Republic of Turkey
Marmara University
Institute of Educational Science
Department of Secondary School Science and Mathematics Education

Palestinian Pre-service Mathematics Teachers' Self-Efficacy Levels and Performance in Statistics

Saja Abed
(Master Thesis)

Advisor:
Assoc. Prof. Dr. Emin AYDIN

Istanbul, 2015

T.C.
Marmara Üniversitesi
Eğitim Bilimleri Enstitüsü
Ortaöğretim Fen ve Matematik Alanları Eğitimi Ana Bilim Dalı
Ortaöğretim Matematik Eğitimi Bilim Dalı

Filistinli Matematik Öğretmen Adaylarının İstatistik Dersi Konularındaki Kavram Yanılgıları Ve İstatistik Dersine Yönelik Öz Yeterlilik İnançları

Saja ABED
(Yüksek Lisans Tezi)

Danışman: Doç. Dr. Emin AYDIN

Istanbul, 2015

APPROVAL

Based on the thesis defense examination held in 20.06.2014, we certify that the thesis entitled "PALESTINIAN MATHEMATICS PRE-SERVICE TEACHERS' PERFORMANCE AND SELF-EFFICACY LEVELS IN STATISTICS" which is completed by Saja Abed who is a candidate for the Master of Education degree is worthy of acceptance

	Signature
SUPERVISOR: Assoc. Prof. Dr. Emin AYDIN	....
JURY: Assoc. Prof. Dr. Ali DELICE	....
JURY: Assoc. Prof. Dr. Serkan Özel	....

T.C.
Marmara Üniversitesi
Eğitim Bilimleri Enstitüsü
Ortaöğretim Fen ve Matematik Alanları Eğitimi Ana Bilim Dalı
Ortaöğretim Matematik Eğitimi Bilim Dalı

ONAY

Saja Abed tarafından hazırlanan "PALESTINIAN MATHEMATICS PRE-SERVICE TEACHERS' PERFORMANCE AND SELF-EFFICACY LEVELS IN STATISTICS" başlıklı bu çalışma, 20.06.2014 tarihinde yapılan savunma bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

İmzalar

TEZ DANIŞMANI: Doç. Dr. Emin AYDIN ....

JÜRİ ÜYESİ: Doç. Dr. Ali DELICE ....

JÜRİ ÜYESİ: Doç. Dr. Serkan Özel ....

CURRICULUM VITAE

2007	Sourif Secondary School
2011	Graduated from Al-Quds University of Mathematics Department
2012	Marmara University Institute of Educational Science, Secondary School Science and Mathematics Education Department

CONTACT INFORMATION

E-mail saja_abed@hotmail.com

Phone Number 00905075263068

ÖZGEÇMİŞ

2007	Sourif Lisesi
2011	Al-Quds Üniversitesi Fen Fakültesi Matematik Bölümünden mezun olma
2012	Marmara Üniversitesi Eğitim Bilimleri Enstitüsü Ortaöğretim Fen ve Matematik Alanları Eğitim Ana Bilim Dalı Ortaöğretim Matematik Eğitim Bilimi Dalı programına giriş

İLETİŞİM BİLGİLERİ

E-Posta	saja_abed@hotmail.com
Telefon	00905075263068

ACKNOWLEDGEMENT

I would never have been able to finish my study without the guidance of my supervisor and support from my family and my fiancé.

I would like to express my deep thanks my advisor, Assoc. Prof. Dr. Emin AYDIN, for his guidance, caring, patience and being abundantly helpful during all phases of the research. I would like to thank the members of the supervisory committee, Assoc. Prof. Dr. Ali DELICE and Assoc. Prof. Dr. Serkan ÖZEL without their knowledge and assistance this study would not have been successful.

I want to extend my thanks to the Turkish government, which gave me the opportunity to get the master's degree by providing the academic and financial support to compete this project successfully.

Special thanks to my parents for caring, understanding and providing endless love through the duration of the study. I would like to extend my deepest thanks to my fiancé, Hatem HAMDAN, who was always there cheering me up and stood by me through the good times and bad.

Finally, I would like to thank my sister, brothers and beloved friends who were always supporting me and encouraging me with their best wishes.

Saja ABED

ÖNSÖZ

Çalışmam boyunca bana yardımcı olan çok değerli hocam Yrd. Doç. Dr. Emin AYDIN 'a ve tez jüri üyesi olarak davetimizi kabul eden ve sundukları görüşlerle çalışmama geri bildirim sağlayan değerli hocalarım Yrd. Doç. Dr. Ali DELİCE' ye ve Yrd. Doç. Dr. Serkan ÖZEL 'a teşekkür ederim. Çalışmam boyunca moral desteği sağlayan ve gerekli zamanı ayırmada büyük fedakârlık gösteren nişanlım Hatem HAMDAN' ye, her türlü manevi desteği sunan ve her zaman yanımda olan anneme ve babama şükranlarımı ifade etmeliyim. Ayrıca bu çalışmayı oluşturmamda büyük katkıları olan Marmara Üniversitesi Matematik Eğitimi Tezsiz Yüksek Lisans öğrencilerine ve teknik konularda destek sağlayan sevgili Işık Emel YENGİN' e teşekkür ederim.

saja ABED

ABSTRACT

PALESTINIAN PRE-SERVICE TEACHERS' CONCEPTUAL KNOWLEDGE AND SELF-EFFICACY LEVELS IN STATISTICS

Misconceptions in statistics and low self-efficacy are common problems for pre-service teachers which need more concern in order to improve the educational system. This study aimed to explore pre-service mathematics teachers' mistakes in statistics and to measure their self-efficacy beliefs.

The present study was conducted at the 2013 – 2014 academic year with a total number of 100 graduate pre-service mathematics teachers who were in their last academic year in the mathematics department of two Palestinian universities; Al-Quds University and Al-Quds Open University. Data was collected utilizing two questionnaires: the Statistics Self- Efficacy Belief Instrument and Statistics Concept Inventory. Descriptive statistics and correlation and were used in data analysis.

Data analysis showed that participants' statistics achievement was moderate. They had some common mistakes in probability, normal distribution, hypothesis testing, interpreting graphics, sampling distributions, and correlation. They had low to moderate level of self-efficacy beliefs in statistics.

Key Words: Statistics Lesson, Pre-Service Mathematics Teachers, Misconception, Self Efficacy Belief.

ÖZET

MATEMATİK ÖĞRETMEN ADAYLARININ İSTATİSTİK DERSİ KONULARINDAKİ KAVRAM YANILGILARI AND İSTATİSTİK DERSİNE YÖNELĞK ÖZ YETERLILIK INANÇLARI

Bu çalışmanın amacı, matematik öğretmen adaylarının istatistik dersi konularındaki kavram yanılgılarının ve istatistik dersine yönelik öz yeterlilik inançlarının incelenmesidir.

Bu çalışmada nitel ve nicel veriler bir arada kullanılmıştır. Araştırmanın örneklemini 2013-2014 öğretim yılı, Filistin Cumhuriyeti'nin Al-Quds ve Al-Quds Open Üniversitesi Fen Fakültesi'nin toplam 100 Matematik öğretmen adayı oluşturmaktadır. Veri toplama araçları olarak İstatistik Kavram Testi ve İstatistik Dersine Yönelik Öz Yeterlilik İnanç Ölçeği kullanılacaktır. Veri analizinde frekans ve yüzde oranları, basit korelasyon kullanılmıştır.

Anahtar Sözcükler: Istatistik Dersi, Matematik Öğretmen Adayları, Öz Yeterlilik Inancı, Kavram Yanılgıları.

ÖZGEÇMİŞ

2007	Sourif Lisesi
2011	Al-Quds Üniversitesi Fen Fakültesi Matematik Bölümünden mezun olma
2012	Marmara Üniversitesi Eğitim Bilimleri Enstitüsü Ortaöğretim Fen ve Matematik Alanları Eğitim Ana Bilim Dalı Ortaöğretim Matematik Eğitim Bilimi Dalı programına giriş

İLETİŞİM BİLGİLERİ

E-Posta	saja_abed@hotmail.com
Telefon	00905075263068

CONTENTS

LIST OF TABLES

SYMBOLS AND ABBREVIATIONS

$\overline{x}$: Arithmetic Mean

S: Standard Deviation

t: t Variable

p: p Variable

F: F Variable

n: Student's Number

I. INTRODUCTION

The aim of this study is to measure the level of the conceptual knowledge of the Palestinian pre service teachers in statistics and also to measure the level of their statistics self-efficacy which allows us to find the relationship between both of these issues and how could both of them affect teacher's performance.

Since mathematics is one of the most important parts of our academic life. Whatever we study, whatever we measure, we cannot accomplish our work without using mathematical tasks. Because of the importance of mathematics and its role in other sciences, we have to be sure about students' level in mathematics in order to qualify them to apply mathematics in a way to make a proper interpretation of the data in a field of study.

Statistics is described as a mathematical body of science that pertains to the collection, analysis, interpretation or explanation, and presentation of data, or as a branch of mathematics concerned with collecting and interpreting data (Kass, Ventura & Brown, 2005).

According to Köksel (1985), statistics is concerned in collecting, interpreting and presenting data in statistical tools (numbers, tables, graphics... etc.) reflecting the values of the characteristics and concepts beyond these tools. As an example the interpretations of the numbers in a statistical study and correlating them with mathematical concepts and characteristics need a special effort, training and mathematical knowledge from the mathematician to understand these data and use them to make decision about the issue of the concerned study.

Many studies indicate that statistics has great effects on other branches of science (Duchastel, 1974; Jolliffe, 1976; Kalton, 1973; Urquhart, 1971). They prove that other sciences cannot be separated from statistical applications at one point. Because of this importance, students in general should have an idea about statistical applications in their domain of study, more specifically mathematics students who are candidates to be teachers must have enough knowledge about statistics and its applications on all of fields of life.

Statistical educators need to know what students want to learn in order to develop their teaching ways according to studies done in this field (Duchastel, 1974; Jolliffe, 1976; Kalton, 1973; Urquhart, 1971) and to assess if these modifications in their teachings are effective in students by monitoring their development of statistical understanding.

Statisticians who are involved in teaching statistics in the class-room have a major concern which is how to ensure that the students understand statistical ideas and are able to apply what they learn to real world situations. These teachers of statistics complain about the difficulties which students have in learning and applying course material (Joan Garfield, 1992).

There are many sources for the difficulties of statistical education from the teacher's side. In my research I focused on two of the most important elements; teachers' self-efficacy levels and their conceptual knowledge in statistics.

According to Tschannen-Moran and Woolfolk Hoy (2007), they explained that teacher's self-efficacy is a teacher's "judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated."

Bandura (1994) noted that self-efficacy is something personal and is different for one teacher to the other. Teachers should have a high level of self-efficacy in order to motivate themselves to do his/her job in a perfect way by delivering information to his/her students, beyond that he/she has to make a challenge with himself to get over the difficulties he/she faces in statistical teaching.

Teachers' confidence reflects on students' confidence and affects their understanding of statistical concepts. Teachers' self-efficacy play a role in the whole educational system, because of that importance of the issue and because of the variety of the obstacles that are resulted from the lack of teachers' self-efficacy a lot of studies focused on that issue and on how to improve it (Bandura, 1997).

The other element which affects learners is the misunderstandings of statistical topics. Teachers pass their misconceptions to their students who will be a teacher one day having the same misconceptions. To clarify these misconceptions, we should focus on the pre-service teachers to avoid misconceptions before starting their professional teaching career (Batanero, Godino, Vallecillos, & Holmes, 1994; Brewer, 1985; Haller & Krauss, 2002).

Better understanding of statistical concepts lead students to understand the skills and tasks of statistics. Understanding of concepts enables the students to be able to link statistical concepts and distinguish them more clearly (Noordman & Vonk, 1998). For a proper understanding of the concepts of statistics we must order these concepts in the form of coherent structures of information in order to make it easy to be understood by learners. (Chi et al. 1981; Wyman and Randel 1998; Kintsch 1998).

On other hand, misunderstanding of statistical concepts which is also named preconception, misuse, or misinterpretation is considered as one of the statistical obstacles in statistics education. According to Eryılmaz and Süremli (2002) every misconception is a mistake but not the opposite. According to them, misconceptions result from arrangement of ideas and concepts in a nonscientific way.

Establishing the link between pre-existing concepts and the new concepts is one of the most effective epistemological methods to understand statistics properly. To understand statistical concepts students must be able to employ their existing knowledge in order to understand the new concepts in a way to make a clear difference between what he had of statistical knowledge, what he/she is learning and what he/she is going to learn (Cansüngü Koray & Bal, 2002; Tekkaya, Çapa & Yılmaz, 2000).

Based on the studies of Bachelard (1938/2002) three resources were said to exist for students' misconceptions: epistemological, psychological and pedagogical. The epistemological misunderstanding related to student's knowledge about statistics, the psychological misconceptions also depend on the student's personality, but the pedagogical misconceptions is related to the teacher and how he/she teaches the concepts of statistics. (Broussea, 1997; Cornu, 1991).

Teachers play a main role in students' misunderstanding of concepts. Teachers, who have a major part in pedagogical system in statistical education, have a strong effect on the students to accept the ideas and concepts in a simple way according to students' level of thinking and understanding. For example, when the teacher has a misconception in a statistical concept and he/she transfer the same misunderstanding to the student.

In the context of talking about statistics in Palestine, and how much the students know about statistics; we have to start our talk from the first lesson that Palestinian students have in their study which is in the ninth grade in the school, which is just one unit in mathematics and it is talking about the basics of statistics like the measures of central tendency and measures of central discrepancy. After that in the tenth, eleventh and twelfth grades there are also just one unit in each grade which also covers the same concepts but in more complex manner compared to the ninth grade. By the end of the twelfth grade the student expected to have some knowledge about statistics and how to use it in his/her daily life events and to build over it if he will join one of the colleges in the university which needs this knowledge to take more advanced courses with more details.

In Palestinian universities in general there is a common course which is called the introduction for statistics. This course is obligatory in most of the colleges e.g. science, medicine, business and mathematics. The aim of the course is that the students will be able to use statistics in the applications of their fields of study and work in their profession later on. Medical students, for example, have a biostatistics course which allows them to be able use statistics in their practices in medicine. On the other hand in the department such as economy and engineering, which need statistics in their practice have more than one course in their curricula. In some case statistics and probability are combined. Finally, mathematical students should have enough and good knowledge in statistics and its applications. There is an introductory statistics course in the first semester and there is a more specified courses in statistics and probability in later semesters. The weight of this course is on probability for the pure mathematics majors and is on statistics for the applied mathematics majors. Students who are studying teaching of mathematics they have courses in statistics but less than those reading in pure and applied mathematics departments.

To talk about pure statistical programs in Palestinian universities, we can talk about one master program in only one university which is Palestine: Annajah University, which has masters program for statistics.

From a general look over the issue of statistics education in Palestinian education it is noticeable that in general, statistics is combined with probability and for most of the students in the STEM subjects it is the only course in which they encounter statistical contents from the elementary school till the end of university education. Palestinian students especially in the mathematics and mathematics teaching departments don't have sufficient number of courses to make them more competent in statistics.

Under the light of these issues, in this study, I aimed to measure the current self-efficacy of Palestinian pre-service teachers by using the questionnaire developed by Finney and Schraw (2003). I also aimed to detect the statistical mistakes that the pre-service teachers have using the Statistics Concept Inventory (SCI) which developed by Allen (2006). In the following section, I will give details on my study problem.

1.1. Study Problem

To solve any problem we have to make a proper diagnosis, in my study I will focus on how to put the finger on the problems that pre-service teachers mainly their self-efficacy in teaching and the misconceptions they have which might affect their way of teaching and delivering the information to students in the classroom, which will give a long term defect in the whole teaching system of mathematics and statistics.

To make sure that mathematics teachers have the enough knowledge and the ability to apply the concepts of statistics in real life, first we have to measure the level of mathematics teachers understanding of statistical concepts and to make sure if they have any misconceptions for using statistical applications, then trying to clarify these misconceptions in order to increase the level of statistical knowledge among mathematics teachers by reducing these misconceptions.

Another issue which may affect mathematics students understanding or applying their statistical knowledge is their self-efficacy. Student's self-efficacy affects student's ability to apply his/her knowledge which he/she already has and to add new information to what he/she has with the ability to make correct connection between them. Moreover, self-efficacy among teachers has a major effect on students' understanding by affecting student's confidence to their teachers. Students need to feel their teacher's knowledge of and self-confidence in the subject they are teaching is satisfactory. That way they would accept the knowledge given by their teacher.

In my study, I tried to adapt a questionnaire and a statistical performance test on pre-service Palestinian teachers. The questionnaire and the test have been applied on pre-service teachers in Turkey by Sevimli (2010). The questionnaire was developed by Finney and Schraw (2003), this questionnaire measuring the self-confidence of the pre-service teachers. The test was developed by Allen (2006) for testing the misconceptions in statistics among pre-service teachers. Hence my major research focus on understanding and self-efficacy levels of Palestinian pre-service mathematics teachers in statistics.

In the study, I discussed these two issues among Palestinian pre-service teachers in the mathematics field. As there are no previous studies handling those issues in Palestine. This gave me a motivation to measure the level of statistics teachers' self-efficacy and their statistical misconceptions to put the figure on the causes of these two problems. This, later on, might motivate researches for doing studies dealing with similar issues.

The study problem can be explained by the following questions:

1. What is the level of performance of Palestinian pre-service statistics teachers?
2. What is the level of self-efficacy of Palestinian pre-service statistics teacher?
3. What are the common mistakes that the Palestinian pre-service teachers have?
4. What is the relationship between students' statistical self-efficacy and their performance in statistics?

1.2. Importance of the Study

Since mathematics is one of the most important parts of our academic life. Whatever we study, whatever we measure, we cannot accomplish our work without using mathematical tasks. Because of the importance of mathematics and its role in other sciences, we have to be sure about students' level in mathematics in order to qualify them to apply mathematics in a way to make a proper interpretation of the data in a field of study. Because of the lack of studies concern with measuring the level of mathematics pre-service teachers, I chose my topic to measure the self-efficacy beliefs and conceptual knowledge in statistics among those teachers to be a first study done in this field in Palestine and to be a base and a reference for future studies to make proper diagnosis for this defect in the educational system in Palestine.

1.3. Theoretical Framework

The first step of self-efficacy started with Bandura (1977) in his social learning theory which renamed social cognitive theory (1986). Self-efficacy defined as how someone motivates himself to do a job or to achieve such a level of performance under a certain circumstances. Self-efficacy is a scale to measure how long someone can persevere to achieve a planned target (Bandura, 1994, 1989).Self efficacy is a measurement of self-knowledge and self-evaluating references (Marsh & Shavelson, 1985).In other words self-efficacy is "self-esteem reactions" when posing a question to measure one's success, such as "How good are you in English?"(Zimmerman,2000). Self- efficacy is to be ready to do the required activities and well measured skills to get over a problem or to accomplish a task (Langenfeld, Thomas & Pajares, 1993; Pajares & Kranzler, 1995; Robert et.al. 2001; in. Gülev, 2008; Zimmerman courtesy, 1995; in ılgar Cosgun, 2004).

Bandura defined self-efficacy as "People's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391). According to Schunk (1991), self-efficacy is a person's judgment of his/her ability to perform a given skill. Efficacy is a process in which persons weigh and combine the tasks that he/she has to accomplish and is a way in which the learner judge his/her efforts and ability to achieve his educational or daily goals (Schunk, 1989).

According to social learning theory, self-efficacy can be described as an individual's understanding of the skills that he/she can offer to the group to which he/she belongs. Social learning theory describes the person as a part of a group he/she learns from them through observation and imitation, so the role of an individual can be described as one of a total of roles affect each other and gain from each other. According to this theory to have a self-efficacy is to understand your role and skills that you can offer to your group (Ormrod, J.E., 1999).

According to self-concept theory, person's success or failure can be described as the way that this person views himself and his/her relationships with others. This theory describes one's self-efficacy as an organized, acquired and dynamic skills in which a person express his/her existence and believe himself to get success and avoid failure throughout his /her life (McAdam, 1986).

Self- efficacy is to be ready to do the required activities and well measured skills to get over a problem or to accomplish a task (Langenfeld, Thomas & Pajares, 1993; Pajares & Kranzler, 1995; Robert et.al., 2001; in. Gülev, 2008; Zimmerman courtesy, 1995; in Coşgun & Ilgar, 2004).

According to Marsh and Shavelson (1985) Self-efficacy can be defined as a scale to measure how long someone can persevere to achieve a planned target. Bandura (1994) defined self-efficacy as how someone motivates himself to do a job or to achieve such a level of performance under a certain circumstances. According to Zimmerman (2000) self-efficacy is "self-esteem reactions" when posing a question to measure one's success.

Schunk (1989) claimed that self-efficacy is not the only influence on behavior. Since behavior is a function of many variables, some other important variables such as skills, outcome expectations, and the perceived value of outcomes, are all of them are important in achievement settings.

Bandura (2000), discussed that self-efficacy emerges from four sources: the first source is one's experiences which allow the person to learn how to be successful and how to avoid failure reasons, the learner will learn that success needs effort. The second source is other's experiences by seeing other similar individuals's performance. A person can strongly influenced by social persuasion which enhances the self-efficacy and lead the person to try hard to get the required result, which is the third source of personal efficacy. The fourth source of self-efficacy is the psychological status of the person which leads him in some cases to judge himself as failure or successful according to his/her mood.

Pajares and Miller (1995) stated that self-efficacy in classroom emerges from three elements: self-efficacy in solving mathematic problems, self-efficacy of being successful in classroom mathematic, and self-efficacy in using daily math calculations properly. So, The one with high sense of self-efficacy is different in his/her way of thinking, emotional response and selecting activities from other people who has less self-efficacy (Çoban & Sanalan, 2002; in. Gülev, 2008). A person who has self-efficacy has the ability to rearranges the surrounding circumstances to get success in spite of the social, psychological or physical difficulties that may face him.

Pintrich and De Groot (1990) noted that self-efficacy relates positively to motivation to employ learning strategies. According to Zimmerman and Martinez-Pons (1990), Efficacy related positively to reported strategy use across domains.

Instructional studies have substantiated the idea that teaching students to use strategies raises self-efficacy and achievement. Schunk and Gunn (1986) showed that modeled strategies enhance self-efficacy and motivation during mathematics instruction.

Ford (1992) claimed that personal motivation or to receive motivation from others is an important aspect to motivate internal self-efficacy to achieve its benefits. The efforts that someone provides to activate his self-efficacy can be an important part to get the benefits of this feature.

Bandura (1986, 1997) noted that one could reinforce his self-efficacy by comparing himself with others who achieved the same or similar activities successfully. A person can expect his/her outcomes by noticing and monitoring the success, failure, rewards or punishments of others, so he can be able to have initial idea about his/her results.

According to social cognitive theory, knowledge acquisition or learning of students directly correlated to the observation of their teachers in the classroom. For that, self-efficacy among teachers is very important in the teaching system as students trust the information delivered by the teacher. Also, well understanding of conceptions from the teachers avoids the students falling in the misconception.

According to attribution theory which focuses on how someone attributes events and how these attributions affect his/her self-perception. This theory describes a person's own perceptions or attributions as to why they succeeded or failed at an activity determine the amount of effort the person will engage in activities in the future. Attribution theory describes self-efficacy as a positive expectancy for events in the future in a way he/she controls his/her feelings even he/she fails to achieve the goals of his/her task and he/she will not consider that as a shame or humiliation (Heider & Fritz,1958).

Self-efficacy is not the only influence on behavior. Since behavior is a function of many variables, some other important variables such as skills, outcome expectations, and the perceived value of outcomes are important in achievement settings (Schunk, 1989). High self-efficacy will not produce the expected success from the learner with a lack in required skills (Bandura, 1989).

Salomon (1984) found that self-efficacy relates to mental effort. Students judge their efficacy for learning from television or from written text, watching a televised film or read the comparable text, judging the amount of mental effort necessary to learn. He found that self-efficacy correlates positively with mental effort and achievement. Meier, McCarthy, and Schmeck (1984) demonstrated that efficacy for writing relates to cognitive processing dimensions.

The one with high sense of self-efficacy is different in his/her way of thinking, emotional response and selecting activities from other people who has less self-efficacy (Çoban & Sanalan, 2002; in. Gülev, 2008). A person who has self-efficacy has the ability to rearranges the surrounding circumstances to get success in spite of the social, psychological or physical difficulties that may face him.

1.4. Operational Definitions

Self-efficacy: is the extent or strength of one's belief in his/her own ability to complete tasks and reach goals (Ormrod J.E., 2006). According to Bandura (2000), self-efficacy is the way in which a person motivates himself to achieve his/her goals. Self-efficacy reflects how someone uses the conditions around him positively to get over his/her obstacles and complete his/her tasks.

Misconceptions: is arrangement of ideas and concepts in a nonscientific way which allows the concepts to be mistaken and to take place of each other (Eryılmaz & Süremli, 2002). These misconceptions resulting from three resources: epistemological, psychological and pedagogical. He explained that epistemological misunderstanding related to student's knowledge about statistics, psychological misconceptions also depending on student's personality, but pedagogical misconceptions related to the teacher and how he teaches the concepts of statistics (Broussea, 1997).

1.5. Limitations of the Study

The participant limits represented by the pre-service mathematics teachers who are graduated or expected to graduate by the end of the academic year 2013-2014. The time limits represented by the years of 2013-2014. The limits resulted from that the study built to measure the level of mathematics teachers only among pre-service teachers also from the lack of studies concern with measuring the level of pre-service teachers in the same factors of my study and the lack of local references studying and discussing self-efficacy and common mistakes among Palestinian pre-service teachers.

II. LITERATURE REVIEW

This chapter contains a review of the literature about self-efficacy and how it could affect students and teachers in understanding mathematics especially in statistical topics. It also contains a literature review of the second topic of my study which is talking about conceptual understanding of statistics topics.

2.1. Statistics Self-Efficacy

Statistics is one of the most important parts of our academic life. Whatever we study, whatever we are measure, we cannot accomplish our studies without statistical tasks. This importance of statistics requires the learner to motivate himself to understand statistical concepts and skills. Like other sciences, to understand statistics, students and also teachers need the self-efficacy in order to get over statistical obstacles and to use statistical skills in its useful way.

Statistical self-efficacy can be explained as an individual's self confidence in his/her ability to understand statistical problems. According to Finney and Schraw (2003), although statistics self-efficacy is a part and shares some characteristics of mathematics self-efficacy, both of them are different enough to be a separate construct.

Recently, researchers developed many scales to measure academic self-efficacy (Bong, 1998; Finney & Schraw, 2003; Forester, Kahn & Hesson-McInnis, 2004; Holden, Barker, Meenaghan & Rosenberg, 1999; Schunk & Pajares, 2002; Silver, Smith, & Greene, 2001). Other two scales where also created. The first scale Environment Scale which developed by Gelso, Mallinckrodt, and Judge (1996) and the other one was Self-Efficacy scale which created by Holden, Barker and Rosenberg (1999).

Finney and Schraw (2003) developed two scales to measure self-efficacy. Statistics Self-Efficacy (CSSE) and Self-Efficacy to Learn Statistics (SELS), both of these scales measures the self-efficacy of undergraduate students after finishing their required statistical courses.

2.1.1. Effects of Self-Efficacy on the Learner

According to Bandura (1986), people regulate their level and distribution of effort in accordance with the effects they expect their actions to have. As a result, their behavior is better predicted from their beliefs than from the actual consequences of their actions.

In educational studies, it was agreed that the self-efficacy has a significant effect on the achievements and the reactions of the learners whether they are students or teachers. Moreover, the confidence of teachers motivates students and increases their confidence and their ability to achieve their educational goals. (Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2001).

Bandura (1994), claimed that people with high self-efficacy have a high ability to face the challenges and to do the possible skills to get over them rather than escaping, when they fail in doing a task they are quick to try again and take the lesson from their mistakes in order not to fall again. These people have less tension than others and they have the ability to control themselves under difficult conditions. They trend to join new projects because of their confidence in their abilities to accomplish them.

Self-efficacy can be affected by many factors: academic success, social achievements, sport activities, choice of profession and many others. According to Cockburn and Haydn (2004), a teacher gains his/her self-efficacy from daily classroom activities, which give him the chance to repeat a task many times and also give him the opportunity to achieve success with his/her students and other active individuals at school.

Many researches have agreed that years of experience, teaching level, gender, teacher's nationality and cultural beliefs of the teacher affect his/her self-efficacy (Klassen, Usher, & Bong, in press; Liu & Ramsey, 2008). Since Wolters and Daugherty (2007), claimed that self-efficacy increases with experience that gained from teaching years and skills, Bandura (1997) proposed that teacher's self-efficacy becomes stable once it's established.

Kooij, de Lange, Jansen and Dikkers (2008) noted that aging may adversely affect one's self-efficacy such that aging may be accompanied by a decline in health and decline in one's confidence in his/her ability to accomplish his/her duties which lead to a decrease in self-efficacy. From another point of view, Wolters and Daugherty (2007) explained that teachers in primary schools have more self-efficacy than teachers who deal with high grade levels.

In the same way that previous factors affect self-efficacy, self-efficacy also reflects on the achievements, the success, and the way of thinking of a person in his/her life. Once a learner established his/her self-efficacy, he/she will achieve more success, more progress and more satisfaction (Klassen & Chiu, 2010).

2.1.2. Studies Related to Self-Efficacy

In recent years there has become a huge trend toward the study of self-efficacy of the learner. As self-efficacy is one of the most effective factors that affect the learner in general and mathematics learner in particular, a lot of researches discussed and talked about self-efficacy. Many researches discussed the effects of self-efficacy on students' success (Denise & O'Neil, 1997; Malpass, et al., 1996; Sewell & George, 2000), other researchers measured the level of self-efficacy among mathematics students and teachers (Andersen, Grene & Loewen, 1988; Huinker & Madison, 1997; Watters & Ginns, 1995), and many other researches discussed and measured the self-efficacy of teachers while using technology and computer skills in mathematics education (Aşkar & Umay, 2001; Busch, 1995; Seferoğlu & Akbıyık, 2009).

1 - International Studies

The study of Miller and Pajares (1994): this study tested the predictive role of self-efficacy beliefs in mathematical problem solving. The study created in a public school in south west of USA. The participants of the study were 350 students of ninth grade students. To answer the research questions the researchers used "Writing Skills Self-Efficacy Scale" created by Shell and others (1989), they also used "Dally-Miller Writing Apprehensive Test" which developed by Dally and Miller (1975), and finally they used "Self-Esteem Scale" by Marsh (1986). According to the researchers, this study supports Bandura's 1986 hypotheses related to the role of self-efficacy contained in his social cognitive theory. The results of this study can be summarized as the following:

a) Gender and prior experience influenced self-concept, perceived usefulness and problem solving.

b) Men had higher performance, self-efficacy and self-concept, in the same time they have lower anxiety.

c) Gender had a direct effect only on self-efficacy and prior experience variables.

The purpose of the study of Pajares and Kranzler (1995): was to test Bandura's (1986) hypotheses which related to self-efficacy in mathematics problem-solving. To achieve the goal of this study, the researcher used Roman's Advanced Progressive Matrices to measure the general mental ability for students. He also used The Mathematics Confidence scale (MSC) which created by Dawling in order to measure the self-efficacy of his sample individuals. And in order to test math anxiety among the participants the researcher used The Mathematics Anxiety Scale which created by Betz. The results of this study can be explained as following:

a) Students' self-efficacy beliefs about their math capability had strong direct effects on math anxiety and on mathematics problem-solving performance even when general mental ability was controlled.

b) The impact of anxiety was a primarily a result of noncausal covariation largely due to the self-efficacy.

c) The high school students were more over confident about their mathematical capabilities than college undergraduate students.

d) There were no gender differences in mathematics performances at the high school level.

e) There were no differences between boys and girls in math self-efficacy but they differed in math anxiety.

The study of John R. Malpass and others (1996) discussed self-regulated learning, self-efficacy worry and learning goal orientation. In this study the researcher focused on the effects of these factors on students' success in mathematics. The sample of this study consists of 144 students, the researcher used sex and M-SAT "the Mathematics-

Scholastic Achievement test" as control variables, he also used Self-Regulation Questionnaire created by O'Neil and others as an instrument to measure the self-report for students. The researcher followed Path analyzing technique to get the following results:

a) Self-regulation was negatively related to worry, and not related to either prior or post mathematics achievement.

b) Self-efficacy mediates the relationship between prior and post mathematics achievement, is related to self-regulation, and is highly and negatively related to worry.

c) Learning goal orientation is positively related to self-regulation and worry, and is not related to self-efficacy or Advanced Placement mathematics achievement. The Math-Scholastic Achievement Test is related to Advanced Placement math achievement.

d) Worry is negatively related to Advanced Placement mathematics achievement.

e) Boys were less worried and had higher self-efficacy than girls.

The study of Schunk (1996) contrasted self-efficacy with related constructs such as perceived control, outcome expectations, perceived value of outcomes, attributions, and self-concept. He also discussed some efficacy research relevant to academic motivation. In this study the researcher studied the effects of goal setting and information processing as person variables on self-efficacy, he discussed also the effects of situation variables such as models, attributional feedback, and rewards on student's self-efficacy. In conjunction with this discussion, he mentioned substantive issues that need to be addressed in the self-efficacy research and summarized evidence on the utility of self-efficacy for predicting motivational outcomes. He also suggested areas for future research.

The study of Hanlon and Schneider (1999) discussed the results of a pilot intervention designed to improve students' mathematics proficiency through self-efficacy training. The sample of the study participated in a five-week summer program that included whole class instruction, small group tutoring, and individual meetings with instructional coordinators. The data from the self-efficacy training intervention were then analyzed using a hierarchical linear model approach. Using hierarchical data analysis approach this study proved that over time, students' achievement scores on a math proficiency exam improved significantly, as did their confidence levels about passing this exam.

The study of Finney and Schraw (2003) aimed to develop two measures of current statistics self-efficacy (CSSE) and self-efficacy to learn statistics (SELS). The study also aimed to address whether statistics self-efficacy is related to statistics performance, and whether self-efficacy for statistics increases during an introductory statistics course. The study included a total of 154 students finished a section of six sections of an introductory statistical methods course. The researchers used their own questionnaires which they developed in this study, Test Anxiety Inventory (TAI; Spielberger, 1980), Statistics test anxiety (modification of Spielberger, 1980), Survey of attitudes toward statistics (SATS; Schau, Stevens, Dauphinee, & Del Vecchio, 1995). Mathematics self-efficacy scale-revised (Kranzler & Pajares, 1997; Pajares & Miller, 1995). In order to analyze the data and to measure the reliability of the questionnaires the researchers used two ways of analysis: descriptive analysis and factor analysis. The results of this study can be summarized as following:

a) The CSSE and SELS were related positively to math self-efficacy and attitudes towards statistics
b) The CSSE and SELS also were related negatively to anxiety.

2 - Studies In The Turkish Literature

The study of Aşkar and Umay (2001) investigated the computer self-efficacy of the freshman, sophomore and junior in the division of Elementary Mathematics Teaching. The sample of this study was the under graduated students of Hacettepe University with a total of 155 student. During the data gathering process, only juniors completed a single "Computer Literacy" course. In this study two instruments were used: a computer self-efficacy scale and a questionnaire for collecting demographic information including experience in computers, access to computers and frequencies of using computers. The result of this study showed that the perceived computer self-efficacy of the students is low and it increases with the access and computer experience.

The study of Umay (2002) aimed to measure the self-efficacy of undergraduate students during the period in which education faculties in Turkey were reconstructed and a new training system became standardized throughout Turkey. The sample of this study was mathematics undergraduate students in Hacettepe University. The researcher created his own instrument to measure students' self-efficacy after measuring the reliability and validity of this instrument. The analyses of the collected data resulted that senior students mathematics self-efficacy is significantly higher than freshman students.

The study of Kesici, Erdoğan and Şahin (2010) examined whether the score of the students in motivation and social comparison predict his mathematics self-efficacy. The sample of this study was the eighth grade students with a total of 173 students. The researchers used the Achievement Motivation Scale which was developed by Umay (2002) in order to measure the achievement motivation of students expected to have a high achievement motivation. They used Social Comparison Scale which developed by Gilbert, Allan and Trend (1991) in order to measure how students perceive themselves when they compare themselves with others. They also used Self-efficacy Scale Toward Mathematics which developed by Umay (2001) to measure self-concept, awareness in behaviors concerning mathematics subjects, and ability to transform mathematics efficacy into life skills. To answer the research questions the researchers used stepwise regression method and the analyses resulted that both social comparison and achievement motivation predicted "mathematics self-efficacy" while achievement motivation predicted "students' awareness of their own behaviors concerning mathematics subjects" and "ability to transform mathematics efficacy into life skills".

3 - Studies In The Palestinian Literature

Searching in the database of An-Najah, Al-Quds, Beirzeit and Al-Quds Open Universities in Palestine resulted that there no studies about self-efficacy in general and mathematics self-efficacy in particular. The majority of the studies are about the reasons of low performance of students in mathematics classroom. Self-efficacy emerged in one study as one of these reasons.: in the study of Hammad and Habbash (2005) the researchers used a questionnaire consists of 46 questions distributed into 4 study fields: students, teachers, curriculum, and students' parents in a public school in Palestine.

They obtained the following findings as a result of the study as causes of the low performance of mathematics students:

a) The high number of students in the same classroom.
b) No cooperation from the parents.

The study of Shawashra (2007) is not about self-efficacy directly but about a related concept. He discussed the effects of motivation on a student's success in mathematics. The participant of this study was a mathematics student who has been monitored during an educational semester. Along this period the student followed a psychological program to increase his motivation in mathematics class. The quantitative and qualitative analysis yielded the following findings:

a) The student's motivation increased with the prepared program.
b) Teaching system in Palestine should be developed to pay attention to the students' psychology.

There are two other studies in the whole database that is about mathematics learning. Firs is the study of Jabr (2007) which was about the impact of using computer skills on students' performance in mathematics class. Second study is the study of Abed (2009) which discussed the impact of training on problem-solving strategies.

2.2. Statistics Misconceptions

Recently, since 1970s there has been a noticeable trend toward studying people's understanding of statistics. These researches included individual's understanding of sampling, randomness, chance and misconceptions (Kahneman, Slovic, & Tversky, 1982; Konold, 1991; Nisbett & Ross, 1980; Shaughnessy, 2003).

Better understanding of statistics concepts lead students to understand the skills and tasks of statistics. Understanding of concepts enables the students to be able to link statistical concepts and distinguish them more clearly (Noordman & Vonk, 1998). For a proper understanding of the concepts of statistics we must order these concepts in the form of coherent structures of information in order to make it easy to be understood by learners. (Chi et al. 1981; Wyman and Randel 1998; Kintsch 1998).

Broussea (1997) noted that students' misconceptions resulting from three resources: epistemological, psychological and pedagogical. He explained that epistemological misunderstanding related to student's knowledge about statistics, psychological misconceptions also depending on student's personality, but pedagogical misconceptions related to the teacher and how he/she teaches the concepts of statistics.

Teachers play a main role in students' misunderstanding of concepts. Teachers, who are a major part in pedagogical system in statistical education, have a strong effect on the students to accept the ideas and concepts in a simple way according to students' level of thinking and understanding. For example, when the teacher has a misconception in a statistical concept and he/she explains it to the students in the same misunderstanding that he/she has, students consequently will have the same misunderstanding.

The incorrect techniques of some teachers in classroom have a direct impact on students' understanding. For example as a teacher to be the only reference for students in classroom will limit their activity and creativity skills. Moreover, the statistical language that the teachers use to deliver an idea to the students must be clear and in a simple scientific way to be easy to be understood.

According to what we have mentioned above about the resources of misconceptions in statistics we can categorize the misconceptions of students in statistics into five main categories:

- Misconceptions in calculations; since statistics depends on calculations which is the way to achieve the goal of the statistical operation and to apply the statistical concepts and skills, the nit is very important to avoid calculating misconceptions which may affect the result of the statistical analysis (Gal, 2000; Chance, 1997).
- Misconceptions in the statistical formulas; a formula is an entity constructed using the symbols and formation rules of a given logical language. For students to get the formula it should be explained well for him from where it came and the ideas beyond that formula. For that, teachers should start with those ideas leading the students to discover the formula, which makes it of sense for teachers to give the formula by the end of the subject not in the beginning (Utts, 1996).
- Misconceptions in communications and expressions; it is very important to find an easy way to communicate in the classroom between teacher and student,

which means to find a language for both to communicate and exchange the ideas and concepts in a statistical discussion. For example, when a teacher wants to deliver an idea or a concept in the class it is better to start with a simple example as an entrance to the main subject then to explain the concept and the concepts which are related to that concept and then to show the places where they are going to use calculations.

- Misconceptions about the sample size; the size of the sample is important in statistical analysis and it should give a picture of the whole space from where it was taken and reflects the characteristics of the space in order to achieve these goals of the sample we should give a lot of care in choosing the members of the sample in a way to be fair to judge the whole space from that sample (Karasar, 2005; Kaptan, 1998).
- Misconceptions in the hypothesis tests; one of the most important factors in any statistical study is the hypothesis test which allow us to understand the results of our study and to prove the validity of our hypothesis. Most of the students don't understand the meanings of some tests that they have applied in their study even they know how to do these tests properly and this related to their misunderstanding of the concepts associated with that test which is a result of accumulated defects in the teaching system of mathematics and statistics (Haller & Krauss, 2002; in. Sotos et al.., 2007).

According to Sotos and his colleagues (2007), statistics misconceptions can be classified into three main categories; misconceptions related to sampling, misconceptions in applying and understanding the hypotheses test and misconceptions related to confidence intervals as shown in table (1)

Table 1: Statistics Misconceptions

<u>Sampling distributions:</u>	
➢ The law of small numbers and sampling variability	-Neglect the effect of sample size on the variance of the sample mean. -Belief in the law of small numbers.
➢ The different distributions	-Confuse the population and the sampling distributions. -Confuse the sample and the sampling distributions.
➢ The central limit theorem	-Belief that the larger the sample size, the closer any distribution is to the Normal. -Inability to justify the use of the central limit theorem and the Normal distribution. -Confusion between the theoretical and the approximated Normal.

Hypotheses test:	
➢ Approaches to hypotheses testing	-Neglect the parallelism between hypotheses test and decision process.
➢ Definition of hypotheses	-Confusion in the definition of null and alternative hypotheses. -Confusion of the null hypothesis and the acceptance region. -Believing that a hypothesis can refer both to a population and a sample.
➢ The conditional nature of significance	-Inverse the conditional of the p-value. -Inverse the conditional of the significance level. -Interpreting the significance level as the probability of one hypothesis. -Interpreting the significance level as the probability of making a mistake. -Interpreting the p-value as the probability that the event happened by chance.
➢ Interpretation of the numerical value of the p-value	-Interpreting the numeric value of the p-value as strength of treatment.
➢ Nature of hypotheses tests	-Considering the test as a mathematical proof. Illusion of probabilistic proof by contradiction.
➢ Evaluation of statistical significance	-Confuse practical and statistical significance.

Confidence intervals :	-Plausible values for the sample mean -Range of individual scores. -Range of individual scores within one standard deviation. -The width of a confidence interval increases with sample size. -The width of a confidence interval is not affected by sample size. -A 90% confidence interval is wider than a 95% confidence interval.

(Sotos, Van hoof, Noortgate, Onghena,2007)

2.2.1. Studies Related to Statistical Misconceptions

Recently, lots of studies and researches discussed mathematics students' misconceptions in statistical topics. Chance, delMas, and Garfield (2004), discussed about students' misconceptions in sample distribution. Cumming, Nordin, Horton and Reynolds (2006), talked about students' misconceptions in confidence interval. Mevarech and Kramarsky (1997), Steinberg and Morris (2001), Liu, Lin, and Tsai,. (2008), Bainbridge, Lasley and Sundre (2003), have discussed about students' misconceptions in statistical correlation.

1 - International Studies

The study of Mevarech and Kramarsky (1997): this study aimed to discuss students' conceptions and misconceptions relating to the construction of graphs. The sample of this study was 92 eight grade students (44 girls and 48 boys), who selected randomly from two Israeli schools) in Israel. Students were tested before and after being exposed to formal instruction on graphing. The researchers used test re-test technique and qualitative and quantitative analysis and they found that the following misconceptions were displayed by less than 10% of the subjects: conceiving a generalized, stereotypic idea of a graph, using arrows or stairs to represent the direction of the covariation, and connecting the ticks on the axes by lines or curves.

The study of Liu and others (2008): this study discussed students' misconceptions in correlation. The study aimed to elucidate the misconceptions held by senior high school students about correlation, to discuss the possible causes of these misconceptions, and to examine the effectiveness, advantages, and limitations of the adopted concept mapping using an interviewing technique for identifying student misconceptions. The sample of this study was twenty-five grade-12 students (13 females, 12 males) from a high schools located in Taiwan. The researchers used concept mapping through interviewing technique was used to collect and analyze data in order to get the following results:

a) Seven misconceptions about correlation were found. From these seven misconceptions, five were newly discovered by this study, while the other two are similar to those found by previous studies
b) Four major factors related to the development of misconceptions about correlation were identified: learning materials, language, daily-life experiences, and existing mathematical concepts.
c) The concept mapping through the interviewing technique adopted in this study was effective in detecting misconceptions about statistics.

2 - Studies in the Turkish Literature

The purpose of the study of Büyüköztürk (2000) was to assess impact of using the SPSS applications on statistics learning. The sample of this study was 20 undergraduate students and they used a scale for assessing attitudes toward statistics. After analyzing data the researcher resulted that the use of the SPSS application and computer skills enhance students' learning in statistics.

The purpose of the study of Doğan (2009) was to test if computer-based teaching of statistics has significant influence on students' achievement levels and their attitudes toward the statistics course. The sample of this study was 71 students from undergraduate students who passed the introduction to statistics course and also passed the computational statistics course. The results of this study showed that use of computers skills, software, internet and programs significantly enhances students' achievement and attitude.

3 - Studies in the Palestinian Literature

The search of the database of An-Najah, Al-Quds, Beirzeit and Al-Quds Open Universities of the Palestinian Universities and a global internet search there are no studies can be found which were about the statistical misconceptions of mathematics students. On except of this is the Abu El-Khair (1990) study in which he investigated the students' misconceptions in the set theory and concluded that students have a lot of misconceptions related to the concepts of set theory.

III. METHODOLOGY

Scientific research is a body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge. To achieve this purpose it is necessary to follow the right approach for scientific research starting from defining the research question and forming the hypotheses finishing with collecting and analyzing the data. This chapter explains the research design, data collection method, population, sample, study instruments and the statistical analysis methods that were applied to get the results of the study.

3.1. Research Design

The research design is a detailed outline of how an investigation will take place. It is the way in how data is to be collected, what instruments will be employed, how the instruments will be used and the intended means for analyzing data collected. In any research, choosing the suitable research design is very important step in order to be able answer the research questions of the study.

3.1.1. Quantitative Paradigm

A research may be classified by tools it uses, by the paradigm it sets itself with and by the form of inquiry it employs. I begin with the paradigm. In this research a quantitative paradigm was chosen. The positivist approach, which is reflection of the quantitative research, is based on the positive verification of the results that gained from the scientific experiences. this approach does not depend on election or intuition. since the aim of this study is to measure the levels of the conceptual understanding and the level of the self-efficacy of pre-service teachers by using a test and and questionnaire. As result, the quantitative method was deemed appropriate as to reach the aim of the study.

3.1.2 Form of Inquiry

Secondly, I will mention the form of inquiry. A commonly used classification distinguishes between exploratory, descriptive, and explanatory (Robson, 1993, p.42). An explanatory purpose for an enquiry seeks an explanation of situation or problem, usually in the form of a casual relationship. A descriptive purpose for an enquiry, on the other hand, portrays an accurate profile of persons, events or situations. A particular study may be concerned with more than one purpose. as a result? descriptive and explanatory purposes predominate in this study.

3.1.3 Reasearch Startegy

When designing an inquiry, it is also important to have a strategy. Strategy refers to the general broad orientation taken in the addressing research question. The strategies which the researcher selects in carrying out research depends on the type of research questions that are posed. The data collection methods or techniques are usually regarded as necessarily linked to particular research strategies. According to Robson (1993, p.40) one simple approach which is widely used distinguishes between three main strategies; experiments, surveys and case studies. A survey is collection of information in standardized form from a specific population, usually by means of questionnaire or interview. A survey strategy was chosen to collect the data for this study. Typical features of surveys that are linked to our study as are following:

- The selection of samples of individuals from known populations,
- The collection of relatively small amount of data in standardised form from each individual,
- Usually employs questionnaire or structured interviews,
- There is normally no attempt tp manipulate variables, or control conditions.

In this study the survey was the data collection method. This method is used to gather quantitative data. While there are types and forms of surveys, they have one common feature: any survey normally involves administering questions to individuals. A survey is a study that obtains data from a subset of a population, in order to estimate population attributes.

As a field of applied statistics, survey methodology studies the sampling of individual units from a population and the associated survey data collection techniques, such as questionnaire construction and methods for improving the number and accuracy of responses to survey. In my study I used two instruments a self-efficacy inventory and a statistics concept inventory with which i was able to collect data from the sample of the study. This method of using survey instruments questionnaires to make statistical inferences about the population being cited as under the survey method of collecting data.

3.2. The Population and the Sample

The population of this study was the mathematics pre-service teacher in the Palestinian universities. Since my study aims to measure the level of self-efficacy of Palestinian pre-service teachers and to discover the mistakes that they have in statistics topics, then the population of the study was the undergraduate students who are in their last year of studying in the Palestinian universities.

To choose the sample of the study the systematic sampling method was used. The randomly selected sample was a subset of a statistical population in which each member of the subset has an equal probability of being chosen.

In order not to be bias in choosing the sample of the study, students were chosen from two Palestinian universities according to their identity card, the students who were selected are the students who had an odd number in the last digit of their identity card. This method allows us to select students randomly and each student had the equal probability of being chosen.

The sample of this study was the pre-service teachers who finished their statistical courses. The sample of the study consists of two parts: first part consists of randomly selected 40 pre-service mathematics teachers as individuals to the initial test. These individuals were selecting from two Palestinian universities ; Al-Quds and Al-Quds Open Universities. Both of the self-efficacy questionnaire and the Statistics Concept Inventory (SCI) test were applied on this sample in order to examine the validity and reliability of the questionnaire and the (sCI) test.

The purpose of the current study is to measure the current self-efficacy levels of Palestinian pre-service teachers using the questionnaire which developed by Finney and Scraw (2003) and to detect the mistakes of these pre-service teachers by employing the Statistics Concept Inventory (SCI) test which developed by Allen (2006).

The second part of the sample consists of another 100 pre-service mathematical teachers who were also randomly selected from the same universities. These teachers were examined using the self-efficacy questionnaire and (SCI) test in order to measure their level of self-efficacy and their statistical knowledge.

3.3. Data Collection Tools

In this study I used an inventory developed by Finney and Schraw (2003) in order to measure the current statistics self-efficacy (CSSE) of the pre-service teachers. I also used The Statistics Concept Inventory (SCI) test which was developed by Allen (2006) in order to detect participants' mistakes in statistics.

The following table summarizes the factors that the study aims to measure and the instrument that used to collect the data of each factor

Table 2: Study Instruments

The factor to be measured	The used instrument
Self-efficacy	current statistics self-efficacy (CSSE) questionnaire
Conceptual knowledge	Statistics Concept Inventory (SCI)

3.3.1. The Statistics Concept Inventory (SCI)

The Statistics Concept Inventory (SCI) has developed by Allen (2006) in order to assess students' conceptual understanding of topics typically encountered in an introductory statistics course. The test consists of 25 multiple choice questions with only one correct answer for each question. The subjects of the test can be divided into four main subjects as shown in table 3. The table also mentioned the questions related to each subject.

Table 3: The Subjects of the (SCI)

The subject	The factors related to the subject
Probability	2, 12, 21, 22, 23.
Descriptive	4, 5, 7, 8, 9, 14, 17, 19, 25.
Inferential	1, 6, 10, 11, 13, 18, 24.
Graphical	3, 15, 16, 20.

The four main subjects of SCI as mentioned in the previous table are: probability, descriptive statistics, inferential statistics, and graphical representation. To be more specific about the SCI subjects, the questions of the SCI can be classified as the following table:

Table 4: The Subjects ot the Items (SCI)

Subjects of the SCI	Items related to the subject
Data Summary and Presentation	
- Importance of data summary	5, 7, 8, 9
- Methods of displaying data	3,16, 20
- Percentiles and quartiles	4, 7
- Measure of variability	17, 19
- Skewness and kurtosis	16
- Stem-and-leaf diagram	3

- Frequency distribution and histogram	3, 16, 20
Probability	
-Sample space and events	12, 23
- Interpretation of probability	12
- Conditional probability	21
- Multiplication and total probability rules	12
- Independence	21
Random variables	
- Expected values	23
Discrete Probability Distributions	
-Discrete uniform distribution	16, 20
- Binomial distribution	2
Continuous probability distribution	
-Normal distribution	14, 20, 22
Parameter Estimation	
-Random sampling	18
-Central Limit Theorem	2, 11, 13, 23, 24
Linear Regression	
-Correlation	15, 25
Confidence Intervals and Hypothesis Testing	
-Inference on the mean of a population	6, 10, 11, 13, 24
-Inference on the means of two normal	

populations -Paired comparisons	11 1

3.3.1.1. The Reliability of the (SCI)

A reliable measuring instrument is the instrument which gives the same measurements and same results when applying it repeatedly on the same unchanged objects. By repeating the test using the same reliable instrument a person expected to get the same score if he/she completes it at two different points of time this way of measuring the reliability called (test-retest) reliability (Carmines & Zeller, 1979).

According to Pearson (1880), cited in Kendall (1955), by applying the test twice or many times if the test is reliable, and the subjects have not changed from time to another, then we should get a high value of correlation coefficient (r). A researcher would be satisfied if the value of r were at least .70 for a research instrument.

Spearman (1904), cited in Daniels (1944), in his "classical measurement theory" he claimed that If a measuring instrument were perfectly reliable, then it would have a perfect positive ($r = +1$) correlation with the true scores. If you measured an object or event twice, and the true scores did not change, then you would get the same measurement both times.

According to Revelle and Zinbarg (2009) the simplest way to measure the reliability of the instrument is to split the data into two halves and calculating the score of the participant based on each half of the scale. As a result, if our instrument is reliable, then the score of the participant on one half will be the same or very similar to his score on the other half. Consequently, across the whole participants, the score of the two halves of the questionnaire should correlate highly. The high correlation between the two halves when using the split half method means that our instrument is reliable.

Cronbach (1951), cited in Cerny & Kaiser (1977), came up with a measure to split the data in every possible way and computing the correlation coefficient for each split. The average of these values is equivalent to Cronbach's alpha, α, which is the most common way to measure the reliability of a measurement scale.

According to Cortina (1993), a researcher should be cautious and follow general instructions while applying his/her scale because Cronbach's α depends on the number of items that on the scale. Kline (1999) notes that the acceptable value of α is 0.8 for cognitive research such as intelligence tests, but for the ability tests a 0.7 or even less can be accepted because of the diversity of constructs being measured. In general we can accept a value of α when it is at least about 0.6- 0.7 to be sure that our scale is reliable.

In this study Cronbach's α technique was used, and the value of Cronbach's α was high with a value of (.914), as shown in Table (5)

Table 5: Relaibility Analysis for the (SCI)

No. of Items	Cronbach's Alpha
25	.914

Source: Developed based on SPSS analysis

3.3.1.2. The Validity of the (SCI)

A measure is valid if it measures what it purports to measure (Creswell & Miller, 2000). A researcher needs to ensure that the questions he or she asks in the instrument related to the construct that he/she intend to measure.

The reliability of the factors emerging from the factor analysis depends on the size of the sample. Since correlation coefficient fluctuated from sample to another, much more in small samples than in large, so there is agreement that there should be more subjects than variables. Gorsuch (1983) proposed that for each analysis there should be at least 100 individual. According to Bryman and Cramer (1997), since the main purpose of a study is to find out what factors underlie a group of variable, it is essential that the sample should be sufficiently large enough to enable this to be done reliably.

The object of factor analysis is to decide the factors to keep in and the factors to exclude from the scale. There are two main criteria that are used to decide which factors to keep and which factors to exclude. The first criterion is Kaiser's criterion, in which we select those factors which have an eigenvalue of greater than one. The second method is scree test, which proposed by Cattell (1966), cited in Revelle & Zinbarg (2009), in this method, a graph drawn of the descending variance accounted for by the factors initially extracted.

Hutcheson and Sofroniou (1999) mentioned that the Kaiser-Mayer-Olkin (KMO) statistic varies between 0 and 1. A value of 0 indicates that the sum of partial correlations is large relative to the sum of correlations. While a value of 1 indicates that patterns of correlation are relatively compact and so factor analysis should yield distinct and reliable factors.

According to Kaiser (1974), the values which are acceptable are the values greater than 0.5 and for the values below this should lead the researcher to either collect more data or to rethink which variables to include. Furthermore, Hutcheson and Sofroniou (1999) mentioned that values between 0.5 and 0.7are mediocre, values between 0.7 and 0.8 are good, values between 0.8 and 0.9 are great and values above 0.9 are superb. The factors analysis of my study sample using the SPSS program as shown in the table 3, shows that the value of the KMO analysis for the whole scale is 0.682 which is an acceptable value for the test to be a valid test and which means not to exclude any of the factors.

Bartlett's measure tests the null hypothesis that the original correlation matrix is an identity matrix. For factor analysis to work, there must be some relationship among variables. In this context, if the R-matrix were an identity matrix then all correlation coefficients would be zero. Therefore, this test must be significant which means to have significance value less than 0.5. A significant test means that the R-matrix is not an identity matrix; therefore, there is a relationship between the variables which lead the researcher to include these variables. The factor analysis of the studying sample by using the SPSS program as shown in table 5, showed that the Bartlett's test is highly significant ($p < .001$), and therefore factor analysis is appropriate to be sure about the validity of the data

Table 6: Kaiser-Mayer-Olkin and Bartlett'e test for the (SCI)

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.682
Bartlett's Test of Sphericity	Approx. Chi-Square	786.7937
	Df	300
	Sig.	.000

Source: Developed based on SPSS analysis

The KMO test showed that there is a relationship among the items of the SCI test, this result of the KMO indicated that the factor analysis method will be an efficient method to measure the validity of the data we get from SCI test.

Table 7 forms one of the products of the factor analysis for the data of the SCI test. The table shows the communalities before and after extraction. The initial assumption of the analysis assumes that all variance is common; therefore, all of variances are 1 before extraction. The other column of the table, which labeled Extraction, shows the common variance among variables. As an example, for the item number 1, for thıs this item 90.8% of the variance associated with this item is common variance.

Table 7: Communalities Table of the items of the (SCI)

Item no.	Initial	Extraction
1	1.000	.908
2	1.000	.901
3	1.000	.773
4	1.000	.797
5	1.000	.674
6	1.000	.778
7	1.000	.895
8	1.000	.844
9	1.000	.846
10	1.000	.839

11	1.000	.887
12	1.000	.840
13	1.000	.875
14	1.000	.882
15	1.000	.820
16	1.000	.911
17	1.000	.886
18	1.000	.915
19	1.000	.955
20	1.000	.902
21	1.000	.868
22	1.000	.901
23	1.000	.925
24	1.000	.925
25	1.000	.899

Source: Developed based on SPSS analysis

Table 8, which is also one of the outputs of the factor analysis of the data we get from the application of the SCI. the table, lists the eigenvalues that associated with each item before extraction, after extraction and after rotation. As shown in the table the first six items have eigenvalues greater than one. The first item explains 38,390% of the total variance while item number 25 explains -2.023E-15 of the total variance. The second part of the table which labeled Extraction Sums of Squared Loadings, lists the items after extraction, the extraction of items leaves us with six items which are the items with eigenvalues greater than one. The last part, which labeled Rotation Sums of Squared Loadings, shows the eigenvalues of the items after rotation. By comparing the results obtained from this part with the results obtained in the first part of this table, we can notice that the first item accounted for 38.390% which is extremely more than the other five items (17.083%, 12.502%, 9.237%, 5.181%, 4.194%, 3.086%, 2.288%, 1.676%). While after rotation, which reduced the difference among variances, the first item accounted for 26.892% of the variance compared to (15.225%, 12.897%, 12.004%, 11.601%, 7.967%).

Table 8: Total Variance Explained for the (SCI) Test

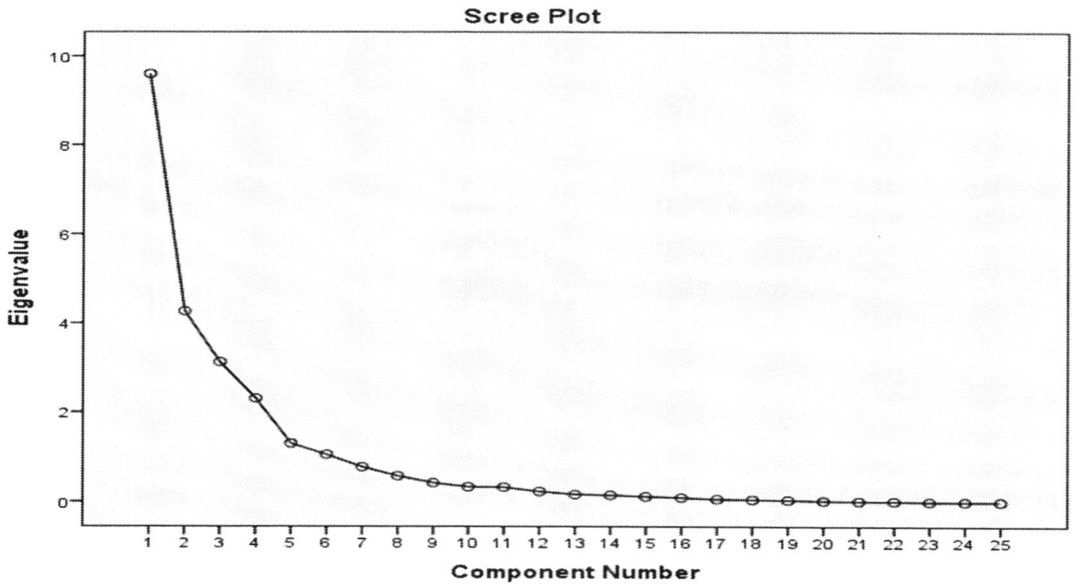
Item no.	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.598	38.390	38.390	9.598	38.390	38.390	6.723	26.892	26.892
2	4.271	17.083	55.474	4.271	17.083	55.474	3.806	15.225	42.118
3	3.125	12.502	67.975	3.125	12.502	67.975	3.224	12.897	55.015
4	2.309	9.237	77.212	2.309	9.237	77.212	3.001	12.004	67.018
5	1.295	5.181	82.393	1.295	5.181	82.393	2.900	11.601	78.619
6	1.048	4.194	86.587	1.048	4.194	86.587	1.992	7.967	86.587
7	.771	3.086	89.672						
8	.572	2.288	91.960						
9	.419	1.676	93.637						
10	.328	1.313	94.950						
11	.323	1.293	96.243						
12	.229	.917	97.160						
13	.163	.652	97.812						
14	.145	.581	98.393						
15	.116	.465	98.858						
16	.091	.363	99.221						
17	.060	.239	99.460						
18	.048	.193	99.653						
19	.038	.153	99.806						

20	.023	.091	99.897						
21	.013	.052	99.949						
22	.012	.048	99.997						
23	.001	.003	100.000						
24	1.015E-16	4.058E-16	100.000						
25	-5.057E-16	-2.023E-15	100.000						

Source: Developed based on SPSS analysis

Figure 1 shows the scree plot which resulted also from the factor analysis of the data of the SCI, the figure shows that the curve of the scree plot began to tail after the sixth factor, so this point is considered to be the point of inflexion on the curve. The figure also shows that the eigenvalue of the factors started to take a value less than 1 after the sixth factor.

Figure 1: The Scree Plot Of the SCI Test



Source: Developed based on SPSS analysis

Table 9 shows the rotated component matrix for the items of the SCI test. This output, which resulted also from the factor analysis for the data, figures out the factors that loading for each item. As shown in the table items number 17, 13, 12, 6, 15, 11, 19, 14 and 23 substantially loaded on factor number 1. Items number 20 and 24 loaded on factor 2 and so on for the other items as I shaded the factors that the items loaded most strongly on.

Table 9: The Rotated Component Matrix for the Items of the (SCI) Test

Item no. \ Factor no.	1	2	3	4	5	6
17	.871	.187	.252		-.136	
13	.856			.268	-.123	.213
12	.834		.128	.259		.234
6	.831	-.161		.246		
15	.798	.250	.302	.106		.134
11	.732		.216	.523		.166
19	.714	-.384	-.447	.137	.277	
14	.607	.557	-.292	.315	-.120	
20	.388	.848		-.118		-.130
2		-.833		-.135	.410	.107
24	-.223	.776		.165		.491
4	.154		.833	.270		
7		-.301	.759	.183	.290	.322
16	.508	.450	.622		.157	.195
18	.130	.468	.613	.101	-.517	.161
10	.388		.264	.738	-.266	
9	.447	.284		.714	.165	.138
8	.372	.276	.358	.704		
5	.131			.660	-.364	.282
23	.395			.230	-.845	
21	.538	-.150		.135	.729	
1		-.462	.346	-.320	.678	
3	.321	-.235	.117	.176		.750
22		.355	.419	.149	-.388	.650
25	.464	.302	.396		.364	.550

Source: Developed based on SPSS analysis

3.3.2. Current Statistics Self-Efficacy Questionnaire (CSSE)

The (CSSE) questionnaire has developed by Sara J. Finney and Gregory Schraw (2003) in order to address whether statistics self-efficacy is related to statistics performance, and whether self-efficacy for statistics increases during an introductory statistics course. The questionnaire yielded reliable and the analysis showed that the CSSE was related positively to math self-efficacy and attitudes towards statistics, but related negatively to anxiety.

In the questionnaire, for all items a six-point Likert scale ranging from 1 (no confidence at all) to 6 (complete confidence) was used. Those points were arranged as following: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence.

The items of the questionnaire were created using the specific terminology of the domain which similar to other specific measures of self-efficacy, such as the mathematics self-efficacy scale revised (MSES-R) which developed by Kranzler and Pajares (1997). To measure the reliability of the questionnaire Finney and Schraw calculated Cronbach's Alpha which had a value of .91.

3.3.2.1. The Reliability of the (CSSE) Questionnaire

Again to measure the reliability of the (CSSE) questionnaire of my study sample I used Cronbach's α technique, and the value of Cronbach's α was (0.809), as shown in table 10. Which means that the questionnaire is reliable since Kline (1999) notes that the acceptable value of α is 0.8 for cognitive research such as intelligence tests, but for the ability tests a 0.7 or even less can be accepted because of the diversity of constructs being measured. In general we can accept a value of α when it is at least 0.6- 0.7 to be sure that our scale is reliable.

Table 10: Reliability Analysis for the (CSSE)

No. of Items	Cronbach's Alpha
14	0.809

Source: Developed based on SPSS analysis

3.3.2.2 The Validity Of The (CSSE) Questionnaire

Again to measure the validity of the (CSSE) questionnaire of my study sample I used the KMO and Bartlett's method. As shown in table 10, The results of the showed that the value of the KMO test was .551 which is a mediocre value since Kaiser (1974), noted that the values which are acceptable are the values greater than 0.5 and for the values below this should lead the researcher to either collect more data or to rethink which variables to include. Furthermore, Hutcheson and Sofroniou (1999) mentioned that values between 0.5 and 0.7are mediocre, values between 0.7 and 0.8 are good, values between 0.8 and 0.9 are great and values above 0.9 are superb.

The analysis also showed that the Bartlett's test is highly significant ($p < .001$) as shown in the table 10. Again this result of the KMO indicated that the factor analysis method will be an efficient method to measure the validity of the data we get from CSSE questionnaire.

Table 11: KMO and Bartlett's test for the (CSSE) Questionnaire

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.511
Bartlett's Test of Sphericity	Approx. Chi-Square	180.889
	Df	91
	Sig.	.000

Source: Developed based on SPSS analysis

Table 12 shows the communalities before and after extraction for the items of the CSSE questionnaire. The initial assumption of the analysis assumes that all variance is common; therefore, all of variances are 1 before extraction. The other column of the table which labeled Extraction shows the common variance among variables. As an example, for the item number 1 as shown in the table 90.8% of the variance associated with this item is common variance.

Table 12: Communalities table for the Items of the (CSSE) Questionnaire

Item no.	Initial	Extraction
1	1.000	.802
2	1.000	.751
3	1.000	.776
4	1.000	.594
5	1.000	.797
6	1.000	.633
7	1.000	.680
8	1.000	.748
9	1.000	.628
10	1.000	.839
11	1.000	.701
12	1.000	.794
13	1.000	.686
14	1.000	.621

Source: Developed based on SPSS analysis

Table 13 lists the eigenvalues that associated with each item before extraction, after extraction and after rotation. As shown in the table the first six items have eigenvalues greater than one. The first item explains 21.787% of the total variance while item number 14 explains .951% of the total variance. The second part of the table which labeled Extraction Sums of Squared Loadings, lists the items after extraction, the extraction of items leaves us with four items which are the items with eigenvalues greater than one. The last part which labeled Rotation Sums of Squared Loadings shows the eigenvalues of the items after rotation. By comparing the results obtained from this part with the results obtained in the first part of this table, we can notice that the first item accounted for 21,787% which is extremely more than the other four items (15.438% , 13.212% , 10.890% , 8.871%). While after rotation, which reduced the difference among variances, the first item accounted for 17.973% of the variance which is slightly more than the other four items (14.467% , 12.951% , 12.604% , 12.203%).

Table 13: Total Variance Explained for the (CSSE) Questionnaire

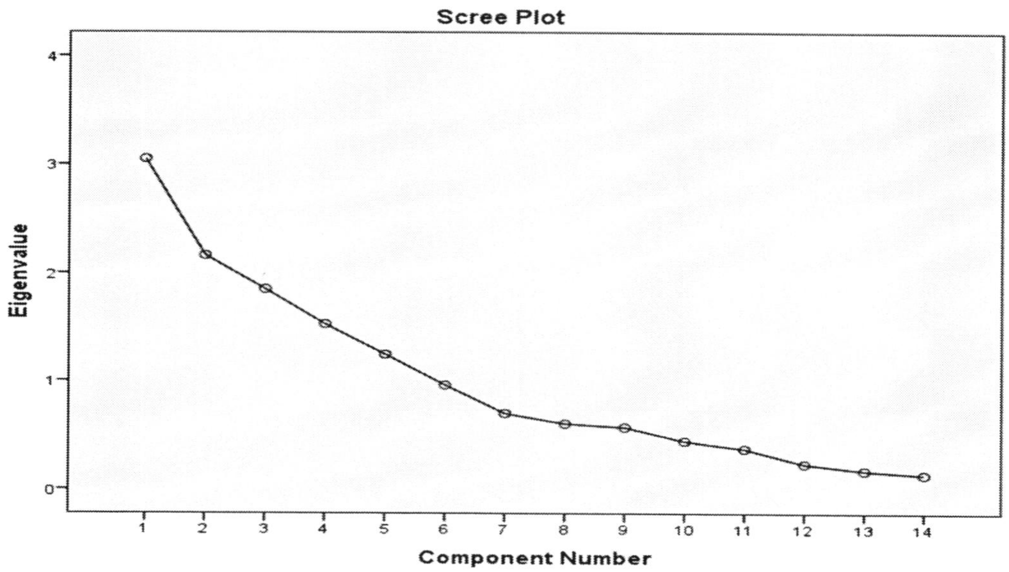
Item no.	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.050	21.787	21.787	3.050	21.787	21.787	2.516	17.973	17.973
2	2.161	15.438	37.224	2.161	15.438	37.224	2.025	14.467	32.440
3	1.850	13.212	50.437	1.850	13.212	50.437	1.813	12.951	45.391
4	1.525	10.890	61.327	1.525	10.890	61.327	1.765	12.604	57.995
5	1.242	8.871	70.199	1.242	8.871	70.199	1.708	12.203	70.199
6	.958	6.842	77.040						
7	.697	4.979	82.019						
8	.602	4.298	86.317						
9	.569	4.063	90.380						
10	.442	3.156	93.536						
11	.371	2.648	96.184						
12	.231	1.653	97.837						
13	.170	1.211	99.049						
14	.133	.951	100.000						

Source: Developed based on SPSS analysis

Figure 2 shows the scree plot which resulted also from the factor analysis of the data of the CSSE Questionnaire, the figure shows that the curve of the scree plot began to tail after the sixth factor, so this point is considered to be the point of inflexion on the curve.

The figure also shows that the eigenvalue of the factors started to take a value less than 1 after the fifth factor.

Figure 2: The Scree Plot Of the CSSE Questionnaire



Source: Developed based on SPSS analysis

Table 14 shows the rotated component matrix for the items of the CSSE questionnaire. This output, which resulted also from the factor analysis for the data, figures out the factors that loading for each item. As shown in the table items number 12, 13, 14, 6, 10 and 7 substantially loaded on factor number 1. Items number 5, 1 and 4 loaded on factor 2 and so on for the other items as I shaded the factors that the items loaded most strongly on.

Table 14: the Rotated Component Matrix for the Items of the (SSE) Questionnaire

Item no. \ Factor no.	1	2	3	4	5
12	.864	.176			-.111
13	.815				-.117
14	.697	-.138	.145		.300
5		.868	-.110	.147	
1	.152	.840		-.143	.226
6	.357	-.566	-.106	.217	.355
3	.221		.811		.247
10	.198		-.745		.144
2	.432		.633	.352	.189
8	-.201	.185		.807	.121
4		.315		-.628	-.301
7	.364	-.204	.315	.574	-.276
11			.171		.815
9				-.411	-.671

Source: Developed based on SPSS analysis

IV. STUDY RESULTS

This chapter summarizes the results of statistical analysis of the study factors in order to answer the research questions that have been addressed in the first chapter of the study. Correlational descriptive and predictive forms of statistical analysis were used to analyze the data. Below I will present the findings in order of research questions and I will start with the first research question.

4.1. Testing the Research Questions

4.1.1 What is the level of performance of Palestinian pre-service statistics teachers?

To answer this research question which aims to measure the performance of level of mathematics teachers in statistics. Descriptive analysis techniques were used to analyze the responses of the participants who were involved in the study. For each question the means and the standard deviation scores for the true and false answers were calculated.

Since SCI consists of 25 multiple-choice questions with only one correct answer for each question, so the highest score of the exam can be 25 and the lowest score can be zero. The result of the analysis as showed in table 15 shows that the item number 1 and the item number 16 had the highest frequency of true answers with a frequency of 78. Relatively, those two items obtained the highest value of mean with a value of 0.78.

Table 15: Descriptive Analysis of (SCI)

Item no.	Min	Max	Mean	SD	True	%	False	%
1	.00	1.00	.78	.42	78	78	22	22
2	.00	1.00	.54	.50	54	54	46	46
3	.00	1.00	.60	.49	60	60	40	40
4	.00	1.00	.55	.50	55	55	45	45
5	.00	1.00	.77	.42	77	77	23	23
6	.00	1.00	.63	.49	63	63	37	37
7	.00	1.00	.29	.46	29	29	71	71
8	.00	1.00	.78	.42	78	78	22	22
9	.00	1.00	.70	.46	70	70	30	30
10	.00	1.00	.67	.47	67	67	33	33
11	.00	1.00	.58	.50	58	58	42	42
12	.00	1.00	.63	.49	63	63	37	37
13	.00	1.00	.50	.50	50	50	50	50
14	.00	1.00	.27	.45	27	27	73	73
15	.00	1.00	.44	.50	44	44	56	56
16	.00	1.00	.58	.50	58	58	42	42
17	.00	1.00	.57	.50	57	57	43	43
18	.00	1.00	.63	.49	63	63	37	37
19	.00	1.00	.33	.47	33	33	67	67
20	.00	1.00	.34	.48	34	34	66	66
21	.00	1.00	.47	.50	47	47	53	53
22	.00	1.00	.33	.47	33	33	67	67
23	.00	1.00	.40	.49	40	40	60	60
24	.00	1.00	.47	.50	47	47	53	53
25	.00	1.00	.25	.43	25	25	75	75
Sum	.00	14.00	13.1	11.91	1310	52.4	1190	47.6

Source: Developed based on SPSS analysis

Performance of teacher candidates was also analyzed in regard to the four subjects of SCI which contains four main subjects: probability, descriptive statistics, inferential statistics, and graphical representation of data. After categorizing each item, the result of analysis as shown in table 16 shows that the probability part obtained 166 true answers with a percentage of 41.5% of the total answers of the probability part which was the lowest percentage among the four main categories. Graphical part follows the probability part with the second lowest percentage of true answers with a 196 true answers with a percentage of 49% of the total answers of the graphical part. Descriptive part took the third place with 451 true answers of the total answers of the descriptive part with a percentage of 50.1%. And the inferential part obtained the highest percentage of true answers with a value 426 true answers of the total answers of inferential part with a percentage of 60.8%.

Table 16: Items Analysis of (SCI) Items According to Subject

The Subject	N	True	%	False	%
Probability	5*100	220	44	280	56
Descriptive	9*100	451	50.1	449	49.9
Inferential	7*100	426	60.8	274	39.2
Graphical	4*100	196	49	204	51

4.1.2. What is the level of self-efficacy of Palestinian pre-service statistics teacher?

To investigate this research question a descriptive analysis was applied which in involves the calculation of frequencies and percentages. The questionnaire consists of 14 items for all items a six-point Likert scale ranging from 1 (no confidence at all) to 6 (complete confidence) was used. Those points were arranged as follows: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence.

The result of the analysis as showed in table 17 shows that item number 4 had the lowest mean among the items with a value of 2.35. On the contrast, item number 6 obtained the highest mean with a value of 4.27. According to the given data in the table, the total score for the mean has been calculated and it was about 2.79. And also the total score for the standard deviation and it took a value of about 1.04.

Table 17: Descriptive Analysis of (CSSE) Questionnaire

Item no.	Min	Max	Mean	SD
1	1	6	2.58	1.13
2	1	5	2.41	1.16
3	1	5	2.61	.87
4	1	5	2.35	1.18
5	1	5	2.57	1.08
6	1	6	4.27	1.13
7	1	5	2.60	1.12
8	1	6	2.86	1.17
9	1	6	2.83	1.23
10	1	6	2.87	1.36
11	1	4	2.81	.84
12	1	5	2.74	.93
13	1	6	3.00	.31
14	1	5	2.54	.99

Source: Developed based on SPSS analysis

The questionnaire consists of 14 items for all items a six-point Likert scale ranging from 1 (no confidence at all) to 6 (complete confidence) .The results showed that the first point (no confidence at all) had a frequency of (186) with a percentage of (13.3%) of the total answers. And the sixth point (complete confidence) had a frequency of (30) with a percentage of (2.2%) of the total answers which was the lowest percentage as shown in table 18. The table also shows that the second point (a little confidence) obtained the highest frequency with a value of (441) and a percentage of (31.5%) of the total answers.

Table 18: Results of (CSSE) Questionnaire Analysis

	1 (no confidence)	2	3	4	5	6 (complete confidence)
Frequency	186	441	411	237	95	30
%	13.3	31.5	29.3	16.9	6.8	2.2

As the questionnaire consists of 14 items for all items a six-point Likert scale ranging from 1 to 6, so the highest score that expected to be obtained could be 84 and the lowest score could be 14. But in this study sample the highest obtained score was 47 which is a mediocre score, and the lowest score was 28 which is a low score. Table 18 shows frequencies and percentages for participants' scores separated into intervals ranging from 14 to 84. Scores between 14 and 37 are considered to be low, scores between 38 and 61 are considered to be moderate, and scores between 62 and 84 are considered to be high.

Table 19: Participant's Scores Evaluation in (CSSE) Questionnaire

The interval	The evaluation	F	%
14 – 37	Low	35	35
38 – 61	Moderate	65	65
62 – 84	High	00	00

Table 19 shows that 35% of the students' scores are ranging between 14 and 37, and the remaining scores are ranging between 38 and 61 with a percentage of 65%. The results also show that there are no scores more than 62 which mean that the Palestinian pre-service teachers' self-efficacy is moderate.

4.1.3 what are the common mistakes that the pre-service teachers have?

The SCI consists of 25 multiple choice questions with only one correct answer for each question, to address the last research question an analysis for each question has been done and the results showed that the participants had different mistakes in different topics. In my study we considered that there is a common misconception related to the subject of an item if more than half of the participants had answered the item wrongly. According to that the most repeated mistakes were sorted by the items' numbers and the mistakes related to those items as shown in the table below:

Table 20: Common Mistakes Done by Staistics Pre-Service Teachers

Item no.	Frequency of false answers	Percentage of false answers	Detected Mitakes
1	22	22%	-Students couldn't differentiate between the t-test and z-test. -Students used the two-sample z-test when they are presented with a question of two products and are required to state which product fullfils a certain condition.
2	46	46%	-Students expected that when incrasing the sample size the probability of an event from that sample will increase. -Students expected that when incrasing the sample size the probability of an event from that sample doesn't change.
3	40	40%	-Students didn't recognize when to apply and how to differentiate between three methods of displaying data: histogram, cumulative frequency, steam and leaf.

4	45	45%	-Students had mistakes related to the concept of the percentile. -Students defined the percentile as a measurement that shows the percent of the total frequency scored are equal to the given value of percentile. -Students defined the percentile as a measurement that shows the percent of the total frequency scored are higher than the given value of percentile.
5	23	23%	-Students had mistakes related to the basics knowledge about the median and the relationship between the median and the items of the series. -Students expected that without changing the median, the items of a series couldn't be changed.
6	37	37%	-Students had mistakes related to the hypothesis testing and how to differentiate between the null hypothesis and the alternative hypothesis. -Students didn't recognize that the null hypothesis is what it is attempted to overturn by the hypothesis test.
7	71	71%	-Students showed mistakes in the concepts of the third quartile and the concept of the outlier and the relationship between these two measurements. -Students expected that the range is the least impact measurement by the extreme outliers. -Students expected that the mean is the least impact measurement by the extreme outliers. -Students expected that the variance is the least impact measurement by the extreme outliers.
8	22	22%	-Students supposed that when calculating the mean for a series we don't have to conclude all the items of that series.
9	30	30%	-Students had mistakes related to the central tendency measures. -Students chose the mean as the most accurately measure to describe the central tendency of a series containing variables ranging from very large variables and variables are very small. -Students chose the standard deviation as the most accurately measure to describe the central

			tendency of a series containing variables ranging from very large variables and variables are very small.
10	33	33%	-Students supposed that for an experiment a 95% confidence interval for the mean means that 95% of the measurements can be considered valid. -Students supposed that for an experiment a 95% confidence interval for the mean means that 95% of the measurements will be between the upper and lower limits of the confidence interval.
11	42	42%	-Students have mistakes about the concepts of standard deviation. -Students supposed that when calculating the mean of a given set it is not important to take the sample size in consideration.
12	37	37%	-Students showed mistakes related to the union of two events (AUB), and they didn't recognize that the union of two events (AUB) occurs if either A or B or both occurs.
13	50	50%	-When performing the same significance test on two samples and these two samples have the same mean and the same standard deviation with a p-value of the first test smaller than the p-value of the second test, students expected that the sample size of the second test is larger than the sample size of the first test.
14	73	73%	-Students expected that the "Age in years of college freshmen" is a statistic could have a normal distribution.
15	56	56%	-Students supposed that the negative value of the correlation indicates that relationship between the variables is considered to be weak. -Students supposed that a correlation coefficient could have a value more than 1.
16	42	42%	-Students supposed to use the normal distribution as a method of displaying data when they are presented with a series with a finite number of values and these values are equally likely to be observed. - Students supposed to use the skewed distribution as a method of displaying data when

			they are presented with a series with a finite number of values and these values are equally likely to be observed. Students supposed to use the bimodal distribution as a method of displaying data when they are presented with a series with a finite number of values and these values are equally likely to be observed.
17	43	43%	-Students supposed that the zero value of the standard deviation of a set means that half of the numbers of that set are above the mean. -Students supposed that the zero value of the standard deviation of a set means that all of the numbers in the set are zero. -Students supposed that the zero value of the standard deviation of a set means that the numbers in the set are evenly spaced on both sides of the mean.
18	37	37%	-Students had mistakes related to the random sampling methods and how to choose a sampling method that will not introduce bias.
19	67	67%	-Students supposed that the negative value of correlation of a set indicates that most of the measurements of the set were negative. -Students supposed that the negative value of correlation of a set indicates that all of the measurements of the set are less than the mean. -Students supposed that the negative value of correlation of a set indicates that all of the measurements of the set were negative.
20	66	66%	-Students didn't have the ability to get a sense of variability in a statistical set by looking at its histogram.
21	53	53%	-Students showed mistakes related to the concept of the conditional probability.
22	67	67%	-Students faced difficulties while dealing with the concepts related to standard deviation and they didn't recognize that standard deviation measures the amount of variation or dispersion from the average.

23	60	60%	-Students supposed that in the case of dealing with small value of the probability of an event, by increasing the number of trials, then the probability value of the event would increase.
24	53	53%	-In the case of dealing with two samples have the same standard deviation and the confidence level is fixed, students expected that the confidence interval for the larger sample would be wider. -In the case of dealing with two samples have the same standard deviation and the confidence level is fixed, students expected that the confidence interval for the smaller sample is equal to the confidence interval for the larger sample.
25	75	75%	-Students supposed that the negative value of correlation means that the relationship between the variables is considered to be weak relationship.

The analysis of the items according to their subjects also took a place. The SCI contains four main subjects: probability, descriptive, inferential, and graphical. After categorizing each item into its subject the results showed that the percentage of true answers of probability part was 18.09% of the total true answers of the exam, the descriptive part formed 34.43%, the inferential part 32.52% and the graphical part 14.96% of the total true answers of the exam.

After extraction the mistakes related to the topics related to each item the results can be summarized as following:

Probability: this part consists of the second, twelfth, twenty first, twenty second and twenty third items. The second item discusses about the binomial distribution especially about the relationship between the sample size and a probability of an event from that sample. A percentage of 54% of the students answered this question correctly, and they expected that when incrasing the sample size the probability of an event from that sample will decrase.

Question 12 showed mistakes related to the union of two events (AUB). To answer this question properly, students need to know that the union of two events (AUB) occurs if either A or B or both occurs. A percentage of 63% of the participants answered this question true.

The item 21, which discusses about the conditional probability, showed mistakes related to the concept of the conditional probability. A percentage of 47% of participants found out that in the conditional probability a probability of an even depends on the occurrence of the previous event and then they answered the question properly.

Related to the results of the item 22, a percentage of 67% of the participants didn't give a correct answer. In this question the majority of students faced difficulties while dealing with the concepts related to standard deviation and they didn't recognize that standard deviation measures the amount of variation or dispersion from the average, which led students to answer the question wrongly.

Question number 23 indicated mistakes related to sample space and events, it also showed mistakes related to expected value and central limit theorem. A percentage of 40% of the students answered this question properly.

Descriptive: this part consists of the fourth, fifth, seventh, eighth, ninth, fourteenth, seventeenth, nineteenth and the twenty fifth items. To answer the fourth question correctly, a participant need to have enough knowledge about the concept of the percentile which is a measurement that shows the percent of the total frequency scored are below the given value of percentile. A percentage of about 55% of the students had the enough knowledge to answer the question correctly.

The fifth question requires students to have the basics knowledge about the median and the relationship between the median and the items of the series. To answer this question properly, students need to know that whatever will be the values of the items of a series, the median is the measurement that separates the higher half of the data from the lower half. A percentage of 77% of the students had the required information to answer this question correctly.

The seventh question showed mistakes in the concepts of of the third quartile and the concept of the outlier and the relationship between these two measurements. Since the outlier is the value which diverges greately overall the collected values, so the third quartile will not impact of the value of the outlier since the third quartile is the middle value between the median ad the highest value of the set of data. A percentage of 29% of the participants answered this question properly.

In the eighth and ninth questions, students in general had good knowledge about the topics related to those two questions. In this case, a percentage of about 78% of the students answered question number eight correctly and they had the enough knowledge about the calculating the mean of a series and a percentage of 70% of the students answered the ninth question properly which means they had the enough knowledge about the central tendency measures. Those high percentages of true answers show that the majority of students have enough knowledge about summary statistics since both of the questions required good information in the topics related to summary statistics.

Question 14 indicated mistakes related to the definition and the applications of normal distribution a percentage of 27% of the students answered this question correctly. This low percentage of true answers reflects that students have mistaken knowledge about the normal distribution.

Question 17 requires students to have the knowledge about the standard deviation which is a measure of the variation of a data set and also they need to know that a low value of standard deviation indicates that the values of the data set are very close to the mean. The participants in this question had mistakes related to these concepts of the standard deviation especially when the value of the standard deviation is zero and they didn't recognize that the zero value of the standard deviation indicates that all of the items of the series are equal. A percentage of 57% of the participants answered this question correctly.

Question 19 also had mistakes in the concepts related to standard deviation. Since the standard deviation measures the variation of a data set, so it couldn't be a minus value. A percentage of 33% of the students answered this question correctly which means that the majority of the students have mistakes in the definition and the calculation of standard deviation.

Question 25 was the last question of the descriptive part. This question aimes to measure student's knowledge about the correlation. While the correlation measures the dependence of a variable on another. Students in this question showed mistakes related to the concept of correlation especially while dealing with a negative value of correlation and they suppose that the negative value means a low value of standard deviation and they didn't recognize that the negative value means an inverse relationship. A percentage of about 25% of students answered this question correctly.

Inferential: this part consists of the first, sixth, tenth, eleventh, thirteenth, eighteenth and twenty fourth questions. In the first question which is talking about statistical hypothesis tests which are: t-test, z-test. This question examines the knowledge of the students about differentiation and usage of these tests. A percentage of 78% of the students had the knowledge to differentiate between the tests and their usage and they recognize choosing the paired comparison t-test when they are presented with a question of two products and are required to state which product fulfills a certain condition.

To answer the sixth question, students need to have enough knowledge about the hypothesis testing and how to differentiate between the null hypothesis and the alternative hypothesis. In this case students need to be able to recognize that the null hypothesis reflects that there will be no observed effect for the experiment and the null hypothesis is what it is attempted to overturn by the hypothesis test. About 63% of the students reflected a good knowledge about the hypothesis testing and they were able to differentiate between both of the null and alternative hypotheses.

Question number ten requires students to have the knowledge about the concepts related to the confidence interval of the mean. Students were expected to know that, as mentioned in the question, for an experiment a 95% confidence interval for the mean means that 95% of the time, the experiment will produce an interval that contains the population mean. About 67% of the students answered the question correctly which means that they have correct knowledge about this topic.

Since question 11 is talking about standard deviation, then students again need to have knowledge about the concepts of standard deviation and they need to know that standard deviation measures the amount of variation from the average. The students who had this knowledge they were able to answer the question correctly, while 42% of the students didn't had that ability and they answered the question wrongly.

Question 13 is talking about the p-value and its relationship with the sample size while fixing the mean and the standard deviation. Half of the students had the enough knowledge about this relationship and they recognized that, according to central limit theorem, when increasing the sample size the test will give better results which means larger p-value.

Question 18 is talking about random sampling. To answer this question correctly, students need to know that when selecting a sample randomly, each member of the sample has an equal chance of being selected and the probability of the member to be selected is unaffected by the selection of other members. The participants also need to know that the selected sample should contain subjects with characteristics similar to population as a whole in order not to introduce bias. About 63% of the students answered this question correctly while the rest had mistakes related to random sampling.

Question 24 is discussing about the relationship between the confidence interval and the sample size. In this case, students expected to know that when increasing the sample size the confidence interval becomes narrower. About 47% of the students answered the question correctly.

Graphical: this part consists of the third, fifteenth, sixteenth and twentieth questions. Question number three is talking about three methods of displaying data: histogram, cumulative frequency, steam and leaf. Students in this question are required to differentiate between these three methods of displaying data and to know how to apply them on a set of data. About 60% of the students answered this question correctly.

The fifteenth question discusses about the correlation between two variables when displaying these variables in a scatter plot. In this question students are required decide whether there is a relationship exists between two sets of data when representing these data in a scatter plot. About 44% of the students in this question had mistakes in determining the correlation between the variable and they had also misunderstanding of the negative value of the correlation.

Question 16 is also talking about the methods of displaying data: normal distribution, discrete uniform distribution, skewed distribution and bimodal distribution. Students in this question are required to differentiate between these methods of displaying data and how they can display these methods on a set of data. About 58% of the students differentiate between these methods and correct the question properly.

Question 20 is the last question of this part. This question is talking about the variability in a histogram. Since it is possible to get a sense of variability in a statistical set by looking at its histogram, then students need to have the enough knowledge about this subject in order to answer the question correctly. About 34% of the students answered this question correctly, which means that the majority of the students have mistakes related to the variability in histograms.

4.1.4. What is the relationship between students' statistical self-efficacy and their performance in statistics?

In order to answer this research question I did a comparison between students' answers in the CSSE questionnaire and their performance in the CSI in the same field of knowledge.

The analysis of students' answers in the CSSE questionnaire has showed that the level of self-efficacy of pre-service teachers was low to moderate. This result leads us to expect that the performance of the students in the SCI will be weak.

After making the analysis of students' answers in the SCI, the analysis showed unexpected results especially in the inferential part of the exam which showed a high percentage of correct answers (about 60.8%), and which was the highest percentage of true answers among the exams fields.

Inferential statistics is concerned with making predictions or inferences about a population from observations and analyses of a sample. That is, a researcher can take the results of an analysis using a sample and can generalize it to the larger population that the sample represents. In order to do this, however, it is imperative that the sample is representative of the group to which it is being generalized.

If we take the items number two and number four of the CSSE into account, these two items had the lowest mean between the questionnaire items. The second item which is measuring student's self-efficacy in interpreting the p-value from a statistical procedure had a mean of a value of 2.41 which was the second lowest mean between the items. If we look at the corresponding question in SCI which is question number thirteen that is talking about the p-value and its relationship with the sample size while fixing the mean and the standard deviation, half of the students answered this question true, which means that the low self-efficacy that students have has no correlation with student's performance in the statistics exam.

Similarly, the fourth item of the CSSE which had the lowest mean with a value of 2.35, the item has measured students' efficacy in selecting the correct procedure to be used to answer a research question. This item measures if a student have the enough self-efficacy in his statistical knowledge about statistical hypothesis tests such as: t-test, z-test, paired comparison t-test and other statistical tests and also measures his self-efficacy to differentiate between those tests and the fields in which they can be applied. By looking at the corresponding items in the SCI which are question number one and question number six, which are talking about the same topics we can notice that they have one of the highest percentages of true answers which means again that the self-efficacy of the students didn't influence their performance in the exam.

In my opinion after analyzing the correlation between self-efficacy and students' performance, a certain relation cannot govern the relationship between the two aspects and both of them don't influence each other.

The results of this study can be summarized as following:

1. According to the descriptive analysis of the Statistics Concept Inventory (SCI), the level of conceptual understanding of statistics of pre-service Palestinian teachers is moderate and they have mistakes related to the following topics:
    - Summary statistics.
    - Percentiles and quartiles.
    - Central limit theorem.
    - Inference on the mean of a population.
    - Normal distribution.
    - Correlation.
    - Measures of variability.
    - Methods of displaying data: discrete uniform distribution, normal distribution, frequency distribution and histograms.
    - Conditional probability and independence.
    - Sample space and events and expected value.
2. The level of current statistical self-efficacy of Palestinian pre-service teacher is low to moderate.
3. The relationship between the self-efficacy and student's performance cannot be governed by a certain relation and both of the them don't influence each other.

V. DISCUSSION AND RECOMMENDATIONS

5.1. Discussion

This study composed of three main parts. The first part which is talking about the conceptual understanding of statistics among pre-service teachers showed that the level of understanding of the teachers in moderate. By asking the participants 25 questions aimed to measure the level of understanding and misunderstanding of statistical topics, the answers of the participants contained different mistakes which reflect that the Palestinian pre-service teachers have common mistakes in different topics in statistics.

As my study is parallel to a Turkish study done by Sevimli (2010), the study showed that the Turkish students have mistakes in graphics and histograms, frequency distribution, the central limit theorem, the basics of probability and also in the concepts related to correlation.

By comparing the results of this study with the results of the study done by Sevimli (2010), both the Palestinian students and Turkish students showed mistakes related to the central limit theory and also in the continuous probability distribution especially in the normal distribution.

From both countries students also showed mistakes in the subjects of summary and presentation of statistical data since students showed low performance while dealing with percentiles and quartiles and also they had common mistakes related to histograms such as steam-and-Leaf diagram and the skewness and kurtosis.

Standard deviation also too a place as a common misconception since students from both countries showed mistakes related to the concept of standard deviation especially when dealing with the negative value of correlation.

The performance of the students in the questions related to the sample space and events, conditional probability, and independence indicated mistaken knowledge and low performance in the concepts related to probability.

Although both of the studies indicated many common mistakes among students from both groups of study, but there was some points of difference between the performance of the students in some subjects such as the hypothesis testing since the performance of Turkish students was lower when dealing with the statistics hypothesis tests: t-test and z-test, as they had difficulties in differentiating between the two tests and when to apply each test to a case of study.

By comparing the results of this study with the results of the study done by Mevarech and Kramarsky (1997), both of the studied showed that the performance of the students from both of the groups of study was the lowest comparing to the other parts of study.

The results of the study which done by Liu and others (2008) showed that students had mistakes related to the correlation and the concepts related to correlation which were common mistakes among the students from the group of this study and the Palestinian pre-service teachers.

This low performance of pre-service Palestinian teachers in the basics concepts in statistics due to the teaching system in the Palestinian universities as my study has been done in two different Palestinian universities and the participants were the mathematics students in their last year of study and since the courses that they had through their study years in these universities didn't give importance to the educational part compared to pure mathematical part, then students focus on high level topics in statistics at the expense of the basics which resulted by many mistakes in basic topics of statistics.

In the second part of this study in which the level of self-efficacy among pre-service Palestinian teachers was measured, the results showed that these pre-service teachers don't have enough self-confidence that makes them to be high-qualified teachers. Based on the data from the results of this study, it is possible to argue that even though the participants have low self-efficacy but they had good scores in the exam. It is possible that low of self-efficacy might have pushed these students to study harder in order to compensate the deficiency that they feel, which could be considered as a positive point of low self-efficacy among students.

Taking in consideration the study that has been done in Turkey by Sevimli (2010). The study showed results contradictory to the results of this study, as it showed that among

Turkish pre-service teachers the levels of self-efficacy was high while students' performance in the exam was weak. Which means that high self-efficacy sometimes gives negative feedback on the performance of the students as they trust themselves more than enough.

Comparing the results that we got from this study with the results of other studies, such as the study of Finney and Schraw (2003) from which they concluded that current statistics self-efficacy is positively related to students' performance. The study of Pajares and Kranzlar (1995) in which they showed that self-efficacy directly affects students' problem- solving performance, these results do not go in line with my results which showed that there is no relationship between self-efficacy and students' performance.

This study pointed out the performance of the pre-service Palestinian teachers in mathematics and the level of their self-efficacy. The study of these two aspects can form a basis from which we can start solving the problems that the pre-service teachers faced in the study of statistics. In order to solve these problems we need basically to point out the reasons which lead to that performance and identifying the factors that affecting students' achievements in statistics.

Starting with the demographic factors and ending with the individual factors, there are many issues affecting students' achievements in mathematics. Gender, socio-economic status and the level of parents' education all of them are demographic factors which may play a role in determining student's achievement in mathematics lesson.

The instructional factors also have a role in determining the level of student's performance in statistics education. One of those is the curriculum which is one of the most important factors that affect the achievements of the students in mathematics. From the results that we got in this study, which measured student's performance in Palestine, those results lead us to think about the Palestinian curriculum as a major element in the educational system.

By making a focus on the Palestinian mathematics book we can clearly find out that statistics occupy a limited part of the mathematics book which is not more than one small chapter starting from high school and extended to two courses for mathematical students in the university.

Besides the curriculum there are other instructional factors playing an important role in determining the educational level of mathematics students. Instructional strategies and methods of the student also play important rules. This means that the student should have the ability to understand and to improve the knowledge that is given in the class in order to achieve the main goal of education. Hence, I believe it is important that there is a system to orient the students in this direction.

Moreover teacher needs to have skills and knowledge to apply their mathematical knowledge in education. School environment is also one of those determining factors by the facilities that the school offers for its students such as their classrooms, libraries ...etc. In Palestine schools could not offer good environment for students as we are still under occupation and building infrastructure for a state.

On the other hand, the individual factors also are determining factors in the educational process. I believe it is important that a mathematics student has the ability to understand the concepts accurately and has the ability to make connections between the concepts, and this is a personal effort of the student which is determining his/her level of achievement. According to Knowles (1975), for the student to achieve the goal of educational process by his own effort he should first to diagnose his needs, formulating the goals that he looking forward to achieve, identifying resources for learning and finally evaluating the outcomes of the learning process.

The skills that the students use to understand the mathematical concepts and apply these concepts in solving real life problems are also important and play a major role in determining students' level in mathematical processes. the skills that the student use in understanding, analyzing and generalizing the mathematical concepts and the way that a student deal with the mathematical problems is also a personal effort that differentiate students from each other.

The teacher also plays a critical role in directing and motivating students to use their skills in solving mathematical problems and offering them the educational environment that they need to achieve the supposed educational goals from the educational process.

Taking the Palestinian student in consideration, their educational situation and the situation in Palestine in general we clearly can notify that those individual factors are variable among student but teachers in general haven't the attitude to direct and motivate students to achieve the required goals in their mathematics education.

Those factors could be determining factors of students' levels of education in the mathematics class and may be they had an effect on the results of this study. to investigate the role of these factors on the education system and achievements levels of students we need further studies to address this role.

5.2. Recommendations

Since this study focused on measuring pre-service teachers' levels of conceptual understanding of statistics and detecting the mistakes that they have in statistical topics, so this study can be a preliminary step in order to overcome the problem of the low self-efficacy levels of pre-service teachers and also to improve their performance in statistics.In this context, additional courses and lessons focusing on topics in which students faced difficulties should be provided to mathematical students at universities before joining schools as teachers which give them the enough knowledge which is very important for their students in order to protect them from falling into the same mistakes.

Training and applying what has been learned is also important. Students need to be enrolled in schools in order to be trained under the supervision of certified teachers in order to qualify them and to increase the level of their self-confidence as it is not high enough as the study showed.

Psychological support to students also plays an important role in promoting students' self-efficacy. In this case the lack of specialists and programs aims to support students' self-efficacy is one of the causes of low self-confidence among students. In order to get over this problem, the researcher suggests that students need to be enrolled into programs to promote their self-efficacy.

REFERENCES

Allen, K., Stone, A., Rhoad, T. R. & Murphy, T. J. (2004). The Statistics Concepts Inventory: Developing a Valid and Reliable Instrument. *Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition.* American Society for Engineering Education.

Allen, K. (2006). The Statistics Concept Inventory: Development And Analysis Of A Cognitive Assessment Instrument In Statistics. Unpublished PHD. The University of Oklahoma, USA.

Anderson R., Gerene, M., & Lowen P. (1988). Relashionships among Teahers' and students' Thinking Skills, Sense of Efficacy and Student Achievement. *Albert Journal of Educational Research, 34*(2), 148-165.

Aşkar, P., & Umay, A. (2001). İlköğretim Matematik Öğretmenliği Öğrencilerinin Bilgisayarla İlgili Öz Yeterlilik Algısı. *Hacettepe Üniversitesi, Eğitim Fakültesi Dergisi, 2*(1),1–8.

Bainbridge, W. L., Lasley, T. J., & Sundre, S. M. (2003). Policy initiatives to improve urban schools: An agenda. *Education and Urban Society 35*(3), 292-299.

Bandura, A. (1977). *Self-efficacy: Toward a Unifying Theory of Behavioral Change, Psychological Review 84(2), 191-215*

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control.* NY: W. H. Freeman and Company.

Bandura, A. (2001). Social Cognitive Theory: An Agentic. *Department of Psychology, Standford University, Annual Rev. Psychology* 52, 1–26.

Bandura, A., & Wood, R. (1989). Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology, 56*, 805-814.

Batanero, C., Godino, J.D., Vallecillos, A., Green, D., and Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics, Education, Science and Technology, 25*(4), 527 – 547.

Belia, S., Fidler, F., Williams, J., & Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods*, 10, 389–396.

Bong, M. M. (1998). Self-efficacy and Self-regulated Learning: The Implication of Research Relatedin Education Engineering. *Journal of Educational Technology,14*(1), 97-118.

Brewer, J. K. (1985). Behavioral statistics textbooks: Source of myths and misconceptions? *Journal of Educational Statistics, 10*(3), 252–268.

Brousseau, G., (1997). *Theory of didactical situations in mathematics* (N. Balacheff, M. Cooper, R. Southland, & V. Warfield (Eds. & Trans.)). Dordrecht, Boston: Kluwer Academic Publishers.

Bryman, A., & Cramer, D., (1997). *Quantitative Data Analysis with SPSS for Windows.* Routledge, London.

Busch, T. (1995). Gender difference in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research, 12*, 147-158.

Büyüköztürk, Ş. (2000). SPSS uygulamalı bilgisayar destekli istatistik öğretiminin istatistiğe yönelik tutumlara ve istatistik başarısına etkisi. *Eğitim Araştırmaları Dergisi, 1*, 13- 20.

Cansüngü Koray, Ö. & Bal, Ş. (2002). İlköğretim 5. Ve 6. sınıf öğrencilerinin ışık ve ışığın hızı ile ilgili yanlış kavramları ve bu kavramları oluşturma şekilleri. Gazi Üniversitesi Gazi Eğitim Fakültesi dergisi, 22(1), 1-11.

Carmines, E.G., Zeller, R.A., (1979). *Reliability and Validity Assessment.* Sage, Newbury Park.

Cerny, B. A., & Kaiser, H. F., (1977). "A study of a measure of sampling adequacy for factor-analytic correlation matrices." *Multivariate Behavioral Research 12*, 43-47.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981) Categorization and representation of physics problems by experts and novices, *Cognitive Science, 5*, 121-52

Cockburn, A. D., & Haydn, T. (2004). *Recruiting and retaining teachers: Understanding why teachers teach.* London, England: Routledge Falmer.

Cornu, B. (1991). Limits. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 153-166). Boston: Kluwer.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98–104.

Coşgun, S. & Ilgar, Z. M. (2004). Rehberlik ve Psikolojik Danışmanlık Deneyimi Çalışmalarının Adayların Öz Yeterlilik Algılarına Etkisi. *XIII. Ulusal Eğitim Bilimleri Kurultayı İnönü Üniversitesi Eğitim Fakültesi Malatya*, 1–14.

Cox, D. R. (1998), "Statistics for the Millennium, Some Remarks On Statistical Education || ", *The Statistician, 47(1):* 211-213.

Creswell, J. W. & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124-131.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika 22*(3), 297-334.

Cumming, J., Nordin, S. M., Horton, R., & Reynolds, S. (2006). Examining the direction of imagery and self-talk on dart-throwing performance and self-efficacy. T*he Sport Psychologist, 20*, 257-274.

Daniels H.E., (1944). The relation between measures of correlation in the universe of sample permutations. *Biometrika 33* (2): 129–135.

delMas, R., Garfield, J., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7*(3).

Denise H., & O'Neil, H.F. (1997). The role of parental expection effort, and self-efficacy in the achievement in the high and low track high school students in Taiwan. *Paper Presented at the Annual Meeting of the American Educational Research Association*, Chicago.

Doğan, N. (2009). Bilgisayar Destekli İstatistik Öğretiminin Başarıya ve istatistiğe Karşı Tutuma Etkisi. *Eğitim ve Bilim, 34*(154), 3–16.

Duchastel, P.C. (1974). Computer applications and instructional innovation: A case study in the teaching of statistics. *International Journal of Mathematics Education in Science and Technology, 5*, 607-616.

Eryılmaz, A. & Sürmeli. E. (2002), Üç AÇamalı Sorularla Öğrencilerin Isı ve Sıcaklık Konularındaki Kavram Yanılgılarının Ölçülmesi. 5. *Ulusal Fen Bilimleri ve Matematik Kongresi* (p.251–257). Ankara: ODTÜ.

Finney, S., & Schraw, G. (2003).Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28*, 161–186.

Ford ME. (1992). *Human Motivation: Goals, Emotions, and Personal Agency Beliefs.* Newbury Park, CA: Sage.

Forester, M., Kahn, J. H., & Hesson-McInnis, M. S. (2004). Factor structures of three measures of research self-efficacy. *Journal of Career Assessment, 12*(1), 3-16.

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. G. J. B. Garfield (Ed.). *The Assessment Challenge in Statistics Education* (pp. 37–51). Netherlands: IOS Press.

Garfield, J. (1992). How Students Learn Statistics. *International Statistical Review 63*(1), 25-34.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*, 44–63.

Geslo, C., Mallinckrodt, B., & Judge, A. (1996). Research training environment, attitudes toward research and research self-efficacy: The revised research training environment scale. *The Counseling Psycology, 24*, 304-322.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative methods. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Thousand Oaks, CA: Sage.

Gülev, D. (2008). *Biyoloji Öğretmen Adaylarının Biyoloji Konularındaki Kavram Yanılgıları, Biyoloji Öğretimine Yönelik Öz yeterlilik İnançları Ve Tutumları.* YayımlanmamıĢ yüksek lisans tezi. Gazi Üniversitesi. Eğitim Bilimleri Enstitüsü. Ankara.

Haller, H., & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research, 7*(1), 1–20.

Heider, Fritz. (1958). *The Psychology of Interpersonal Relations*. New York: John Wiley & Sons.

Holden, G., Barker, K., Meenaghan, T. & Rosenberg, G. (1999). Research self-efficacy: A new possibility for educational outcomes assessment. *Journal of Social Work Education, 3*, 463-476.

Huinker, D., & Madison, S. K. (1997). Preparing efficacious elementary teachers in science and mathematics: The influence of methods courses. *Journal of Science Teacher Education, 8*, 2, 107-126.

Hutcheson, G. D., & Sofroniou, N. (1999). The Multivariate Social Scientist: an introduction to generalized linear models. Sage Publications.

Joliffe, F. R. (1976). A continuous assessment scheme for statistics courses for social scientists. *International Journal of Mathematics Education in Science and Technology, 7*, 97-103.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* England: Cambridge University Press.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika,* 39, 31-36.

Kalton, G. (1973). Problems and possibilities with teaching introductory statistics to social scientists. *International Journal of Mathematics Education in Science and Technology, 4,* 7-16.

Kaptan, F., (1998). Fen Öğretiminde Kavram Haritası Yönteminin Kullanılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 14,* 95 - 99.

Kendall, M.G., (1955). Rank Correlation Methods, second ed. Charles Griffin & Co, London.

Kesici S., Erdogan A., & Şahin I. (2010). Mathematics anxiety according to middle school students' achievement motivation and social comparison. Education, 131:54-63.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, UK: CambridgeUniversity Press.

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*(3), 741–56.

Klassen, R. M., Usher, E.L., & Bong, M. (2010). Teachers' collective efficacy, job satisfaction, and job stress in cross-cultural context. *The Journal of Experimental Education,* 78, 464-486.

Kline, P. (1999). The handbook of psychological testing (2nd ed). London: Routledge.

Knowles, M. S. (1975). *Self- Directed Learning: A Guide for Learners and Teachers.* New York: Cambridge Book Co.

Konold, C. (1991), "Understanding Students' Beliefs About Probability," in *Radical Constructivism in Mathematics Education,* ed. E. v. Glasersfeld, Amsterdam: Kluwer, pp. 139-156

Kooij, D., de Lange, A., Jansen, P., & Dikkers, J. (2008). Older workers' motivation to continue to work: Five meanings of age. *Journal of Managerial Psycology, 23,* 364-394.

Köksal, B. A. (1985). *İstatistik Analiz Metotları,* İstanbul: Çağlayan Kitabevi.

Krauss, S., & Wassner, C. (2002). *How significance tests should be presented to avoid the typical misinterpretations.* In Proceedings of the sixth international conference on teaching statistics. Voorburg, The Netherlands: International Statistical Institute.

Langenfeld, T. E., & Pajares, M. F. (1993). The Mathematics Self-Efficacy Scale (MSES): Refining the construct. Paper presented at the meeting of the American Educational Research Association, Atlanta.

Liu, T., Lin, Y., & Tsai, C. (2008). Identifying Senior High School Students' Misconceptions About Statistical Correlation, And Their Possible Causes: An Exploratory Study Using Concept Mapping With Interviews. *International Journal of Science and Mathematics Education* (2009) 7: 791Y820 National Science Council, Taiwan.

Liu, x., & Ramsey, J. (2008). Teachers' job satisfaction: analyses of the Teacher Follow-up Survey in the United States for 2000-2001. *Teaching and Teacher Education, 24*, 1173-1184.

Luszczynska, A., & Schwarzer, R. (2005). Social cognitive theory. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (2nd ed. rev., pp. 127-169). Buckingham, England: Open University Press.

Malpass, J. & Others (1996). Self-regulation, goal orientation, self efficacy, and mathematics achievement. Paper Presented at the Annual Meeting of the American Educational Research Association, New York.

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted hierarchical structure. *Educational Psychologist, 20*, 107-125.

McAdam, E. K. (1986). Cognitive behavior therapy and its application with adolescents. *Journal of Adolescence 9*, 1, 15.

Meier, S., McCarthy, P., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research, 8*, 10-120.

Mevarech, Z. R., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. Educational Studies in Mathematics, 32, 229-263.

Miller, N. E., & Dollard, J. (1941). *Social Learning and Imitation.* New Haven: Yale University Press

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review 102*, 246-268

Nicolaou, A. A., & Philippou, G. N. (2004). Efficacy beliefs, ability in problem sing, and mathematics achievement, *Proceedings of the 3rd International Biennial SELF Research Conference, Self-Concept, Motivation and Identity: here to from here?* Berlin, Germany.

Nisbett, R., & Ross, L. (1980). *Human inferences: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, N.J: Prentice-Hall.

Noordman, L. G. M., & Vonk, W. (1998). Memory-based processing in understanding causal information. *Discourse Processes, 26*, 191-212.

Ormrod, J. E. (1999). *Human learning* (3rd ed.). Upper Saddle River, NJ: Merrill.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving in gifted students. *Contemporary Educational Psychology, 21*, 325–344.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental-ability in Mathematical problem-solving, *Contemporary Educational Psychology, 20*, 426–443.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86*, 193–203.

Pajares, F., & Miller, M. D. (1995). Mathematics Self-Efficacy and Mathematics Performance: The Need for Specificity of Assessment. *Journal of Counseling Psycology, 42* (2), 190-198.

Pajares, F., & Schunk, D. H. (2001). *Self beliefs and school success; Self efficacy, self concept and school achievement*. London: Ablex Publishing.

Pintrich, P. R., & De Groot E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), pp. 33-50.

Pintrich, P. R., & Schunk, D. H. (2002), *Motivation in education: Theory, research, and applications* (2nd ed.), Prentice Hall Englewood Cliffs, NJ.

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika, 74*, 145-154.

Robert E Kass, Valerie Ventura, and Emery N Brown. Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology, 94*(1):8–25, 2005.

Salomon, G. (1984). Television is “easy” and print is “tough”: The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76(4)*, 647658.

Schunk, D. H., & Gunn, T. P. (1986). Self-efficacy and skill development: Influence of task strategies and attributions. *Journal of Educational Research, 79*, 238-244.

Schunk, D. H., (1991). Self-efficacy and academic motivation. *Education psychologist, 26* (3 & 4), 201-231.

Schunk, D. H., & Pajares, F. (2005). Competence perceptions and academic functioning. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 85-104). Guilford Press, New York.

Seferoğlu, S. S. & Akbıyık, C. (2009). *Bilişim teknolojilerinin okullarda kullanımı: Öğretmenlerin teknolojiyi kullanma durumları. 18.* Eğitim Bilimleri Kurultayı, Ege Üniversitesi Eğitim Fakültesi 1-3 Ekim 2009, Sürmeli Efes Otel, Selçuk/Kuşadası - İZMİR.

Sevimli, N. (2010). *Matematik Öğretmen Adaylarının İstatistik Dersi Konularindaki Kavram Yanılgıları; İstatistik Dersine Yönelik Özyeterlilik İnançları Ve Tutumlarının İncelenmesi.* Unpublished master's thesis. Marmara University.

Sewell, A., & St George, A. (2000). Developing efficacy beliefs in the classroom. *Journal of Educational Enquiry, 1*(2), 58-71.

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*, 611–625.

Shaughnessy, J. M., Ciancetta, M., & Canada, D. (2003). Middle school students' thinking about variability in *repeated trials: A cross -task comparison. In N. Pateman, B. Dougherty, & J. Zillah (Eds.). Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education*: Vol. 4. Honolulu, HI: University of Hawaii.

Shultz, K. S., & Koshino, H. (1998). Evidence of reliability and validity for Wise's Attitude toward Statistics scale. *Psychological Reports.82* (2), 27–31.

Siegle, D. (2003). Influencing student mathematics self-efficacy through teacher training. *Paper Presented at the Annual Meeting of the American Research Association*, Chicago, IL.

Silver, B. B., Smith, E. V., Jr., & Greene, B. A. (2001). A study strategies self-efficacy instrument for use with community college students. *Educational and Psychological Measurement, 61*(5), 849-865.

Smith, C. H., (1992). Using writing assignments in teaching statistics: An empirical study. *Mathematics and Computer Education, 26*, 21–34.

Steinberg, L., & Morris, A. (2001). Adolescent Development. *Annual Review of Psychology, 52*,83–140.

Stone, A., Allen, K., Rhoads, T. R., Murphy, T. J., Shehab, R. L. & Saha, C. (2003). The statistics concept inventory: a pilot study, *Paper presented at the 33rd ASEE/IEEE Frontiers in Education Conference, Boulder* (p.251–256). University of Oklahoma, USA.

Sotos, A. E. C., Vanhoof, S., Noortgate, W. V., & Onghena, P. (2007) Students‘ misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review 2*(3), 98–113.

Tekkaya, C., Çapa, Y., & Yılmaz, Ö. (2000). Biyoloji öğretmen adaylarının genel biyoloji konularındaki kavram yanılgıları. *Hacettepe Eğitim Fakültesi Dergisi,* 18: 37-44.

Tschannen-Moran, M. & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. Teaching & Teacher Education, 23(6), 944-956.

Urquhart, N.S. (1971). Nonverbal instructional approaches for introductory statistics. *American Statistician, 25*(2), 20-25.

Utts, J. (1996). *Seeing Through Statistics*, Belmont, CA: Duxbury Press.

Watters, J. J., & Ginns, I. S. (1995). *Origins of and changes in preservice teachers' science teaching self efficacy.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology, 99*, 181–193.

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *The Academy of Management Review, 14*(3), 361–384.

Wyman, D. (1994). Shrubs and Vines for American Gardeners. *The Macmillan Co.*, London

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego: Academic Press.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to SE and strategy use. *Journal of Educational Psychology, 82(1)*, 51-59.

Appendices

Appendix 1

Statistics Concept Inventory

Question Guidelines: For the following test questions, mark the correct answer according to the given options.

Test time: 40 minutes

1. A certain diet plan claims that subjects lose an average of 20 pounds in 6 months on their plan. A dietitian wishes to test this claim and recruits 15 people to participate in an experiment. Their weight is measured before and after the 6-month period. Which is the appropriate test statistic to test the diet company's claim?

a) two-sample Z test

b) paired comparison t test

c) two-sample t test

2. Which would be more likely to have 70% boys born on a given day: A small rural hospital or a large urban hospital?

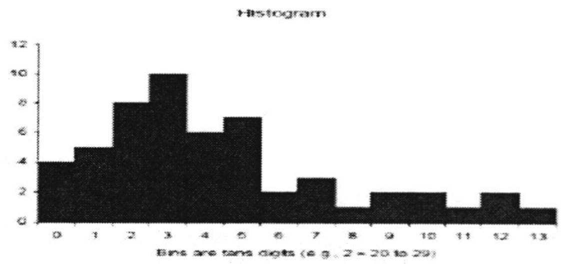
a) Rural

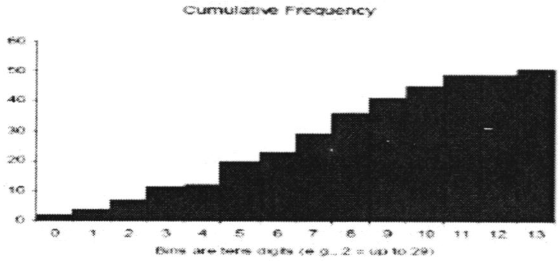
b) Urban

c) Equally likely

d) Both are extremely unlikely

3. Two of the following are graphical presentations of the same set of data. Which graph is of a different data set?





Stem	Leaf
0	65
1	79
2	124
3	1356
4	6
5	0023567[illegible]
6	179
7	033348
8	0136799
9	00358
10	2679
11	1455
12	
13	22

a) Histogram

b) Cumulative Frequency

c) Stem and Leaf

4. A student scored in the 90th percentile in his Chemistry class. Which is always true?

a) His grade will be an A

b) He earned at least 90% of the total possible points

c) His grade is at least as high as 90% of his classmates

d) None of these are always true

5. The following are temperatures for a week in August: 94, 93, 98, 101, 98, 96, and 93. By how much could the highest temperature increase without changing the median?

a) Increase by 8°

b) Increase by 2°

c) It can increase by any amount

d) It cannot increase without changing the median

6. A bottling company believes a machine is under-filling 20-ounce bottles. What will be the alternate hypothesis to test this belief?

a) On average, the bottles are being filled to 20 ounces.

b) On average, the bottles are not being filled to 20 ounces.

c) On average, the bottles are being filled with more than 20 ounces.

d) On average, the bottles are being filled with less than 20 ounces.

7. Which of the following statistics is least impacted by extreme outliers?

a) Range

b) 3rd quartile

c) Mean

d) Variance

8. A student attended college A for two semesters and earned a 3.24 GPA (grade point average). The same student then attended college B for four semesters and earned a 3.80 GPA for his work there. How would you calculate the student's GPA for all of his college work? Assume that the student took the same number of hours each semester.

a) $\frac{3.24 + 3.80}{2}$

b) $\frac{3.42\ (2) + 3.80\ (4)}{2}$

c) $\frac{3.24\ (2) + 3.80(4)}{6}$

d) It is not possible to calculate the student's overall GPA without knowing his GPA for each individual semester.

9. For the following set of numbers, which measure will most accurately describe the central tendency? 3, 4, 5, 5, 6, 8, 10, 12, 19, 36, 83

a) Mean

b) Median

c) Mode

d) Standard deviation

10. A researcher conducts an experiment and reports a 95% confidence interval for the mean. Which of the following must be true?

a) 95% of the measurements can be considered valid

b) 95% of the measurements will be between the upper and lower limits of the confidence interval

c) 95% of the time, the experiment will produce an interval that contains the population mean

d) 5% of the measurements should be considered outliers

11. The mean height of American college men is 70 inches, with standard deviation 3 inches. The mean height of American college women is 65 inches, with standard deviation 4 inches. You conduct an experiment at your university measuring the height of 100 American men and 100 American women. Which result would most surprise you?

a) One man with height 79 inches

b) One woman with height 74 inches

c) The average height of women at your university is 68 inches

d) The average height of men at your university is 73 inches

12. A meteorologist predicts a 40% chance of rain in London and a 70% chance in Chicago. What is the most likely outcome?

a) It rains only in London

b) It rains only in Chicago

c) It rains in London and Chicago

d) It rains in London or Chicago

13. You perform the same two significance tests on large samples from the same population. The two samples have the same mean and the same standard deviation. The first test results in a p-value of 0.01; the second, a p-value of 0.02. The sample mean is equal for the 2 tests. Which test has a larger sample size?

a) First test

b) Second test

c) Sample sizes equal

d) Sample sizes are not equal but there is not enough information to determine which sample is larger

14. Which statistic would you expect to have a normal distribution? I) Height of women II) Shoe size of men III) Age in years of college freshmen

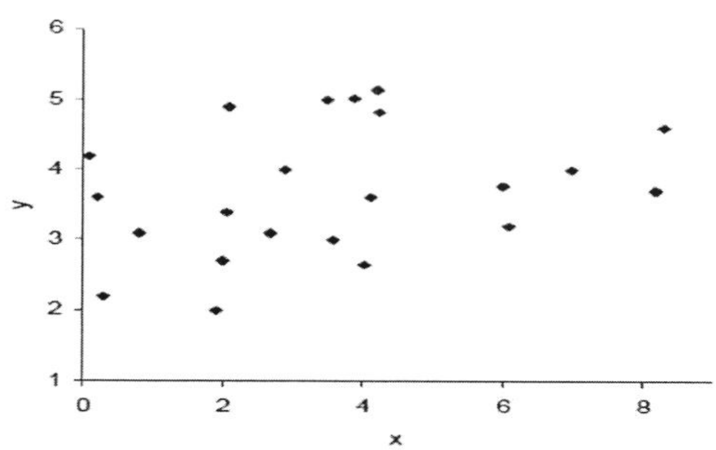
a) I & II

b) II & III

c) I & III

d) All 3

15. Estimate the correlation coefficient for the two variables X and Y from the scatter plot below.



a) -0.3

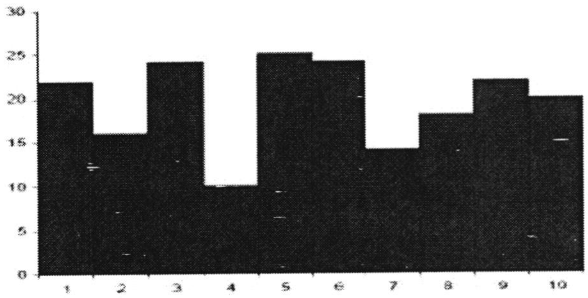
b) 0

c) 0.3

d) 0.9

e) 1.6

16. Consider the sample distribution below. This sample was most likely taken from what kind of population distribution?



a) Normal

b) Uniform

c) Skewed

d) Bimodal

17. You have a set of 30 numbers. The standard deviation from these numbers is reported as zero. You can be certain that:

a) Half of the numbers are above the mean

b) All of the numbers in the set are zero

c) All of the numbers in the set are equal

d) The numbers are evenly spaced on both sides of the mean

18. In order to determine the mean height of American college students, which sampling method would not introduce bias?

a) You randomly select from the university basketball team

b) You use a random number table to select students based on their student ID

c) You roll a pair of dice to select from among your friends

d) None of the methods will have bias

19. A scientist takes a set of 50 measurements. The standard deviation is reported as -2.30. Which of the following must be true?

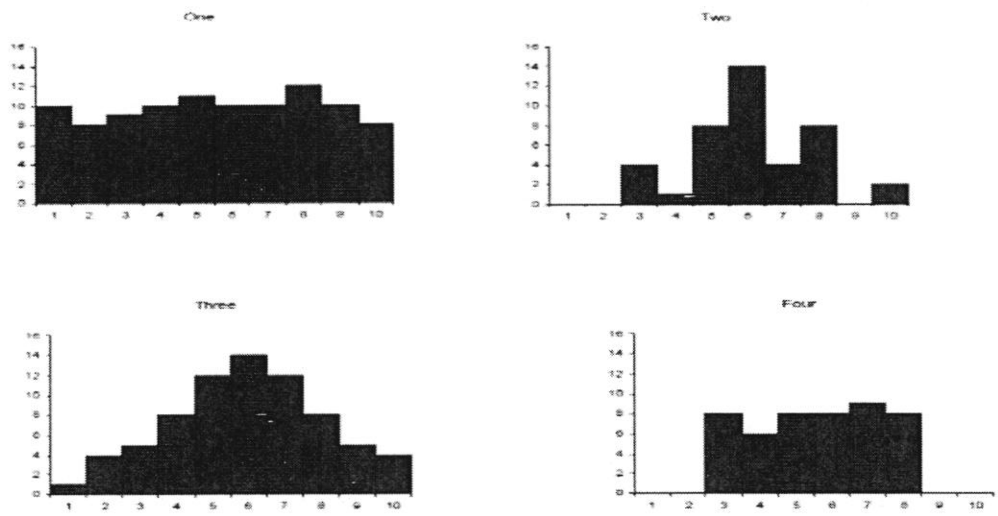
a) Most of the measurements were negative

b) All of the measurements are less than the mean

c) All of the measurements were negative

d) The standard deviation was calculated incorrectly

20. The following are histograms of quiz scores for four different classes. Which distribution shows the most variability?



a) I

b) II

c) III

d) IV

21. In a manufacturing process, the error rate is 1 in 1000. However, errors often occur in groups, that is, they are not independent. Given that the previous output contained an error, what is the probability that the next unit will also contain an error?

a) Less than 1 in 1000

b) Greater than 1 in 1000

c) Equal to 1 in 1000

d) Insufficient information

22. For the past 100 years, the average high temperature on October 1 is 78° with a standard deviation of 5°. What is the probability that the high temperature on October 1 of next year will be between 73° and 83°?

a) 0.68

b) 0.95

c) 0.997

d) 1

23. You are rolling dice. You roll 2 dice and compute the mean of the numbers rolled, then 6 dice and compute the mean, then 10 dice and compute the mean. Under which scenario would you be most surprised to find a mean of at least 4.5?

a) Rolling 2 dice

b) Rolling 6 dice

c) Rolling 10 dice

d) There is no way this can happen

24. Two confidence intervals are calculated for two samples from a given population. Assume the two samples have the same standard deviation and that the confidence level is fixed. Compared to the smaller sample, the confidence interval for the larger sample will be:

a) Narrower

b) Wider

c) The same width

d) It depends on the confidence level

25. Information about different car models is routinely printed in public sources such as Consumer Reports and new car buying guides. Data was obtained from these sources on 1993 models of cars. For each car, engine size in liters was compared to the number of engine revolutions per mile. The correlation between the two was found to be -0.824. Which of the following statements would you most agree with?

a) A car with a large engine size would be predicted to have a high number of engine revolutions per mile.
b) A car with a large engine size would be predicted to have a low number of engine revolutions per mile.
c) Engine size is a poor predictor of engine revolutions per mile.
d) Engine size is independent of revolutions per mile.

Appendix 2

Current statistics self-efficacy

Please rate your confidence in your current ability to successfully complete the following tasks. The item scale has six possible responses: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence. For each task, please mark the one response that represents your confidence in your current ability to successfully complete the task.

	No confidence at all					complete confidenence
1. Identify the scale of measurement for a variable	1	2	3	4	5	6
2. Interpret the probability value (p-value) from a statistical procedure	1	2	3	4	5	6
3.Identify if a distribution is skewed when given the values of three measures of central tendency	1	2	3	4	5	6
4.Select the correct statistical procedure to be used to answer a research question	1	2	3	4	5	6
5.Interpret the results of a statistical procedure in terms of the research question	1	2	3	4	5	6
6.Identify the factors that influence power	1	2	3	4	5	6

7.Explain what the value of the standard deviation means in terms of the variable being measured	1	2	3	4	5	6
8.Distinguish between a Type I error and a Type II error in hypothesis testing	1	2	3	4	5	6
9.Explain what the numeric value of the standard error is measuring	1	2	3	4	5	6
10.Distinguish between the objectives of descriptive versus inferential statistical procedures	1	2	3	4	5	6
11.Distinguish between the information given by the three measures of central tendency	1	2	3	4	5	6
12.Distinguish between a population parameter and a sample statistic	1	2	3	4	5	6
13.Identify when the mean, median and mode should be used as a measure of central tendency	1	2	3	4	5	6
14. Explain the difference between a sampling distribution and a population distribution	1	2	3	4	5	6