers' critical thinking skills and knowledge and analyze the extent to which exams prepared in teacher education programs assess critical thinking skills and knowledge.

In this circumference, researches regarding teacher education in critical thinking focused on the teachers' perceptions of critical thinking and teachers' integration of specific methods into classroom environments to develop students' critical thinking.

For example, in one of the studies regarding teachers' perceptions, Innabi and Sheikh (2007) examined secondary mathematics teachers' perceptions of critical thinking and improvement in their perception after 15 years of educational reform in Jordan. The participants of the study consisted of 24 teachers selected from twelve secondary schools in 1988 and 23 teachers from the same school in 2004. Data were collected through interview with the participants on the basis of three main questions: understanding of critical thinking concept, role of critical thinking in mathematics learning, and teaching strategies to develop students' critical thinking. The findings of the study indicated that teachers could not reflect comprehensive understanding of critical thinking. Half of the teachers have difficulty in offering a teaching situation supporting critical thinking development. Another finding was that three quarter of the teachers was aware of the teaching critical thinking skills; but, most of them could relate critical thinking skill with mathematics learning and teaching. The third finding was that teachers suggested different strategies for improvement of critical thinking. However, most of them were not closely related to critical thinking; mostly related to general teaching strategies such as discussion and collaborative learning. The last main finding arising from the study was that there were no significant change in perception of critical thinking between years of 1988 and 2004. In this circumference, they recommended that teacher education programs improve understanding of critical thinking and train teachers as to teaching critical thinking skills. Supportively, Allazi (2008) found similar findings regarding perception of critical thinking with secondary school social studies teachers.

With regard to development of teachers' critical thinking, Recalde (2008) examined the relationships between technology integration, critical thinking skills on the basis of teaching practices. The study was conducted with 36 middle school mathematics teachers, 12 teachers for each grade of 6^{th}, 7^{th} and 8^{th}. Three instruments were used to examine the aim of the study; California Critical Thinking Test (CCTST) developed by Facione and his colleagues (2002), Educational Technology Survey Test (ETS), and Survey of Enacted Curriculum (SEC). The CCTST included six

scores; overall score of critical thinking skills and five subscales of analysis, inference, evaluation, deduction, and induction. One of the results arising from the study was that teachers indicated high level critical thinking skills for each five sub-scales. Another result was that teachers' critical thinking levels did not significantly changed across grade levels. The third result was that there was negative correlation between technology use (calculators, math games, graphing tools etc.) and critical thinking skills, which would be due to the fact that traditional methods such as memorization and repetition were mostly preferred among middle school teachers (Recalde, 2008).

From a different perspective, Osana and Seymour (2004) integrated critical thinking into statistics education and designed a rubric on the basis of argumentation theory and statistical reasoning in order to evaluate pre-service teachers' critical thinking. The rubric included three main components; conceptions and use of evidence, conceptions and use of research as a decision-making tool, and considering alternative perspectives (Osana, 2004, p. 476). They implemented a cognitive apprenticeship intervention during five weeks of a course about school community and society to measure pre-service teachers' critical thinking about complicated educational issues before and after the intervention. Intervention included three phases; modeling, coaching and scaffolding and fading. In the modeling phase, students were introduced by instructor a discussion piece and asked to discuss and evaluate the information in the discussion piece. In the coaching phase, students worked in the groups about new discussion pieces, which contained reasoning fallacies. They worked collaboratively with discussion with the other students. In the last phase of intervention, scaffolding and fading, students presented their discussion pieces and collaboratively critiqued each other reasoning in a whole-class discussion. They were evaluated before and after intervention on the basis of three criteria. Qualitative analysis of the data indicated that after the intervention, students improved their thinking skills as they recognize the evidences and judge their qualities in a text or use research to make decision about a complex issue.

To conclude, critical thinking should be a primary goal of education. There is need for learning and teaching for critical thinking to be as efficient citizenships in the society. Critical thinking could be integrated into a specific subject matter in realistic situations. Moreover, in the overview of 56 studies in critical thinking, Cotton (1991) pointed out that most of the studies were conducted with regard to elementary school students' development of critical thinking. Another interesting finding was that training teacher in teaching critical thinking is a significant factor that affects students' achievement. Kennedy et. al. (1991), in the review of the studies in relation to critical thinking, also stated that most of the researches proposed working within groups, cooperative learning, and questioning methods widely to develop critical thinking. In addition, they manifested the agreement among the researchers that knowledge of subject matter in which people engage with the contextual tasks have important role on their thinking performances.

2.2.4. Research Regarding Critical Thinking in Turkey

In a theoretical study, Seferoğlu and Akbıyık (2006) overviewed the studies about critical thinking. To participate actively in the society, students' critical thinking skills should be improved in various situations. To develop critical thinking skills, students could be encouraged to interrogate, or question though open-ended problems and solve problems with different strategies. Moreover, it is necessary that teachers provide opportunity students to transfer critical thinking skills to different contexts. Thus, there is a need for in-service and pre-service training for teachers in the critical thinking.

In this circumference, studies conducted in Turkey, mostly focused on investigating relationships between factors such as gender, critical thinking dispositions, experience, grade level in different subject domains such as mathematics (Türnüklü & Yesildere, 2005), social sciences (Demirkaya, 2003), Turkish (Türkmen-

Dağlı,2008), physics (Şengül & Üstündağ, 2009), foreign language education (Tufan, 2008). For example, in mathematics education, Türnüklü and Yesildere (2005) investigated pre-service elementary mathematics teachers' critical thinking dispositions with a descriptive analysis. The participants of the study consisted of 277 pre-service elementary mathematic teachers in the third and fourth year. Data collection tools were California Critical Thinking Dispositions Inventory (CCTDI) developed by Facione and his colleagues (1995) and Mathematical Critical Thinking Problems (MCTP) developed by the researchers. The results of the study indicated that in general pre-service mathematics teachers have positive critical thinking dispositions. However, critical thinking dispositions are low in the subscales of self-confidence and truth seeking. In MCTP, total score was not high in general when they were asked to give appropriate answer to five mathematical problems.

Güven and Kürüm (2008) investigated relationship between pre-service teachers' learning styles and critical thinking dispositions. The participants of the study included 251 teacher candidates who enrolled to the Faculty of Education in a university. The results of the study indicated that there is significant relationship between learning styles and critical thinking dispositions. In addition, according to California Critical Thinking Dispositions Test results, pre-service teachers have low critical thinking dispositions in general, supportively with Şengül and Üstündağ (2009). More specifically, they reflected low level in the subscales of open-minded, confident in reasoning, truth-seeking, and systematic whereas they showed middle level in the inquisitive subscale and high level in the analytical subscale. In the further study aimed to investigate pre-service teachers critical thinking dispositions, Alper (2010) supportively reached such a conclusion that pre-service teachers reflected low critical thinking tendency regarding truth-seeking.

Moreover, some experimental studies were also conducted to investigate specific instructional strategies such as constructivist approach, discussion method, 7E

learning cycle method on students' development of critical thinking skills (Mecit, 2006; Uysal, 1998).

From a different perspective, in qualitative study, Türkmen-Dağlı (2008) investigated how teachers teach would support or inhibit the development of students' critical thinking skills in the three phases of instruction; planning, implementation, and reflection. Three teachers who are teaching Turkish course at fourth grade from three different primary schools were selected to conduct a comparative case study. Data were collected through classroom observations, interviews with both students and teachers, and, students' logs and documents. The findings of the study indicated that teachers' restriction of channels between students could constrain their development of critical thinking. In this circumference, teachers should be able to give opportunities students to express their ideas, decision-making process with writing or oral language. Another conclusion was that there is a need for school guideline for critical thinking instruction including sample tasks that include significant and interesting issues and encourage critical discussion environment. The study also pointed out that teachers favored their answers to the questions in the text rather than encourage students' thinking. The other interesting finding arising from the study was that teachers presented informal fallacies such oversimplification, personal considerations unwittingly. Teachers who restrict argumentation between students prevented students to check and argue their arguments with their friends, which could be an obstacle for the improvement of critical thinking. Teachers' reflections regarding their classroom practices indicated that students' emotions as critical thinking disposition were not necessarily considered to develop students' critical thinking like cognitive abilities such as inference, analysis, and interpretation. Most importantly, this study concluded that teachers could confuse the differences between cognitive critical thinking skills. For example, one of the teachers thought that she encouraged students to make predictions; but in fact, she had students found the main idea of the text. Although finding main idea of the text is one of the most

common thinking skills, teachers reflected inconsistencies with regard to main idea of the texts, which caused to confusion in students' mind. Finally, this study also pointed out the importance of classroom culture on the improvement of critical thinking.

In the further study, Kanik (2010) conducted a phenomelogical study to investigate teachers' conceptions of critical thinking and their practices for development of critical thinking at seventh grade level. Seventy teachers teaching Turkish, social studies, science and technology, and mathematics courses participated to the study. Major data of the study were collected through interviews with the participants. The findings of the study revealed that there were cognitive skills, dispositions and criteria that teachers perceived to relate to critical thinking. The results of the study indicated that few teachers could make conceptualization of critical thinking in detail and appropriately. Most of the teachers clarified the concept of critical thinking in terms of either skills or dispositions by giving examples from their classroom experiences. Teachers have a clear agreement regarding integration of critical thinking into instruction. They also expressed that current curriculum prevented them to promote students' critical thinking since it covers the content superficially rather than in depth. Finally, the researcher suggested that there is a need for teachers' assessment of critical thinking with alternative assessment methods.

From a different aspect, Sapancı (2007) addressed the problematic issue of widespread usage of internet as main research tool and the need for critical thinking as an educational goal. He mentioned about possible students' tendencies in using internet such as accepting all information or claims without questioning. To avoid from dangers of internet, students should be educated as citizens to criticize the reliability of the sources, seek reasons or evidences behind the sources. In the further study, Öztürk, Bıçak and Sabancı (2009) developed a scale to investigate pre-service teachers' critical thinking with regard to e-knowledge sources in primary

school, social studies, Turkish teacher education programs. They propose that there be further study to apply such a scale to investigate pre-service teachers in different domains.

To conclude, researches regarding critical thinking in Turkey were conducted mostly with quantitative methods. More specifically, while some researchers contributed to the literature in critical thinking by investigating relationships between factors such as gender, critical thinking dispositions, experience, grade level, the others used experimental design to investigate the effects of specific instructional strategies on the development of students' critical thinking skills. Differently from these, studies, Türkmen-Dağlı(2008) and Kanik (2010) are among the studies which investigated teachers' conception of critical thinking by utilizing the qualitative method and suggested qualitative studies to gain more insight into teachers' thinking. It is also noteworthy point that there is no enough study regarding development of critical thinking skills of students and teachers in mathematics education.

2.3. Summary of the Literature

Summary of the literature was presented in relation two main topics; statistical literacy and critical thinking. Firstly, review of the studies related to statistics and probability in context indicated that the researchers have a consensus on the importance of understanding, critically evaluation and communicate the ideas related to statistical and probabilistic terms reported in social and scientific texts in order to be efficient citizens in the society. It is clear that statistics and probability is everywhere in our life and for this reason, there has been a change from traditional teaching of probability and statistics based on computational procedures to be consumers of statistical and probability in context rather than producers as Shaughnessy (2007) asserted that "teaching probability goes well beyond "balls in urn" problems." (p. 933). In the light of these ideas, recent researchers and many

curriculum documents (NCTM, 2000; AEC, 1991) suggested that students should be educated as efficient citizenship with critical appreciation of statistical and probabilistic statements given in social and scientific reports. To achieve this goal, teachers have a potential role on education of students in statistics and probability contexts as they implement the curriculum, design the instructional environments, or decide tasks, activities. However, there has been no enough study related to teachers in the domain of statistics and probability (Shaughnessy, 2007). Studies related to teachers have developed towards their statistical content knowledge and pedagogical knowledge. On the other hand, there has not been enough study related to teachers' interpretation of statistical and probabilistic information in social texts, especially in Turkey.

Review of the literature regarding critical thinking indicated that there are several common points arising from these studies. First of all, critical thinking has been an educational goal to nurture efficient citizens in the society. To accomplish this goal, various approaches were observed among the researchers. One of these approaches to develop teaching of critical thinking is that critical thinking is transferable to the different contexts and could be integrated it into specific subject matter. Teaching critical thinking and transferability of the critical thinking to the other subject matters are curial issues. There has been an agreement regarding the usage of real-world contexts, problems or popular media texts such as newspaper articles, magazine articles, advertisements or television programs in the classroom environments in teaching of critical thinking. Moreover, researches regarding teacher education in critical thinking focused on the teachers' perceptions of critical thinking and teachers' integration of specific methods into classroom environments to develop students' critical thinking.

In Turkey, studies regarding critical thinking were conducted mostly with quantitative methods to examine relationships between factors such as gender, critical thinking dispositions, experience, grade level, and to investigate the effects

of specific instructional strategies on the development of students' critical thinking skills. There is a limited study conducted with qualitative methods to investigate teachers' critical thinking, especially in mathematics education. This studies manifested inconsistent results regarding teachers' critical thinking skills, which indicate the need for further investigation of teachers' critical thinking skill in detail. It is also noteworthy point that there is no enough study regarding investigation of critical thinking skills of students and teachers in mathematics education.

In conclusion, there is need for investigating pre-service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts.

CHAPTER 3

METHOD

The purpose of the study is to investigate pre-service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts. This chapter describes underlying background of the study, which includes overall design of the study, data collection and data analysis procedures. Then, it continues with the trustworthiness of the study. At the end of the chapter, ethical issues and limitations of the study are discussed.

3.1 Overall Design of the Study

The purpose of the study is to investigate pre-service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts. In line with the purpose of the study, the qualitative research design was followed by the researcher since the qualitative research aims to investigate how human thinks and what meanings are formed (Bogdan&Biklen, 1998; Merriam, 2009).

In this study, case study design was used as a form of the qualitative research since case study examines a single entity such as a single individual, one particular group, or one particular setting in detail (Merriam, 2009; Stake, 2005; Yin, 2009). Merriam (2009) defines a case study as "an in-depth description and analysis of a bounded system" (p. 40). This bounded system specifies the boundary of the case; that is, it shows the situation not to be studied (Miles &Huberman, 1994). In a similar way, Stake (2005) remarks the importance of identification of the case as a unit of analysis and its bounded system. Since the purpose of the study is to investigate pre-

service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge, the case is delimited as "senior pre-service middle school mathematics teachers". To set boundaries of the study, senior pre-service middle school mathematics teachers were selected since they are on the brink of graduation from elementary teacher education program and have completed almost all fundamental courses such as Statistics and Probability, Research Methods, and Method of Teaching Mathematics. Another boundary of the study, participants who have tendency to use valid quantitative procedures and mathematical language, and critically evaluate newspaper article were selected in order to investigate their critical thinking process in detail. The case was investigated in the context of Elementary Mathematics Education Program which was explained in the following part.

3.2 Context

Elementary Mathematics Teacher Education Undergraduate Program (EME) in a public university constitutes the context in which the case of the study was investigated. Pre-service middle school mathematics teachers are educated to be qualified mathematics teachers in schools from grade level 5 to 8. The EME program is a four year undergraduate program in which pre-service middle school mathematics teachers take courses ranging from mathematics, science, statistics and probability, educational science, mathematics teaching courses to history, language and elective courses. In the first and second year, they take courses mostly regarding mathematics and science while in the last two years taking mostly general educational courses and mathematics teaching courses. Moreover, they take 'Introduction to Probability and Statistics I and II' in the second year. In the third year, the course of 'Methods of Teaching Mathematics I and II' was offered in the program. In addition, they take 'Research Methods' course in the seventh semester.

3.3 Participant Selection

Sampling strategies for qualitative research have been discussed by many researchers (Merriam, 2009; Miles &Huberman, 1994; Patton, 2002; Stake, 2005). Merriam (2009) mentioned about two fundamental sampling strategies, namely probability and nonprobability sampling. In a qualitative study, it is appropriate to use nonprobability sampling rather than probability sampling since the purpose of the qualitative research is not to make a generalization from sample to population (Merriam, 2009). In this study, purposeful sampling strategy was used as a form of non-probability sampling. Purposeful sampling strategy provides in-depth understanding in line with the purpose of the study due to the selection of sample that gives rich information, not generalization, which is the strength of the purposeful sampling (Patton, 2002). Patton (2002) proposed several strategies for purposeful sampling. Of these strategies, criterion and convenience sampling strategies were used purposively in this study. Patton (2002) explained that "the logic of criterion sampling is to review and study all cases that meet some predetermined criterion of importance." (p. 238). Similarly, Merriam (2009) stated that "to begin purposive sampling, you must first determine what selection criteria are essential in choosing the people or sites to be studied." (p. 77). Because of convenience of location and time, the participants of the study were selected from the university in which the researcher works as a research assistant. At the same time, they consist of pre-service middle school mathematics teachers who took courses of Statistics and Probability I and II and Researcher Methods; that is, senior pre-service middle school mathematics teachers. In Elementary Mathematics Education Program during 2011-2012 academic year, there were about 45 senior pre-service middle school mathematics teachers.In order to select participants, a questionnaire (see in Appendix A) involving newspaper article about autism was implemented to 38 senior pre-service middle school mathematics teachers, who were volunteered to participate to the study. The selection of participants among 38

pre-service middle school mathematics teachers was based on predetermined three criteria as follows:

- *Using valid quantitative procedures and mathematical language:* It was examined whether participants use mathematical procedures and terminology effectively and accurately both in mathematical and real life context and explain their ideas to someone else clearly by using appropriate mathematical language**.**

- *Critical Evaluation:* To critically evaluate, it is essential to have the ability of questioning the statistical arguments, claims including misleading information and bias. That is, they should be able to detect errors, deficiencies in the context to make proper justification.

- *Giving Rich Information:* It was examined whether participants express and clarify their ideas, opinions, feelings while examining the newspaper articles. It is useful to determine the participants who have given rich information and deep description about their thinking process.

The purpose such a criterion based selection is to gain an early insight into participants' critical thinking process, which requires using valid quantitative procedures and mathematical language (Watson &Callingham, 2003), and evaluate critically (Gal, 2004; Facione, 1990; Watson &Callingham, 2003), since it is difficult to examine this process with extremely low critical thinkers.

The analysis of questionnaire was conducted with a holistic scale developed by the researcher (see in Appendix A). Based on the rubric, pre-service middle school mathematics teachers performances were scored for each criterion; using valid quantitative procedures and mathematical language, critical evaluation, and giving rich information. The analysis revealed that the participants' scores ranged from 0 to

4 (see in Appendix B). The participants with scores above average (score 3 and 4) were considered for the examination of critical thinking process with regard to statistical and probabilistic knowledge in detail. After examination of the questionnaire, ten students had scores 3 or 4. Then, the researcher consulted their instructor. Based on instructor's view, eight of them were selected to interview. At the end, seven students (3 males and 4 females) were interviewed since one of them was not willing to participate to the study. After interviews, the researcher selected four students (one male and three females) who gave more rich data during the data collection process in order to gain insight into their thinking process.

3.4 Participants of the Study

In this part, participants of the study are described on the basis of pre-interview and post-interview (see in Appendix C), which includes open-ended questions about the essence of statistics and probability, usage of statistics and probability in real life, the importance of statistics and probability in real life, and their strengths and difficulties confronted with the concepts of statistics and probability. There were four participants with pseudonyms Ali, Meltem, Melek, and İrem. Participants' opinions regarding statistics and probability are explained respectively in the following.

Ali: Hetook courses of *Introduction to Statistics and Probability I and II* in the second year with grades of AA. He registered to minor program in the Statistics Department in the third year. He also completed *Research Methods* course with a grade of AA. He appreciates the role of statistical and probabilistic knowledge on real life. For example, he explained many real life examples such as football statistics, credit interest, the frequency of stores' intensivity, and salary raise. He also expressed that he did not have difficulty in interpretation of such examples reported in the newspaper article because they are almost based on the percentages. However, he also underlined many researches reported in the newspaper articles can

be misleading with ambigous language, bias in sampling process. He feel qualified to teach students statistics and probability though he admitted lack of knowledge regarding how to conduct a research. He also asserted that pre-service teachers have lack of statistical and probabilistic knowledge and lack of interpretation of statistical and probabilistics concepts although statistics and probability is one of the content standards in the elementary mathematics curriculum.

Meltem: She took courses of *Introduction to Statistics and Probability I and II* with grades of AA. At the same time, she started to minor program in the Statistics Department. She also completed *Research Methods* course in the summer school of third year with the grade of AA. However, she thinks that she could not adapt statistical and probabilistic knowledge to real life although she asserts statistics and probabilisty is everywhere in real life such as in wheather forecasts, in evaluation of articles, in education. She appreciates the need for statistical and probabilistic knowledge while reading newspaper article. For example, she gave an example regarding background of rating record determination and stressed that it is essential to know background of such researches.

Melek: She completed courses of *Introduction to Statistics and Probability I and II* with grades of DC and CC respectively and *Research Methods* with a grade of AA She did not registered any minor program related to mathematics education. She thinks statistics and probability is an integral part of everyday life. She gave examples such as interpretation of weather forecasts, making decisions about work, survey research in schools, and assessment of students in educational settings. However, she claimed that courses in undergraduate program did not facilitate to interpret such real life examples. She also stated that she have difficulty in understanding probabilistic language given in the probabilistic problems and in interpretation of researches reported in popular media texts since they could be reported with ambiguous language. She appreciated the role of knowledge regarding percentages, and increase in risk on interpretation of media texts.

İrem : Shetook courses of *Introduction to Statistics and Probability I and II* with grades of CB and DC respectively. He started to minor program in the Mathematics Department. She also completed *Research Methods* course with a grade of AA. She is aware of usage of statistics and probability in real life. In addition, she does not believe in the nature of statistics and probability in contrast to deterministic nature of mathematics although they are closely related domains. She thinks that it is meaningless to conduct studies, which include generalization from a sample to population. She asserted that to believe conclusions of such studies, they should be conducted with every people in the population. These studies do not play efficient role on her decisions in real life. However, she also expressed statistics and probability somewhat prevents people to reach wrong generalization about the phenomenon reported in the media. Moreover, she appreciated the necessity of knowledge of statistics and probability for mathematics tachers; but, she do not feel confident in teaching statistics especially, histograms, standard deviations, which are new concepts in the Turkish Curriculum, although she feel sufficient herself in teaching of probabilistic problems.

3.5 Data Collection Procedure

Data collection process involves procedures of development of data sources, pilot study, and interviews with the participants. In Table 1, time schedule for data collection was presented. The major data sources of the study consisted of interviews with pre-service middle school mathematics teachers regarding two newspaper articles. Interview protocols regarding newspaper articles, pre-interview and post-interview questions, and a questionnaire were developed with different purposes. In the following parts, their contents, purposes of development, and how they were developed were described and other data collection procedures were explained.

Table 1

Time Schedule for Data Collection

Date	**Data Collection Procedures**
October 2011 – December 2011	Development of Interview Protocols and Questionnaire
January 2012 - February 2012	Pilot Study of Interview Protocols
March 2012- April 2012	Interviews with Participants

3.5.1 Questionnaire

The questionnaire was developed on the basis of questions in the interview protocols regarding newspaper articles, which are major data source of the study. The purpose of the development of the questionnaire is to select participants to be interviewed and get rich information about the case of the study. The questionnaire was prepared regarding a newspaper article about health issue, autism (see in Appendix D). It includes five main open-ended questions. The questions were developed to select participants on the basis of pre-determined three criteria; u*sing valid quantitative procedures and mathematical language, critical evaluation, and giving rich information.*

3.5.2 Interview Protocols Regarding Newspaper Articles

The purpose of development of interview protocols regarding newspaper articles is to investigate pre-service middle school mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts. Development of interview protocols includes three phases: the selection of the newspaper articles, improvement of interview questions for each of the selected news and pilot study of interview protocols regarding newspaper articles.

The process of selecting newspaper articles were carried out in three phases. In the first phase, the newspaper articles were searched in two ways: using the search engine with the keys words such as “according to the research” or searching daily popular news in Turkey from the internet or buying the daily news. In the second phase, the news was grouped under different topics such as health, social, sport, economic, education, cultural. After the grouping, in the third phase, news in the different groups were selected with the criteria in which news includes rich and fundamental statistical and probabilistic information and is comprehensible, not including technical terms for the participants. According to these criteria of news selection, one newspaper article was selected related to health (see in Appendix E). Its Turkish and English versions were added to Appendix E.1 and Appendix E.2 respectively. The second newspaper article was found through research articles in the literature, which is related to social life (see in Appendix F). Since the language of the article is English, it was translated into Turkish language by the researcher. Its Turkish and English versions were added to Appendix F.1 and Appendix F.2. It was considered that different topics could address different participants’ interest in the study. The first one was about babies with Down syndrome whereas the second was related to cheating in young couples. Moreover, both of them include mostly probabilistic statements underlying especially conditional probability so that the researcher could gain insight into participant's thinking process.

Construction of semi-structured interview questions for each of the news was carried through examining the studies in the literature of statistics and mathematics education based on the purpose of the study. Firstly, semi-structured interview questions were developed for the first newspaper article. Then, pilot study of the news was conducted with a doctoral student. The purpose of the pilot study for only the first news was to develop interview questions in the other newspaper article appropriately. After the pilot study, some questions were reviewed and other interview questions were developed in this way. Then, the actual pilot study of all

interview protocols constructed was carried out with a senior pre-service middle school mathematics teacher.

The pilot study is useful for reorganization of the data collection process related to content and procedures followed in this process (Yin, 2009). In light of this information, the purpose of the pilot study was to review selected news, the interview questions related to news, and procedures used during the data collection. The pilot case was selected with the criteria of convenience access. After the pilot study, interview protocol for news were reviewed and refined. Then, opinions of five experts were requested; one graduate student in statistics department, one graduate student and two instructors in elementary mathematics education department. Data collection process was refined at the end of the pilot study and by expert opinion. These reviews and recommendations are categorized in three main topics as follows;

- *Formal structure of questions:* The changes were made with regard to clarity and order of the questions. With respect to clarity of language, there were some ambiguous words that lead different readers to make a variety of interpretations. Furthermore, the order of the questions was revised. One of the examples is that the question of "To extent what do you rely on the conclusions of the study?" was moved from part of 'general questions' to the end of the other questions for all newspaper articles. The reason of such a change was that when this question was asked at the first part of interview protocols, the pilot case had a tendency of believing there is a mistake in this newspaper. This would restrain participant to make critical evaluation related to newspaper article.

- *Content of questions:* There have been three main changes with respect to content of questions. The first change was that questions became more structured. Secondly, after the expert opinion, more probes and follow-up

questions was added. These were used as alternative questions when the researcher encountered during the interview. Thirdly, interview questions were added to the beginning and the end of the interviews in order to understand what participants think about articles and usage of articles in elementary mathematics education program and schools.

- *The selection of news*. Of the selected newspaper articles, second newspaper article with the topic of "Fast music increases their risk of having an accident twice" was removed from the study. Instead, another newspaper article with the topic of "cheat radar better tuned in men" was added. This newspaper article was shown in Appendix F. It was translated into Turkish language by the researcher with the permission and pilot study was conducted again. The reason of such a change was that there were ambiguous expressions (e.g. twice, more than twice) that prevented the participant from focusing to the study.

To conclude, interview protocols were reviewed and retested in light of these recommendations and criticism made by pilot case and experts. In this way, the last version of interview protocols was formed (see in next part). Interview protocols include two parts; general questions and particular questions for each newspaper article. The first part of interview protocols, 'General Questions', have same structure in two newspaper articles. The second part of the interview protocols consists of specific questions about newspaper article. The last version of the interview protocols are presented below:

Newspaper Article I:

The first newspaper article is about blood test asserted to detect a pregnancy with Down syndrome, published on internet website of NTV (Doğuş Publication Board) in Turkey (see in Appendix E). The language of the newspaper article is Turkish. It

was implemented to the participants without any change in its structure. Participants were asked to answer questions regarding Newspaper Article I (see Turkish and English versions in Appendix E.1.1 and Appendix E.2.1). The duration of interview was approximately forty five minutes and interviews were video-recorded.

Newspaper Article II:

The second newspaper article is about cheating partners, published in the Mercury newspaper in Tasmania (see in Appendix F). Watson (2011) proposed this newspaper article to be used in educational settings. The newspaper article was translated into Turkish language by the researcher and revised by an expert in language program. Participants were asked to answer questions regarding Newspaper Article I (see Turkish and English versions in Appendix F.1.1 and Appendix F.2.1). The duration of interview was approximately forty five minutes and interviews were video-recorded.

3.5.3 Pre Interview and Post Interview Questions

Pre-interview and post-interview questions were developed in order to get information about subjects' feelings, ideas, thoughts in five main topics: about the essence of statistics and probability, usage of statistics and probability in real life, the importance of statistics and probability in real life, suggestions related to elementary mathematics curriculum and teacher education program, the strengths and difficulties confronted with the concepts of statistics and probability. Post-interview questions also include the discussion of selected newspaper article regarding how to use in middle schools and teacher education programs (see in Appendix C). Pre-Interview questions were implemented at the beginning of the interview protocols in relation to news (see Turkish and English versions in Appendix C.1.1 and Appendix C.1.2). Post-Interview questions were implemented at the end of the interview protocols (see Turkish and English versions in Appendix

C.2.1 and Appendix C.2.2). These pre interview and post interview questions provide information about the participants and provide qualified discussion of the findings.

3.5.4 Interview with Participants

In this study, the researcher interviewed each participant in the privacy rooms individually. They were interviewed at two sessions. At the beginning of the interviews the purpose of the study was explained. Then, it was reminded that there was no time limit to think deeply about newspaper articles. Each interview regarding newspaper articles took approximately forty five minutes. Moreover, interviews were video-recorded. In the first interview, pre-interview questions were asked in order to get information regarding what they think about statistics and probability, or usage of statistics and probability in real life, their suggestions related to curriculum and elementary mathematics program. Then, the first newspaper article was given and related interview questions were asked to the participants. The other interviewing was conducted one week later. In the second interviewing, the second newspaper article was investigated by the each participant and then post-interview questions were asked. The duration of pre-interview and post-interview was about thirty minutes.

3.6 Data Analysis Procedure

Data analysis is the process of regulation, putting in order of the data and making sense of collected data (Merriam, 2009). In a similar way, Bogdan and Biklen (1998) stated that "analysis involves working with data, organizing them, breaking them into manageable units, synthesizing them, searching for patterns, discovering what is important and what is to be learned, and deciding what you will tell others." (p. 157). In this study data analysis was conducted in order to examine pre-service middle school mathematics teachers' critical thinking processes with regard to

statistical and probabilistic knowledge in the newspaper articles. Thus, data were analyzed on the basis of two dimensions, namely statistical and probabilistic knowledge, and critical thinking skills. Coding system for two dimensions was explained respectively.

Firstly, coding system regarding the first dimension, statistical and probabilistic knowledge, was developed mostly on the basis of Gal's studies (2004; 2005) and Watson's studies (1997; 2006). Gal (2004) developed a model for statistical literacy which includes five knowledge bases: Literacy skills, statistical knowledge, mathematical knowledge, context knowledge, and knowledge of critical questions while interpreting the claims. More specially, statistical knowledge base was explained under five subdomains; "knowing why data are needed and how data can be produced, familiarity with descriptive statistics, familiarity with graphical and tabular displays, understanding basic notions of probability, and knowing how statistical conclusions or inferences are reached."(Gal, 2004, p. 58). In the further study, Gal (2005) also developed model of probabilistic literacy which includes five knowledge bases; big ideas of probability, the terminology and language of probabilities, critical questions, and context knowledge. From a different perspective, Watson (2006) examined statistical literacy on the basis of six statistical and probabilistic concepts (sampling, average, data representation, chance, inference and variation); but, it shares common points with Gal's models. (See more detail in Chapter 2). These models were revised on the basis of data collected and organized under three categories; base of reported findings, reported statistics and generalizability of reported finding (See in Appendix G).

The process of such categorization was explained as follows: The first category *base of reported findings* mostly was constituted on the basis of Gal' (2004) study. Gal mentioned about knowledge regarding origin of the study as "knowing why data are needed and how data can be produced" and "knowing how statistical conclusions or inferences are reached". Watson (2006) also mentioned about the understanding

notions of sampling and variation, which require understanding role of sampling in inference from sample to population, representativeness, purpose of the sampling, bias in sampling method, and considering unusual values in the data set such as outliers, the existence of variation in mean values with regard to sample size. These are compiled as a research process including *sampling*, *data collection*, *data analysis*, and *results/ conclusions*, which constitute the base of reported findings.

The second category *reported findings* includes basic statistical and probabilistic notions reported in the newspaper articles. Gal (2004) mentioned about the familiarity with descriptive statistics, graphical and tabular displays, and understanding basic notions of probability, which are among the statistical knowledge bases for statistical literacy. Moreover, Gal (2005) also identified big ideas of probability, and the terminology and language of probabilities as a knowledge base for probabilistic literacy. In a similar way, Watson (2006) addressed the need for knowledge regarding notions of chance, average, and data representation in context. These are compiled as knowledge regarding the notions of *probabilistic statements*, *percentage*, *measures of central tendency*, and *graphical and tabular displays*, which are represented as *reported findings* in the newspaper articles.

The third category *generalizability of the reported findings* is compiled on the basis of Gal's study (2004). Gal (2004) addressed generalizability of results under the topics of "knowing why data are needed and how data can be produced" and "knowing how statistical conclusions or inferences are reached".

Secondly, coding regarding the second dimension, critical thinking process, was conducted on the basis of critical thinking framework prepared with a consensus among experts in the field (Facione, 1990). However, it is important to note that this study was limited to skill aspect of critical thinking rather than dispositional aspect. Categories of critical thinking skills consist of six primary skills; interpretation,

analysis, evaluation, inference, explanation, and self-regulation, presented in Table 2 (See in detail in the literature part). It is also important to emphasize that these critical thinking skills were interpreted on the basis of nature of this study.

Table 2

Critical thinking skills proposed by Facione (1990)

Core Critical Thinking Skills			
Interpretation	Categorization	Decoding Significance	Clarifying Meaning
Analysis	Examining Ideas	Detecting Arguments	Analyzing Arguments
Evaluation	Assessing Claims	Assessing Arguments	
Inference	Querying Evidence	Conjecturing Alternatives	Drawing Conclusions
Explanation	Stating Results	Justifying Procedures	Presenting Arguments
Self-Regulation	Self-Examination	Self-Correction	

Each code was analyzed on the basis of both statistical and probabilistic knowledge and critical thinking by the researcher. The researcher coded the same data three times for each participant to reach consistent codes. To increase dependability of the study and reduce bias in data analysis, a second coder also analyzed the data in terms of both critical thinking and statistical and probabilistic knowledge. Second coder informed about coding systems for two dimensions, statistical and probabilistic knowledge and critical thinking. Inconsistent codes between them were discussed at two different times up to reach fully consensus. Moreover, coding for the first newspaper article was matched to coding of the second newspaper article to provide consistency between codes and reduce the bias in the study.

3.7 Trustworthiness

Trustworthiness of a qualitative study is established through several strategies, credibility, transferability, dependability, and confirmability (Merriam, 2009). Firstly, Merriam (2009) explained internal validity "… deals with the question of how research findings match reality" (p. 213). Merriam (2009) identified five strategies to be used in order to increase the credibility of a qualitative study, namely, triangulation, member checks, adequate engagement in data collection, researcher' position, peer review. In this study, in order to increase the credibility of the study, the strategies of adequate engagement with the participants, researcher position, member checks and peer review was considered. In terms of adequate engagement, four participants were selected and each participant was interviewed at two different sessions since it is important to provide the saturation of findings (Merriam, 2009). Researcher also reflectively stated her role in the research. The findings of the study were tested with participants that I interviewed about whether there is misleading interpretation of the findings. In addition, the comments of my supervisor and thesis committee members; was taken into consideration.

Secondly, one of the criteria to provide trustworthiness in a qualitative study is transferability referring to external validity. External validity is pertinent to the generalization of the findings in a study (Merriam, 2009). Guba (1981), however, stated that "it is not possible, they [naturalists] believe, to develop "truth" statements that have general applicability; rather, one must be content with statements descriptive or interpretative of a given context…" (p. 86). Similarly, Merriam (2009) mentioned the differences between qualitative and quantitative study in terms of generalizability and emphasized that rich and thick description should be presented to the reader so that they could understand and transfer the findings of the study to the other situations. In this study, detailed description related to context, participants, data collection methods, analysis of the findings was presented.

The third criterion to facilitate trustworthiness is dependability referring to the reliability or consistency. Reliability deals with the question of the extent to which the findings of the study is consistent with the data (Merrriam, 2009). In this study, all techniques used to ensure credibility increase dependability of the study. Moreover, data collection procedures, process of data analysis, the way of the researcher reach the findings was explained in detail.

Finally, the last criterion to provide trustworthiness of a qualitative study is confirmability referring to objectivity. Confirmability is concerned with the question of "How can one establish the degree to which the findings of an inquiry are determined by the subjects (respondents) and conditions of the inquiry and not by the biases, motivations, interests, or perspectives of the inquirer? (Lincoln &Guba, 1985, p. 290). In this study, detailed description was presented with regarding procedures used through the study and researcher's role and assumptions in the study was addressed in order to establish confirmability.

3.8 Limitations

There are limitations and possible bias issues in this study. The first limitation of the study is researcher experience. To reduce this bias, I reported reflexivity regarding my role in the study since human is the instrument in qualitative studies (Johnson, 1997; Merriam, 2009). I was not experienced in interviewing and qualitative study. To reduce this bias, I have gained experience with the pilot study of interviews and working with my supervisor and consulted experts in statistics and mathematics education through the research. Furthermore, I explained, in detail, the research context, how I conducted to the study and reached the findings of the study. Participants were also stranger to the setting of the interview. Some participants might behave differently in the first interview. Thus, the participants were interviewed at two different times, which increase also quality of the study. The second limitation regarding credibility of the study is to use one type of data source,

interview protocols regarding newspaper articles. It might be more useful to use more data sources to increase quality of the study. To reduce this limitation, prolonged engagement with the participants was provided. The third limitation is the selection of newspaper articles. That is, the study is limited to newspapers. There could include more different media texts such as journals, magazines, different contexts such as economics, political, social contexts to address participants' interest and different statistical and probabilistic concepts such as graphs, charts, and variability.

CHAPTER 4

RESULTS

The purpose of this study is to investigate to what extent pre-service middle school mathematics teachers use critical thinking skills with regard to statistical and probabilistic knowledge. Participants' interpretations of Newspaper Article I and Newspaper Article II were analyzed based on two dimensions; namely, critical thinking skills and statistical and probabilistic knowledge. Newspaper Article I includes health news regarding Down syndrome, while Newspaper Article II includes social context regarding cheating of partners. At the end of the chapter, conclusions with respect to Newspaper Article 1 and Newspaper Article 2 are summarized.

4.1 Newspaper Article I

The findings regarding Newspaper Article I are organized into three main topics; the base of reported statistics, reported statistics, and generalizability of the reported findings in the newspaper article.

4.1.1 Bases of the Reported Findings in the Newspaper Article

This part includes participants' thoughts about the bases of the findings reported in the newspaper article with regard to their statistical and probabilistic knowledge. The bases of the reported findings in the newspaper article consist of the background of the study in the newspaper article; how the sample could be selected, how the data could be collected and analyzed, and how the statistical conclusions reported in the newspaper could be reached. Thus, participants' critical thinking

processes regarding the bases of the reported findings in the newspaper article are presented under four main headings, namely sampling, data collection, data analysis, results and conclusions.

4.1.1.1 Sampling

This part includes the participants' critical thinking processes with respect to the sampling of the study reported in the newspaper article. Participants made use of critical thinking skills by using their knowledge of sampling, which includes identification of sample characteristics, sampling method, role of sample on inference from sample to population, and possible biases in sampling. Analysis of participants' critical thinking processes regarding sampling is summarized in Table 3. Then the extent to which participants made use of critical thinking is explained in detail.

Table 3

Critical thinking processes used by participants regarding sampling

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM DEC	CM	CM	CM
Analysis	EI	-	EI	EI
Evaluation	AC AA	AA	-	AC
Inference	CA	CA	CA	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; DEC: decoding significance; EI: examining ideas; AC: assessing claims; AA: assessing arguments; CA: conjecturing alternatives.

Analysis of participants' knowledge regarding sampling indicated that all participants focused on the notion of sample size as a characteristic of sample. They restated the following statements reported in the newspaper article:

> This method, which was experimented on 753 pregnant women, was found to be successful in its detection yielding no incorrect results. All of the volunteer women chosen for the test were selected from among those who held much higher risks than normal in terms of giving birth to a baby with Down syndrome.

All of them *interpreted* these two statements by paraphrasing the characteristic of sample and restating the condition of "pregnant women at high risk in terms of giving birth to a baby with Down syndrome" to *clarify the meaning* of them. In the *interpretation* process, three of the participants (Ali, Melek, and İrem) *analyzed* these two statements by *examining ideas* with the application of proportional reasoning. Ali, for example, expressed the following: "We found 86/753 as the answer. Does this equal, well, approximately 1.877 1/9...they have indeed been chosen from among those who carry a much higher risk than average in terms of giving birth to a baby with down syndrome." (Q1). In this process, he identified closely related statements regarding characteristic of sample and compared them on the basis of proportional reasoning.

Differently, Ali was also in the process of *decoding significance* while saying, "They have been chosen from among people carrying a much higher risk than average, there seems to be bias here." (Q2). When Ali read the newspaper article at first, he attempted to detect the author's directive intention regarding sample and was suspicious that there could be a bias in the sample selection.

From a different perspective, Melek attempted to *conjecture alternatives* regarding the selection of the sample. In this process, she formulated multiple alternatives with regard to how a sample could be selected and appreciated the role of outliers, which requires understanding outliers' effect on the results of the study. In this regard, she postulated removement of possible outliers (pregnant women not having a risk or certainly at risk) from the study. She reported as in the following:

> 753 people is an odd number, they [researchers] may have eliminated them from the data group due to, for example, some problems, some constraints. Or well

> here it says they are chosen from among those carrying a higher risk, they [sample] may have eliminated those with no risk at all or have a 100% risk of the down syndrome incidence... maybe there were individuals who didn't want to continue participating in the study. Or as I said, errors were made in the tests. They may have seen that the outputs had very different results when they did a check test and, thus, thinking that they would change the course of the study, they may have eliminated those data. In this way, 753 people remained. (Q3)

Another result of the study was that two of the participants (Ali and İrem) attempted to *evaluate* the sample size of the study to *assess* the credibility of the *claims*. To illustrate, İrem was not sure in her evaluation process and could not give evidence supporting her evaluation regarding sample size although she recognized sample size as an important factor in the evaluation of the credibility of the study. She said, "I suppose 750 is good. That is, I don't think it a very small sample... I don't know according to what but I associated this 'one out of 800'" (Q4). She indicated immediate decision regarding sample size. Ali, however, made a comprehensible *evaluation* when compared with İrem. He *assessed argument* which included the following claim: *"The new method based on a simple blood test indicates with a high degree of precision whether or not the fetus in the mother's womb has Down syndrome."* in the newspaper article and the reasons underlying the claim as stated below:

> Reason 1: "Actually, two percent of the babies not having Down syndrome were incorrectly diagnosed to have Down syndrome."
>
> Reason 2: "This method, which was experimented on 753 pregnant women, was found to be successful in its detection yielding no incorrect results."

In the process of *assessing argument*, he focused on the notion of sample size as a crucial factor to evaluate the credibility of the reported study. He reflected different critical thinking skills by *assessing claims* and *conjecturing alternatives* regarding sample size to assess the argument. He recognized sample size as a factor which makes the reported study credible, while assessing the degree of credibility of the given information, as presented in the following quotation:

Ali: 753 – numbers are very important or when I read a news article I pay attention to numbers
Researcher: You said numbers are important. Why do you think so?
Ali: Because it shows the quality of the study. For example, here if they had chosen 35 pregnant women instead of 753, the results of the study may not have been reliable. Now if 3753 people were chosen instead of 753, as it would be much more comprehensive, there would be much fewer errorss; I mean to arrive at a more firm judgement, [this] data seems to be more convenient; that is, in terms of representing the population this sample.
Reseacher: Why? What kind of an impact can number of people have?
Ali: If they check more people, there will be fewer errors...(Q5)

In this evaluation process, it is clear that Ali appreciated the essential role of the sample, which requires understanding the need tı make inferences from the sample to the population and the need for samples to be representative. Moreover, during the evaluation of the study, he attempted to make use of the *inference* skill by *conjecturing two alternatives* with respect to sample size; 35 and 3753 people. He, then, projected advantages and disadvantages which are likely to result if those alternative sample sizes were selected. He intuitively has a sense of variability in sampling and stability of variation in the increase of sample size. On the other hand, he could not assess to what extent sample is representatively structured.

Similar to Ali's critical thinking process, Meltem attempted to *evaluate* the study by *assessing* the given *argument*. She made evaluation by means of *conjecturing alternatives* regarding sample size. She considered three different situations (studies with 753, 1000, or 10000 people) to give a credible picture with respect to the reported study. In this process, she expressed that there a need for more people participating to the study in order to educe proper conclusions from the study.

In conclusion, all the participants focused on the notion of sample size by *clarifying meaning*. They also made use of different skills, such as *analysis*, *inference*, and *evaluation* with respect to the sampling process of the study. They attempted to make evaluation mostly on the basis of sample size as a sample characteristic, rather

than other characteristics such as their age and country. Only one of the participants detected missing information regarding the reason of sample selection and appreciated the role of sample in a study, which requires understanding the need to make inferences from the sample to the population and the need for sample to be representative. None of the participants could assess to what extent sample size could be enough to reach a proper conclusion. This would be due to the fact that they were not familiar with health statistics or due to their lack of statistical and probabilistic knowledge regarding sampling.

4.1.1.2 Data Collection

This part includes participants' critical thinking processes with respect to the data collection procedure of the study reported in the newspaper article. The participants made use of critical thinking skills by using their knowledge regarding data collection, which includes knowing at least intuitively how data can be produced on the basis of research design, the contribution of a well-designed research for data production for the possibility of answering specific questions, and the possibility of bias in data collection. Analysis of participants' critical thinking processes regarding data collection is summarized in Table 4. Then, the extent to which the participants made use of critical thinking is explained in detail.

Table 4

Critical thinking processes used by participants regarding data collection

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	-	CM	-	-
Analysis	-	-	-	-
Evaluation	-	AC	-	-
Inference	-	-	-	QE CA
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; AC: assessing claims; QE: querying evidence; CA: conjecturing alternatives

Whereas Ali and Melek did not mention the data collection procedure, Meltem and İrem attempted to make use of critical thinking processes with regard to data collection. When asked how the researchers could reach such conclusions reported in the newspaper article, Meltem utilized *interpretation* skill by *clarifying meaning* of the contextual information regarding Down syndrome. She restated in a simple way as follows: "Here (refers to the first three paragraphs) there is already information about the test; well, because DNA is released into the mother's plasma, perhaps I think they get and trace the liquid from the DNA to see if there is an extra copy of the 21st Chromosome"(Q6). She also tried to recognize important factors, such as contextual explanations regarding data collection, which are essential to *assess the claims* reported in the newspaper article.

> In fact it's good information is given here; I mean it in a way explains how it occurs, what is looked at...If it had told me only this [fourth paragraph of the article], I mean in that case I would have been overwhelmed with too many numbers and my reading this part [first three paragraphs] showed me how the test could be done[...] (Q7)

Moreover, İrem differs from Meltem by her *conjecturing alternatives* and *querying evidence* to make *inference* regarding the data collection process of the reported study, as presented in the following quotation:

> *İrem*: They were chosen from among women carrying the highest risk (referring to the second last paragraph); then, all the other conditions are equal; I mean no treatment whatsoever was made to any of them... if there is an impact, it will not be due to the test; they cannot generalize the test.
> *Researcher*: For example, what kinds of factors may have been controlled here.
> *İrem:* Well, they are all under equal conditions; for example, perhaps none them have smoken. I mean whatever there are affecting down syndrome, they have all kept distant from it... is the test done before birth, when is it done...Of course I don't know what affects down syndrome; I should know them; for instance, it says that it emerges when an extra copy of the 21st Chromosome occurs; but when does it occur, what are there that affect it. (Q8)

In this process, she postulated alternatives regarding the design of the study. In particular, she assumed that data would be collected from the sample with equal characteristics under equal conditions. That is, she appreciated the contribution of a good design for data production in a study and controlling several varibles, such as smoking or not smoking, to prevent errors in the study. However, she did not raise questions about whether this reported study could include such a proper design. Moreover, she *querred evidence* regarding the data collection procedure (time for test implementation) and contextual information regarding occurance of Down syndrome.

To conclude, when the participants were asked how this study could be conducted, two of the participants focused on the data collection procedure. They attempted to make use of *interpretation* and *inference* skills. In this process, they mostly focused on the importance of contextual information regarding Down syndrome. This could be due to the fact that they were not familiar with such a health context. One of the participants also appreciated the role of a good design by controlling confounding variables. However, she could not question whether this study could actually include such a proper design.

4.1.1.3 Data Analysis

This part includes participants' critical thinking processes with respect to the data analysis of the study reported in the newspaper article. Participants made use of critical thinking skills by using their knowledge regarding data analysis, which requires them to have a sense of how data were analyzed and detect at least informally possible problems in the process of data analysis. The analysis of the participants' critical thinking processes regarding data analysis is summarized in Table 5. Then, the extent to which participants made use of critical thinking is explained in detail.

Table 5

Critical thinking processes used by participants regarding data analysis

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	-	-	-	CM
Analysis	-	-	-	-
Evaluation	-	-	-	-
Inference	-	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning

None of the participants, except for İrem, attempted to make use of critical thinking processes regarding data analysis. İrem tried to make *interpret* contextual information by restating with different words even though she indicated a weak critical thinking process regarding data analysis. When asked how this study could be conducted, she mentioned the data analysis of the study as follows: "The test was conducted on these 753 people and at the end they [the test results] were checked and matched to see if they were accurate or inaccurate." (Q9) However, this interpretation is very limited to made comments regarding data analysis of the study.

To conclude, it is clear that almost none of the participants could utilize the critical thinking skills regarding data analysis. It could be due to the fact that participants are not familiar with data analysis procedures in statistics regarding health or due to lack of information regarding data analysis of the research study in the newspaper article.

4.1.1.4 Results and Conclusions

This part includes participants' critical thinking processes with respect to how results and conclusions reported in the newspaper article were reached. The participants made use of critical thinking skills by using their knowledge, which requires understanding that the differences between the groups may not be large enough or stable to draw a reliable conclusion or can be caused by chance processes and that attention needs to be paid to the size of the groups, to the quality of research process, which have a potential effect on the results of the study.

The analysis of the participants' critical thinking processes regarding the results and conclusions reported in the newspaper article is summarized in Table 6. Then, the extent to which participants made use of critical thinking is explained in detail.

Table 6

Critical thinking processes used by participants regarding results and conclusions

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	DEC	DEC	DEC	DEC
Analysis	EI	-	-	-
Evaluation	AA	-	AA	-
Inference	-	DC	-	DC
Explanation	-	-	-	-
Self-Regulation	-	-	SE	-

Note. DEC: decoding significance; EI: examining ideas; AA: assessing arguments; DC: drawing conclusions; SE: self-examination.

All the participants reflected the *interpretation* skill by *decoding significance*. They interpreted the main idea of the text and distinguished subordinate ideas from the author's main claim. For example, Ali stated as follows:

> The message here is that a recently emerged blood test, a blood test implemented during the period of pregnancy gives information to the parents about whether the

> baby will be born with down syndrome. What's more, as it claims, it has a very high rate of accuracy. (Q10)

Meltem also pointed out the last paragraph of the newspaper article by decoding its significance by stating, "In fact this part [refers to the last paragraph] is also important.... It mentions that the test is actually not all that perfect." (Q11)

Another result of the study was that two of the participants (Ali and Melek) attempted to *evaluate* the credibility of the study considering the given *argument* (see in pages 93-94) in the newspaper article. Ali raised questions regarding sampling and assessed the credibility of the study if it included pregnant women at normal risk in having a baby with down syndrome in the following way:

> Well, it seems like if it were implemented on normal women, it would yield errors; I mean it reduces my confidence [in it], I acquire such biases... I wonder, isn't this blood test able to detect [it] in a normal individual, isn't it able to understand whether it has down syndrome or not...if normal individuals were selected, in this way, they reached success, I would trust much this test. (Q12)

Toward the end of interview, he also evaluated the premises of the following argument in the newspaper article: *"The new method based on a simple blood test indicates with a high degree of precision whether or not the fetus in the mother's womb has Down syndrome."* He thought that there is a contradiction between the two premises of the given argument by stating, "It says it never makes an incorrect detection, but then it says 2% error. There is no such thing, I think this part is conflicting, I mean when I read the news article, this attracts my attention."(Q13). However, his evaluation was based on wrong assumptions. This may be due to the fact that he has difficulty in understanding the difference between two conditional probabilistic statements because he could not *analyse* them by *examining* different *statements.*

Differently, Melek also attempted to examine her reasoning process regarding the question of what would happen if she read this article in daily life. That is, she self-consciously examined her thinking process by making use of the *self-examination* skill in the process of *evaluation*, toward the end of the interview as follows:

> Normally, if I had read this in the newspaper in [my] daily life, I would not have done any of these calculations. And it was going to be like the guys have talked with numbers, they must be telling the truth, they have proven it, it is a highly reliable test, and that is why this risk of 2% must be small; perhaps sometimes people can arrive at such opinions. (Q14)

Moreover, Meltem and İrem *drew* inappropriate *conclusions* from the reported statistics; thus, they could not evaluate the article in terms of credibility. İrem, for example, stated as follows:

> In fact it seems quite accurate; in my opinion there is little margin of error...I mean it looks like it's going to be favourable... So what if 2% of the babies that did not have down syndrome were detected as having the disorder, when actually they didn't. Well, the test detected all those having down syndrome. If I am having a test done and if I am told that I am going to have a baby with down syndrome, then I won't believe in this 2%. (Q15)

To conclude, all participants were in the process of *interpretation* by *decoding* the *significance* of the article. They differed from each other by making use of different critical thinking skills. It could be concluded that insufficient reflection of some critical thiking skills such as *analysis* or *inference* might affect proper evaluation of the results or conclusions reported in the newspaper article. Two of the participants attempted to evaluate the results or conclusions in a general way. They had difficulty in evaluating the base of the research process to reach such results. For example, they did not raise questions regarding the results caused by chance processes nor did they pay attention to the quality of the research process, which may have effect on the results of the study. This would be due to their lack of knowledge with regard to how results or conclusions could be reached in a research study or their unfamiliartiy with the context.

4.1.2 Reported Statistics

This part includes participants' critical thinking processes regarding the reported statistics in the newspaper article. Reported statistics refers to summary or descriptive statistics reported in popular media texts such as percentages, measures of central tendency, graphs or probabilistic statements etc. Newspaper Article I includes probabilistic statements and percentages as reported statistics. Thus, participants' critical thinking processes were examined in detail on the basis of probabilistic statements and percentages.

4.1.2.1 Probabilistic Statements and Percentage

In this part, participants' critical processes during their examination of the probabilistic statements reported in the newspaper article are presented. What statistical and probabilistic knowledge they made use of in this process is another finding of this part. There are two main probabilistic statements in the newspaper article. This part is organized in terms of these statements respectively. Moreover, there are four situations related to these probabilistic statements underlying conditional probability statements. Participants' critical thinking processes regarding these situations were also examined in detail.

Statement 1:

Statement 1 reported in the newspaper article requires participants to recognize expressions of likelihood, compare probabilistic statement, and apply them to the context, as shown below:

> **While the ratio of giving birth to a baby with Down syndrome is normally one out of 800, 86 of 753 women had given birth to a baby with this disease, and the new test could diagnose all of these before birth.**

The analysis of critical thinking processes with regard to Statement 1 is summarized in Table 7. Then, the extent to which the participants made use of critical thinking is explained in detail.

Table 7

Critical thinking processes used by participants regarding Statement 1

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM	CM	CM	CM
Analysis	-	-	EI	-
Evaluation	-	-	-	-
Inference	-	-	QE	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; EI: examining ideas; QE: querring evidence.

With regard to *Statement 1*, all the participants made use of the *interpretation* skill by *clarifying* its *meaning*. In the *interpretation* process, Ali and İrem made explicit the statement of "…the ratio of giving birth to a baby with Down syndrome is normally one out of 800..." by appreciating the role of variation among samples selected from the same population and the concept of average. For example, Ali said; "From the 800 mother candidates, from those who became pregnant in the same period, only 1 on average, oh some can be 2 out of 800, zero out of 800, that is, these on average one out of 800 mother candidates"(Q16).

Melek and Meltem, however, were differentiated from the other participants in using their statistical and probabilistic knowledge. They just restated this statement by using different expressions in a simple way.

From a different perspective, Melek also made use of the *inference* and *analysis* skills. In particular, she compared two probabilistic statements (1 of 800 and 86 of 753) by *examining* the *ideas*. Then, she *querred evidence* regarding the base rate for

pregnant women at high risk in having a baby with Down syndrome in order to clarify the statement and make reasonable interpretations and inferences. She uttered the following statements:

> [...]this test was done on people at risk, well, if we had seen this ratio among people at risk, then in my view, I frankly feel that we could do more logical and rational interpretations...these mothers are apparantely not normal...That's why I believe that we need to compare it [86 of 753] with that [the base rate of pregnant women at high risk in having a baby with down syndrome]. (Q17)

In summary, all the participants attempted to *interpret* Statement 1 by *clarifying* its *meaning*. However, they differed from each other by using different statistical and probabilistic knowledge such as variation among samples, average concept, and the need for base rate. Melek was differented from the others by her making use of different critical tihking skills such as *inference* and *analysis*.

Statement 2:

Statement 2, as stated below, reported in the newspaper article requires the participants to know the concept of percentage, concept of conditional probability, and apply them to the context: *The test is not yet completely perfect. Actually, two per cent of the babies no having Down syndrome were incorrectly diagnosed to have Down syndrome.* The analysis of critical thinking processes with regard to Statement 2 is summarized in Table 8. Then, the extent to which the participants made use of critical thinking is explained in detail.

Table 8

Critical thinking processes used by participants regarding Statement 2

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM	CM	CM	CM
Analysis	EI	EI	EI	EI
Evaluation	-	-	-	-
Inference	-	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; EI: examining ideas.

The analysis of the participants' expressions with respect to *Statement 2* indicated that all the participants made use of the *interpretation* and *analysis* skills by *clarifying* its *meaning* and *examining* closely related *statements* reported in the newspaper article. For example, when Meltem was asked what she understood from Statement 2, she reflected interpretation and analysis skills in the figure 2;

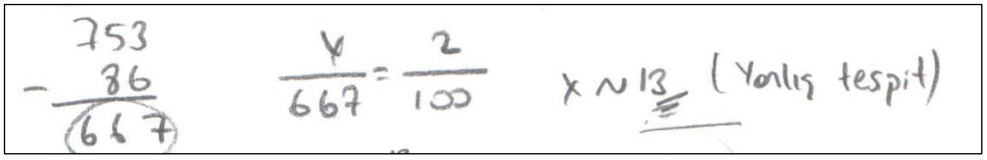


Figure 2. Meltem's interpretation and analysis process

In the interpretation process, she recognized the condition of "babies without Down syndrome" and reasoned proportionally to make clarification regarding the meaning of the statement. She also reflected the analysis skill by recognizing closely related statements (Statement 1 and Statement 2). In a similar way, other participants were in the same process by reasoning proportionally.

Another finding was that all participants attempted to *draw conclusions* to make *inference* from the reported findings when asked to find the accuracy rate of the test.

When the participants were asked to find the accuracy rate of the test, Ali and İrem reached the same conclusions while Melek and Meltem drew same conclusion from the newspaper article. For example, İrem's inference process is presented in the figure 3:

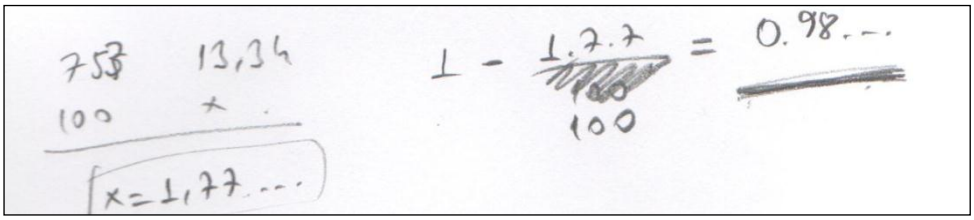


Figure 3. İrem's inference process

Differently, Melek reached the following conclusion: "86 / 99. Because, as a matter of fact, all our events amount to 99 people. The 86 people that the test found as having the disorder, it becomes an event, but in fact there are 13 more people that are overlooked...it becomes 87%." (Q18).

Similar to Melek's conclusion, Meltem reached the same result. However, she also made use of different critical thinking skills. In particular, she attempted to utilize the *evaluation* and *self-regulation* skills. This process is presented in her utterance as follows:

> [...] when I first read the news, I thought that it could really be of benefit and that it could eliminate some things but after I got a pen and paper and did these things, I realized that this percentage [margin of error] was high...when I read this margin of error, I perceived it to be 2% but according to what I did now, I found a higher percentage... you know there is this thing, in elections they play with the bar graphs to show higher higher [percentages], well that's how it seemed, by expressing it with a percentage [2%], I felt like it was lower. (Q19)

In the evaluation process, Meltem endeavored to assess the following argument reported in the newspaper article: *"The new method based on a simple blood test indicates with a high degree of precision whether or not the fetus in the mother's womb has Down syndrome."*

She tried to judge the strength of the argument and the truthfulness of its premise reported in the newspaper as "Actually, two per cent of the babies not having Down syndrome were incorrectly diagnosed to have Down syndrome". She recognized misleading numbers leading readers to incorrect or inappropriate decisions. Futhermore, in the evaluation process, she also tried to explain her reasoning process by reflecting *self-examination* and *self-correction* skills. In the self-regulation process, she attempted to reflect on change in her mind to reach an evaluation and correct her first misperception regarding the newspaper article.

The researcher consciously asked, firstly, the accuracy rate of the study to examine if the participants recognized different conditions for the accuracy rate of the study, which are explained in detail later. Some of them were categorized into four main situations and the participants' thoughts were presented in this circumference. The analysis of the participants' expressions regarding these four situations is summarized in Table 9. Then, the extent to which the participants made use of critical thinking is explained in detail.

Table 9

Critical thinking processes used by participants regarding conditional statements

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CAT	-	CM	CAT
Analysis	EI	EI	EI	EI
Evaluation	AA	-	-	-
Inference	QE CA DC	QE CA DC	QE DC	QE DC
Explanation	-	-	-	-
Self-Regulation	SE	-	SE SC	-

Note. CAT: categorization; CM: clarifying meaning; EI: examining ideas; AA: assessing arguments; QE: querying evidence; CA: conjecturing alternatives; DC: drawing conclusions; SE: self-examination; SC: self-correction.

Situation 1: *How likely is the test to detect the absence of Down syndrome in a baby without Down syndrome?*

Situation 1 requires participants to understand the conditional statement and recognize the condition in the statement (babies without Down syndrome) to find the probability, represented by the researcher with a shaded area in Table 10. They were asked to find the probability of the test detecting the absence of Down syndrome in a baby without Down syndrome, which is called as specificity value (∑True Negative / ∑Down Syndrome Negative) in health statistics. The purpose of the research is to examine whether they recognize different conditions in finding the accuracy rate of the test and misleading language in conditional statements in the newspaper article, not to find the correct answer regarding the question in Situation 1.

Table 10

Accuracy rate of the test regarding specificity

		Actual Result	
		Down Syndrome Positive	Down Syndrome Negative
Test Result	Test Result Positive	True Positive	False Positive
	Test Result Negative	False Negative	True Negative

The analysis of the participants' expressions regarding Situation 1 indicated that three of them (Ali, Melek, and İrem) tried to make *inference* by *drawing conclusions.*

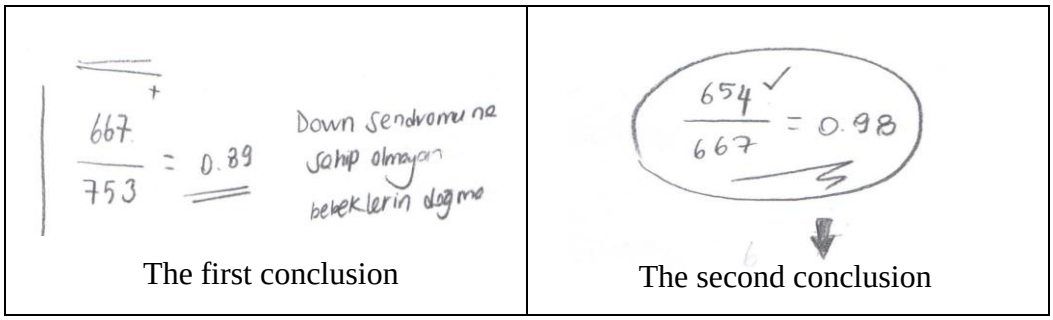
Ali educed a proper conclusion by using relevant information. He recognized the condition in the question. Then, he made calculation based on the condition and found the probability given in the condition as follows: "It doesn't have Down

syndrome, healthy, and the test says that, yes, your child will be born healthy, 654/667 and that will be exactly 98%, it may be a bit too high; 654/667 is 98.05, approximately 98." (Q20)

Melek reached a similar result. She, however, differed from Ali's inference process by reflecting different critical thinking skills; *interpretation* and *self- regulation*. When asked the question presented in Situation1, she recognized the condition in the probabilistic statement, saying "Uumm, we are thinking about the condition of detecting them [babies without Down syndrome]." (Q21). Before make an inference, she needed to *clarify the meaning* of the results reported in the newspaper article. She found the number of the fetuses without down syndrome and then calculated the probability based on the number of all pregnant women. She made wrong inferences regarding Situation 1 although she had recognized the condition in the question at first.

Then, she noticed her incorrect inference regarding Situation 1 when the researcher asked Melek to clarify the meaning of the probability found as 0.89. In this self-regulation process, she attempted to reflect on her inccorrect recognization of the condition and corrected her mistake by making use of the *self-examination* and *self-correction* skills. The following statements represent this process:

> [...]One minute, it occured to me while I was writing; now it's the accuracy of the test you are mentioning, right? Then it is going to be 654/667; now I understood what you said; now our test indicated that 667 individuals would give birth to healthy babies, but with the margin of error, actually it indicated that 654 would give birth to healthy babies...while the overall conclusion that we presume is this [shows 667], the datum we obtain is this [shows 654],..if we evaluate the accuracy of this [prediction], if we evaluate the test, then this time a high ratio of 98% emerges. (Q22)



The first conclusion | The second conclusion

Figure 4. Melek' inference process regarding Situation 1

İrem and Meltem, however, are differentiated from the other participants. While İrem reached a wrong inference regarding Situation 1, Meltem could not *draw* any *conclusion*. İrem, for example, found the probability asked in Situation 1 as 100% because she understood the condition wrongly identifying the "test result as negative". İrem's calculation is summarized by the researcher in the following table:

Table 11

Representation of İrem's inference process formed by the researcher

		Actual Result		
		Down Syndrome Positive	Down Syndrome Negative	**Total**
Test Result	Test Result Positive	86	13	99
	Test Result Negative	0	654	654
	Total	86	667	753

Unlike other participants, Meltem did not attempt to *query* any *evidence* to make inference regarding Situation 1. Besides, she just claimed that there was no information regarding Situation 1 to draw conclusion and it is not possible to find the probability in Situation 1. The reasons behind her claims are that she believed that 86 was not a general number to reach a percentage and appreciated variation

among different samples drawn from the same population. The following statements clarify Meltem's thinking process:

> Here, this 86 is not something overall, it is not a number that indicates a percentage to me because in this sample the number of women giving birth to babies with down syndrome is important; here it was 86 for example. Then, I pose myself this question: if it were 700 and if it detected all 700... now I will do a calculation, I will use this 86 in that calculation but I will use 700. (Q23)

Situation 2: *How likely is the test to detect the presence of a Down syndrome in a baby with Down syndrome?*

Situation 2 requires participants to understand the conditional statement and recognize the condition (babies with Down syndrome) to find the probability, represented by the researcher with shaded area in Table 12. They were asked to find the probability of the test detecting the presence of Down syndrome in a baby with Down syndrome, which is called as sensitivity value (∑True Positive / ∑Down Syndrome Positive) in health statistics. The purpose of the research was to examine whether they recognized different conditions in finding the accuracy rate of the test and the misleading language of the conditional statements in the newspaper article, not to find correct the answer regarding the question in Situation 2.

Table 12

Accuracy rate of the test regarding sensitivity

		Actual Result	
		Down Syndrome Positive	Down Syndrome Negative
Test Result	Test Result Positive	True Positive	False Positive
	Test Result Negative	False Negative	True Negative

The analysis of the participants' expressions regarding Situation 2 indicated that three of them (Ali, İrem, and Melek) attempted to make *inference* by *drawing conclusions* and *analysis* by *examining ideas.* However, three of them reached different conclusions. Ali drew a proper conclusion without help regarding Situation 2. Based on his previous representation, shown in figure 5, he found the probability as 100%. In this process, he used the following information in the newspaper article: "86 of the 753 women had given birth to a baby with this disease, and the new test could diagnose all of these before birth." (Q24). He also *queried evidence* in this process regarding information about the situation of babies born with down syndrome when the test predicted as health and then states as "...but when it says "without making a faulty detection", I understand something like this; all the ones claimed to be born with down syndrome all have down syndrome." (Q25)

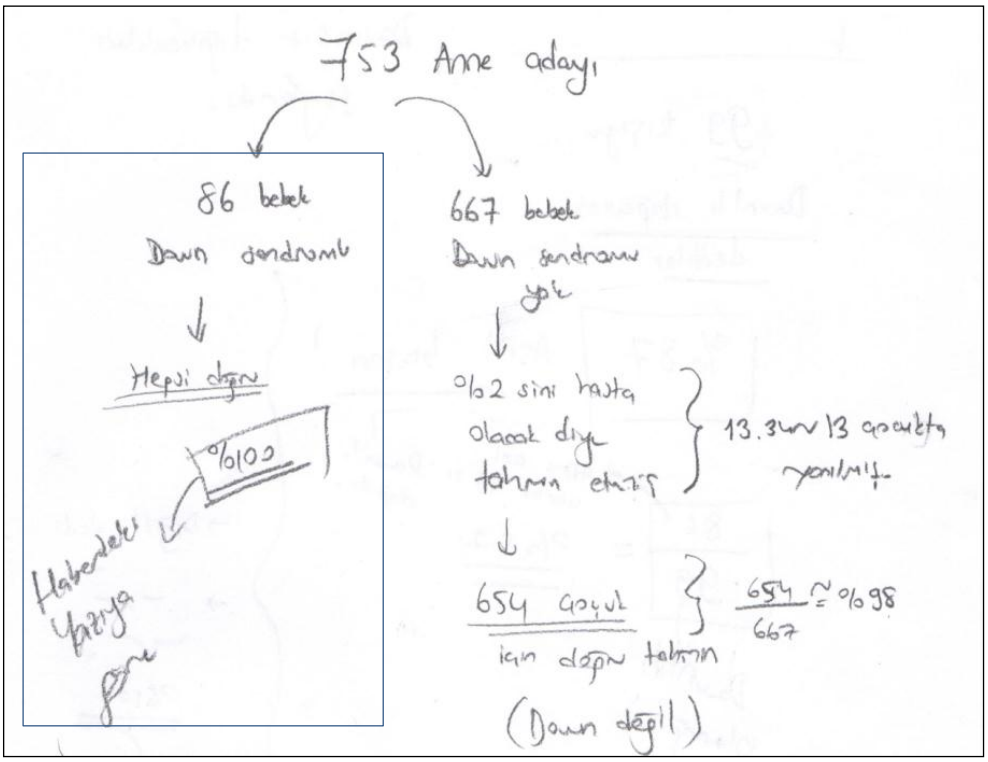


Figure 5. Ali's representation in the inference process

When he was asked to relate the previous conclusions regarding the accuracy rate of the test with this new inference, he could not compare and contrast these situations and recognized the difference among these conditional probabilities. Thus, he failed to *analyse* these situations becuse he could not *examine* different *ideas* in different conditional statements.

Similarly, Melek and İrem could not *analyse* these situations. They also claimed that these were the same situations. Therefore, they reached the same conclusions with their first conclusions regarding accuracy rate of the test, 87% and 98%, respectively.

Moreover, Meltem reflected a different process from the other participants. She attempted to *conjecture alternative* by formulating a new sample with 700 of the 753 women who had given birth to a baby with Down syndrome; but she claimed that it is not possible to find such a percentage since she appreciated variation among different samples drawn from the same population, similar to that Situation 1. Thus she could not *draw* any *conclusion* and recognize the conditions in Situation 2.

Situation 3: *How likely is a baby with a positive test result to actually have the Down syndrome?*

Situation 3 requires participants to understand the conditional statement, underlying *Bayes Theorem*, and recognize the condition (babies with a positive test result) to find the probability, represented by the researcher with the shaded area in Table 13. They were asked to find the probability of babies with a positive test result actually having Down syndrome, which is called as positive predictive value (∑True Positive / ∑Test Result Positive) in the health statistics. The purpose of the research was to examine whether they recognized different conditions in finding the accuracy rate of the test and misleading language of the conditional statements in the

newspaper article, not to find the correct answer regarding the question in Situation 3.

Table 13

Accuracy rate of the test regarding positive predicted value

		Actual Result	
		Down Syndrome Positive	Down Syndrome Negative
Test Result	Test Result Positive	True Positive	False Positive
	Test Result Negative	False Negative	True Negative

The analysis of the participants' expressions regarding Situation 3 indicated that all of them attempted to *draw conclusion* and they reached the same result. However, they differed from each other in terms of their reasoning process. For example, Ali attempted to make use of *analysis*, *inference*, and *self-regulation* skills. More specificially, he *examined related results* of the study in the analysis process and proposed a new alternative for the accuracy rate of the test. He was aware of different interpretations of the accuracy rates of the test by *conjecturing alternative* to find the probability in Situation 3. In this *inference* process, he also reflected his thinking process self-consciously.

> 86 out of 753 is correct, and there were 13 errors. They either told 99 people that they would give birth to a baby with down syndrome...now, if it were me, I would have another assessment; to what extent is this test successful. Here comes another calculation... the success rate of this test is actually 87%, the real success, at the moment I interpreted it in that way, I brought a new perspective on my own hook. (Q26).

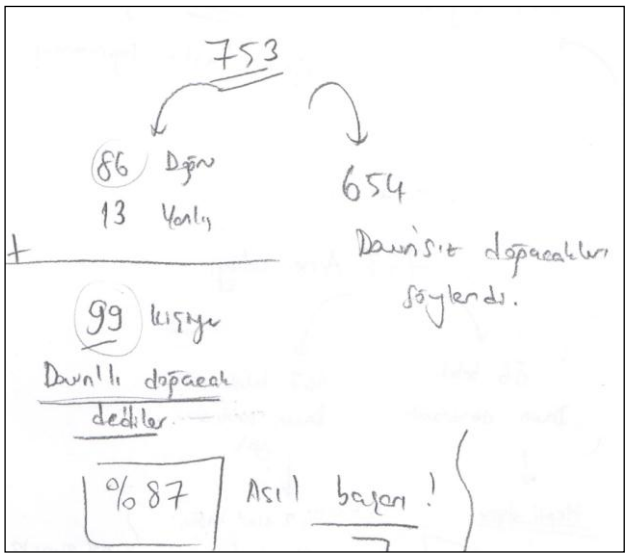


Figure 6. Ali's representation in inference process regarding Situation 3

When asked the relationship between the previous situations and new conclusion, he also attempted to make use of the *evaluation* skill by *assessing* the following *argument* reported in the newspaper article: *"The new method based on a simple blood test indicates with a high degree of precision whether or not the fetus in the mother's womb has Down syndrome."* In the *evaluation* process, he could notice the misleading number in the newspaper article, which could lead readers to reach a wrong conclusion and make improper decisions regarding their life. Moreover, he could not make *analysis* by examining the relatioship between these different four situations. The following paragraph exemplifies this process:

> now since it [newspaper article] said here "without error", according to the news article I said if they are not making error, saying that babies to be born with down syndrome will have down syndrome is 100%, but saying that babies without down syndrome will have down syndrome will be 2%, these are 98%...87% of real success, this becomes a deception, a very high ratio to be deceived by. (Q27)

In a similar way, the other participants reached the same result. However, they could not identify the difference between the four situations. Thus, their conclusions would emerge by chance because they could not reflect the *analysis* skill by

examining the ideas within the different conditions in the situations. Although Meltem just expressed the difference between Situation 2 and Situation 3, none of them were aware of the different situations underlying conditional probability.

Situation 4: *How likely is a baby with a negative test result to actually not have Down syndrome?*

Situation 4 requires participants to understand the conditional statement, underlying *Bayes Theorem*, and recognize the condition (babies with a negative test result) to find the probability, represented by the researcher in Table 14. They were asked to find the probability of a baby with a negative test result actually not having Down syndrome, which is called as negative predictive value (∑True Negative / ∑Test Result Negative) in health statistics. The purpose of the research was to examine whether they recognized different conditions in finding the accuracy rate of the test and the misleading language of conditional statements in the newspaper article, not to find the correct answer regarding the question in Situation 4.

Table 14

Accuracy rate of the test regarding negative predicted value

		Actual Result	
		Down Syndrome Positive	Down Syndrome Negative
Test Result	Test Result Positive	True Positive	False Positive
	Test Result Negative	False Negative	True Negative

The analysis of the participants' expressions regarding Situation 4 indicated that all of the participants attempted to make *inference* by *querying evidence* and *drawing conclusions*. In the process of inference, Ali, for example, made use of the *analysis* skill by *examining closely related results* and *queried evidence* regarding Situation

4, and examined self-consciously his reasoning by rereading relevant information in the newspaper article if he overlooked some important information in the process of inference. By making use of the *categorization* skill, he also organized the results of the study, which were related with each other.

> Uumm, it said "it isn't", I mean, it yielded a negative result; saying, "is this 100 percent, no it isn't"; there is no information in the news, let me look again for a minute... there isn't the necessary information. Well, I think it is 100% because there is no information in the news. It says uumm it [the test] did not make error in the 86 births with down syndrome, it said that they were all going to have down syndrome, and they all did. .. Now let's interpret it well (draws the diagram below)[...] (Q28)

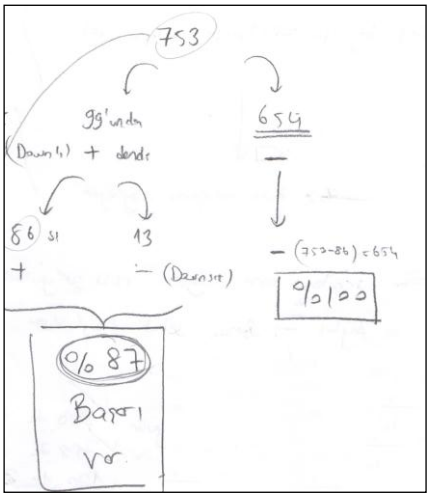


Figure 7. Ali's categorization in the process of inference regarding Situation 4

Similarly, İrem attempted to make *inference* by *drawing conclusions* and *querying evidence* regarding Situation 4, and *analyzed closely related results*. She also tried to *categorize* the results of the study regarding Situation 4. However, she could not organize them in a clear way when compared to Ali's categorization. She queried evidence and drew conclusions as stating "…there is no example of those detected

to be healthy being born with Down syndrome... so if this test says that it is healthy, you can rely on it 100%...” (Q29).

Similar to İrem and Ali, Meltem utilized the same critical skills in the inference process. However, she differed from the others from a different perspective. More specifially, she expressed that there was a missing information regarding the number of pregnant women with a negative test result diagnosed as down syndrome. Thus, she drew a more general and uncertain conclusion compared with the other participants:

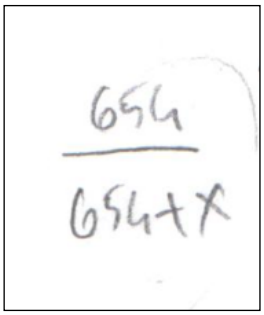


Figure 8. Meltem’s conclusion regarding Situation 4

4.1.3 Generalizability of the Reported Findings

This part includes participants’ critical thinking processes with respect to generalizability of the reported statistics in the newspaper article. Analysis of participants’ critical thinking processes regarding generalizability is summarized in Table 15. Then the extent to which participants made use of critical thinking is explained in detail.

Table 15

Critical thinking processes used by participants regarding generalizability of the reported statistics

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	-	-	-	-
Analysis	-	-	-	-
Evaluation	AA	AA	AA	-
Inference	CA	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. AA: assessing arguments; CA: conjecturing alternatives

Analysis of participants' critical thinking processes regarding generalizability revealed that three of them (Ali, Melek, and Meltem) attempted to *assess* the main *argument* reported in the newspaper article**.** In this *evaluation* process, they recognized the relevant factors to decide if a given argument was applicable to the other situations. Participants, however, differed from each other by focusing on different factors, presented in the Table 16:

Table 16

Factors regarding generalizability focused by participants

	Participants			
Factors	Ali	Meltem	Melek	İrem
Representativeness	x			
Cultural Factors			x	
Confounding Variables				x
Sample Characteristics	x		x	
Chance Variability	x	x		

Of the participants, Ali considered more factors regarding generalizability compared to the others. More specifically, he put emphasis on the notions of sample size, representativeness, and chance variability to determine whether the study was generalizable. Ali perceived the notion of generalizability as an inference from the

sample to the population. In this process, he also *conjectured alternatives* regarding sample size (100.000 or 25 people) by projecting difficulties or benefits in different situations to evaluate the results of the study. Ali stated as follows:

> [...]for instance, there are 100000 people living in a city; if they do this research by examining all the pregnant women among the 100000 people of that year, you would say 'yes, the result obtained is correct' because they implemented it [the test] on everybody, but if they said only 25 people were selected and implemented on, you may not be able to trust it; it may not represent everyone since because it represents the population, the more people it represents, the closer our predicted probability gets to the existing probability. (Q30)

Toward the end of the study, he also considered the factor of chance in the generalizability of the results, stating:

> Now the test is 60% successful; you will go, it will yield a correct result, you will come across 3 out of 5 people, but the test is 90% successful, you will chance on the 10% part; in other words, there is no need to create an atmosphere in which there is excessive confidence or too many biases or a pessimistic outlook. (Q31)

Similar to Ali, Melek focused on the sample characteristic and appreciated the role of the sample to make a generalization from the reported study. She also recognized cultural factors in the *evaluation* of reported statistics' generalizability as follows:

> Of course the group of people selected has to be limited...if it is done all round the world or on all the women in that country, we will obtain reliable results; ...but you know these 753 women were in the higher risk group; now if we generalize, we will be eliminating the group with women at lower risks but, in fact, there is such a group...That is why I think the test [result] should not be generalized...for all people, to be able to generalize [the results] to the entire country, I think, well, 800 people is a small number. For instance, I'm thinking of my own country; in my opinion, it would be problematic to make a decision saying it is definitely accurate or it is definitely inaccurate since it is tried on 800 women....We cannot generalize it [to Turkey]. We are talking about the DNA[...] (Q32)

In the evaluation process, Meltem focused on the notion of chance variability, which requires appreciation of the variation among samples selected from the same population, stating; “To investigate the reliability of this, let me get another 753, and another...that’s why I can’t detect its accuracy, it seems like I can’t, I will say something wrong; it will be more general in this way, it looks like it will be more right.” (Q33)

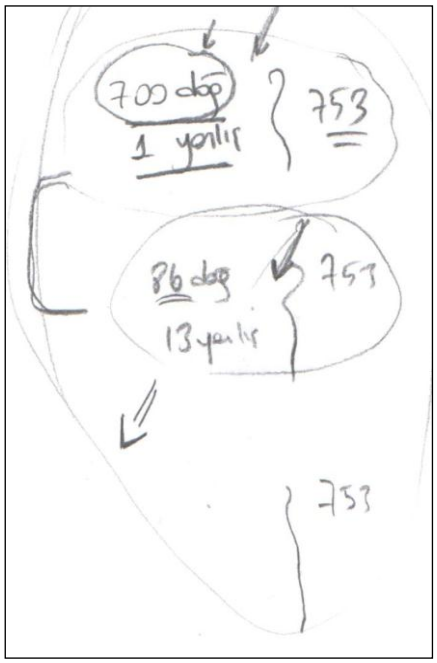


Figure 9. Meltem’s expressions in the evaluation process

From a different perspective, she recognized the factor of confounding variables. She states that the study could be generalizable provided that all conditions such as no smoking is the same for all participants of the reported study on the factor of confounding variable that has potential effect on the result of the study.

In summary, most of the participants attempted to *assess the argument* reported in the newspaper article to evaluate the generalizability of the reported statistics. However, they recognized different factors, such as sample size, sample characteristics, cultural factors, and chance variability in the process of evaluation.

4.2 Newspaper Article II

The findings regarding Newspaper Article II are organized into three main topics; bases of the reported statistics, reported statistics, and generalizability of the reported findings in the newspaper article.

4.2.1 Bases of the Reported Findings in the Newspaper Article

This part includes participants' thoughts about the bases of the findings reported in the newspaper article with regard to their statistical and probabilistic knowledge. Bases of reported findings in the newspaper article consist of background of the study in the newspaper article; how the sample could be selected, how data could be collected and analyzed, and how statistical conclusions reported in the newspaper could be reached. Thus, participants' critical thinking processes regarding the bases of the reported findings in the newspaper article are presented under four main headings respectively; sampling, data collection, data analysis, results and conclusions.

4.2.1.1 Sampling

This part includes participants' critical thinking processes with respect to the sampling of the study reported in the newspaper article. Participants made use of critical thinking skills by using their knowledge of sampling, which includes identification of sample characteristics, sampling method, role of sample on inference from sample to population, and possible biases in sampling. Analysis of participants' critical thinking processes regarding sampling is summarized in Table 17. Then the extent to which participants made use of critical thinking is explained in detail.

Table 17

Critical thinking processes used by participants regarding sampling

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM	CM	CM	CM
Analysis	-	-	-	-
Evaluation	AA	-	-	-
Inference	CA	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; AA: assessing arguments; CA: conjecturing alternatives.

The analysis of participants' knowledge regarding sampling showed that all of them focused on the notion of sample size as a characteristic of sample. They restated the phrase of "203 young couple" reported in the newspaper article by using different words. In other words, all participants were in the process of *clarifying meaning* by paraphrasing the sample size of the study. Meltem, for example, stated the sample size of the study as "[…]In the research, there are 203 young couples. Here, firstly, 203 young couples, that is, there are 203 women and 203 men […] (Q34)". Similarly, Melek paraphrased sample size by saying, "Now, calculations was conducted regarding 203 young couple. So, 406 people participated to the research. (Q35)"

Ali, however, differed from a different perspective when compared with the other participants. More specifically, Ali made use of the *evaluation* and *inference* skills besides the *interpretation* skill. By *assessing arguments* in the process of evaluation, he recognized sample size as a factor which makes the reported study credible while assessing the degree of credibility of the given information as in his accounts presented below:

> […] 203 couples, actually, the number is good. It gives considerable information, that is, it is neither 1 nor 2, 200 couples… Well, I thought, 203 couples,… if this study had been conducted with 20 couples, the answers would not have been all

> that good, the extreme properties of those 20 people would have emerged [if 20 people were selected], but as the sample size increases, the power of representativeness of the sample also increases. If this research had been conducted with 2000 people, we would state that it was more generalizable and more accurate […] (Q36)

In this evaluation process, it is clear that Ali appreciated the essential role of the sample, which required understanding the need to make inference from the sample to the population and the need for samples to be representative. Moreover, during the evaluation of the study, he attempted to make use of the *inference* skill by *conjecturing two alternatives* with respect to sample size; 20 couples and 2000 people. He, then, projected advantages and disadvantages which are likely to result if those alternative sample sizes were selected. On the other hand, he could not assess to what extent sample was representatively structured. This may be due to the fact that accessible population of the study was not reported in the newspaper article or that the participant lacked knowledge in sampling processes.

To conclude, all of the participants were in the process of *interpretation* with regard to the sample size of the study. They, however, did not interrogate how the sample was selected or whether it was biased and could not detect the fact that other characteristics of the sample such as age, the number of years married and their area of residence were not reported in the newspaper article, which prevents people from making proper judgements regarding the research reported in the newspaper.

4.2.1.2 Data Collection

This part includes participants' critical thinking processes with respect to the data collection procedure of the study reported in the newspaper article. Participants made use of critical thinking skills by using their knowledge regarding data collection, which includes knowing at least intuitively how data can be produced on the basis of research design, contribution of a well-designed research for data production to answer specific questions, possible bias in data collection, what was

measured by instrument, or the relationship specific questions asked in the instrument and the reported findings. The analysis of the participants' critical thinking processes regarding data collection is summarized in Table 18. Then, the extent to which participants made use of critical thinking is explained in detail.

Table 18

Critical thinking processes used by participants regarding data collection

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM	CM	CM	CM
Analysis	-	-	-	-
Evaluation	AA	-	-	-
Inference	QE	-	-	-
Explanation	PA	-	-	-
Self-Regulation	-	-	SE	-

Note. CM: clarifying meaning; AA: assessing arguments; QE: querying evidence; PA: presenting arguments; SE: self-examination.

When the participants were asked how the study published in the newspaper could have been conducted, all the participants made comments about the data collection procedure of the study by *clarifying the meaning* of the statement reported in the newspaper article: "Researchers at Virginia Commonwealth University in Richmond gave confidential questionnaires to 203 young couples, asking them whether they had ever strayed, and they suspected or knew their partner had." in the newspaper article. In other words, all the participants were in the process of *interpretation* with an attempt to clarify the meaning of the statement. İrem, for example, made clarification by restating the statement as "203 young couples were asked questions, like did you cheat on your spouse, do you know that your spouse is cheating on you, and are you suspicious of your spouse" (Q37). Meltem also explained her lack of knowledge regarding questions asked to the subjects in the data collection instrument. This may be due to lack of contextual information reported in the newspaper article regarding items in the data collection instrument.

On the other hand, Melek made a strong clarification compared to Meltem and İrem. She made explicit the meaning of the statement by removing its ambiguous language. Moreover, she attempted to make use of *self-examination* skill by self-consciously monitoring her own reasoning regarding clarification of questions reported in the newspaper as shown below:

> For example, at first, misconception was that… I mean, thinking of cheating in the past, [I thought that it was] a question such as 'have you ever cheated in daily life". It could have just said like this: the young couple could have been informed that this study was about their current relationship or, in the same way, in men's current relationships, a statement that currently 29% of them cheat [on their spouses] could have been given. (Q38)

Likewise, Ali reflected different critical thinking processes besides *interpretation*. He made use of the *evaluation, inference* and *explanation* skills in his comments about the statements reported in the newspaper article. To illustrate, he recognized the possible bias in the measurement by *assessing the argument* reported in the newspaper article as "The results, published in New Scientist, show 29 per cent of men admitted that had cheated compared with 18.5 per cent of women.". He indicated a healthy concern about the logical strength of the author's arguments on the basis of its premises, stating "Are men better confessors or deceive [their spouses] more, unclear… some deceive [their spouse] and say they didn't. They, for example, refrain from the researcher; that's why, I think, this may not give an idea about who deceives more." (Q39). In this *evaluation* process, he raised a question regarding the possible threat of subject characteristics to the internal validity of the study, which could emerge in the data collection process. Moreover, he examined the article with a large standpoint towards the end of the interview by making use of the *inference* and *explanation* skills. He judged what background information would be useful to make the claim "…they [males] are only more suspecting because they are more likely to cheat" a persuasive argument by *questioning evidence* regarding questions asked to the subjects in the study. He also attempted to *present an argument*, stating "I think different questions should be

asked to [make this inference]". In his *explanation* processes, he supported reason with a counter evidence to the author's claim by stating "…a man who did not cheat his partner would also be suspicious unnecessarily".

In conclusion, all the participants focused on the issue of what was measured and how it was measured by *clarifying the meaning* of the statements reported in the newspaper article. Two of them differed from the others by reflecting different critical thinking processes such as *self-regulation*, *inference*, *evaluation* and *explanation*. In addition, they had a sense of survey design even though they did not state the term *survey design*. None of the participants, on the other hand, discussed to what extent this instrument measured their characteristics in a reliable way and the errors which may have arose during the data collection.

4.2.1.3 Data Analysis

This part includes participants' critical thinking processes with respect to the data analysis of the study reported in the newspaper article. Participants made use of critical thinking skills by using their knowledge regarding data analysis, which requires them to have a sense of how data were analyzed and detect at least informally possible problems in the process of data analysis. The analysis of participants' critical thinking processes regarding data analysis is summarized in Table 19. Then, the extent to which participants made use of critical thinking is explained in detail.

Table 19

Critical thinking processes used by participants regarding data analysis

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CT	CT	CT	CT
Analysis	EI	EI	EI	EI
Evaluation	-	-	-	-
Inference	-	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CT: categorization; EI: examining ideas.

The analysis of participants' critical thinking processes regarding data analysis revealed that all of them attempted to make use of the *interpretation* and *analysis skills* when they were asked to explain how the researcher could manage data regarding the study reported in the newspaper article. The participants gave almost similar answers in this process. Melek, for example, commented on the statement in the newspaper article as "The results, published in New Scientist, show 29 per cent of men admitted they had cheated compared with 18.5 per cent of women". She attempted to judge how data could be organized and broken up into more manageable groups in the processes of *categorization* and *examining ideas* as stated below:

> First of all, I think they divided them into 2 main groups as men and women, and they they must have examined the two groups in the same way...for instance, in a question like 'have you ever cheated [on your spouse]', it is [found] that 29% of men and 18.5% of women do; they made tables accordingly; I think that they have made inferences such as women cheat more or less. (Q40)

Moreover, all the participants made similar explanations regarding the statement reported as "Men were more likely to catch out a cheating partner, picking up on 75 per cent of the reported infidelities compared with 41 per cent discovered by women." They made comments on the data analysis of the study in terms of the statement. In this process, they made use of the *categorization* and *examining ideas*

skills. They described data management and analysis procedures by providing appropriate categorization (women versus men) and determined how data could be more manageable by examining and comparing the answers of the subjects in the reported study. To illustrate, Meltem stated as below;

> […]for instance, women mark whether or not they have cheated [on their husbands]; for example, there is 'yes' or 'no'; what percentage of women marked 'yes', and they can do uumm; as there are couples, for example, does his/her spouse know that he/she is cheating [on him/her], they must have checked this as well from there. They may have checked couple by couple. (Q41)

İrem, on the other hand, gave rich information about the data analysis of the study compared to the other participants, as stated in the figure 10. She coded correct inferences as "1", incorrect inferences as "0".

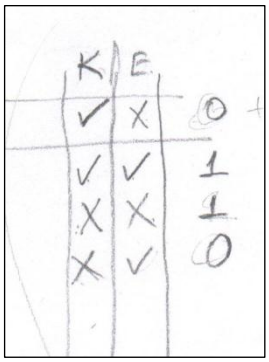


Figure 10. İrem's categorization of the findings reported in the newspaper article

As can be seen in Figure 15, İrem formed two contingencies; women versus men and cheating versus not cheating and related the raw data to percentages as a summary statistics. In other words, she attempted to classify and compare data or findings based on their attributes such as cheating and not-cheating.

In summary, all the participants possess some sense of how to summarize data such as using percents, graphs or tables. In addition, all of them attempted to use the

analysis and *interpretation* skills with respect to how data could be analyzed, in particular how to summarize data by using percentage.

4.2.1.4 Results and Conclusions

This part includes the participants' critical thinking processes with respect to how the results and conclusions reported in the newspaper article were reached. Participants made use of critical thinking skills by using their knowledge, which required understanding that differences between groups may not be large enough or stable to draw reliable conclusions or they could be caused by chance processes and that attention needs to paid to the size of the groups, to the quality of research process, which have a potential effect on the results of the study.

The analysis of participants' critical thinking processes regarding results and conclusions reported in the newspaper article is summarized in Table 20. Then, the extent to which participants made use of critical thinking is explained in detail.

Table 20

Critical thinking processes used by participants regarding results and conclusions

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	DEC	DEC	DEC	DEC
Analysis	-	-	-	-
Evaluation	AA	AA	AA	AA
Inference	-	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	SE	SE

Note. DEC: decoding significance; AA: assessing arguments; SE: self-examination.

The analysis of the participants' critical thinking processes regarding results and conclusions reported in the newspaper article indicated that all participants had a tendency to *interpret* and *evaluate* the reported conclusions. Moreover, some

participants made use of the *self-regulation* skill. The findings of this part were organized into three main results. Firstly, all participants attempted to detect the author's indirect intentions by making use of the *decoding significance* skill when the participants were asked what they thought about the author's intention or purposes of informing the readers about the study. Meltem, for example, stated as follows:

> [...]It looks like normally as if the journalist also thought that women noticed this better but that wasn't the case; as if the journalist is saying look here there is a study, that's how I felt the attitude of the journalist. I mean as if it were a new thing for the journalist as well...he/she comes by saying stuff like 'women be careful' to draw attention quickly[...] (Q42)

Beyond what she stated above, she also described the main idea as the last paragraph of the article that author intended to express, by distinguishing it from subordinate conclusions reported in the newspaper article.

None of the participants, however, could detect and describe the authors' one-sided arguments that included only men's correct inferences by comparing them with women, rather than looking at the conclusions of the study.

Secondly, all the participants attempted to *assess the arguments* reported in the newspaper article. However, three of them (Ali, Meltem, and Melek) could not assess the acceptability of the given conclusions, which required them to realize that observed differences may not be statistically significant or may not be large enough or consistent or can occur by chance. In the following example, Melek stated;

> […] If numerical data are compared, for instance, it was found that 80 of women's inferences were correct but 94% of men were right in these inferences. There is 14 % difference, below [last paragraph of the article] there is a much higher difference. It can have a difference of 75%; in the other one it can detect 41%, so that's why I thought the test is really reliable. (Q43)

İrem, on the other hand, recognized the importance of sample size for each cell or category to assess the degree of credibility of the results or conclusions and detected the fact that the number of participants for each category or condition such as cheating or not cheating and being right or wrong may not be enough for the reliability of the conclusions drawn. For instance, she stated that if the number of people cheating [on their spouses] had been higher than 29%, the ratio of 41% would change possibly. Moreover, she detected inequality of sample sizes for each category, represented in figure 16, as stating there are much more people in the category of "women/men not cheating and their partner inferences right." (the third row in the tables).

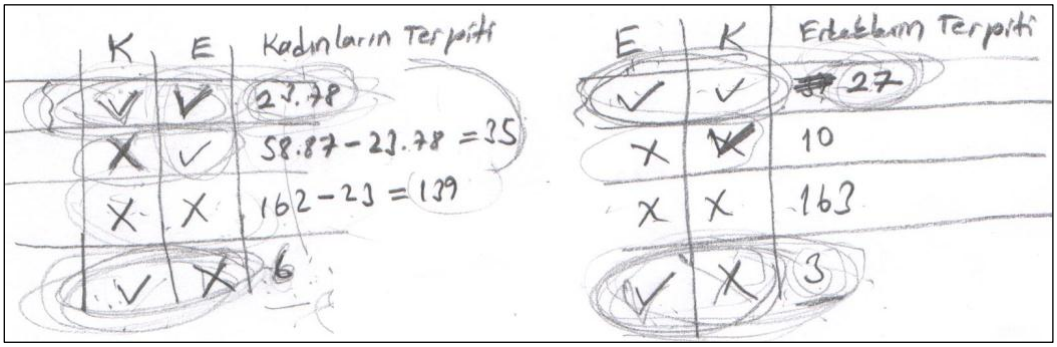


Figure 11. İrem's categorization of the results reported in the newspaper article

The last main finding was that Meltem and Melek tried to make an objective assessment regarding their reasoning process at the end of the interview by somewhat the reflecting *self-examination* skill. Melek, for example, stated the following:

> For example, if it says 29% of men cheat [on their wives] more, then yes I think so too…Then, for example, I would not have thought men were better at detecting this [cheating]…when assessing the test, I think I'm adding my own opinions a little too; but, for instance, when considering its reliability, I'm looking at the claims made at the beginning and the numbers below, I'm comparing them. So even if I am not doing calculations, maybe because it fits my line of thought a little, I mean I believe in it more. (Q44)

To conclude, all the participants were in the processes of decoding significance and assessing the arguments. Meltem and Melek differed from the others by their self-consciously monitoring their thinking process at the end of the interview.

4.2.2 Reported Statistics

This part includes the participants' critical thinking processes regarding the reported statistics in the newspaper article. The reported statistics refers to summary or descriptive statistics reported in popular media texts such as percentages, measures of central tendency, graphs or probabilistics statements etc. Newspaper Article II includes probabilistic statements and percentages as reported statistics. Thus, the participants' critical thinking processes were, in detail, examined on the basis of probabilistic statements and percentages.

4.2.2.1 Probabilistic Statements and Percentages

In this part, the participants' critical processes during their examination of probabilistic statements reported in the newspaper article are presented. What statistical and probabilistic knowledge they made use of in this process is another finding of this part. There are four probabilistic statements in the newspaper article. This part is organized in terms of these statements respectively.

Statement 1:

Statement 1 reported in the newspaper article requires the participants to know the concept of percentage and apply it to the context, as shown below. The analysis of the critical thinking processes with regard to Statement 1 is summarized in Table 21.

The results, published in New Scientist, show 29 per cent of men admitted they had cheated compared with 18.5 per cent of women.

Table 21

Critical thinking processes used by participants regarding Statement 1

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM	CM	CM	CM
Analysis	-	-	-	-
Evaluation	-	-	-	-
Inference	-	-	-	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning.

The analysis of the participants' expressions with respect to Statement 1 indicated that all of the participants attempted to *interpret* the statement by *clarifying* its *meaning*. Their interpretations, however, differed from each other by applying different ways of clarification. Three of them (Ali, Melek, and İrem) clarified the meaning of the statement by restating the authors' claim using different words. Melek, for example, interpreted it by reasoning proportionally as in the following figure:

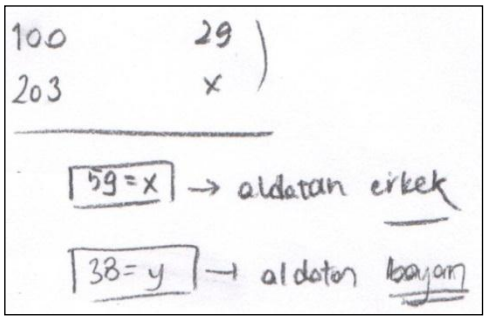


Figure 12. Melek's interpretation of Statement 1

Meltem, however, tried to make clarification by removing the ambiguous language in the statement. She expressed that the number of female cheating was 203 and the

number of females admitting they cheated their partner was 29 per cent of 203 since she misunderstood the statement. She said;

> Fewer women have confessed but is it because they didn't cheat [on their husbands]?, no they do cheat too but I guess some lied, they didn't confess, considering the word "confess", that's what I understand…only 18% of the women confessed cheating, the others didn't confess…then it is not clear how many cheat and how many don't, so I thought having given this ratio is illogical and that's why I thought as cheating women. (Q45)

Statement 2:

Statement 2 reported in the newspaper article requires participants to know the concept of percentage as an expression of likelihood and apply it to the context, as shown below. The analysis of critical thinking processes with regard to Statement 2 is summarized in Table 22.

> **Researcher Paul Andrews said men were better at judging fidelity than women. "Eighty per cent of women's inferences about fidelity or infidelity were correct, but men were even better, accurate 94 percent of the time" Dr. Andrews said.**

Table 22

Critical thinking processes used by the participants regarding Statement 2

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CM	CM CAT	CM	CM
Analysis	-	-	EI	EI
Evaluation	-	-	-	-
Inference	-	-	QE CA	-
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; CAT: categorization; EI: examining ideas; QE: querying evidence; CA: conjecturing alternatives.

The analysis of the participants' critical thinking processes with respect to *Statement 2* indicated that all of the participants made an attempt to comprehend the meaning of the statement by *clarifying the meaning* of the statement. Meltem, for example, explained the statement through an example:

> […]In other words, they asked, 'do you think your spouse is cheating on you' and the woman said 'no'. Actually, they looked and he wasn't cheating. Then someone else said 'yes', they looked and saw that he was cheating. That is, the percentage of knowing correctly whether it is 'yes' or 'no' is 80. (Q46)

When asked how to calculate the number of women who made correct inferences in relation to fidelity or infidelity, Meltem had ability of *clarifying meaning* and *categorization* while *interpreting* the statement.

> …I asked a woman: 1st person, is your spouse cheating on you; let's say, she said 'yes'; we asked her husband, he said, 'no'. Now, this, for instance, did not fit my range. Then, it is not within this 80%, the inference about him is not correct. The 2nd one said 'yes' and this said 'yes' too; this percentage is within 80, "no - no" this is too within 80%... So they asked in this way to 203 couples, there are 203 data at hand; apparently, the couples who answered as yes -yes, no- no were 80%. (Q47)

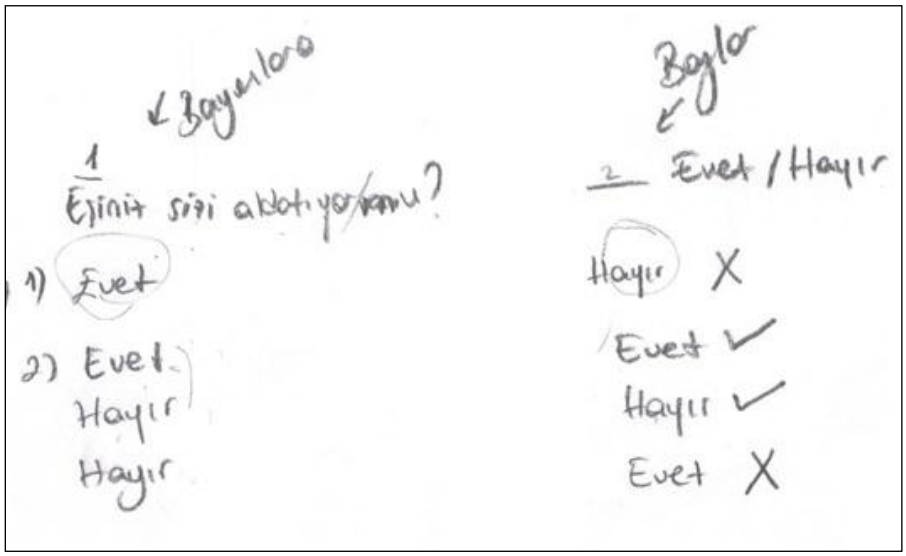


Figure 13. Meltem's categorization of all possible outcomes

Melek and İrem, on the other hand, were differentiated with a different thinking process compared to the other participants. They had difficulty in c*larifying the meaning* of the statement and confused the meanings of infidelity and fidelity when they first read the newspaper. Meltem, for example, said;

> Now, 80% of women's inferences concerning their spouse's fidelity or infidelity is correct. And here it says, women detected 41% of cheating spouses. These two, what, then are they different...I mean 41% of the women who said their spouse cheated on them predicted correctly. So what is this 80%? Then it's [80%] an inference regarding infidelity. I thought this part was related to something else. According to women, what was better? fidelity or infidelity?; I don't know, then I did not understand the concept of fidelity. (Q48)

Likewise, Melek could not distinguish the difference between the Statement 2 and Statement 3. Thus, she could not interpret the statement when she first read the newspaper article, stating "I couldn't understand these last two statistics mentioned (reads the last two paragraphs)... I felt like statements that mean the same, that are stated in the same way are given different values... I perceived that there is perhaps a conflict here." (Q49).

It can be concluded that Melek and İrem had difficulty in analyzing those statements reported in the newspaper since they could not compare and contrast the statements by *examining the ideas*. This may be due to the fact that they could not clarify the ambiguous language in the statement due to the meaning of fidelity or infidelity concepts.

In the further process, when asked what they understood from Statement 2, Ali, Meltem and İrem *clarified the meaning* of the statement in a similar way and made the statement more explicit by reasoning proportionally. İrem, for example, stated; "…162 women predicted correctly whether or not their spouse cheated on them. And I understood that 190 men accurately predicted whether their spouse cheated on them." (Q50)

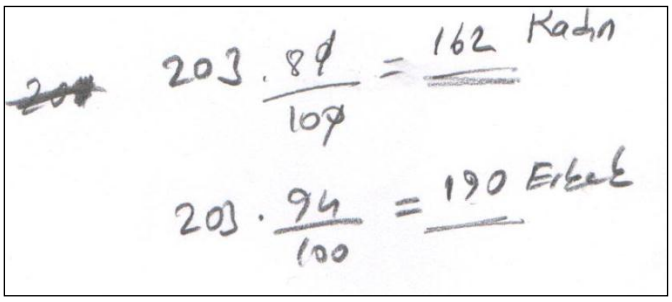


Figure 14. İrem's clarification of the Statement 2

Before the process exemplified above, İrem hesitated if she understood the exact meaning of the concepts infidelity or fidelity and represented all possible outcomes in a table by *categorizing*, which is presented below:

> […] I think I don't know the meaning of the concept 'fidelity'… I can't distinguish these two conditions…I think predicting fidelity correctly means when they say that they don't think their spouse cheated on them and actually they hadn't; and predicting infedility correctly means when they say that their spouse definitely must have cheated on them and the spouse had done so. (Q51)

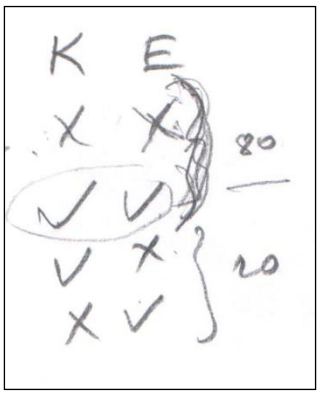


Figure 15. İrem's categorization of all possible outcomes

Melek, however, had a different thinking process when compared to the others. In this process, she attempted to *clarify the meaning* of the statement, *question the evidence*, and *conjecture alternatives* in order to resolve the ambiguity of the language. She interpreted the concepts of infidelity or fidelity as idea of cheating or

idea of not cheating, respectively; though the author describes infidelity as cheating and fidelity as not cheating. Due to misunderstanding the authors' claim, she questioned what background information would be useful to draw a conclusion. To resolve this problem, she also *conjectured an alternative* stating that the author should state the base rate of the fidelity and infidelity for both women and men. It could be concluded that misperception of the infidelity and fidelity concepts prevented her from making further interpretations regarding the statement and from making it explicit. This indicates that newspaper articles may include ambiguous terms and how statements are perceived could lead participants to think differently.

Statement 3:

Statement 3 reported in the newspaper article requires participants to know the concept of conditional probability and apply it to the context, as shown below. The analysis of critical thinking processes with regard to Statement 3 is summarized in Table 23.

> **"Men were more likely to catch out a cheating partner, picking up on 75 per cent of the reported infidelities compared with 41 per cent discovered by women."**

Table 23

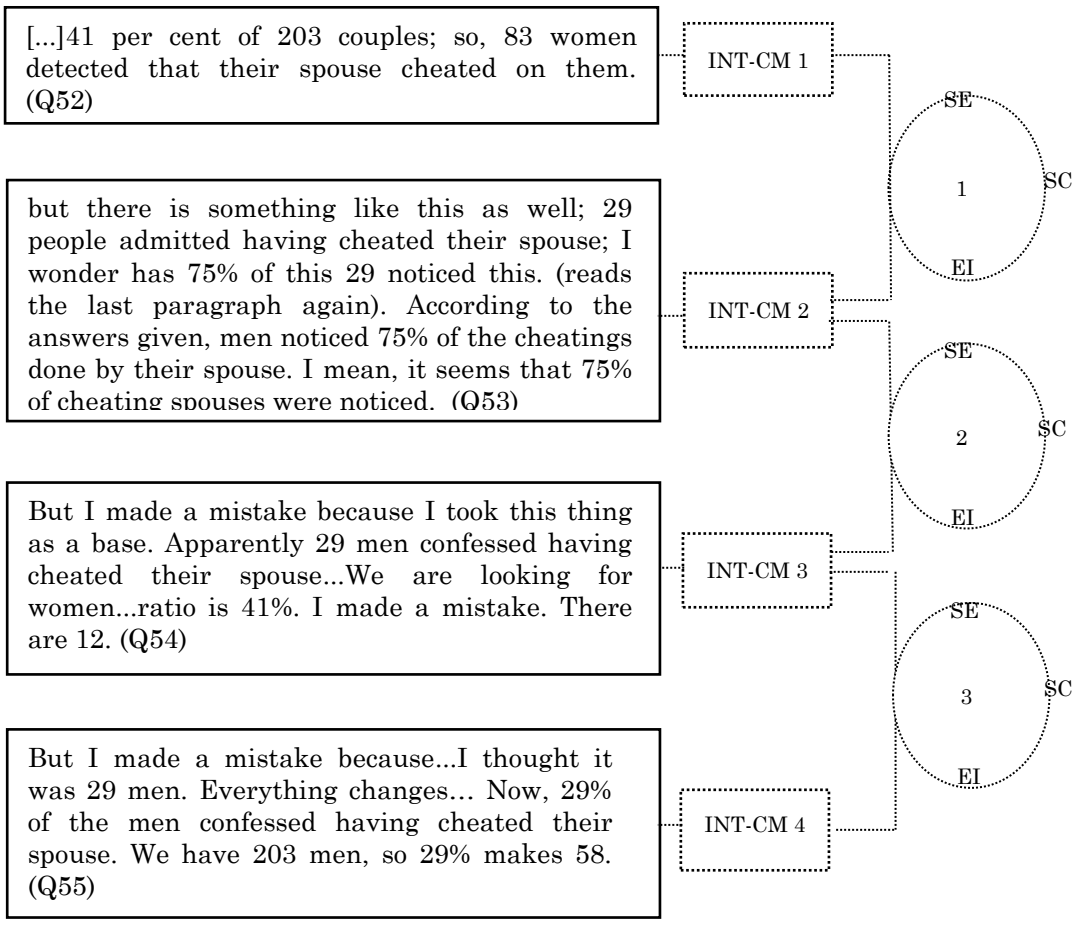
Critical thinking processes used by participants regarding Statement 3

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	CAT CM	CAT CM	CAT CM	CAT CM
Analysis	EI	-	EI	EI
Evaluation	-	-	-	-
Inference	DC	CA DC	DC	DC
Explanation	-	-	-	-
Self-Regulation	SC	SE SC	SE SC	SE SC

Note. CAT: categorization; CM: clarifying meaning; EI: examining ideas; DC: drawing conclusions; CA: conjecturing alternatives; SE: self-examination; SC: self-correction.

The analysis of the participants' critical thinking processes with respect to Statement 3 indicated that all of the participants were in the processes of *interpretation, inference,* and *self-regulation.*

When the participants were asked how the researcher could reach such a conclusion as expressed through Statement 3, participants firstly tended to *interpret* the statement by *clarifying its meaning.* Three of them (Ali, Melek, and İrem) had a similar thinking process. This thinking indicates a recursive process. One example including İrem's critical thinking process was presented in four steps, as shown in the figure below:



Note. INT-CM: interpretation clarifying meaning; SE: self-examination; SC: self-correction; EI: examining ideas

Figure 16. İrem's critical thinking process regarding Statement 3

In figure 16, İrem made an error in the first three steps, which may be due to the fact that İrem did not notice the condition of cheating in the statement and misunderstood the statement. However, in the fourth step, İrem achieved to *examine* two related *statements* (Statement 1 and Statement 3) reported in the newspaper by making use of the *analysis* skill. After correctly interpreting the statement for women, she made an appropriate interpretation of the probabilistic statement for men. In the same way, Ali and Melek have such a recursive process of thinking.

Meltem, on the other hand, noticed the conditional situation in Statement 3 even though she made incorrect interpretations. She *clarified the meaning* of Statement 2 by finding an example and reasoning proportionally although she misunderstood the number of cheating men and women, explained in Statement 1. Such an improper interpretation may be due to lack of the comprehending Statement 1;

> […] among 152 couples, the women said they cheated [on their spouse] and the men said they noticed being cheated. For women, for instance, they asked the men if they cheated [on their spouse], they said 'yes'; they asked the women whether they noticed this and they said 'yes'. And in the same way if this is 41%…Apparently there were 83 couples like this that noticed it. (Q56)

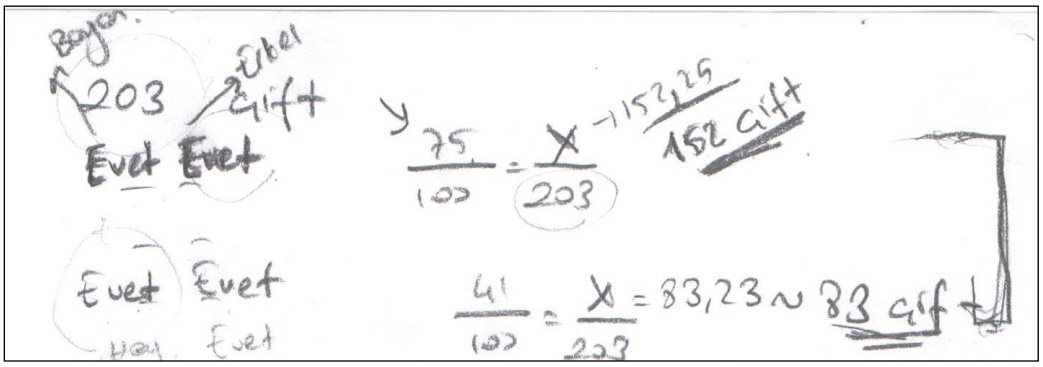


Figure 17. Meltem's interpretation of the Statement 3

An especially interesting finding regarding Meltem's critical thinking process was that she *drew* a new *conclusion* regarding Statement 2. More specifically, when Meltem was asked to explain her reasoning process during the computation process, she stated as in the vignette below:

> […]We think of these two [152 couples ve 83 couples] like a set, now they conflict, but actually they shouldn't have been conflicting... Then they didn't do it like this. I'm thinking whether there is any point that doesn't conflict. Why did it emerge like this? (1*) [reads again] Meltem: ...But the questions pose are different, they don't need to be the same because two-way questions, for instance, they ask men if their spouse cheated on them, then they ask the woman if her spouse cheated on her. You know we were surprised that the same ratio did not emerge because the questions are different, I mean they don't have to be the same. (2*)…I mean, it could actually be the intersection of these [152 couples and 83 couples]. That is, the women may have cheated and her husband may be aware of it; her husband may have cheated on her in a marriage... now, among 152 couples, men know that their spouses are cheating on them. And among 83 couples, women know that their spouses are cheating on them. Now, if these did not intersect, I mean if there weren't this mutual noticing [the cheating], it would be 235 in total. Since this is higher than 203 couples, there must definitely be couples noticing each other's cheatings. (3*) (Q57)

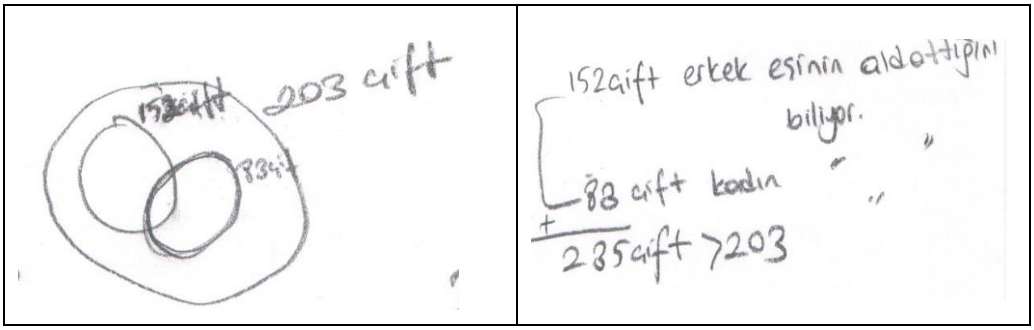


Figure 18. Meltem's inference about Statement 3

In the above vignette, quotations referring to 1* and 2* indicated that Meltem was in the process of *self-regulation*. 1* refers to the *self- examination* process since she monitored her own thinking process and reread the statement to make sure that she did not notice important key points. 2* refers to *self-correction* as she corrected the problem where self-examination reveals misunderstanding. 3* refers to the

inference process in which she *drew* a new c*onclusion* by considering relevant information. When compared to the others, Meltem reached such an interesting conclusion. However, it is also important to emphasize that she drew a conclusion on the basis of misinterpretation of Statement 1.

Secondly, all the participants organized reported conclusions in the newspaper article by *categorizing* and *clarifying* the *meaning* of the information. When asked how Ali could evaluate the newspaper article, for example, he stated the following;

> […] Here there are 3 questions, then (refers to the 4th paragraph from the beginning) 3 different ratios are given. 1st ratios, 2nd ratios, and 3rd ratios (5th, 6th and 7th paragraphs), but in my first reading, I did not understand that. I mean, this could have been presented more clearly…Well, I could present it with a number. 190 of the 203 men participating in the study predict correctly whether they cheat on them or not; I would word it in this way directly. (Q58)

He classified the reported conclusions regarding the questions asked to the subjects in the study and claimed that the author made the conclusions explicitly by expressing the results directly.

Ali and İrem advanced their *categorization* by *examining ideas* and *drawing conclusions*. On the other hand, Meltem and Melek could not go further since they could not make explicit the differences between the conclusions (Statement 2 and Statement 3) reported in the newspaper explicit. This may be due to the ambiguous language in the newspaper article or the participants' lack of analysis skill. İrem and Ali's categorizations of the results reported in the newspaper article were illustrated as follows;

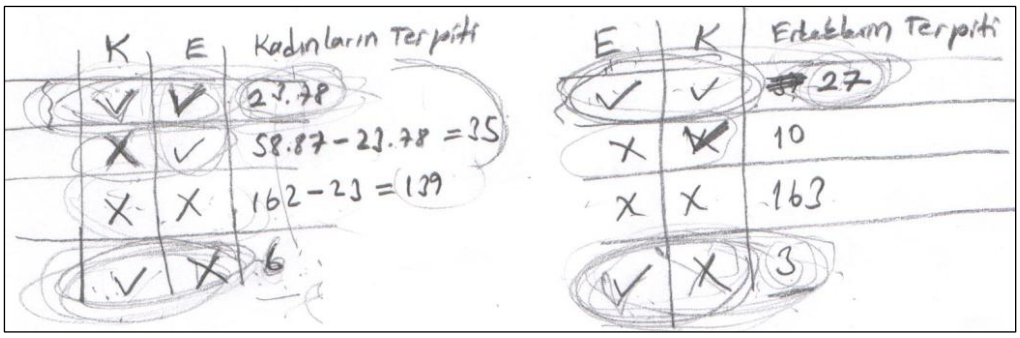


Figure 19. İrem's categorization of the results reported in the newspaper article

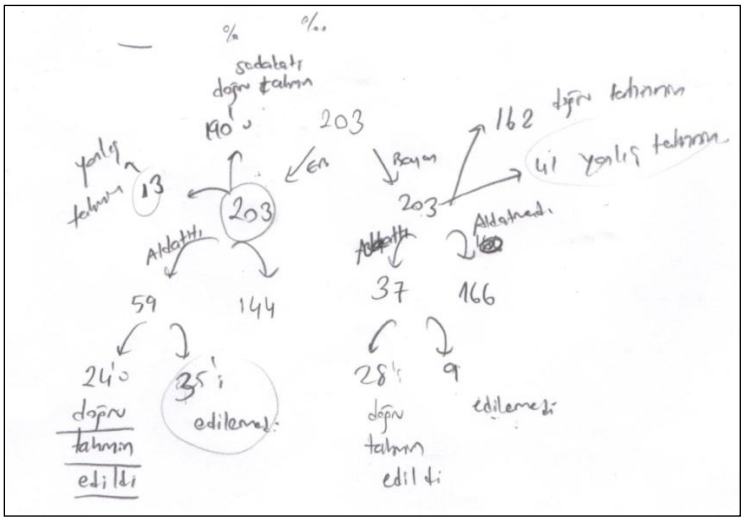


Figure 20. Ali's categorization of the results reported in the newspaper article

Drawing these figurative expressions and tables, Ali and İrem display the ability to *analyze* and make *inference* from the results and conclusions published in the newspaper article besides the ability to make *interpretation*s. They reached similar conclusion on the basis of the reported findings, as in Figure 25. Ali, for example, examined related results and educed proper conclusions as follows:

(1) I would draw 203 people. 203 to 203, 59 men confessed their cheating, and what happens to the remaining; 144 did not cheat. Of the 203 here, 37 cheated and 166 didn't.
(2) Then, 24 of those who cheated were predicted correctly by their spouses. 28 of the 37 who cheated were predicted correctly. Well, 9 couldn't [be predicted].

(3) Then, this correct prediction of their spouse's fidelity made by the 190 spouses of the 203; 13 of them wrong predictions; and of these 203, 162 correct predictions, 14 wrong predictions.
(4) Then, something else emerged, now 41 of the women make a wrong prediction, that 35 made a wrong prediction claiming that their spouse did not cheat on them. So there must be 41 wrong predictions and 6 think that their spouses are cheating on them, but this is not true, they did not cheat [on them]. Let's do the same thing for this here. There are 13 wrong predictions. 9 of them make a wrong prediction claiming that their spouses did not cheat on them; 4 of them, 4 men accused [their spouses] unnecessarily. (Q59)

In Figure 25, Ali identified key findings of the study published in the newspaper article and determined the conceptual relationships of these findings with each other by *examining ideas*. Then, he reached several conclusions not stated in the newspaper article. The researcher organized Ali's inferences (Q59) into two-way tables as shown in Table 24 and Table 25:

Table 24

Male inferences if their partners cheating or not cheating

	Male wrong inferences	**Male right inferences**	**Total**
Female cheating	9	28	37
Female not cheating	4		166
Total	13	190	203

Table 25

Females inferences if their partners cheating or not cheating

	Female wrong inferences	**Female right inferences**	**Total**
Male cheating	35	24	59
Male not cheating	6		144
Total	41	162	203

Thirdly, when asked the probability of men's incorrect inferences about infidelity or fidelity, all participants attempted to *draw conclusion*s from the three statements.

Meltem and İrem, for example, reached a proper conclusion by considering relevant information as follows:

> [...] the probability of men making incorrect predictions about their spouses was less than that of women's…. Their chance of making an error is lower, when compared with women…if men's ratio for correct prediction is 94%, it is 6% for predicting incorrectly…Let's calculate these for women as well. Was it 80% for women. Yes. 20% cannot predict correctly. (Q60)

When the participants were asked what is the probability of incorrect inferences of men whose partner is not cheating, about their partner. Ali and İrem drew inappropriate conclusion as 2% and 1.4% respectively.

However, they misinterpreted the statements since they could not realize the conditional statement. In fact, the correct calculation of the probability should be as follows:

P(Male wrong inferences | Female not cheating)= 3/165= 1.8 %
P(Female wrong inferences| Male not cheating)= 6/144= 4.2 %

On the other hand, Meltem and Melek could not *draw* any *conclusion* since they claimed that there is no enough information like numerical values regarding infidelities that do not exist. Meltem, for example, stated as "I can't calculate that... if it [newspaper article] had indicated the ratio of men who thought that their spouses were cheating on, while actually they weren't, then I could calculate its percentage." (Q61)

Meltem attempted to judge what background information would be useful by *querying evidence* and *conjectured an alternative* by proposing a specific percentage to draw a conclusion, though there is no need for such evidence.

Melek attempted to *draw conclusions* when asked the probability of correct inferences of women whose partners are not cheating, about their partners. She

could not find asked probability since she misinterpreted the probabilistic statement reported in the newspaper. Although 59 per cent corresponds to the probability of women' incorrect inferences if their partner cheating, she interpreted it as the probability of women's incorrect inferences if their partner not cheating. In fact, the researcher asked Melek to the probability of women's correct inferences if their partner not cheating. This confusion indicated that she had difficulty in applying the conditional probability to the social context.

Statement 4:

Statement 4, which reported the conditional statement verbally in the newspaper article, required participants to know the concept of conditional probability and apply it to the context as shown below. The analysis of critical thinking processes with regard to Statement 3 is summarized in Table 26.

> **Men are better at detecting a cheating partner than females, and they are more likely to suspect infidelities that do not exist**.

Table 26

Critical thinking processes regarding Statement 4

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	-	CM	CM	CM
Analysis	-	EI	EI	EI
Evaluation	AC		-	AC
Inference	-		-	CA
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. CM: clarifying meaning; EI: examining ideas; AC: assessing claims; CA: conjecturing alternatives.

The analysis of the participants' critical thinking processes with respect to *Statement 4* indicated that Meltem, Melek, and İrem interpreted the statement by *clarifying the*

meaning and analyzed the statement *by examining closely related statements* reported in the newspaper when asked the meaning of Statement 4. For example, Melek states as "I am thinking of the data in the last paragraph; while men noticed 75%, 3 out of 4, almost all of the cheating spouses, women have not been able to notice half of this ratio. Well, the numerical equivalence of this sentence is this [associates the last paragraph with the title of the news article]." (Q62).

Furthermore, Ali and İrem made use of different critical thinking skills regarding this statement when compared with the others. Ali, for example, attempted to *assess the claims*, but his assessment of claims regarding acceptability level was wrong because his evaluation was on the basis of his wrong conclusion previously drawn, as explained before. He said, "[the percentage] in cheating do not exist was found as 2%. It is wrong according to the numbers [in the newspaper article]…Are women more suspicious...umm in fact it is found so… I think women are more suspicious]" (Q63).

In a similar way, İrem tried to *assess the claim* though she was worried about her reasoning (Q64). In this *evaluation* process, İrem dealt with the *clarification of the Statement 4*. In order to clarify, she *conjectured* two *alternatives* with respect to the concept of "being suspicious" in the following:

1. Male / Female says that his /her partner cheating
2. Male / Female says that his/her partner cheating if his /her partner is not cheating in real world.

İrem claimed that the second alternative would possibly be true although the author's claim was correctly on the basis of the first alternative. This ambiguity led İrem to draw an improper conclusion and partially evaluate the claims.

In fact, the statement of "*they [men] are more likely to suspect infidelities that do not exist"* would lead readers to interpret the statement as İrem interpreted. That is, the given statement corresponds to "The probability of males' wrong inference if their partners do not cheat is more than the probability of females' wrong inference if their partners do not cheat", as symbolized below:

P(Male wrong inferences|Female not cheating) > P(Female wrong inferences|Male not cheating)

When the probabilities are calculated by using the relevant information and conclusions, they do not match with the claim of the author, as follows:

P (Male wrong|Female not cheating) = 3/165=1.8%
P (Female wrong|Male not cheating) = 6/144=4.2%

In conclusion, the *Reported Statistics* part can be summarized in terms of four main conclusions. One of the main conclusions was that the participants had a tendency to *interpret* all the statements reported as descriptive statistics. In particular, they dealt with clarifying meanings of Statement 2 and Statement 3 by making explicit the difference between them. This may be related to participants' language skills or vaguenessof the meaning of the terms expressed in the newspaper article.

The second main conclusion that can be drawn is that the analysis of their expressions indicated that they attempted to make use of different critical thinking skills, especially regarding Statement 3 and 4, ranging from *interpretation* to *self-regulation*. This may be due to the complex nature of conditional probability statements or due to the difficulty they experience in dealing with conditional probability. However, they less frequently reflected the *explanation* and *self-regulation* skills. This could be due to the limitations of the research, which give limited opportunity to the participant to apply these skills or due to the participant's

weak ability of *explanation* and *self-regulation*. Moreover, they indicated inadequate ability in *assessing* the reported *claims*, especially regarding claims which require understanding of the conditional probability.

Another conclusion is that how the statements reported in the newspaper article were perceived could have participants to think differently since newspaper articles may include ambiguous language.

The last main conclusion is that participants had difficulty in dealing with the statements including conditional probability. They had a predisposition to confuse conditional events as a result of which they drew incorrect conclusions. This may pertain to the difficulty in noticing the conditional event or to the grammatical expressions of the statements including ambiguity or vagueness or misuse of conditional probability in the newspaper article. In addition, regarding Statement 3 including conditional probability, participants with the ability to construct a table or diagram to summarize the findings of the study made comments and interpretations more easily by clarifying meaning, analyzing the statements, and drawing conclusions, unlike the others.

4.2.3 Generalizability of the Reported Statistics

This part includes participants' critical thinking processes with respect to generalizability of the reported statistics in the newspaper article. The analysis of the participants' critical thinking processes regarding generalizability is summarized in Table 27. Then the extent to which participants made use of critical thinking is explained in detail.

Table 27

Critical thinking processes used by participants regarding generalizability of the reported statistics

	Participants			
Critical Thinking Process	Ali	Meltem	Melek	İrem
Interpretation	-	-	-	-
Analysis	-	-	-	-
Evaluation	AA	AA	AA	AA
Inference	QE	-	-	QE
Explanation	-	-	-	-
Self-Regulation	-	-	-	-

Note. AA: assessing arguments; QE: querying evidence

The analysis of participants' critical thinking processes regarding generalizability revealed that all the participants attempted to *assess the arguments* presented as summary statistics in the newspaper article when asked what they thought about generalizability of the study. In this *evaluation* process, they recognized the relevant factors to decide if given arguments or claims were applicable to the other situations. Participants, however, differed from each other with their focusing on different factors, presented in the Table 28:

Table 28

Factors regarding generalizability focused by participants

	Participants			
Factors	Ali	Meltem	Melek	İrem
Representativeness	x			
Cultural Factors	x		x	
Sampling Method	x			x
Sample Characteristics	x	x		x
Chance Variability				x

Of the participants, Ali and İrem, also, *questioned evidence* to make a sound generalization regarding generalizability of the published study. Ali, for example,

considered more factors regarding generalizability compared to the others. More specifically, he put emphasis on the notions of sample size, sampling method, representativeness, and cultural factors to determine whether the study was generalizable. Ali perceived the notion of generalizability as applicability of the results to the other situations, as shown below:

> Well, it hasn't been mentioned from where the research was taken… the researchers in Virmond Virginia University did not indicate where they found the couples they selected from. Did they [researchers] choose them from the same place [where the research was conducted]; maybe it is unique to that region; the detection of the men in that place.. this research representes a place in Australia. In other words, if this research had been done on a more comprehensive basis, with more effort, with many more participants having different personalities, I mean if they had received more support from more places, it would have been more reliable. So I think we cannot generalize it to Turkey because the men in Turkey, I think would be higher than 29%, well […] (Q65)

When assessing the generalization of the study, he also made use of the *querying evidence* skill regarding sample characteristics such as their residences and cultural properties. He discussed whether the study was conducted at a large scale by applying informal knowledge of stratified sampling to the social context. He, however, did not consider the term of generalizability referring to the fact that similar results would be obtained from the population if the study had been carried out with everyone in the population. Whereas Ali and Meltem decided that the study could be generalizable, Melek and İrem stated it could not be generalized. İrem, for example, discussed the generalizability as shown below:

> Well, I don't know if 203 couples are enough. I don't think it can be generalized. My usual opinion, you can't imagine something big from a small sample. If I ask each of the 203 men or if I get 58 men, in this case, will only 23 of 58 of all their wives predict correctly? It seems that this will not be correct all the time…You know, different results will be obtained from different samples; well, here the 203 couples don't have any characteristic features anyway. I mean, where do they live, in which country, I don't know how long they have been married; maybe there are many influential factors. Uumm, it has only mentioned that they are young couples[…] (Q66)

İrem *assessed the arguments* reported in the newspaper article in terms of the generalizability of the study. She perceived generalizability as obtaining similar results from the population if the study had been carried out with everyone in the population. During the evaluation process, she also questioned what background information regarding sample characteristics would be useful to make a sound argument regarding the study by *querying evidence* during the *inference* process. However, she had doubts as to whether sample size was enough large and had a tendency to think deterministically, stating "[…] you can't imagine something big from a small sample [...]". An especially interesting finding was also that she considered the term of chance variability, which requires understanding variability among repeated samples selected from the same population.

On the other hand, Meltem evaluated the conclusions of the study as generalizable and focused on the notion of sample size when assessing the reported statistics' generalizability.

> Generalizability, well, after all, it is done with 203 young couples, 203 is actually a good number; in statistics when we, for example, carry out a study, we say it's a good result when it is over 30, or 100, for instance. Well, 203, compared with that, is good, that's why.. it can be generalized, I think, because everything is clear […] (Q67)

Nevertheless, she stated weak expressions in the process of *evaluation*. She only recognized the factor of sample size and considered the sample size of the study enough to generalize the conclusions and tended to make a relationship between the generalizability of the study and the clarity of the reported study.

In summary, all participants attempted to *assess the arguments* reported in the newspaper article to determine if it was generalizable. However, they recognized different factors such as sample size, sampling method, or cultural factors in the process of evaluation.

4.3 Summary

Summary of the results and conclusions of this study are presented in two main sections: Newspaper Article I and Newspaper Article II. The results are separately summarized for each newspaper article because they include different contexts.

Newspaper Article I

The results regarding Newspaper article I are summarized in Table 29. Then the extent to which participants made use of critical thinking is explained under three sub-headings: base of reported statistics, reported statistics, and generalizability of the reported statistics.

Table 29

Participants' critical thinking processes through statistical and probabilistic knowledge in the Newspaper Article I

			Critical Thinking Skills					
			Interpretation	Analysis	Evaluation	Inference	Explanation	Self-Regulation
Statistical and Probabilistic Knowledge	Base of Reported Statistics	- Sampling	x	x	x	x		
		- Data Collection	x		x	x		
		- Data Analysis	x					
		- Results and Conclusions	x	x	x	x		x
	Reported Statistics	- Probabilistic Statements and Percentage	x	x	x	x		x
	Generalizability of Reported Statistics				x	x		

Base of Reported Findings

The summary regarding base of reported findings is presented under four main sections: sampling, data collection, data analysis, and results and conclusions.

Sampling

Analysis of the participants' expressions regarding sampling indicated that all participants focused on the notion of sample size by *clarifying meaning*. They also made use of different skills such as *analysis*, *inference*, and *evaluation* with respect to the sampling process of the study. They attempted to make evaluation mostly on the basis of sample size as a characteristic of sample, rather than on the basis of other characteristics such as their age and country. Only one of the participants detected missing information regarding the reason of sample selection and appreciated the role of sample in a study, which requires understanding the need for inference from sample to population and the need for sample to be representative. None of the participants could assess to what extent sample size could be sufficient to reach a proper conclusion. This could be due to the fact that they were not familiar with health statistics or their lack of statistical and probabilistic knowledge regarding sampling.

Data Collection

Analysis of the participants' expressions regarding data collection indicated that two of the participants, when asked how this study could be conducted, focused on the data collection procedure. They attempted to make use of *interpretation* and *inference* skills. In this process, they mostly focused on the importance of contextual information regarding Down syndrome. This would be due to the fact that they were not familiar with such a health context. One of the participants also appreciated the role of a good research design by controlling confounding variables.

However, she could not, in fact, question whether this study could include such a proper design.

Data Analysis

Analysis of the participants' expressions regarding data analysis revealed that almost none of the participants could utilize the critical thinking skills regarding data analysis. Only one of the participants attempted to make basic interpretation regarding data analysis. It could be due to the fact that they were not be familiar with data analysis procedures in statistics regarding health or lack of information regarding data analysis of the research study in the newspaper article.

Results and Conclusions

Analysis of the participants' expressions regarding results and conclusion showed that all participants were in the process of *interpretation* by *decoding the significance* of the article. They differed from each other by making use of different critical thinking skills. It could be concluded that insufficient reflection of some critical thiking skills such as *analysis* or *inference* might have an impact on the proper evaluation of the results or conclusions reported in the newspaper article. Two of the participants attempted to generally evaluate the results or conclusions. They had difficulty in evaluating the base of research process to reach such results. For example, they did not raise questions regarding results caused by chance processes nor pay attention to the quality of the research process, which could impact the results of the study. This could derive from their lack of knowledge with respect to how results or conclusions could be reached in a research study or their unfamiliarity with the context.

Reported Statistics

Probabilistic Statements and Percentage

This part is summarized in terms of eight main findings. One of the findings was that participants mostly had a tendency to *interpret* probabilistic statements and percentages by *clarifying* their *meanings*. This could relate to participants' language skills or vague meanings of the terms expressed in the newspaper article. In the interpretation process, they mostly made use of the analysis skill by examining related statements in the newspaper article.

The second finding was that in their critical thinking processes, they made use of different statistical and probabilistic knowledge such as variation among samples, average concept, and base rate of a given situation.

The third finding was that they reflected different thinking skills in the interpretation process. That is, they were intertwined. For example, during the interpretation process, they made an analysis regarding closely related statements or made inferences. This may be due to the complex nature of critical thinking process.

The fourth finding was that reasoning proportionally might be a factor in making valid interpretations, inferences or evaluations because they could analyze statements by examining close relationships between them, and by comparing and contrasting them.

The fifth main finding was that the analysis of their expressions indicated that they attempted to make use of different critical thinking skills ranging from *interpretation* to *self-regulation*. This may be due to complex nature of conditional probability statements or they had difficult in dealing with conditional probability. However, they less frequently reflected *explanation* and *self-regulation* skills. This would be

due to the research's limitations, which give limited opportunity participant to apply or participant's weak ability of *explanation* and *self-regulation*.

The sixth finding was that all participants attempted to make inference regarding the four situations which require understanding conditional statements. However, they could not identify different situations of the accuracy rate of the test. They had difficulty in dealing with the statements including conditional probability. They had a predisposition to confuse conditional events as a result of which they drew incorrect conclusions. This could pertain to the difficulty in noticing the conditional event or grammatical expressions of the statements including ambiguity or vagueness or misuse of conditional probability in the newspaper article. In addition, unlike other participants, those with the ability to construct a table or diagram to summarize the findings of the study made comments and interpretations more easily by clarifying meaning, analyzing the statements, and drawing conclusions.

Another finding was that some of the participants made correct inferences regarding conditional statements in the newspaper article, which required understanding the Bayes Theorem. They could make proper calculation without knowing the rule regarding Bayes Theorem. Moreover, their categorizations could be beneficial in this process.

The last finding was that some of the participants could recognize the misleading numbers and language in the newspaper article toward the end of the study. The reason underlying such a change could be their mathematical thinking and in-depth examination of the article, rather than superficial evaluation of the text. Thus, they could sufficiently reflect evaluation skill in this process.

Generalizability of the Reported Statistics

Analysis of the participant's expressions regarding generalizability of the reported statistics indicated that most of the participants attempted to *assess the argument* reported in the newspaper article to evaluate the generalizability of the reported statistics. However, they recognized different factors such as sample size, sample characteristics, cultural factors, and chance variability in the process of evaluation.

Newspaper Article II

The results regarding Newspaper article II are summarized in the Table 30. Then the extent to which participants made use of critical thinking is explained under three sub-headings: base of reported statistics, reported statistics, and generalizability of the reported statistics.

Table 30

Participants' critical thinking processes through statistical and probabilistic knowledge in the Newspaper Article II

			Critical Thinking Skills					
			Interpretation	Analysis	Evaluation	Inference	Explanation	Self-Regulation
Statistical and Probabilistic Knowledge	Base of Reported Statistics	- Sampling	x		x	x		
		- Data Collection	x		x	x	x	x
		- Data Analysis	x	x				
		- Results and Conclusions	x		x			x
	Reported Statistics	- Probabilistic Statements and Percentage	x	x	x	x		x
	Generalizability of Reported Statistics				x	x		

Base of Reported Findings

The summary regarding base of reported findings is summarized in four main sections: sampling, data collection, data analysis, and results and conclusions.

Sampling

Analysis of participants' expressions regarding sampling indicated that all of the participants were in the process of *interpretation* with regard to sample size of the study. They, however, did not interrogate how the sample was selected or whether it was biased and could not detect the fact that other characteristics of the sample such as age, the number of years married and their area of residence were not reported in the newspaper article, which prevents people from making proper judgements regarding the research reported in the newspaper. Only one of the participants attempted to make evaluation and conjectured alternatives regarding the sample size of the study in which he appreciated the essential role of sample and the need for samples to be representative to reach credible results regarding a research study.

Data Collection

All participants focused on the issue of what was measured and how to be measured by *clarifying the meaning* of the statements reported in the newspaper article. Two of them differed from the others by reflecting different critical thinking processes such as *self-regulation*, *inference*, *evaluation* and *explanation*. In addition, they had a sense of survey design even though they did not state the term *survey design*. None of the participants, on the other hand, discussed to what extent this instrument measured the characteristics of cheating in a reliable way, nor did they consider the errors which are possible to arise during data collection.

Data Analysis

Analysis of the participants' expressions regarding data analysis revealed that all participants possess some sense of how to summarize data such as using percentages, graphs or tables. In addition, all of them attempted to use *analysis* and *interpretation* skills with respect to how data could be analyzed, in particular how to summarize data by using percentages. They made more comments regarding the data analysis procedure of the research study in Newspaper Article II when compared to that of Newspaper Article I. The reasons behind such a difference could derive from their unfamiliarity with heath statistics or their knowledge regarding descriptive studies such as surveys.

Results and Conclusions

Analysis of the participants' expressions regarding results and conclusions showed that all participants were in the processes of decoding significance and assessing arguments. They attempted to detect the author's indirect intentions and the main idea of the text. None of them, however, could detect the author's one-sided claims in the newspaper article. Moreover, two of the participants differed from the others by self-consciously monitoring their thinking process at the end of the interview. One of the participants was differentiated from the others because she made use of statistical knowledge, which requires realizing that observed differences may not be statistically significant or may not be large enough or consistent or can occur by chance. In addition, she recognized the importance of sample size for each cell or category (sample composition for each group) whereas the others tried to evaluate the results of the study holistically.

Reported Statistics

Probabilistic Statements and Percentage

This part is summarized in terms of five main findings. One of the main findings was that participants had a tendency to *interpret* all the statements reported as descriptive statistics by *clarifying their meaning*. In particular, they made an attempt to clarify the meanings of Statement 2 and Statement 3 by making explicit the difference between them. This could relate to the participants' language skills or to the vague meaning of the terms expressed in the newspaper article.

The second main finding was that the analysis of their expressions indicated that they attempted to make use of different critical thinking skills, especially regarding Statements 3 and 4, ranging from *interpretation* to *self-regulation*. This may be due to the complex nature of conditional probability statements or the difficulty participants experienced in dealing with conditional probability. However, they less frequently reflected *explanation* and *self-regulation* skills. This might be due to the limitations of the research, which provide the participant with limited opportunity to apply them or due to the participant's weak ability of *explanation* and *self-regulation*. Moreover, they indicated inadequate ability in *assessing* the reported *claims*, especially regarding claims which require understanding the conditional probability.

Another finding was that some participants perceived the statements reported in the newspaper article differently, which lead them to think differently since newspaper articles may include ambiguous language.

Another finding was that some critical thinking skills consisted of recursive process. For example, there of the participants were in the circular process of interpretation,

self-examination, self-correction, and examining ideas respectively in their interpretation process of the Statement 3.

The last main finding was that participants had difficulty in dealing with the statements including conditional probability. They had a predisposition to confuse conditional events as a result of which they drew incorrect conclusions. This may pertain to the difficulty in noticing the conditional event or grammatical expressions of the statements including ambiguity or vagueness or misuse of conditional probability in the newspaper article. In addition, regarding Statement 3, which included conditional probability, unlike the other participants, those with the ability to construct tables or diagrams to summarize the findings of the study made comments and interpretations more easily by clarifying meaning, analyzing the statements, and drawing conclusions.

Generalizability of the Reported Statistics

Analysis of the participants' expressions regarding the generalizability of the reported statistics indicated that all participants attempted to *assess the arguments* reported in the newspaper article to determine if they were generalizable. On the other hand, they differed from each other in terms of recognizing different factors such as sample size, sampling method, and cultural factors in the process of evaluation. They perceived the term *generalizability* differently. For example, whereas one of the participants perceived the term *generalizability* as obtaining similar results from the population if the study had been carried out with every individual in the population, another participant perceived it as the applicability of the results to other situations. Moreover, two of them attempted to conjecture alternatives regarding sample characteristics in the evaluation process.

CHAPTER 5

CONCLUSION, DISCUSSION, AND IMPLICATIONS

The purpose of the study is to investigate to what extent pre-service middle school mathematics teachers make use of critical thinking skills through their statistical and probabilistic knowledge in the context of popular media texts. This chapter includes main conclusions in line with the purpose of the study and discussions with regard to previous studies. At the end of the chapter, implications for further researcher studies and educational practices are also addressed.

5.1. Critical Thinking Skills through Statistical and Probabilistic Knowledge

The findings of the study with regard to *sampling* indicated that most of the participants *interpreted* samples of studies by *clarify*ing their properties paying attention to the preserving meaning of the author's statements in the newspaper articles. They also made use of different critical thinking skills; *analysis*, *inference*, and *evaluation*. However, they mostly focused on the notion of sample size in these processes, which is actually parallel with the findings regarding informal level, which is one of the six hierarchical levels of statistical literacy identified by Watson and Callingham (2003). The fact that the participants mostly focused only the notion of sample size corresponds to the informal level in which students consider a single aspect of statistical and probabilistic concepts and present lack of engagement with the context. Moreover, few of them detected missing information regarding the reason of sample selection and appreciated the role of sample in a study, which requires understanding the need for inference from sample to population and the need for sample to be representative, which are essential in *evaluating* the contextual

information (Facione, 2011a) and to be statistically literate person (Watson, 2006; Gal, 2004). That is, most of them generally made comments on the basis of existing information rather than missing information such as sampling method and other sample characteristics apart from sample size, which is consistent with Watson's study (1997) and the study of Watson and Moritz (2000b) in which students could not make any criticism about a sampling task from the media.

Moreover, in different contexts of newspaper articles, participants engaged with the sampling of the study in a different way by means of critical thinking skills. In particular, in Newspaper Article I, related to health problem, participants reflected a more complex thinking process in terms of critical thinking compared to the second article. There could be several reasons underlying such a difference between critical thinking processes in the engagement with different newspaper articles. One of the reasons might be that they could be more unfamiliar with the health context compared to a social context like cheating. This is consistent with previous studies (Watson, 2006; Gal, 2004), which identified *context* as a potential factor for the statistical literacy construct. Another possible reason behind such a difference is that participants were skeptical about the characteristic of sample because the study had consisted of only pregnant women at high risk in having a baby with Down syndrome, rather than at low or normal risk. This skepticism might have lead participants to exercise a more complex process of critical thinking and make use of a variety of critical thinking skills. These findings are parallel with several research studies on critical thinking (Ennis, 1985; Facione, 1990; Jones et. al, 1995; Siegel, 1988). These studies have emphasized that it is essential to have not only critical thinking skills, but also dispositions toward effortful thinking to be critical thinkers in the society. Furthermore, their skepticism might derive from their lack of knowledge regarding purposeful sampling as an alternative sampling method in a research study. This could be evidence unfolding the intertwined relationship between statistical knowledge, critical thinking dispositions, and critical thinking skills.

The findings with regard to *data collection* and *data analysis* revealed that the participants' usage of statistical and probabilistic knowledge regarding data collection and data analysis was differentiated in Newspaper Article I and Newspaper Article II. More specifically, they made more comments regarding data collection and data analysis procedures in Newspaper Article II compared with the other article. In this circumference, their critical thinking processes were also differentiated with respect to the articles. For example, almost none of the participants focused on the data analysis procedure of the study in Newspaper Article I whereas the same participants possessed some sense of how to summarize data such as using percentages, graphs, or tables, and reflected *analysis* and *interpretation* skills. The reason behind such a difference could be that participants have lack of statistical knowledge regarding data analysis procedures in the context of health statistics while they are more familiar with knowledge of descriptive statistics in social studies. However, it is essential for individuals to have some sense of how data can be analyzed in different contexts (Gal, 2004). Moreover, this finding presented the essential role of context in the development of statistical literacy as proposed in the statistical literacy models of Gal (2004) and Watson (2006).

In addition, only one of the participants reflected a more comprehensible thinking process because of her high level of *categorization* skill. She transformed information in the newspaper article to meaningful knowledge regarding data analysis more comprehensibly utilizing the categorization skill. She represented two contingencies in a table to analyze the data and related raw data to percentages as a summary statistics. This interesting finding arising from the study confirms previous studies in terms of the role of the representation skill in critical thinking and mathematical thinking (McKendree, Small, Stenning, & Conlon, 2002; Pape & Tchoshano, 2001). While McKendree et. al. (2002) addresses the role of representational skill in critical thinking, Pape et. al. (2001) supportively pointed out

the necessity of representation in mathematical learning because representations provide learners with opportunities to make manipulations or tranformations of information during the critical thinking process, providing flexible cognive process in the transformation of raw information to meaningful knowledge for themselves, and helps them organize their workings in an engagement with a problem in mathematical learning.

Compared to *data analysis*, participants were more familiar with the *data collection* procedures. They attempted to use a variety of critical thinking skills such as *interpretation*, *inference*, *evaluation*, or *self-regulation*. However, they reflected particularly weak expressions regarding *evaluation* of *data collection procedures*. Only one of the participants, for example, appreciated the contribution of a good design for data production in a study and controlling of confounding variables in the research study reported in Newspaper Article 1. Additionally, none of the participants discussed the extent to which the instrument measured the participants' characteristics in a reliable way and possible data collection errors in Newspaper Article II. Moreover, two participants reflected the *self-regulation skill* by self-consciously examining their thinking process. Their familiarity with survey design may have allowed them to apply more comprehensible skills although this self-regulation process is not strong enough to be a critical thinker. This could be due to the limitation of this study which did not require students to write their opinions, ideas, or arguments. Writing could be a potential factor in the development of critical thinking skills since both of them involve recursive processes (Olson, 1984). Halpern (2003) also emphasized the role of writing in the critical thinking process stating that "working memory is the term used for the "place" in which we consciously think. When learning is difficult, we need to reduce the load on working memory. This can be done by writing information on paper, making it more automatic or attending to the information to-be-learned." (p. 84).

The findings regarding *results and conclusions* reported in the newspaper article indicated that all the participants attempted to make *interpretation* by *decoding the significance* of the text in both newspaper articles. They also reflected *analysis*, *evaluation*, *inference*, and *self-regulation skills*. One of the interesting findings arising from this process is that the participants who tried to make *evaluation* had difficulty in judging the base of the research process to reach results or conclusions reported in the newspaper articles. They did not raise questions regarding the results caused by chance processes or pay attention to the quality of the research process, which can impact the results of the study in Newspaper Article I. On the other hand, regarding Newspaper Article II, only one of the participants evaluated the acceptability of the given conclusions. She recognized the importance of sample size for each cell or category to assess the degree of the credibility of the results and detected the fact that the number of participants for each category or condition, such as cheating versus not cheating or right versus wrong inferences, may not be sufficient for the reliability of the conclusions drawn. Watson (2011) supportively made similar criticisms regarding this newpaper article, which are required knowedge for mathematics teachers to implement it in the classroom environments. Moreover, this study presents lack of knowledge regarding the base of results/conclusions reported in the media texts, although it is essential for adults to have some sense of how the conclusions could be reached (Gal, 2004).

Moreover, toward the end of the interview with the participants, some of them reflected a somewhat *self-regulation* skill even if it were not completely successful. They self-consciously monitored their thinking process and overviewed their interpretations and evaluations as a process from the beginning to the end of the interview. This finding confirms Facione's (2011a) conceptualization of critical thinkers' characteristic in the self-regulation process, which includes monitoring all the processes (interpretation, analysis, evaluation, inference, and explanation) of critical thinking in a recursive way. People with strong critical thinking evaluate their own strengths and weaknesses by asking themselves questions of "Do I have

any unclear status regarding the issue", "Is my evidence enough good", or "What is the missing point I overlooked?" (Facione, 2011a). Moreover, it could also be related to whether the time allowed to the participants to think about their expressions regarding the media text was sufficient (Watson, 2006). Supportively, Osana and Seymour (2004) pointed out that enough engagement with complex issues in their cognitive apprenticeship intervention could improve pre-servie teachers' critical thinking skills.

The findings of the study with regard to *probabilistic statements and percentages* indicated that participants had difficulty in dealing with the statements including conditional probability in social context. They had a tendency to confuse conditional events; as a result of which they drew incorrect conclusions or could not reach conclusion at all. This conclusion was in accordance with various studies indicating that people had difficulty in defining conditional event (Carnell, 1997; Falk, 1986; Kramer & Gigerenzer, 2005). Incorrect or insufficient inferences regarding conditional statements could pertain to the difficulty in noticing the conditional event, the complicated nature of the conditional probability (Huerta, Cerdan, Lonjedo, & Edo, 2011) or grammatical expressions of the statements including ambiguity or misuse of conditional probability in the newspaper article (Gal; 2005). In this regard, it is essential that pre-service teachers overcome confusions regarding the conditional probability and its language to teach this concept to the students (Contreras et al., 2011). On the other hand, none of the participants appreciated the inverse relationship between P(A|B) and P (B|A) in both heath and social life contexts. This finding is paralell with Watson and Nathan's study (2010b), in which it was indicated that pre-service mathematics teachers could not identify the difference between conditional probabilistic statements. Roca and Batanero (2006) also pointed out pre-service teachers' confusions regarding conditional probabilities in relation to two-way tables.

Another conclusion is in relation to the finding that few of the participants made correct inferences regarding conditional statements in the newspaper article, which required understanding Bayes' Theorem. They could make proper calculations without knowing the standard rule regarding Bayes' Theorem. This could stem from their *categorization* skills with which they represented the reported findings in a diagram or table. This study also indicated that pre-service elementary mathematics teachers could have intuitive understanding of conditional probability and Bayes' Theorem and they could explore Bayes' theorem in two-way tables, parallel with Rossman and Short's study (1995) on students even though they did not notice big idea behind probabilistic statements as conditional probabilities underlying Bayes' theorem, they could have intuitive understanding of the theorem by means of two-way tables.

In addition, a variety of interpretations of the probabilistic statements reported in the newspaper article could lead participants to make use of different critical thinking skills since newspaper articles may include ambiguous or vague language. For example, one of the participants tried to clarify the meaning of the probabilistic statement and conjectured alternatives during the evaluation process. This indicates that critical thinking processes were intertwined rather than hierarchical as stated by Ennis (1993) and Facione (1990). In addition, the misinterpretation of the statements in the newspaper article could be a potential factor in applying the other critical thinking skills. For example, some participants could not clarify the statements; thus, they could not educe a conclusion or examine the statements in the newspaper article supporting that some skills may be prerequisite for others (Facione, 1990). Their lack of interpretation skills may be due to the their language skills, which is a critical factor in statistical literacy (Gal, 2004; 2005; Watson, 2006).

Most of the participants reflected *proportional reasoning* especially in examining related statements in the newspaper article by making use of the *analysis* skill, which indicated the role of mathematical knowledge in statistical literacy (Gal,

2004). Participants who were successful in the process of analysis by reasoning proportionally would make further inferences or evaluations. This is parallel with the findings regarding critical-mathematical level, which is the highest stage among the six hierarchical levels of statistical literacy identified by Watson and Callingham (2003). The fact that the participants reflected advanced critical thinking skills on account of proportional reasoning corresponds to the critical-mathematical level in which students engage with the context critically and make appropriate inferences on the basis of quantitative reasoning.

The findings regarding *generalizability of the reported statistics* indicated that all of the participants mostly attempted to make *evaluation* when they made comments with regard the generalizability of the reported statistics. On the other hand, they differed from each other by recognizing different factors such as sample size, sample characteristics, cultural factors, and chance variability in the process of evaluation. The present finding confirms the skills required for evaluation in the inductive reasoning such as making generalizations or inferences from the research studies, identified by Facione (2011a) and Halpern (2003). It is clear that some participants appreciated the role of increasing the sample size in the observation of the stability in the variation in the sample (Bakker & Gravemeijer, 2004), or variation among the different samples selected from the same population, referred as chance variability by Gal (2004). However, some participants made immediate comments on the basis of their personal ideas about sample size, which is parallel with the study of Innabi (2006) on secondary school students' misconceptions regarding generalizability on the basis of sampling in which it was suggested that students are educated in an environment in which they think critically about the generalizability of the statistics reported in the newspaper article. However, their *immediate* comments regarding sample size as "it is good" or "not enough for generalizability of the reported statistics" contradicts with the nature of critical thinking, which requires skeptical, and effortful thinking, and conscious inquiry on the basis of evidence (Dewey, 1933; Facione, 1990; Siegel, 1988).

Another finding was that some participants mostly attempted to make use of the *inference* skill by *querying evidence* or *conjecturing alternatives* in the process of evaluation. This situation mostly emerged when they did not find useful information in the newspaper article to support their judgment about the generalizability of the study. This could be due to the fact that newspaper articles may not, most of the time, include sufficient information regarding the research process of the reported study (Utts, 2005; Gal, 2004).

In summary, the participants of the study, who have the tendency to use valid quantitative procedures and mathematical language reflected different critical thinking skills and made use of different statistical and probabilistic knowledge in their thinking processes. Even though the participants have such a tendency, they reflected in some conditions (eg. health contexts, conditional probability statements) partial and improper statistical and probabilistic knowledge and utilize critical thinking skills inappropriately.

5.2. Implications and Recommendations

The findings of this study display many possible questions for future research into critical thinking in popular media texts. Five key questions arise from this study, which are presented below:

The study indicated that there are dispositional elements, namely, dispositional skills toward critical thinking and dispositions toward statistics and probability, which would be a potential factor in the critical thinking process and statistical literacy while reading media texts. The first question arises from this finding: What is the role of dispositional skills in the critical thinking process while reading media texts?

Another conclusion of the study was that there are several elements, such as mathematical knowledge, language skills, and context knowledge, which all have a crucial role in the development of statistical literacy and all are closely intertwined with statistical and probabilistic knowledge, which is supportive with the statistical literacy model of Gal (2004). This suggests the second question for further research: What is the role of these elements in the development of statistical literacy and their interactions with critical thinking skills while reading popular media texts?

This study also revealed that pre-service middle school mathematics teachers' critical thinking processes' relevant to statistical and probabilistic knowledge could be different in different contexts, particularly in health and social life contexts underlying mostly probabilistic statements. This raises the third question for further research: To what extent do pre-service middle school mathematics teachers make use of critical thinking skills while reading different media contexts such as economics, political, transportation, or other social contexts underlying different statistical and probabilistic concepts like average, data representation, and risk assessment?

This study investigated the nature of pre-service middle school mathematics teachers' critical thinking processes in popular media texts. However, there is also need for examination of their improvement in critical thinking in a period of time, which gives rise to the fourth question: How could pre-service middle school mathematics teachers develop critical thinking within a special environment including such as discussion, cooperative learning, or cognitive apprenticeship intervention in the statistics education?. This raises another question of assessment in critical thinking: How could they be assessed in terms of critical thinking relevant to statistical literacy?

Finally, this study investigated pre-service teachers' critical thinking processes with regard to statistical and probabilistic knowledge. It is also important to investigate

middle school students' and in-service mathematics teachers' critical thinking processes regarding their statistical and probabilistic knowledge in popular media texts.

The findings of this study could also be an attempt for learners in various issues relevant to educational practice in critical thinking with regard to statistical and probabilistic knowledge, which could lead to the reorganization of learning and teaching environments in statistics education.

Possible developments in statistics education from the points of students, teachers, teacher educators, curriculum designers, or policy makers could be expected as follows: Teachers and teacher educators should develop open-ended items including media tasks in the assessment of both middle school students and pre-service middle school mathematics teachers' critical thinking in context including statistical and probabilistic information; pre-service middle school mathematics teachers should be allowed to engage with popular media texts in both method courses of teaching mathematics, and statistic and probability courses with the integration of critical thinking since it allows them to communicate with each other and transfer their knowledge to the real world. Teacher educators should overview of the content of teaching methods of mathematics and statistics courses in teacher education programs and pre-service teachers should be provided with the opportunity to use and develop efficient and appropriate tasks or activities including critical questions from the media. In particular, teacher educators could design a learning environment in which pre-service teachers are encouraged to think critically in a real life context including statistical and probabilistic information. To achieve this goal, they could apply the cognitive apprenticeship model, which is "the use of an apprentice model to support learning in the cognitive domain" (Dennen, 2003, p. 813). Moreover, instructors may also integrate technology into such a learning environment so that pre-service teachers think critically and reflectively through the interactive nature of technological tools such as TinkerPlots and Fathom. These technological tools could

be mediators to design a learning environment in which pre-service teachers are not exposed to learn just computational procedures of statistics and probability. They could encourage them to focus on the interpretation and evaluation of the data regarding contextual situations. In this process, pre-service teachers could be encouraged to think about both proper and improper examples of newspaper articles, or magazines, advertisements and write their reflections and critiques regarding such media texts on the basis of statistical and probabilistic knowledge. These popular media texts should include such diverse contexts such as heath, politics, forecasts, economy, and social life, which they could address the need and interest of pre-service mathematics teachers.

Another suggestion emerges from the point of view of journalists and researchers. Newspaper articles could include misleading language related to statistical and probabilistic information. This could be due to the facts that journalists might have lack of statistical and probabilistic knowledge or researchers may not be sharing their publications in an understandable way with the journalists and public. This issue leads to the need for strong communication between journalists and researchers in order to release credible information to the public. Due to difficulty in interpreting raw data and technical terms, researchers should also release their publications with appropriate interpretation of results and share them with journalists.

REFERENCES

Akkaş, E. N. (2009). *6.-8. Sınıf öğrencilerinin istatistiksel düşüncelerinin incelenmesi*. Unpublished master's thesis. Abant İzzet Baysal Üniversitesi, Bolu.

Alazzi, K. F. (2008). Teachers' perceptions of critical thinking: A study of Jordanian secondary school social studies teachers. *The Social Studies, 99*(6), 243-248.

Alper, A. (2010). Critical thinking disposition of pre-service teachers. *Education and Science, 35*(158), 14-27.

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

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Carlton, Vic.: Author.

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147-168). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Batanero, C. (2002). The role of models in understanding and improving statistical literacy. *International Statistical Review, 70*(1), 37-40.

Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben_Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3-15): Kluwer.

Ben-Zvi, D., & Garfield, J. (2008). Introducing the emerging discipline of statistics education. *School Science and Mathematics, 108(8),* 355-361.

Bogdan, R.C. & Biklen, S.K. (1998). *Qualitative research for education: An introduction to theory and methods.* Boston: Allyn and Bacon.

Brown, A. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist, 52*(4), 399-413.

Bulut, S. (2001). Investigation of performances of prospective mathematics teachers on probability. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *20*, 33-39.

Bulut, S., Yetkin, İ. E., Kazak, S. (2002). Matematik öğretmen adaylarının olasılık başarısı, olasılık ve matematiğe yönelik tutumlarının cinsiyete göre incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *22*, 21-28.

Callingham, R. (2006). *Assessing statistical literacy: A question of interpretation*. Paper presented at at the the Seventh Congress of International Conference on Teaching Statistics. Bahia, Brazil

Carmichael, C., Callingham, R., Hay, I., &Watson, J. (2010). Statistical literacy in the middle school: The relationship between interest, self-efficacy and prior mathematics achievement. *Australian Journal of Educational & Developmental Psychology, 10*, 83-93.

Carmichael, C. S., Callingham, R., Watson, J. M., & Hay, I. (2009). Factors influencing the development of middle school students' interest in statistical literacy. *Statistics Education Research Journal*, *8*(1), 62-81.

Carnell, L. J. (1997). *Characteristics of reasoning about conditional probability*. Unpublished doctoral dissertation, University of North Carolina, Greensboro.

Contreras, J. M., Batanero, C., Díaz, C., & Fernandes, J. A. (2011). *Prospective teachers' common and specialized knowledge in a probability task.* Paper presented at the Seventh Congress of the European Society for Research in Mathematics Education. Rzeszów, Poland.

Cotton, K. (1991).Teaching thinking skills. School improvement research series. Retrieved from from www.nwrel.org/scpd/sirs/6/cu11ktml.

delMas, R. C. (2002). Statistical literacy, reasoning, and learning: A commentary. *Journal of Statistics Education*, *10*(3). Retrieved from, www.amstat.org/publications/jse/v10n3/delmas_discussion.html.

Demirkaya, H. (2003). Eleştirel düşünme kuramının coğraya programı üzerindeki etkileri. *Ankara Üniversitesi Eğitim Bilimleri Dergisi, 36*(1-2), 97-106.

Dennen, V. P. (2003). Cognitive apprenticeship in educational practice: research on scaffolding, modeling, mentoring, and coaching as instructional strategies. In D. H. Jonassen (Ed.),*Handbook of Research on Educational Communications and Technology*. Mahwah, NJ: Lawrence Erlbaum, Inc.

Dewey, J. (1933). *How we think–A restatement of the relation of reflective thinking to the educative process*. Boston, MA: Heath.

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

Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership*, *43*(2), 44–48.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice*, *32*(3), 179–86.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. *Research findings and recommendations.* ERIC, ED 315423.

Facione, P. A. (2011a). *Think critically.* New York: Pearson Education, Englewood Cliffs.

Facione, P. A. (2011b). Critical thinking: What it is and why it counts. Millbrae, CA: Insight Assessment, California Academic Press. Retrieved from, www.insightassessment.com/CT-Resources/Independent-Critical-Thinking-Research/pdf-file/Critical-Thinking-What-It-Is-and-Why-It-Counts-PDF.

Falk, R. (1988). Conditional probabilities: Insights and difficulties. In *Proceedings of the Second International Conference on Teaching Statistics***.** R. Davidson and J. Swift (Eds.). Victoria, B.C.: University of Victoria.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K–12 curriculum framework*. Alexandria, VA: American Statistical Association.

Gal, I. (2002). Adult's statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-25.

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. (2005). Towards "probability literacy" for all citizens: building blocks and instructional dilemas. In G. Jones (Ed.), *Exploring probability in schools: challenges for teaching and learning* (pp. 39-63). New York: Springer.

Gibson, G. (1995). Critical thinking: implications for instruction. *Reference & User Services Quarterly*, *35*(1), 27–35.

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal, 327***,** 741-744.

Guba, E. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Technology Research and Development, 29*(2), 75-92.

Güven, M., & Kürüm, M. A. D. (2007). Teacher candidates' learning styles and critical thinking dispositions. *Elektronik Sosyal Bilimler Dergisi, 6*(21), 60-90.

Halpern, D. F. (1997). *Critical thinking across the curriculum*. Mahwah, NJ: Lawrence Erlbaum.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53*(4), 449–455.

Halpern, D. (2003). *Thought and knowledge*, 4th ed. Mahwah, NJ: Erlbaum.

Huerta, M. P., Cerdan, F., Lonjedo, M. A. (2011). *Assessing Difficulties of Conditional Probability Problems*. Paper presented at the Seventh Congress of the European Society for Research in Mathematics Education. Rzeszów, Poland.

Innabi, H. (2006). *Factors considered by secondary students by secondary students when judging the validity of a given statistical generalization*. Paper presented at the the Seventh Congress of International Conference on Teaching Statistics. Bahia, Brazil.

Innabi, H. & Sheikh, O. E. (2006). The change in mathematics teachers' perceptions of critical thinking after 15 years of educational reform in Jordan. *Educational Studies in Mathematics, 64*(1), 45-68.

Jones, E. A., Hoffman, S., Moore, L. M., Ratcliff, G., Tibbetts, S., & Click, B. L. (1995). *National Assessment of College Student Learning: Identifying the College Graduate's Essential Skills in Writing, Speech and Listening, and Critical Thinking*. Washington D.C.: National Center for Educational Statistics, US Department of Education, Office of Educational Research and Improvement. OERI Publication NCES 93-001.

Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education, 118*(2), 282-292.

Joram, E., Resnick, L. & Gabriele, A. J. (1995). Numeracy as a cultural practice: An examination of numbers in magazines for children, teenagers, and adults. *Journal for Research in Mathematics Education, 26*(4), 346-361.

Kanik, F. (2010). *An assessment of teachers' conceptions of critical thinking and practices for critical thinking development at seventh grade level*. Unpublished doctoral dissertation. Middle East Technical University, Ankara.

Kazak, S. (2008). Öğrencilerin olasılık konularındaki kavram yanılgıları ile öğrenme zorlukları. M. F. Özmantar, E. Bingölbalı & H. Akkoç (Eds.) *Matematiksel Kavram Yanılgıları ve Çözüm Önerileri* (pp. 121-150). PEGEMA Yayınevi.

Kazak, S. (2009). Olasılık Konusu Öğrencilere Neden Zor Gelmektedir?. In E. Bingölbalı & M. F. Özmantar (Eds.) *İlköğretimde Karşılaşılan Matematiksel Zorluklar ve Çözüm Önerileri* (pp. 216-239). PEGEMA Yayınevi.

Kennedy, M, Fisher, M.B. & Ennis, R.H. (1991). Critical thinking: Literature review and needed research. In L. Idol & B.F. Jones (Eds.), *Educational values and cognitive instruction: Implications for reform* (pp. 11-40). Hillsdale, NJ: Erlbaum.

Krämer, W., & Gigerenzer, G. (2005). How to confuse with statistics or: The use and misuse of conditional probabilities. *Statistical Science, 20*(3), 223–230.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*(2), 16-25.

Kvatinsky, T. and Even, R. (2002). Framework for teacher knowledge and understanding about probability. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

McKendree, J., Small, C., Stenning, K., & Conlon, T. (2002). The role of representation in teaching and learning critical thinking. *Educational Review, 54*(1), 57-67.

Mecit, Ö. (2006). *The effect of 7E learning cycle model on the improvement of fifth grade students' critical thinking skills*. Unpublished doctoral dissertation. Middle East Technical University, Ankara.

Memnun, D. S. (2008). Olasılık Kavramlarının öğrenilmesinde karşılaşılan zorluklar, bu kavramların öğrenilememe nedenleri ve çözüm önerileri. *İnönü Üniversitesi Eğitim Fakültesi Dergisi, 9*(15), 89–101.

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

Merriam, S. B. (2009). Qualitative research: A guide to design and implementation. San Francisco, CA: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.). Newburg Park, CA: Sage.

Ministry of National Education (MNE). (2005), İlköğretim Okulu Ders Programları: Matematik Programı 6-7-8 (Elementary Curricula Programs: Mathematics Curricula Program for Middle Grades). Ankara: Milli Eğitim Basımevi.

Mooney (2002). A Framework for Characterizing Middle School Students' Statistical Thinking. *Mathematical Thinking And Learning*, *4*(1), 23-63.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Olson C. B. (1984). Fostering critical thinking skills through writing. *Educational Leadership*. *42*(3), 28-39.

Osana, H. P., & Seymour, J. R. (2004). Critical thinking in preservice teachers: A rubric for evaluating argumentation and statistical reasoning. *Educational Research and Evaluation*, *10*(4-6), 473–498.

Öztürk, F., Bıçak, B., Sapancı, A. (2009). E-bilgi kaynaklarına yönelik eleştirel düşünme ölçeği geliştirme çalışması. *Sosyal Bilimler Dergisi*, *6*(2), 133-150.

Pape, S. J., & Tchoshanov, M. A. (2001). The role of representation(s) in developing mathematical understanding. *Theory Into Practice, 40*(2), 118-127.

Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.).Thousand Oaks, CA: Sage Publications, Inc.

Paul, R. (1984). Critical thinking: Fundamental for education in a free society. *Educational Leadership, 42*(1), 4–14.

Paul, R., & Elder, L. (2007). *Miniature guide to critical thinking: Concepts and tools*. Retrieved from, http://www.d.umn.edu/~jetterso/documents/CriticalThinking.pdf.

Paul, R. W., Elder, L, & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendation*. Sacramento: California Commission on Teacher Credentialing.

Recalde, X. D. (2008). *Teachers' educational technology practices: Understanding teachers' critical thinking skills and teachers' cognitive demand practices in the classroom*. Unpublished doctoral dissertation. Northen Illinois University, Illinois.

Ridgway, J., Nicholson, J., & McCusker, S. (2011). Developing statistical literacy in students and teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education: A Joint ICMI/IASE study: The 18th ICMI study*, (pp. 311-322). London: Springer.

Roca, A. S., & Batanero, C. D. (2006). Computing probabilities from two-way tables: An exploratory study with future teachers. *International Conference on Teaching Statistics*. Retrieved from, http://www.stat.auckland.ac.nz/~iase/publications/17/C413.pdf

Rossman, A. J., & Short, T. H. (1995). Conditional probability and education reform: Are they compatible?. *Journal of Statistics Education* (electronic journal), *3(*2).

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

Sapancı, A. (2007). İnternete dayalı araştırmalarda eleştirel düşünme sorunsalı: 21. Yüzyıl skolastiğine doğru. *Sosyal Bilimler Araştırmaları Dergisi*, *2*, 56-71.

Scheaffer, R.L. (2003). Statistics and quantitative literacy. In B.L. Madison & L.A. Steen (Eds.),*Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 145-152). Princeton, NJ: The National Council on Education and the Disciplines.

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

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

Seferoğlu, S. S., & Akbıyık, C., (2006). Eleştirel düşünme ve öğretimi. *Hacettepe Üniversitesi Dergisi*, *30,* 193-200.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and education*. New York: Routledge.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In *Handbook of Research on Mathematics Teaching and Learning.* D.A. Grouws (Ed.). pp. 465-494. New York: Macmillan.

Shaughnessy, J. M. (2007). Research on statistics learning and reasonig. In F. K. Lester (Eds), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Charlotte: Information Age Publishing.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of Qualitative Research (pp. 443-466). Thousand Oaks, CA: Sage Publications.

Steen, L. A. (Ed.) (2001). *Mathematics and democracy: The case for quantitative literacy.* National Council on Education and the Disciplines. Princeton: Woodrow Wilson Foundation.

Stohl (2005). Probability in teacher education and development. In G. Jones (Eds.), *Exploring probability in schools: Challenges for teaching and learning* (pp. 345- 366). Dodrecht: Kluwer.

Şengül, C. & Üstündağ, T. (2009). Fizik öğretmenlerinin eleştirel düşünme eğilimi düzeyleri ve düzenledikleri etkinliklerde eleştirel düşünmenin yeri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *36,* 237-248.

Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: Teaching strategies. *Learning and Instruction*, *14*(4), 359-379.

Tufan, D. (2008). *Critical thinking skills of prospective teachers: Foreign language education case at the Middle East Technical University*. Unpublished master's thesis. Middle East Technical University, Ankara.

Türkmen Dağlı, M. (2008). Integrating critical thinking skills into planning and implementation of teaching Turkish: A comparative case study of three teachers. Unpublished doctoral dissertation. Middle East Technical University, Ankara.

Türnüklü, & Yesildere (2005). A profile from Turkey: Critical Thinking Dispositions and Abilities of Pre-Service Mathematics Teachers of 11-13 Year. *Ankara Üniversitesi Eğitim Bilimleri Dergisi*, *38*(2), 167-185.

Ucar, Z., & Akdogan, E. N. (2009). Middle School Students' Understanding of Average. *İlköğretim Online, 8*(2), 391-400.

Ulutaş, F. & Ubuz, B. (2008). Matematik eğitiminde araştırmalar ve eğilimler: 2000 ile 2006 yılları arası. *İlköğretim Online, 7*(3), 614-626.

Utts, J. (2003). What Educated Citizens Should Know About Statistics and Probability. *The American Statistician, 57*(2) , 74–79.

Utts, J. (2005). *Seeing Through Statistics*. Pacific Grove, CA: Duxbury, 4th ed.

Uysal, A. (1998). *Sosyal bilimler öğretim yöntemlerinin eleştirici düşünme gücünün geliştirilmesindeki rolü*. Unpublished master thesis. İnönü Üniversitesi, Malatya.

Wallman, K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1–8.

Williams, R.L. (2005). Targeting critical thinking within teacher education: The potential impact on society. *The Teacher Educator, 40*(3), 163-187.

Watson, J. M. (1997). Assessing statistical literacy through the use of media surveys. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). Amsterdam, The Netherlands: International Statistical Institute/ IOS Press.

Watson, J. M. (1998). The role of statistical literacy in decisions about risk: Where to start. *For the Learning of Mathematics, 18*(3), 25-27.

Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: *The case of chance and data. Journal of Mathematics Teacher Education, 4*(4), 305-337.

Watson, J. M. (2004). Quantitative literacy in the media: An arena for problem solving. *Australian Mathematics Teacher, 60*(1), 34-40.

Watson, J. M. (2006). *Statistical literacy at school: Growth and goals.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Watson, J. M. (2011). Cheating partners, conditional probability and contingency tables. *Teaching Statistics, 33*(3), 66-70.

Watson, J. M., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3–46.

Watson, J., Callingham, R., & Nathan, E. (2009). Probing teachers' pedagogical content knowledge in statistics: "How will Tom get to school tomorrow?" In R. Hunter, B. Bicknell, & T. Burgess (Eds.), Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia (Vol. 2, pp. 563-570). Adelaide: MERGA.

Watson, J. M., & Kelly, B. A. (2008). Sample, random and variation: The vocabulary of statistical literacy. *International Journal of Science and Mathematics Education, 6*(4), 741-767.

Watson, J. M., & Moritz, J. B. (2000a). Developing concepts of sampling. *Journal for Research in Mathematics Education, 31*(1), 44-70

Watson, J. M., & Moritz, J. B. (2000b). Development of understanding of sampling for statistical literacy. *Journal of Mathematical Behavior, 19*(1), 109-136.

Watson, J. M., & Moritz, J. (2002). Quantitative Literacy for pre-service teachers via the Internet. *Mathematics Teachers Education and Development, 4* ,42-55.

Watson, J., & Nathan, E. (2010a). Biased sampling and PCK: The case of the marijuana problem. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education. Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 610–617). Fremantle, WA: MERGA.

Watson, J., & Nathan, E. (2010b). *Assessing the interpretation of two-way tables as part of statistical literacy*. Paper presented at the Eighth International Conference on the Teaching of Statistics, Ljubljana, Slovenia. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_5E1_WATSON.pdf

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, California: Sage Publications.

Yolcu (2012). *An investigation of eighth grade students' statistical literacy, attitudes towards statistics and their relationship*. Unpublished master's thesis, Middle East Technical University, Ankara.

APPENDICES

APPENDIX A HOLISTIC SCALE FOR PARTICIPANT SELECTION

	2 points	1 point	0 point
Using valid quantitative procedures and mathematical language	Applying appropriate procedure by using data given in the news and correct usage of mathematical terminology	Applying appropriate procedure by using created data in the news and correct usage of mathematical terminology or, Applying inappropriate procedure, but mathematical terminology is used correctly or, Applying appropriate procedure, but there is partial answer or incorrect usage of mathematical terminology	Applying inappropriate procedure and incorrect mathematical terminology or, Not applying any quantitative procedure
Critical Evaluation	Detecting at least three errors or deficiencies in the newspaper article and make justification or decision based on statistical and probabilistic knowledge	Detecting at least three errors or deficiencies in the newspaper article and not making justification or decision based on statistical and probabilistic knowledge or, Detecting one or two errors or deficiencies in the newspaper article and making justification or decision based on statistical and probabilistic knowledge	Detecting one or two errors or deficiencies in the newspaper article and not making justification, decision based on statistical and probabilistic knowledge or making inappropriate statistical judgment or, Not detecting anything, believe everything in the news
Giving Rich Information	Making both formal and informal meaningful explanations in detail and having at least three different points of views compared to the others	Making partial formal explanation and informal explanation and having one or two different points of views compared to the others	Making just superficial informal explanation and having same ideas or justifications with the others

APPENDIX B ANALYSIS OF QUESTIONNAIRE

	Criteria of Participant Selection			
Student	Using valid quantitative procedures and mathematical language	Critical Evaluation	Giving Rich Information	Total
S1	1	1	2	**4**
S2	1	1	2	**4**
S3	1	1	2	**4**
S4	1	0	2	**3**
S5	1	0	2	**3**
S6	1	0	2	**3**
S7	1	0	2	**3**
S8	0	1	2	**3**
S9	1	0	2	**3**
S10	1	1	1	**3**
S11	1	0	1	**2**
S12	1	0	1	**2**
S13	1	0	1	**2**
S14	0	0	1	**2**
S15	0	1	0	**2**
S16	0	1	0	**2**
S17	0	0	1	**2**
S18	1	0	1	**2**
S19	1	0	1	**2**
S20	0	0	1	**1**
S21	0	0	1	**1**
S22	0	0	0	**1**
S23	1	0	0	**1**
S24	0	1	0	**1**
S25	1	0	0	**1**
S26	0	0	0	**1**
S27	0	0	1	**1**
S28	0	0	1	**1**
S29	1	0	0	**1**
S30	0	0	0	**0**
S31	0	0	0	**0**
S32	0	0	0	**0**
S33	0	0	0	**0**
S34	0	0	0	**0**
S35	0	0	0	**0**
S36	0	0	0	**0**
S37	0	0	0	**0**
S38	0	0	0	**0**

APPENDIX C.PRE AND POST INTERVIEW QUESTIONS

APPENDIX C.1 PRE-INTERVIEW QUESTIONS

APPENDIX C.1.1 TURKISH VERSION OF PRE-INTERVIEW QUESTIONS

1. İstatistik ve olasılık dersini ve araştırma yöntemleri dersini aldıktan istatistik ve olasılığın günlük yaşamda kullanımına dair düşüncelerinde nasıl bir değişme meydana geldi?
2. Kendi hayatında istatistik ve olasılığın kullanımının önemine dair ne düşünüyorsun?
3. Bir birey için istatistik ve olasılık kavramlarını bilmenin gerekliliği hakkında ne düşünüyorsun?
4. Medyada yer alan araştırmaları nasıl değerlendirip yorumluyorsun?
    - Yorumlarken karşılaştığın zorluklar nelerdir?
    - İstatistik ve olasılık kavramlarını biliyor olmak medyada yer alan bilgileri (haber, bilgi, reklam) yorumlarken bir oynar mı?
        - Evet ise; Nasıl? / Neden?
            - Ne tip haberlerde/reklam/bilgilerde rol oynar?
        - Hayır ise; Neden?
5. İstatistik ve olasılık kavramlarının ilköğretim matematik derslerinde öğretiminin gerekliliğine dair ne düşünüyorsun?
6. İstatistik ve olasılığı öğrenme konusunda kendini güçlü ya da zayıf hissettiğin yönler nelerdir?
    - İstatistik ve olasılığı öğretme konusunda kendini güçlü ya da zayıf hissettiğin yönler nelerdir?
7. İstatistik ve olasılık öğrenme alanında hangi konuları kendin öğrenirken güçlük çekeceğini düşünüyorsun?
    - İstatistik ve olasılık öğrenme alanında hangi konuları öğretirken güçlük çekeceğini düşünüyorsun?
8. Stajda kaçıncı sınıfların dersini gözlemliyorsun?
    - İstatistik ve olasılık dersini gözlemledin mi?
    - İstatistik ve olasılık dersini anlatmayı ister misin? Neden?

APPENDIX C.1.2 ENGLISH VERSION OF PRE-INTERVIEW QUESTIONS

1. What changes occurred in your thoughts regarding the usage of statistics and probability in real life after you took the courses of Statistics and Probability and Research Methods?
2. What do you think about the importance of statistics and probability in your real life?
3. What do you think about the necessity of statistical and probabilistic knowledge for an individual in the society?
4. How do you interpret and evaluate the research studies published in the media?
    - What difficulties do you face in interpreting newspaper articles?
    - Does knowing statistical and probabilistic concepts have a role in interpretation of information (news, advertisements) in the media?
        - If yes, How? and Why?
        - In which type of news/ advertisements/ information does it have a role in?
        - If no, Why not?
5. What do you think about the necessity of teaching statistics and probability concepts in elementary mathematics education?
6. What strengths and difficulties do you have in learning the concepts of statistics and probability?
    - What strengths and difficulties do you have in teaching the concepts of statistics and probability?
7. What topics do you think you will have difficulty in while learning statistics and probability?
    - What topics do you think you will have difficulty in while teaching statistics and probability?
8. Which grade are you observing in your practice course?
    - Did you observe the lesson on statistics and probability in your school practice?
    - Do you want to teach statistical and probabilistic concepts in your school practice? Why?/Why not?

APPENDIX C.2 POST-INTERVIEW QUESTIONS

APPENDIX C.2.1 TURKISH VERSION OF POST-INTERVIEW QUESTIONS

1. Kendi hayatında istatistik ve olasılığın kullanımının önemine dair ne düşünüyorsun?
2. Haberleri matematiksel olarak değerlendirilebiliyor olmanın bir birey için önemi nedir?
3. Yeterince ayrıntılı değerlendirebildiğini düşünüyor musun?
    - Daha ayrıntılı değerlendirmek için ne gerekiyor?
4. Bu haberleri değerlendirirken yaşadığın zorluklar nelerdir?
5. Medyadaki haberlerin eğitim ve öğretimde kullanılmasına dair fikirlerini merak ediyorum. Haberlerin hangilerinin ilköğretim matematik derslerinde ya da öğretmen yetiştirme kurumlarında kullanılmasını tavsiye ediyorsun?
    - İlköğretim matematik derslerinde istatistik ve olasılık kavramlarının öğretilmesinde kullanılmalı mıdır?
        - Evet ise; Neden? Kullanımına yönelik önerilerin nelerdir?
        - Hayır ise; Neden?
        - İstatistiki bilgi düzeyi açısından nasıl değerlendirirsin?
        - İçerik açısından nasıl değerlendirirsin?
    - Öğretmen yetiştirme kurumlarında istatistik ve olasılık kavramlarının öğretilmesinde kullanılmalı mıdır?
        - Evet ise; Neden? Kullanımına yönelik önerilerin nelerdir?
        - Hayır ise; Neden?
        - İstatistiki bilgi düzeyi açısından nasıl değerlendirirsin?
        - İçerik açısından nasıl değerlendirirsin?
6. Yapmış olduğumuz görüşmeden sonra istatistik ve olasılığın kullanımının günlük yaşamda kullanımına yönelik düşüncelerinde bir değişiklik olduğunu düşünüyor musun?
    - Yeni fikirler verdiğini düşünüyor musun?
    - Bu çalışmayla fark ettiğin kavramlar, kullanımlar oldu mu?
7. Son olarak istatistik ve olasılığın ilköğretim matematik derslerinde öğretilmesine yönelik önerilerin nelerdir?
    - İstatistik ve olasılığın öğretmen yetiştirme kurumlarında öğretilmesine yönelik önerilerin nelerdir?

APPENDIX C.2.2 ENGLISH VERSION OF POST-INTERVIEW QUESTIONS

1. What do you think about the importance of statistics and probability in your real life?
2. What is the importance of evaluating newspaper articles mathematically for a citizen in the society?
3. Do you think you evaluate newspaper articles in enough detail?
    - What is needed to evaluate newspaper articles in more detail?
4. What difficulties do you encounter in the evaluation of newspaper articles?
5. I wonder about your opinions regarding the application of newspaper articles to educational environments. Which newspaper articles do you suggest using in elementary mathematics education or elementary mathematics teacher education programs?
    - What do you think about the necessity of newspaper articles for teaching statistical and probabilistic concepts?
        - If yes, Why? What do you suggest regarding how to use them in learning environments?
        - If no, Why not?
        - How do you evaluate the appropriateness of the content of these newspaper articles in terms of statistical and probabilistic knowledge for elementary school students?
        - How do you evaluate the context of the newspaper articles for elementary school students?
    - What do you think of the necessity of newspaper articles in teaching statistical and probabilistic concepts in teacher education programs?
    - If yes, Why? What do you suggest regarding how to use them in teacher education programs?
    - If no, Why not?
    - How do you evaluate the appropriateness of the content of these newspaper articles in terms of statistical and probabilistic knowledge for pre-service elementary mathematics teachers?
    - How do you evaluate the context of the newspaper articles for pre-service elementary mathematics teachers?
6. After the interview, do you think there is a change in your opinions regarding the usage of statistics and probability in real life?
    - Do you think you have gained a different point of view regarding learning and teaching of statistics and probabilistic in mathematics education after the study?
    - Is there any new usage of statistical and probabilistic concepts or new concepts you recognized after the study?
7. Lastly, what are your suggestions regarding teaching statistical and probabilistic concepts in elementary mathematics?
    - What are your suggestions regarding teaching statistics and probabilistic concepts in teacher education programs?

APPENDIX D QUESTIONNAIRE

Otizm riskini arttıran ilaçlar

California'da yapılan bir araştırmaya göre hamileliği esnasında anti-depresan ilaç kullanan kadınların çocukları kullanmayanlara göre daha çok otizmli olma riski taşıyor.

300 civarında otizmli çocuğun dahil olduğu 1800 çocuk üzerinde yapılan araştırmaya göre hamileyken ya da öncesinde anti-depresan kullanan kadınların çocuklarının otistik olma ihtimali kullanmayanların 2 katından daha fazla.

Hamileliğinin ilk 3 ayında anti-depresan kullananlar için ise durum daha da vahim.

Bu dönemde anti-depresan kullananların çocukları diğerlerinin 4 katından daha fazla otizme ya da otizmle bağlantılı bir bozukluğu yaşıyor.

Çalışma serotonin fluoxetine, paroxetine ve sertraline maddelerini içeren ilaçlar üzerine odaklanan bir çalışmaydı.

Araştırmanın SSRIs üzerine odaklanmasının sebebi ise serotonin hormonunun otizmin gelişimi üzerinde bir etkisi olduğuna dair kuvvetlenen görüşler.

Daha önce yapılan bazı araştırmalar otizm teşhisi konulan çocukların kanlarında diğerlerine göre serotonin miktarının daha yüksek olduğunu ortaya koydu.

Ancak araştırmada kafa karıştıran bir nokta ise şöyle;

Otizme sebep olan bahsi geçen anti-depresan ilaçlar mı yoksa anti-depresan kullanmaya sebep olan depresyon mu? Henüz bu soru tam olarak yanıt bulamasa da uzmanlar kadınların tedavilerine doktor kontrolünde devam etmelerini öneriyor...

KAYNAK:http://www.ozelegitimsitesi.com/otizm-ve-egitimi/otizm-riskini-arttiran-ilaclar.html

Aşağıdaki çalışma, gazete haberlerinde yer alan istatistiksel bilgileri yorumlamaya yöneliktir. Haberi okuduktan sonra aşağıdaki soruları cevaplayınız. Bütün sorulara cevap vermeniz araştırma açısından önem taşımaktadır.

Araştırmaya katılımınızdan dolayı teşekkür ederiz.

1. Haberde verilen araştırmada bilim adamlarının ulaşmış olduğu sonuç nedir ve bu sonuçtan nasıl bir çıkarım yaparsınız?

2. Haberin başlığı da dahil olmak üzere haberde yanlış olarak ifade edildiğini ya da eksik olduğunu düşündüğünüz ifadeler var mı? Yanlış ya da eksik bir ifade varsa nasıl bir düzenleme yaparsınız? Nedenini açıklayınız.

3. Haberde yer alan **"Hamileyken ya da öncesinde anti-depresan kullanan kadınların çocuklarının otistik olma ihtimali kullanmayanların 2 katından daha fazla."** ifadesinden ne anlıyorsunuz? Bu sonuca ulaşmak için araştırmacı nasıl bir uygulama yapmış olabilir? Sayısal örneklerle açıklayınız.

4. Haberde yer alan **"Hamileliğinin ilk 3 ayında anti-depresan kullananlar için ise durum daha da vahim.Bu dönemde anti-depresan kullananların çocukları diğerlerinin 4 katından daha fazla otizme ya da otizmle bağlantılı bir bozukluğu yaşıyor."** ifadesinden ne anlıyorsunuz? Verilen ifadeyi farklı şekilde nasıl ifade edersiniz?

5. Hamileliği sırasında anti depresan kullanan bir yakınınıza yaşayabileceği riskli durumları matematiksel olarak nasıl ifade edersiniz? Haberi tekrar inceleyerek antidepresan kullanımına yönelik nasıl bir öneride bulunursunuz?

APPENDIX E INTERVIEW PROTOCOL REGARDING NEWSPAPER ARTICLE I

APPENDIX E.1 TURKISH VERSION OF NEWSPAPER ARTICLE I

Kan testiyle Down sendromu teşhisi

Basit bir kan testine dayanan yeni yöntem, anne karnındaki bebeğin Down sendromu taşıyıp taşımadığını büyük oranda doğru gösteriyor.

Derisinin delinip de kanının küçük bir metal boru içine çekilmesinden kimse pek hoşlanmaz. Fakat kan testleri tıbbi teşhis açısından oldukça önemli ve diğer bir çok yönteme göre çok daha hızlı ve daha az riske sahip. Bu nedenle çoğu hastalığın anlaşılabilmesi için kan testlerinin kullanılmasına yönelik çalışmalar yoğun bir şekilde sürdürülüyor.

Yeni bir kan testi yöntemi, Down sendromuna sahip fetüsleri henüz anne karnındayken doğru bir şekilde teşhis edebilme iddiasında. Yöntem anneden alınan kanın analizi ve bebeğin bir ultrason görüntüsünün bileşimini içeriyor.

Down sendromu, bebeğin 21. kromozomunun fazladan bir kopyasının oluştuğu durumda ortaya çıkıyor. Fetüsün DNA'sı annenin plazmasına karıştığı için doktorlar bu fazladan kromozomun izini anne kanında takip edebilirler.

753 hamile kadın üzerinde denenen yöntem hiç bir yanlış saptama yapmaksızın başarılı olmuş. Test için seçilen gönüllü kadınların tamamı, Down sendromlu bebek dünyaya getirme yönünden normalden çok daha fazla risk taşıyan kişiler arasından seçilmiş. Normalde Down sendromuna sahip bebek doğurma oranı 800'de birken, 753 kadının 86'sı bu hastalığa sahip bebek dünyaya getirmiş ve yeni test bunların tamamını doğumdan önce tespit edebilmiş.

Test henüz tam olarak mükemmel değil. Aslında hastalık taşımayan bebeklerden yüzde ikisini de Down sendromlu olarak tespit etmiş. Fakat daha düşük risk grubuna sahip bayanlar arasında bu yanlış sonuca varma oranının daha düşük olacağı öngörülüyor.

Kaynak: http://www.ntvmsnbc.com/id/25172270/

17 Ocak. 2011 Pazartesi

APPENDIX E.1.1 TURKISH VERSION OF INTERVIEW PROTOCOL REGARDING NEWSPAPER ARTICLE I

Genel Sorular:

1- Haberde anlatılmak istenen nedir?
2- Haberde verilen araştırmada bilim adamlarının ulaşmış olduğu sonuç nedir?
3- Araştırmanın sonucundan nasıl bir çıkarımda bulunursun? Bu sonucu nasıl yorumlarsın?

Habere Özgü Sorular:

1. Haberde yer alan "**Normalde Down sendromuna sahip bebek doğurma oranı 800'de birken, 753 kadının 86'sı bu hastalığa sahip bebek dünyaya getirmiş.**" İfadesinden ne anlıyorsun?
    - "Down sendromuna sahip bebek doğurma oranı 800'de bir" ifadesini matematiksel ifadeler kullanarak nasıl ifade edersin?
2. **"Hastalık taşımayan bebeklerden yüzde ikisini de Down sendromlu olarak tespit etmiş."** ifadesinden ne anlıyorsun?
    - Yanlış olarak tespit edilen çocukların sayısı bulunabilir mi?
    - Verilen cümleyi matematiksel ifadeler kullanarak farklı şekilde nasıl anlatırsın?
    - Araştırmacı bu ifadeye nasıl ulaşmış olabilir?
3. Anne karnındaki bebeğin Down Sendromu olup olmadığını belirlemeye çalışan, testin doğruluk oranı hakkında ne düşünüyorsun?
    - Bu testin Down sendromlu bir fetüsün varlığını doğru olarak teşhis etme olasılığı bulunabilir mi?
        - ✓ Testin doğruluk oranını bulmada etkili olabilecek bütün olası durumlar hakkında ne düşünüyorsun? Hangi bilgilere ihtiyaç olduğunu düşünüyorsun? Yeterince veri var mı bu olasılığı bulmak için?
        - ✓ Bu olası durumları tablo, ağaç diyagramı ya da farklı bir temsil biçimi kullanarak gösterebilir misin?
    - Bu testin down sendromlu olmayan bir fetüsün varlığını doğru olarak teşhis etme olasılığı bulunabilir mi?
        - ✓ Bu iki olasılık arasında bir ilişki var mıdır? Aynı şeyi mi ifade ediyor?
        - ✓ Hangi matematiksel kavram ile ilişkili olabilir?
    - Test aracılığıyla down sendromuna sahip olduğu belirlenen bir fetüsün gerçekte bu sendroma sahip olma olasılığı hakkında ne düşünüyorsun? Önceden sormuş olduğum iki olasılıksal ifade ile bu olasılık arasında ilişki var mıdır?
        - ✓ Araştırmada bahsedilen kan testine göre bebeğinde Down Sendromu ortaya çıkan bir kişiye önerin nedir?

- ✓ Matematik ifadelerle anlaşılır bir şekilde nasıl anlatırsın?

4. Gazete haberinde verilen araştırmanın sonuçlarını güvenilirlik açısından nasıl değerlendirirsin?
   - ✓ Sonuç hakkında çıkarımda bulunurken ve sonuçların güvenirliğini değerlendirirken nelere dikkat edersin?
     (i) Bu kriterlere dikkat etmesen neler ortaya çıkardı?
   - ✓ Haberde daha farklı şekilde ifade edilmesini düşündüğün ifadeler var mı? Varsa nasıl düzenlersin?
5. Bu haberi kullanarak öğrencilerine istatistik ve olasılıkla ilgili nasıl bir soru sorabilirsin?

APPENDIX E.2 ENGLISH VERSION OF NEWSPAPER ARTICLE I

Down syndrome diagnosis by blood tests

The new method based on a simple blood test indicates with a high degree of precision whether or not the fetus in the mother's womb has Down syndrome.

No one quite likes a hole to be made in one's skin and then have his/her blood drawn into a tube. Yet, blood tests are very important for medical diagnosis and are much quicker and less risky when compared to many other methods. For this reason, for the diagnosis of most diseases, studies on the use of blood tests are intensively carried out.

A new blood test method has the potential to diagnose with precision fetuses with Down syndrome when they are still in the mother's womb. The method involves a combination of the analysis of blood drawn from the mother and the examination of the fetus' ultrasound image.

Down syndrome occurs when the fetus has an extra copy of the 21st chromosome. Because the fetal DNA is found in the mother's plasma, physicians can trace this additional chromosome in the mother's womb.

This method, which was experimented on 753 pregnant women, was found to be successful in its detection yielding no incorrect results. All of the volunteer women chosen for the test were selected from among those who held much higher risks than normal in terms of giving birth to a baby with Down syndrome.While the ratio of giving birth to a baby with Down syndrome is normally one out of 800, 86 of the 753 women had given birth to a baby with this disease, and the new test could diagnose all of these before birth.

The test is not yet completely perfect.Actually, 2 per cent of the babies not having Down syndrome were incorrectly diagnosed to have Down syndrome. However, the ratio of arriving at an incorrect result among the women in the lower risk group is predicted to be low.

Reference: http://www.ntvmsnbc.com/id/25172270/

17 January 2011 Monday

APPENDIX E.2.1 ENGLISH VERSION OF INTERVIEW PROTOCOL REGARDING NEWSPAPER ARTICLE I

General Questions

1. What is the main idea of the newspaper article?
2. What conclusions did researhers reach?
3. What conclusions could you draw from the text?

- How do you interpret result or conclusions reported in the newspaper article?

Specific Questions

1. What do you understand from the following statement reported in the newspaper article: "While the ratio of giving birth to a baby with Down sydrome is normally one out of 800, 86 of the 753 women had given birth to a baby with this disease"?

- How do you mathematically restate the following expression: "the ratio of giving birth to a baby with Down sydrome is normally one out of 800."

2. What do you understand from the following statement reported in the newspaper article: "Actually, 2 per cent of the babies not having Down syndrome were incorrectly diagnosed to have Down syndrome."

- Is it possible to find the number of children who were wrongly diagnosed in this situation?
- How can you rephrase the given statement mathematically in a different way?
- How could the research draw such a conclusion in the statement?

3. What do you think about the accuracy rate of the test which determine whether or not the fetus in the pregnant women has Down syndrome?

- It is possible to find the accuracy of detecting fetuses without Down syndrome?
- What do you think about all possible conditions that have a role in finding the accuracy rate of the test?
    - What information would be useful?
    - Is there enough information to find the accuracy rate?
    - How do you represent all possible conditions in a table, tree diagram or any other graphical display?
- It is possible to find the accuracy rate of detecting fetuses with Down syndrome?
    - How do you relate these two accuracy rates to each other? Is there a difference between them?
    - Which mathematical concept are they related to?
- What do you think about the probability of babies with a positive test result actually having the Down syndrome?

- What is the relationship between this question and the previous two questions?
- What would you advice people with a positive test result?
- How would you explain it by using mathematical expressions in a clear way?

4. How do you evaluate the results and conclusions reported in the newspaper article in terms of their credibility?

- What factors do you recognize for the assessment of the credibility of the claims or arguments reported in the newspaper article?
- What would happen if you did not recognize these factors?
- Are they any statements in the newspaper article that need to be revised or restated in a different way? If yes, how would you revise them?

APPENDIX F INTERVIEW PROTOCOL REGARDING NEWSPAPER ARTICLE II

APPENDIX F.1 TURKISH VERSION OF NEWSPAPER ARTICLE II

Yapılan araştırmaya göre, erkekler eşinin aldattığını daha iyi tespit ediyor

Kadınlar dikkat. Yeni yapılan bir araştırmaya göre, aldatan eşi tespit etmede erkekler kadınlardan daha iyi ve erkekler, var olmayan aldatmalardan şüphelenmeye daha eğilimli.

Amerika'da çiftlerle ilgili yapılan çalışma, erkeklerin daha şüpheci olduğunu buldu. Fakat, cinsiyetler üzerine çalışan Avustralyalı bir araştırmacı, daha çok aldatma eğiliminde oldukları için erkeklerin daha şüpheci olduklarını söylüyor.

Sydney'den terapist Rosie King "Buradaki durum, 'tencere dibin kara, seninki benden kara'nın açık bir örneğidir." diyor.

Richmond'daki Virginia Commonwealth Üniversitesi'nden araştırmacılar, 203 genç çifte ait cevapların gizliliği korunacak şekilde hazırlanmış anketler yoluyla eşlerini hiç aldatıp aldatmadıklarını, eşlerinin aldattığını bilip bilmediklerini ya da eşlerinden şüphelenip şüphelenmediklerini sordu.

New Scientist dergisinde yayımlanan sonuçlara göre erkeklerin yüzde 29'u aldattığını itiraf ederken, kadınlarda bu oran yüzde 18.5.

Araştırmacı Paul Andrews, "Kadınların eşlerinin sadakat ya da sadakatsizliği hakkında çıkarımlarının yüzde 80' i doğru. Ancak, erkekler yüzde 94'lük doğruluk oranıyla kadınlara göre daha iyi." diyerek erkeklerin, eşlerinin sadakatini yargılamada kadınlardan daha iyi olduğunu vurguluyor.

Verilen cevaplara göre, erkekler eşlerinin aldatmalarının %75' ini fark ederek, aldatmayı büyük bir olasılıkla tespit ederken, kadınlar aldatan eşlerin % 41' ini tespit ettiler.

Kaynak: Australian Associated Press. (2008). Cheat radar beter tuned in men, study finds. *The Mercury*, 30 October, p.3

APPENDIX F.1.1 TURKISH VERSION OF INTERVIEW PROTOCOL REGARDING NEWSPAPER ARTICLE II

Genel Sorular:

1. Haberde anlatılmak istenen nedir?
2. Haberde verilen araştırmada ulaşılan sonuç nedir?
3. Araştırmanın sonucundan nasıl bir çıkarımda bulunursun?

Habere Özgü Sorular:

1. Araştırmacı bu çalışmayı nasıl yapmış olabilir?
    - ✓ Araştırmacı çalışmasını kaç kişiyle yapmıştır?
    - ✓ Araştırmacı katılımcılara ne sormuş olabilir?
    - ✓ Verilen cevapları araştırmacı nasıl organize etmiş olabilir?
    - Bütün olası durumları belirleyebilir misin?
    - Bu olası durumları tablo, ağaç diyagramı ya da başka bir temsil biçimiyle nasıl gösterirsin?
2. "Verilen cevaplara göre, erkekler eşlerinin aldatmalarının %75' ini fark ederek, aldatmayı büyük bir olasılıkla tespit ederken, kadınlar aldatan eşlerin % 41' ini tespit ettiler." İfadesinden ne anlıyorsun?
    - Verilen durumu matematiksel ifadelerle nasıl ifade edersin?
    - Olasılık içeren bu durum nasıl elde edilmiş olabilir?
    - Verilen değerleri oluşturduğun yapmış olduğun temsil biçimiyle nasıl ilişkilendirebilirsin?
3. "Araştırmacı Paul Andrews, "Kadınların eşlerinin sadakat ya da sadakatsizliği hakkında çıkarımlarının yüzde 80' i doğru. Ancak, erkekler yüzde 94'lük doğruluk oranıyla kadınlara göre daha iyi." diyerek erkeklerin, eşlerinin sadakatini yargılamada kadınlardan daha iyi olduğunu vurguluyor." ifadesinden ne anlıyorsun?
    - Verilen durumu matematiksel ifadelerle nasıl ifade edersin?
    - Araştırmacı bu sonuca nasıl ulaşmış olabilir?
4. Çalışmaya katılan çiftlerin birbirleri hakkında yanlış karara varma olasılıkları hakkında ne düşünüyorsun?
    - Erkeklerin, gerçekte eşi aldatmadığı halde eşinin aldattığını düşünerek yanlış karara varma olasılığı bulunabilir mi? Bulunabilir ise nasıl hesaplarsın?
        - ✓ Bulunabilir ise nasıl hesaplarsın?
        - ✓ Bulunamaz ise, bulunabilmesi için hangi bilgiler gerekli?
    - Kadınların, gerçekte eşi aldatmadığı halde eşinin aldattığını düşünerek yanlış karara varma olasılığı bulunabilir mi? Bulunabilir ise nasıl hesaplarsın?
        - ✓ Bulunabilir ise nasıl hesaplarsın?
        - ✓ Bulunamaz ise, bulunabilmesi için hangi bilgiler gerekli?

5. “Yeni yapılan bir araştırmaya göre aldatan eşi tespit etmede erkekler kadınlardan daha iyi ve erkekler, var olmayan aldatmalardan şüphelenmeye daha eğilimli.” İfadesini nasıl değerlendirirsin?
    - Haberi tekrar inceledikten sonra erkeklerin ve kadınların eşlerinin aldatıp aldatmadığını hakkındaki doğru ve yanlış yargılara varmaları konusunda ne düşünüyorsun?
    - Haberde daha farklı şekilde ifade edilmesini düşündüğün ifadeler var mı? Varsa nasıl düzenlersin?
6. Gazete haberinde verilen araştırmanın sonuçlarını nasıl değerlendirirsin?
    - Sonuçlara ne ölçüde güvenirsin?
        - ✓ Sonuç hakkında çıkarımda bulunurken ve sonuçların güvenirliğini değerlendirirken nelere dikkat edersin?

APPENDIX F.2 ENGLISH VERSION OF NEWSPAPER ARTICLE II

Cheat radar better tuned in men, study finds

WOMEN beware. New reseach shows men are better at detecting a cheating partner than females, and they are more likely to suspect infidelities that do not exist.

A U.S study of heterosexual couples found men are more suspicious, but an Australian sex researcher says they are only more suspecting because they are more likely to cheat.

“What we have here is a clear case of the pot calling the kettle black,” said Sydney thearapist Rosie King.

Researchers at Virginia Commonwealth University in Richmond gave confidential questionnaires to 203 young couples, asking them whether they had ever strayed, and whether they suspected or knew their partner had.

The results, published in *New Scientist*, show 29 per cent of men admitted they had cheated compared with 18.5 per cent of women.

Researcher Paul Andrews said men were better at judging fidelity than women.

“Eighty per cent of women’s inferences about fidelity or infidelity were correct, but men were even better, accurate 94 per cent of the time,” Dr. Andrews said.

Men were more likely to catch out a cheating partner, picking up on 75 per cent of the reported infidelities compared with 41 per cent discovered by women.

AAP

MERCURY-3

Thursday, October 30, 2008

APPENDIX F.2.1 ENGLISH VERSION OF INTERVIEW PROTOCOL REGARDING NEWSPAPER ARTICLE II

General Questions

1. What is the main idea of the newspaper article?
2. What conclusions did researhers reach?
3. What conclusions could you draw from the text?

- How do you interpret result or conclusions reported in the newspaper article?

Specific Questions

1. How could the researcher conduct the study reported in the newspaper article?

- What is the number of participants in the study?
- What questions could the researcher ask to the participants to collect data?
- How could the researcher organize data?
- Could you identify possible conditions?
- How do you represent these possible conditions in a table, tree diagram, or any other representation?

2. What do you understand from the following statement: "Men were more likely to catch out a cheating partner, picking up on 75 per cent of the reported infidelities compared with 41 per cent discovered by women."

- How can you express the statements mathematically?
- How could such a conclusion be reached?
- How do you related your conclusions with your representation?

3. What do you understand from the following statements: "Researcher Paul Andrews said men were better at judging fidelity than women. "Eighty per cent of women's inferences about fidelity or infidelity were correct, but menwere even better, accurate 94 per cent of the time," Dr. Andrews said."

- How can you express the statements mathematically?
- How could such a conclusion be reached?

4. What do you think about the probability of couples' arriving at incorrect inferences about each other?

- Is it possible to find incorrect inferences of men whose partners do not cheat them, about their wives?
    - If yes, how would you calculate it?
    - If no, what information would be needed to find the probability?
- Is it possible to find incorrect inferences of women whose partners do not cheat them, about their husbands?
    - If yes, how would you calculate it?
    - If no, what information would be needed to find the probability?

5. How do you evaluate the following statement: “New research shows men are better at detecting a cheating partner than females, and they are more likely to suspect infidelities that do not exist.”?

- What do you think about couples’ incorrect and correct inferences for each other whether or not cheating exists after reading the newspaper article again?
- Are there any statements needed to restate it in a different way? If yes, how would you revise them?

6. How would you evaluate the results or conclusions stated in the newspaper article?

- To what extent would you rely on the conclusions?
- What factors would you consider while assessing the newspaper article and drawing conclusions from the article?

APPENDIX G CODING SYSTEM FOR STATISTICAL AND PROBABILISTIC KNOWLEDGE

1.Base of Reported Findings: To know at least intuitively background of the study reported in the newspaper article; how sample could be selected, how data could be collected and analyzed, and how statistical conclusions reported in the newspaper could be reached.

- **Sampling:** To know sampling processes of the researches reported in the newspaper articles, which requires identification of sample characteristics, sampling method, role of sample on inference from sample to population, and possible biases in sampling. *For example*: Knowing what the sample is and its purpose, what characteristics (sample size, properties such age, gender etc.,) it has; Knowing essential role of sample in a study, which requires understanding the need inference from sample to population and the need for samples to be representative; Knowing the relationship between the notion of representativeness and the variation among samples, which emerges from repeated sampling from the same population; or bias in convenience sampling or power of probability sampling; bias in sample size for each group or category in the study.

- **Data Collection**: To know at least informally how data can be produced in the researches reported in the newspaper articles such as survey, experimental, or correlational study. *For example:* Knowing what was measured and how to be measured; the relationship specific questions asked in an instrument and the reported findings; the need for definition of concepts to be measured especially in social studies; possible biases in data collection (exp. to what extent reliably measured); design of the study that has powerful role on controlling errors.

- **Data Analysis**: Knowing at least informally how data would be analyzed. *For example:* Knowing how the researcher could analyze the data to reach summary statistics such as percentages, means, or medians reported in the newspaper article; Knowing, even not formal, how deficiencies or errors could affect data analysis process.

- **Results and Conclusions**: This part includes results and conclusions reported in the newspaper articles. To know the ways to determine significance of the results or differences between groups reported in the newspaper. *For example:* Knowing what factors (exp. sample size for each group, quality of sampling process) could be effective in the determination of the significance of the differences among the groups; Knowing the difference between associational and causational statements reported in the media; knowing difference among groups may not be enough large or stable or may emerge by chance even the difference could be observed among them.

2. Reported Findings: This part includes summary or descriptive statistics reported in the popular media texts such as percentages, means, medians etc and probabilistic estimates or risks which are reported through percentages, odds, ratios, or verbal estimates in the newspaper articles.

- **Probabilistic Statements:** Probabilistic statements can be reported as in either verbally or numerically in the newspaper article. *For example:* Knowing that newspaper articles should include base rate of a phenomenon because the interpretation of the probabilistic estimates differ in terms of base rate of that phenomenon; difference between inverse relationship between conditional probability

statements P(A|B) and P(B|A); conditional probability statements underlies Bayes' Theorem to detect erroneous probabilistic claims or arguments; probabilistic statements expressed through relative risk or absolute risk; probabilistic statements in the newspaper articles are based on formal or subjective estimates which affect the credibility of the study published in the newspaper.

- **Percentages:** Percentages as a summary statistics are used with different expressions such as a number, statistical value, the probability of an event in the newspaper articles. *For example:* Percentage could be used for both mathematical and statistical values and make computations regarding different types of percentages stated as bigger than 100%, percentage of percent, margin error.

- **Measures of Central Tendency:** Measures of central tendency as a summary statistics are used to show the center of data set. *For example:* To know in which conditions measures of central tendency could be misleading (exp. mean is more sensitive value affected by extreme values than median; measures); measures of central tendency could be misleading when the distribution of data is very uneven or bimodal (not normal), or when sample is not representative of the population from which it was selected.

- **Graphical and Tabular Displays:** To know that graphical and tabular displays are used to organize information and observe the general picture of data. *For example:* To know what conventions we need to notice in the construction of graphical and tabular displays and detect misleading conventions; graphs can be constructed to mislead people to believe general trend which does not exist in reality.

3.Generalizability of the Reported Findings: To know intuitively the extent to which reported findings can be generalize and possible factors (confounding variables, sample size, design of the study, cultural and social factors, genetic factors) affect the generalizability of the study; limitations of the study. *For example:* Knowing that sampling method, or sample size for groups has potential role on generalization from sample to population; need for sufficient information in describing sample and population so that people can determine applicability of the reported findings to other contexts.

APPENDIX F

TEZ FOTOKOPİSİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü	
Sosyal Bilimler Enstitüsü	x
Uygulamalı Matematik Enstitüsü	
Enformatik Enstitüsü	
Deniz Bilimleri Enstitüsü	

YAZARIN

Soyadı : Özen
Adı : Mehtap
Bölümü : İlköğretim Fen ve Matematik Eğitimi

TEZİN ADI (İngilizce) : Investigation of pre-service mathematics teachers' critical thinking processes through statistical and probabilistic knowledge in the context of popular media texts

TEZİN TÜRÜ : Yüksek Lisans [x] Doktora [ ]

1. Tezimin tamamından kaynak gösterilmek şartıyla fotokopi alınabilir. [x]
2. Tezimin içindekiler sayfası, özet, indeks sayfalarından ve/veya bir bölümünden kaynak gösterilmek şartıyla fotokopi alınabilir. [ ]
3. Tezimden bir bir (1) yıl süreyle fotokopi alınamaz. [ ]

TEZİN KÜTÜPHANEYE TESLİM TARİHİ: