EFFECTS OF GENDER AND SOCIOECONOMIC STATUS TOWARDS FORM FOUR PUPILS' STATISTICAL THINKING ABILITIES

ONG JING HUI

FACULTY OF EDUCATION
UIVERSITI MALAYA
KUALA LUMPUR
2020

EFFECTS OF GENDER AND SOCIOECONOMIC STATUS TOWARDS FORM FOUR PUPILS' STATISTICAL THINKING ABILITIES

ONG JING HUI

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF MATHEMATICS EDUCATION

FACULTY OF EDUCATION
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Ong Jing Hui

Matric No: PGM130002

Name of Degree: Masters of Education (Mixed Mode)

Title of Dissertation: Effects of Gender and Socioeconomic Status on Statistical Thinking Abilities of Form Four pupils

Field of Study: Master of Mathematics Education

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ABSTRACT

Statistical thinking abilities of Form Four students are not known as there is a lack of studies that measures it. Despite the fact that the students were exposed to the same curriculum and pedagogy, the thinking abilities of students varied in terms of gender. There is also uncertainty in whether socio-economic status is a cause that affects statistical thinking abilities due to lack of research in that particular area.

This study aimed to examine the effects of gender and socio-economic status on statistical thinking abilities of Form Four students. To answer the three research questions of this study, survey descriptive design and causal comparative design were employed. 180 Form Four students (90 males, 90 females) of three socio-economic status groups mainly low socio-economic status group (N = 65), middle socio-economic status group (N = 98), and high socio-economic status group (N = 17) from several secondary schools in the area of Shah Alam, Selangor participated in this study.

Statistical thinking abilities consist of three levels mainly ability to organize and reduce data, ability to represent data, and ability to analyze and interpret data. An instrument called the Statistical Thinking Abilities Test (STAT) was used to measure Form Four students' statistical thinking abilities levels. The result of analysis of descriptive statistics showed that most Form Four students were at level 1 (42.2%) and level 2 (46.1%) of statistical thinking abilities levels.

The result from the Mann-Whitney U test revealed that the level of statistical thinking abilities of male students (*Mean rank* = 95.61) was different with female students (*Mean rank* = 85.39). The result indicated no significant differences of students' statistical thinking abilities by gender, $U = 3590$, $p = .148$. The result from Kruskal Wallis H test shown significant differences in students' statistical thinking abilities by socio-economic status, $\chi^2(2) = 34.591$, $p = .001$, with mean rank statistical

thinking abilities low socio-economic status (60.07), middle socio-economic status (101.06), and high socio-economic status (145.97).

This study has implication to teachers, researchers, curriculum planners, and teaching educators. Teachers are able to utilize the rubric to gauge the statistical thinking abilities levels of the students and prepare suitable lesson plans accordingly. Researchers can pursue further study by looking at other factors that affect statistical thinking abilities and the interaction effect among them. Curriculum planners benefit from this study through better understanding of the statistical thinking abilities to incorporate them into the mathematics curriculum. Teaching educators also benefit from this study by utilizing the ideas and operations of statistical thinking abilities to monitor students' development pace of statistical thinking abilities from various socio-economic status backgrounds.

In conclusion, past researches' findings shown gaps in students' statistical thinking in comprehending statistical concepts, which prompted the study to measure the statistical thinking abilities levels of Form Four students and the factors that affect it. The findings shown evidence of significant effect of socio-economic status on statistical thinking abilities of Form Four students where most of them are at the first and second level of statistical thinking abilities.

KESAN JANTINA DAN STATUS SOSIOEKONOMI TERHADAP KEBOLEHAN PEMIKIRAN STATISTIK MURID TINGKATAN EMPAT

ABSTRAK

Kebolehan pemikiran statistik murid Tingkatan Empat adalah tidak diketahui disebabkan kekurangan kajian yang mengukur kebolehan tersebut. Walaupun murid-murid terdedah kepada kurikulum dan pedagogi yang sama, ternyata kebolehan pemikiran mereka adalah tidak sama dari segi jantina. Terdapat juga ketidakpastian sama ada status sosio-ekonomi merupakan faktor yang mempengaruhi kebolehan pemikiran disebabkan kekurangan kajian dalam bidang tersebut.

Kajian ini bertujuan untuk mengkaji kesan jantina dan status sosio-ekonomi terhadap kebolehan pemikiran statistik murid Tingkatan Empat. Untuk menjawab tiga soalan kajian ini, reka bentuk kajian tinjauan deskriptif, dan reka bentuk kausal-komparatif telah digunakan. 180 murid Tingkatan Empat (90 lelaki, 90 perempuan) untuk tiga kumpulan status sosio-ekonomi iaitu status sosio-ekonomi rendah (N = 65), status sosio-ekonomi pertengahan (N = 98), dan status sosio-ekonomi tinggi (N = 17) daripada beberapa sekolah menengah di kawasan Shah Alam, Selangor mengambil bahagian dalam kajian ini.

Kebolehan pemikiran statistik mengandungi tiga tahap iaitu kebolehan untuk mengatur dan mengurangkan data, kebolehan untuk mewakili data, dan kebolehan untuk menganalisis dan mentafsir data. Instrumen yang dikenali sebagai STAT telah digunakan untuk mengukur kebolehan pemikiran statistik murid. Keputusan analisis statistik deskriptif menunjukkan bahawa kebanyakkan murid Tingkatan Empat mempunyai kebolehan pemikiran statistik pada tahap 1 (42.2%) dan tahap 2 (46.1%).

Keputusan ujian Mann-Whitney U menunjukkan bahawa tahap kebolehan pemikiran statistik murid lelaki (*Pangkat min* = 95.61) adalah berbeza daripada murid perempuan (*Pangkat min* = 85.39). Keputusan ujian menunjukkan tiada perbezaan yang signifikan antara kebolehan pemikiran statistik murid dengan jantina, $U = 3590$, $p = .148$. Keputusan daripada ujian Kruskal Wallis H menunjukkan perbezaan yang signifikan antara kebolehan pemikiran statistik murid dengan kumpulan status sosio-ekonomi, $\chi^2(2) = 34.591$, $p = .001$, dengan pangkat min kebolehan pemikiran statistik status sosio-ekonomi rendah (60.07), status sosio-ekonomi pertengahan (101.06), dan status sosio-ekonomi tinggi (145.97).

Kajian ini mempunyai implikasi terhadap para guru, penyelidik, perancang kurikulum, dan pendidik pengajaran. Para guru dapat menggunakan rubrik untuk mengukur tahap kebolehan pemikiran statistik murid-murid dan menyediakan rancangan pengajaran yang sesuai. Penyelidik dapat melakukan penyelidikan lanjut dengan merujuk kepada faktor-faktor lain yang mempengaruhi kebolehan pemikiran statistik dan kesan interaksi antara mereka. Perancang kurikulum mendapat manfaat daripada kajian ini melalui pemahaman kebolehan pemikiran statistik yang lebih mendalam untuk menerapkannya ke dalam kurikulum matematik. Pendidik pengajaran memanfaat daripada kajian ini dengan menggunakan idea dan operasi kebolehan pemikiran statistik untuk memantau kadar perkembangan kebolehan pemikiran statisik di kalangan murid-murid daripada pelbagai latar belakang status sosio-ekonomi.

Kesimpulannya, hasil kajian lepas menunjukkan jurang dalam kebolehan pemikiran murid-murid dalam memahami konsept statistik, seterusnya mendorong kajian untuk mengukur kebolehan pemikiran statistik murid Tingkatan Empat dan faktor-faktor yang mempengaruhi kebolehan tersebut. Hasil kajian menunjukkan bukti daripada kesan status sosio-ekonomi terhadap kobolehan pemikiran statistik murid Tingkatan Empat di mana kebanyakkan murid berada pada tahap satu dan dua

ACKNOWLEDGEMENT

First of all, I would like to express my heartfelt gratitude to God for giving me the blessings and health to complete this research. Despite having ups and downs during the journey, a number of kind and helpful people have aided me with their continuous support till the completion of the research.

I am very grateful and would like to express my sincere thanks to my dissertation supervisor, Mr. Norjoharuddeen Bin Mohd Nor for his time, patience, support, and guiding me on the right track of writing quality dissertation. His encouragement and motivation keeps me going and push myself to put in extra effort to complete this research.

Finally, I would like to express my profound gratitude to my parents for providing me with unwavering love and continuous support in accompanying and motivating me throughout the completion of journey. At point of difficult times during the writing of the dissertation, their encouragement kept me going and not giving up. This accomplishment would not have been possible without them.

CONTENTS

Chapter 1: Introduction

Chapter 2: Literature Review

Chapter 3: Methodology

Chapter 4: Findings

Chapter 5: Discussion and Conclusion

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

ANOVA	:	Analysis of Variance
CRCT	:	Criterion-Referenced Competency Tests
IBM	:	International Business Machines
ICC	:	Intraclass Correlation Coefficient
KSSM	:	Standard-Based Curriculum for Secondary Schools
MANCOVA	:	Multivariate Analysis of Covariance
MODEST	:	Modeling to Elicit Statistical Thinking
MOE	:	Ministry of Education
M3ST	:	Middle School Students' Statistical Thinking
PAW	:	Planning Analytics Workplace
SOLO	:	Structured of the Observed Learning Outcome
SPSS	:	Statistical Package for the Social Sciences
STAT	:	Statistical Thinking Abilities Test
STEM	:	Science, Technology, Engineering, and Mathematics
STLE	:	Statistical Thinking Learning Environment
TIMSS	:	Trends In International Mathematics and Science Study

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Statistics is presented in the modern world of the Information Age and era of big data in science, industry, health, and business. The continuous growth of big data further promotes the vital needs of data analysis and statistics. Societies widely approach the study of statistics in order for them to analyze statistics products by reacting reasonably to the quantitative information and making wise decisions (Ben-Zvi & Garfield, 2004). All education levels required the introduction and implementation of statistics instruction as it is a necessity for students to learn statistics (Garfield & Ben-Zvi, 2018).

The question comes to what is statistics? Statistics is the science of studying from data which to be exact it is the collection, analysis, interpretation, and making inferences from data. Besides, statistics are sets of mathematical equations that are used to analyze the world around us using information. Statistics is also defined as the methods that are crucial and intellectual which applies to variation, data, and probability everywhere at the same time in modern life (Moore, 1998).

Statistics is introduced in the secondary school curriculum in the country of Malaysia as one of the topics for Form Four Mathematics subject. It consisted the learning of class interval's concept, central tendency measures' concepts such as mean, median, and mode, graph representation and data interpretation's concepts of histograms and frequency polygons, dispersion measures' concepts, and cumulative frequency concepts.

Many years back, statistics barely involved in students' school experiences where students were only exposed to collect data from science laboratory lessons and perhaps calculating mean of set of numbers from mathematics lesson. However, nowadays students were exposed to statistics at early ages where they encountered statistical tests, and distributions in their lesson, without the exclusion of central tendency and dispersion as well (Le, 2017). Its connection with literacy, science, health, and social science grew stronger to the extent of statistics teaching being implemented across the curriculum with consideration of relationship with mathematics (Usiskin & Hall, 2015). Statistics' importance was felt highly as large number of statistical implications were implemented throughout all education levels of various countries' school curriculum. The implementation was a necessity to foster students' ability in utilizing statistics to manage the quantitative information efficiently (Garfield & Ben-Zvi, 2018). The increase of students' admission on statistics courses in various departments of colleges shown the acknowledgement of the importance of statistics (Scheaffer & Stasny, 2004).

The importance of statistics was well appreciated as statistics was utilized in research theory to explain, predict, and understand phenomena by specifying relations among variables (Shmueli, 2010). Statistics was used to compare the methods of data collection necessary to estimate treatment effects and to test hypotheses (Massey & Miller, 2006). In addition, statistics allowed the usage of descriptive statistics to assist theory to describe the sample's behavior. The summarizing of the probability distributions of sample can be done to summarize the relationships between the quantities measured with the sample. Last but not least, the relationship among variables was summarized using inferential statistics such as regression analysis (Ding, 2006).

Students nowadays were surrounded by quantitative information which required them to analyze, make decision, interpret, and solve problem (Scranton, 2013). The ability to process information was a necessity for them to cope with the volatile era of quantitative information. An understanding of the statistics learnt was required to apply them to everyday life which enable the students to analyze everything around them effectively (Groth, 2003). This was where statistical thinking played an important aspect to aid them in their thought process to understand, sort and apply the data and information.

What is Statistical thinking? Chance, delMas, and Garfield (2003) defined statistical thinking as a comprehension of the process behind the execution of statistical inquiries and its concepts. These concepts consisted of the diversity and the application of suitable data analysis' methods such as graphical representations and numerical display of data along with the best execution plan (Chance, delMas, & Garfield, 2003). Furthermore, statistical thinking engaged a comprehension of the essence of sampling, inferring of samples to populations' methods, and reasoned for the need of designed experiments to establish causation. Chance, delMas, and Garfield (2003) also stated that statistical thinking included ability to grasp a situation problem to investigate and make conclusions, through the ability of comprehending the whole flow of data collection and data analysis. Statistical thinkers also had the ability to assess and evaluate the outcome of a solved problem after analysis of the whole problem.

The importance of statistical thinking was seen in our daily lives in news media on large range of coverage on global and nation news, which included environment, health, sports, entertainment, science, politics, and jobs (Watson, 1997). All of these used data that required the need of decision making to extract meaning and insight about the real life situations and context (Pfannkuch, 2008). On most occasions where statistics were presented correctly, there was only need of statistical thinking to grasp and decipher the data to conclude and make wise choices (Watson, 1997). The diversity of data available required the expertise of professionals and business manager that were proficient in statistical thinking to intelligently collect, analyze, and explain data to assist their decision making. Their expertise was required to ensure that they can arrive with the best solution for all the issues faced, making firm decisions, and reduce guesswork.

Statistical thinking also played an important role in education where it assisted students to further understand and apply the terminology and concepts of the statistics learnt in class instead of just developing basic understanding of them (Watson, 1997). Statistical thinker students were capable to question and probe issues arisen instinctively of certain situations from time to time with the learnt knowledge (Chance, 2002).

The implementation of Standard-Based Curriculum for Secondary Schools (KSSM) in phases throughout all the secondary schools in Malaysia encompassed the Learning Standards, Content Standards, pedagogical and assessment approach. It was introduced with aim to improve the secondary school curriculum's quality by preparing and training students to become competent persons to face the uprising challenges in the 21st Century. Standard Document of Curriculum and Assessment revealed that elements of higher order thinking abilities had been incorporated in the curriculum where the notable abilities identified were problem solving, critical thinking, reasoning, creativity, decision making, and strategic thinking (Ministry of Education Malaysia, 2018). In fact, there were similarities in the characteristics of those abilities in comparison with statistical thinking. No doubt statistical thinking shown its presence in the new Standard-Based Curriculum for Secondary Schools now that it was revealed that its primary focus was to prepare students with the necessary higher order thinking abilities in this 21st century to compete globally.

Gender in academic event had been reported to reveal discrepancies across levels of socio-economic status (Bécares & Priest, 2015). Gender and socio-economic status were popular factors that affected students in the educational area of research. Gender was a common variable that was often reported in various research that linked different characteristics of boys and girls in education disparities. There were evidences that shown a relationship in gender and thinking (Kohan-Mass, 2016; Kawashima & Shiomi, 2007). In addition, there was awareness of socio-economic status' importance on students. Parents had strong influence in ensuring that intellectual stimulation, physical, and psychological safety were provided by a home for the children (Ewumi, 2012). Blums, Belsky, Grimm, and Chen (2017) supported the claim by stating that there was strong evidence that a family's socio-economic status during children's early childhood played a key role in affecting the cognitive ability of children. The intriguing effect of these two variables in playing their part to affect students' thinking abilities grabbed the utmost attention and therefore selected to be studied on and prompted the research on the effects.

1.2 Statement of Problem

Statistics in the secondary school curricula was relatively new a decade back and the research in this area had started to accumulate throughout the years (Groth, 2003). The research base of statistics education started focusing onto statistical thinking for the past decades (Ben-Zvi & Garfield, 2004). For instance, Pfannkuch and Horring (2004) conducted study on the developing of statistical thinking in secondary school. However, most findings shown that there were gaps in students' statistical thinking in understanding concepts of statistics (Groth, 2006). Students spent more time creating representations rather than interpreting or analyzing those constructed representations (Scranton, 2013). As a matter of fact, the extent of secondary school students' understanding and levels of statistical thinking were not known. This was due to the fact that there was still a lack of studies that peeked into the measuring of secondary school students' statistical thinking levels.

Gender is a variable that was being related in various research to study the effect on the differences. Over the years, there was a significant increase in gender gap where girls outperformed boys at every level (Ministry of Education Malaysia, 2013). Gender's biological differences or stereotypes might be the reason behind the contrast identified in the scholastic performance of the students from two different groups of gender (Klein, 2004). Individuals such as students were expected to have the same thinking abilities given that they were exposed to the same curriculum and pedagogy in the secondary school environment. Despite that, the thinking abilities of students varied among each other.

The research on gender in terms of cognitive processes and intellectual abilities were lacking (Ewumi, 2012). The researches on the effect of gender on statistical thinking were even lesser where the focus were on either thinking abilities or gender only. There were several researches that study on statistical thinking or gender, but there was lack of research that discussed about whether gender was a factor that affected statistical thinking ability. As such, the gender's effect on statistical thinking was uncertain. Langrall and Mooney (2002) conducted study on characterizing middle school students' statistical thinking. Martin, Hughes, and Fugelsang (2017) conducted a study that shown the issue of male and female in statistical reasoning. Based on these past studies, it was hypothesized that there might be a gender effect on statistical thinking that showed the differences in gender and statistical thinking abilities. Thus, comparison between gender mainly male and female was necessary to identify the differences in their level of statistical thinking abilities.

Socio-economic status is a multidimensional construct that involved measures of material wealth, social prestige, and even education. Individuals can be affected by socio-economic status in terms of both brain structures and cognitive functions across their development (Hackman, Farah, & Meaniey, 2010). Individuals entered schools with inequalities imposed by home environment and low socio-economic status individuals often prevail compared to other peers. The lack of expertise and school funding in low socio-economic area further worsen it. Ministry of Education Malaysia (2013) mentioned that there was clear evidence that shown the gap where students from lower socio-economic status group were not on par with the students from other two socio-economic status groups in terms of achievements. Ewumi (2012) supported the claim by stating that low socio-economic status areas schools' students often have low academic achievement test scores.

There were various researches that looked into socio-economic status and its effect on variable such as mathematics achievement (Caro, 2009; Ford, 2013). However, there was a lack of research discussing about whether socio-economic status was a cause that had effects on statistical thinking abilities. Therefore, the effect was uncertain. Gustafsson, Nilsen, and Hansen (2016) found that there was issue of students' socio-economic status on mathematics achievement. Jankowska and Karwowski (2019) conducted study that shown the effect of students' socio-economic status on creative thinking development. Based on these past studies, it was hypothesized that there might be differences in students' socio-economic status and statistical thinking abilities. Thus, comparison among levels of socio-economic status was vital to determine whether they affect students' statistical thinking abilities.

This study will look in depth into these variables to see how each of the variables will bring changes to statistical thinking abilities of students. It is intriguing to look at the effect of these variables as it is hypothesized that they will affect and explain the variation of the levels of students' statistical thinking abilities.

1.3 Research Aim and Objectives

The study aimed to determine the effects of gender and socio-economic status on statistical thinking abilities of Form Four pupils. To achieve this aim, the study intended to accomplish the following objectives:

a) Measure statistical thinking abilities of Form Four pupils.
b) Determine the difference of statistical thinking abilities of Form Four pupils by gender.
c) Determine the difference of statistical thinking abilities of Form Four pupils by socio-economic status.

1.4 Research Questions

The research questions were formulated according to the objectives of this study:

a) What are the statistical thinking abilities of Form Four pupils?
b) Is there any significant difference in the statistical thinking abilities of Form Four pupils by gender?
c) Is there any significant difference in the statistical thinking abilities of Form Four pupils by socio-economic status?

1.5 Research Hypothesis

The research hypothesis of this study was stated as below:

a) H_0: There is no difference between statistical thinking abilities by gender.

   H_1: There is difference between statistical thinking abilities by gender.

b) H_0: There is no difference between statistical thinking abilities by socio-economic status.

   H_1: There is difference between statistical thinking abilities by socio-economic status.

1.6 Definition of Terms

1.6.1 Statistical Thinking Abilities

Statistical thinking abilities are defined as the ability to comprehend data-handling concepts that are complicated and develop over time (Langrall & Mooney, 2002). It consists of three statistical thinking processes which are organizing and reducing data, representing data, and analyzing and interpreting data as stated by Langrall and Mooney (2002). In this study statistical thinking abilities are operationally defined as ability measured by the Statistical Thinking Abilities Test (STAT) instrument based on three statistical thinking processes mainly ability to organize and reduce data, ability to represent data, and ability to analyze and interpret data.

1.6.1.1 Ability to Organize and Reduce Data

It involves three sub-processes of categorizing data, describing data using measures of center, and describing data using measures of spread as defined by Langrall and Mooney (2002). In this study, ability to organize and reduce data is operationally defined as ability measured using the rubric with two sub-processes mainly expressing data using measures of center, and expressing data using measures of spread.

1.6.1.2 Ability to Represent Data

It involves three sub-processes namely creating a data display for a given data set, and assessing the effectiveness of a given data sets as defined by Langrall and Mooney (2002). In this study, ability to represent data is operationally defined as ability measured using the rubric with 2 sub-processes mainly creating a data representation for a given data set, and creating a different representation for the same data set.

1.6.1.3 Ability to Analyze and Interpret Data

Analyzing and interpreting data involves recognizing trends and forming inferences or predictions about the data (Langrall & Mooney, 2002). These levels resulted four sub-processes described by Langrall and Mooney (2002) as comparing within data sets or data displays, comparing between data sets or data displays, generating conjectures from a given data set or data display, and apply comparative reasoning. Ability to analyze and interpret data is operationally defined as the ability measured using the rubric with two sub-processes mainly making comparison within data sets or data displays, and making comparison between data sets or data displays.

1.6.2 Gender

Ewumi (2012) defined gender as a male or female's psychological and socio-cultural dimensions.

1.6.3 Socio-Economic Status

Taylor and Yu (2009) defined socio-economic status as the proxies of income, education, or occupation. Socio-economic status is operationally defined in this study as the median of household income where the criteria used for determining the levels of socio-economic status are of three groups mainly high socio-economic status (T20), middle socio-economic status (M40) and low socio-economic status (B40). The T20 group is defined with criteria of family income with median of RM13,148; M40 group defined with criteria of family income with median of RM6,275; and B40 group defined with family income with median of RM3,000 (Department of Statistics Malaysia, 2017).

1.7 Research Paradigm

A research paradigm is about how issues need to be addressed and understood by scientists through the sharing of common beliefs and agreements between them. My research paradigm for this study is the positivism paradigm as I am a positivist. The positivism paradigm states that observation of actual events can be done first handedly and justified with rational analysis. The criterion for assessing a scientific theory's validity is whether the knowledge claims are on par with information obtained using senses (Kaboub, 2008).

This study is based on positivism research paradigm because the study is based on the following assumptions of the ontology and epistemology assumptions. The ontology assumption of positivism is that there is one observable reality or truth while the epistemology assumption of positivism states that reality can be measured through valid and reliable tools. Since this study is based on positivism paradigm, therefore the study will be measuring using quantitative research methodology.

1.8 Limitation and Delimitation of the Study

There were several limitations identified for this study. The three limitations involved research design, data collection procedure, and sampling. On the other hand, three delimitations in this study involved psychological construct, mathematical construct, and research participants.

This study utilized survey descriptive research as the research design. The selected research design has its weakness of its inability to do theoretical generalization. It is more limited to do statistical generalization of sample to population.

This study had a limitation on the procedure of data collection whereby the test cannot be administered to the students at the same time. Due to the large number of participants, instructors had to administer the test depending on classes' timetable, classroom environment, and students' characteristics. This uncontrolled aspect of test administration might result in the lack of generalizability of the test scores.

There was also a limitation of the sampling where the administration of instruments was only done to participants of only certain schools in Selangor. The sample of the study was only limited to the subpopulation of Form Four students in several schools and not the whole population of Form Four students throughout the country.

The delimitation of this study were psychological construct, mathematical construct, and research participants. This study only focused on measuring students' statistical thinking abilities and not on other constructs such as creative thinking abilities. The construct of statistical thinking abilities was selected due to its importance where students are exposed to statistics curriculum at early ages nowadays.

This study also emphasized on the mathematical construct of statistics. Other contents in Mathematics curriculum such as quadratic equation was not included in this study. The mathematical construct of statistics was selected due to its importance and wide usage in the current society surrounded by products of statistics.

Last but not least, this study focused on measuring Form Four students' statistical thinking abilities. Form One, Form Two, Form Three, and Form Five students were not selected as the research participants for this study. Form Four students were selected due to the reason that they were exposed initially to statistics in the Mathematics curriculum.

1.9 Significance of the Study

This study is beneficial to stakeholders of mathematics education such as researchers, school mathematics curriculum planners, and mathematics educators. The result is believed to be beneficial to researchers. Researchers need proper measurement data in order to understand and conduct their research. The result of the measurement from this study will provide insight to researchers so that they can further pursue in depth study on the statistical thinking abilities. For instance, they can look into other factors which might affect the students' statistical thinking abilities. This study serves as a guideline for researchers to identify the gap and further improve the current study. The instruments in this study can be adapted or adopted as well for other educators or researchers to utilize them in measuring statistical thinking abilities of students.

This study will benefit curriculum planners as they are able to understand thinking abilities better and it will be easier for them to incorporate them into the mathematics curriculum to suit students' abilities with the availability of measurements. The result of the study hopefully will provide insight for mathematics curriculum planners to make improvement to the curriculum planning for area of statistical thinking abilities.

Mathematics educators also benefit from the result of the study in a way that they are able to identify the necessary pedagogy to be implemented on different gender and socio-economic status students. Educators could identify the causes that affect the development of statistical thinking abilities and thus they are able to pay attention to every student's needs according to these factors. Mathematics educators can understand the elements of thinking abilities better and use the knowledge to monitor the students in classroom and guide them in their learning. From there on they can access the level of the students' abilities to ensure students develop thinking abilities according to their own pace to think critically to solve mathematical problems.

1.10 Summary

This chapter highlighted the idea of statistics, statistical thinking, their importance and factors affecting the students' statistical thinking abilities. There were issues that needed to be addressed. The study aimed to identify whether gender and socio-economic status were factors that affect Form Four pupils' statistical thinking abilities. Research questions were formulated as well. The terms of statistical thinking abilities, socio-economic status, and gender were conceptually and operationally defined and discussed as well in this chapter. The limitations and delimitations of this study were explained thoroughly. Last but not least, the significance of this study was highlighted on the importance of this study in benefiting various organizations. The study went on to explain the literature review in the Chapter Two, the methodology in Chapter Three, result of study in Chapter Four, and discussions, conclusions, and research implications in Chapter Five. All the references were placed in the reference section while the supporting materials were displayed under Appendices section.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter explained the four main components of literature review mainly the theory, the variables which were statistical thinking abilities, gender, socio-economic status, review of past literatures on statistical thinking, gender, socio-economic status, and lastly the conceptual framework.

2.2 Statistical Thinking Framework Theory

The statistical thinking framework for this study was adapted from research of Langrall and Mooney (2002), who constructed framework to explain the statistical thinking of students in middle school. Their framework consisting of four statistical thinking processes aimed to describe the students' cognitive levels. The identified four processes as explained by Langrall and Mooney (2002) were describing, organizing and reducing, representing, and analyzing and interpreting data. There were descriptors for every process to describe characteristics of different levels of thinking.

Describing data was an important and complex process entailing various levels of composure in thinking where students become skillful in this process in terms of explaining data from statistical representations (Groth, 2003). Groth (2003) highlighted the importance of this process by stating the necessity of students to have strong ability in reading from various types of graphs to ensure that the later data analysis was purposeful.

The process of describing data involved the process of reading graphs and defined by Friel, Curcio, and Bright (2001) as graph readers' ability to understand and obtain information out of the created graphs. Friel, Curcio, and Bright (2001) mentioned that there were several reasons that had influences towards graph comprehension. The first factor was the purpose behind the selection of graphs to be used. The usage of graphs that were intended for data analysis purpose were more likely to be more meaningful compared with graphs that were shown in traditional textbook instruction. Next, the second factor was the tasks' characteristics that linked to the appearance of graphs. There was research conducted which pointed out that the comprehension of graphs was influenced by whether they came with context or not. The third factor was the representation of the data sets' characteristic. Graph comprehension can be influenced by the data sets in terms of the data type, data size, data spread and variation. Lastly, the fourth factor was the characteristics of the learner. The graph comprehension was affected by the logical and abstract thinking abilities, exposure to various types of graphs, and mathematical knowledge of the students who were trying to understand the graphs.

On the other hand, Langrall and Mooney (2002) defined describing data as the reviewing of data demonstrated in charts, tables, and graphical representations precisely. The ability to skim different types of graphical displays allowed students to perform higher level processes after the describing data process such as organizing, representing, and analyzing the data (Groth, 2003). Jones et al. (2000) viewed describing data as the process of precise glancing of tables, graphs, or charts to go through presented raw data. The "reading of data" from study conducted by Curcio (1987) was similar to this process.

Mooney (2002) initially identified four sub-processes of describing data, namely: (a) demonstrating alertness of exhibited features, (b) differentiate same data in several data presentations, (c) evaluate the efficiency of the data presented, (d) distinguishing data values units. Later on, Langrall and Mooney (2002) revised the sub-processes of describing data to only demonstrating awareness for exhibited features, and distinguishing data values units.

The research on the process of organizing and reducing data had wider pool of resources in comparison to the research on describing data. It was clearly shown in the focus of various past researches that concentrated on identifying students' understanding and difficulties in regards to the measure of center and spread from scope covering specific to general (Groth, 2003).

Students in the secondary school level were exposed to the process of organizing and reducing data through the learning of measures of central tendency such as mean, median, and mod. They were being taught on the concepts of each of the measures of central tendency. However, they tend to make bad decisions when it came to decide which measures of central tendency was suitable to be used when describing data (Zawojewski & Shaugnessy, 2000). Groth (2003) agreed by stating that most students often use mean rather than median to describe data, even when the data sets were consisted of outliers that could make the mean inaccurate as an indicative of center. Student's inadequate familiarity with statistic explained the lack of usage of median. Their inability to distinguish the best options of central tendency's measures shown that they were still lacking in terms of the process of organizing and reducing data. However, there were also research that shown that some students did portray high level of thinking in the process of organizing and reducing data, by efficiently utilizing both measures of center and spread.

Organizing and reducing data comprised of intellectual activities of aligning, grouping, and compiling data (Groth, 2003; Langrall & Mooney, 2002). Classifying and reducing data was an important ability to be learnt in order to evaluate and decipher data (Langrall & Mooney, 2002). Central tendency and dispersion measures assisted the classification of data efficiently for data analyzing purpose through the identification of mean, median, mode, standard deviation, and range (Council & Committee, 2001).

Mooney (2002) stated the three sub-processes of organizing and reducing data: (a) arranging or classifying data, (b) outline data through usage of central tendency measures, and (c) expressing the measures of spread. Langrall and Mooney (2002) revised the sub-processes to: (a) classifying data, (b) describing data by measures of center and (c) outlining data using measures of spread.

The third process which was representing data involved portraying data in the format of graphs and like the other two processes earlier, was vital in evaluating and deciphering data (Langrall & Mooney, 2002). Students needed to understand how to group data and furthermore labeling and scaling to create data representations (Council & Committee, 2001). Students were able to grasp the underlying essence of data and use visual representations to convey data analysis' results to others once they learnt to construct the visual representations (Groth, 2003).

Nowadays, there were many graphical representation tools that could assist students to come out with various types of graphs to represent the raw data. The advancement in technology enable students to learn statistics using computer aided instruction and further producing graphical representations from utilizing them. Students' abilities to represent data had been studied upon by researchers either with assistance of computer software or without (Groth, 2003).

Mooney (2002) stated the three sub-processes for representing data mainly: (a) creating a data set's presentation, (b) finishing a fractional built deviant data presentation, and (c) designing a different data presentation for data shown in a given display. Langrall and Mooney (2002) revised the sub-processes to: (a) creating a data set's presentation, and (b) analyzing the data presentations' effectiveness.

Analyzing and interpreting data was the fourth process and was the most complex of all the four processes. Langrall and Mooney (2002) labeled it as trends recognition and data inference and prediction. Curcio (1987) described two levels mainly "reading between the data" and "reading beyond the data" as the base for analyzation and interpretation of data.

Groth (2003) highlighted that the analyzing data process was heavily focused by the research community as it was the most complex process out of the four statistical thinking processes. The research was focused on three strands where the first strand was the focus on students' cognition during their engagement in informal data analysis; the second strand was the focus on students' classroom social environment during the informal data analysis; and the third strand being the focus on exploring on students' impression about formal inference and its perceptive foundations. Research shown that students were able to analyze and draw conclusions of data sets from assistance of instruction, but those conclusions were mostly unrefined due to the weak process of analyzing the data sets (Groth, 2003).

Analyzing and interpreting data had sub-processes which were identified by Mooney (2002) as followed: (a) comparing inner data sets or presentations, (b) comparing among data sets or presentations, and (c) inferencing of data set or presentation. The revision made by Langrall and Mooney (2002) was the addition of an additional sub-process which is: (d) using proportional reasoning. Once the previous abilities of three processes of reviewing data displays, data organization and reduction, and data exhibition were learnt, ability to analyze and interpret data was developed upon them (Langrall & Mooney, 2002).

In this study, statistical thinking process of describing data was removed due to the reason that the measurements involved were in quantitative form. In order to observe this process effectively, the process of describing the data had to be done with questions and followed up questions by the instructor qualitatively. Students might not able to convey their thoughts with a rigid quantitative form of measurement. Revision was done to the sub-processes as well to fulfill the criteria of quantitative measurement.

The remaining statistical thinking processes were adapted and renamed accordingly as ability to organize and reduce data, ability to represent data, and ability to analyze and interpret data. Two sub-processes under the process of data organization and reduction ability were adapted which were: (a) describing data using measures of center, and (b) describing data using measures of spread.

In the process of ability to represent data, two sub-processes were adapted which were: (a) constructing a data display, and (b) evaluating the effectiveness of data displays. No revision is made on the sub-processes.

The process of ability to analyze and represent data had adaptation of two sub-processes after the revision which were: (a) making comparisons within data sets or data displays, and (b) making comparisons between data sets or data displays.

The statistical thinking framework constructed by Langrall and Mooney (2002) was established basing upon the cognitive model of Structured of the Observed Learning Outcome (SOLO) by Biggs and Collis (1982). This SOLO cognitive model was mentioned by Biggs and Collis (1982) as the theory of neo-Piagetian that explained hierarchy of levels of the increase of complexity in students' understanding.

This model was applicable to measure learning outcomes achieved in various cognitive areas and subjects, among students with different levels (Biggs, 2011). Several researchers had utilized the model to study the statistical thinking's levels. Based on Biggs and Collis' model, Jones et al. (2000) formulated the framework for the development of statistical thinking of elementary school students; Mooney (2002) constructed Middle School Students Statistical Thinking (M3ST) framework; and Groth (2003) formulated high school statistical thinking framework. The model for this study will be further described in regards to relevant parts of the model to suit framework for secondary school students.

The Biggs and Collis (1982) model involved five development modes: sensorimotor (since birth), ikonic (about 18 months old), concrete symbolic (about 6 years old), formal (about 14 years old), and post-formal (about 20 years old). Three mental levels which were unistructual, multistructural, and relational, leaned upon reprocessing to act in place for complication shifts of students' rationale in each mode. The five development modes function in the sense of its relation among each other where there was an ongoing development of previous modes. As such, two cognitive levels which were the pre-structural and extended abstract were related to the five development modes where earlier mode was linked to pre-structural and later mode was linked to extended abstract. The model emphasizes on humans using lower development modes' thinking abilities to link to higher development modes' thinking abilities to solve tougher tasks. (Biggs, 1991). The five modes of development and SOLO levels are shown in Table 2.1.

Table 2.1

Bigg and Collis SOLO Model's Modes of Development and Levels

Modes of development	Levels
Sensorimotor	Pre-structural
Ikonic	Unistructural
Concrete Symbolic	Multistructural
Formal	Relational
Post-formal	Extended Abstract

Biggs and Collis's theory of development coincided with Piaget's in the way of the three developments levels in every node. Unistructural levels was the first development level where feedback to tasks were not complete due to the reason that students use relevant aspect of the given task that they first thought of. The next level was the multistructural level where students aware and consider various aspects of task but cannot assimilate them and therefore cause incomplete responses to tasks. Students were only able to understand relationship among aspects of tasks to solve problems when they grasped the relational level (Piaget, 1983).

Langrall and Mooney (2002) considered the ikonic and concrete symbolic modes to be best suited for students from middle schools, equivalent to secondary school students. The statistical thinking framework of Langrall and Mooney associated with SOLO model where the idiosyncratic level of statistical thinking was linked to the ikonic mode's prestructural level. In addition, transitional, quantitative, and analytical levels were the second, third, and fourth statistical thinking levels and were linked to the concrete symbolic mode's levels of unistructural, multistructural, and relational.

Biggs and Collis (1982) also claimed that the occurrences of unequal functioning across modes which also known as "decalages", happened during engagement of tasks and were more common than Piaget's work. This occurrence also happened across statistical thinking processes as shown in the research report by Mooney (2002) where there was appearance of uneven statistical thinking levels pertaining to each processes. A specific level of thinking linked to a certain statistical thinking process observed from a student working on tasks will not automatically display that particular level of thinking across all processes (Groth, 2003). It was therefore important to analyze students by referring to task that integrated different statistical thinking processes to identify a more accurate level of thinking (Mooney, 2002).

2.3 Statistical Thinking Ability

The researches on statistics education had been growing tremendously over the past decade that focused on higher order thinking and one of them was none other than statistical thinking. The question comes to what is statistics? There were various definitions by researchers. "Statistics is the mathematical study of probabilities and chance events and the scientific attempt to draw conclusions from data in the face of inevitable error and imprecision. Modern statistics does not simply apply mathematical results to determine the properties of particular statistical methods; it includes a concern for discerning, describing, and confirming patterns and relationships in data" (Gattuso & Ottaviani, 2011). Moore (1998) defined statistics as crucial and important methods that applied to variation, data, and probability everywhere at the same time in modern life. Statistics can be considered as a technique for investigating collective phenomena quantitatively.

The modern society required citizens to be competent and critical in their everyday life to analyze and interpret the phenomena surrounding them to make informed decisions and solve problems. When there was quantitative data and information required to be analyzed, statistics and statistical thinking came to mind with their important role (Gattuso & Ottaviani, 2011).

Students were exposed to statistics nowadays at young age where various statistical activities can be identified in their curriculum. Usiskin and Hall (2015) highlighted the incorporation of statistics with Mathematics in the curriculum due to growth of literacy, science, health, and social science. Statistics was introduced as one of the topics for Form Four Mathematics in the secondary school curriculum in the country of Malaysia. It consisted of the learning of class interval's concept, central tendency measures' concepts such as mean, median, and mode, graph representation and data interpretation's concepts of histograms and frequency polygons, dispersion measures' concepts, and cumulative frequency concepts.

Statistics was recognized for its importance as it brought about various benefits to students. Gattuso and Ottaviani (2011) stated that statistics utilized the posing of questions, analyzing, representing, and conveying quantitative information to trigger motivation and promote problem solving abilities. Students' involvement in the learning of mathematics was related to how well statistics was able to bring out students' interests. Furthermore, the involvement of questioning and finding ways to answer them in statistics allowed students to go through the process of data collection and analysis which will enrich students' thinking even at elementary level. Schwartz (2006) agreed by stating that students as young as kindergarten level can also exercise the ability to work out a question and be analytical about it.

Statistics' effects to the school curriculum can also be seen in the statistical analysis process it contained. The processes of comparing between groups, analyzing the distribution of the data, determining the clusters, outliers, investigating the measures of center such as medians, means, mode, and measures of spread such as range and standard deviation allow students to analyze the data thoroughly by going back and forth in finding solutions to the problems. Last but not least, the representation of the data in statistics was vital in assisting students to visualize statistical distributions and prove the existence of relationship among the variables studied. The representations will reveal how well the students understanding of the distribution (Gattuso & Ottaviani, 2011). This is where the concept of statistical thinking came into the picture to get a clearer picture on the analysis and interpretation of the data.

Statistical thinking's role in modern days cannot be overlooked as the digitalization era involves circulation of information in the form of data where it was necessary for students to acquire it for information interpretation. The huge demand and call for statistical thinking abilities to be incorporated in statistics education was backed by arguments that the traditional methods of teaching statistics were overly focused on procedures, skills, and computations which did not promote development of students' statistical thinking (Ben-Zvi & Garfield, 2004). Students' learning of statistics was pretty straightforward, where most of the time the results were exposed to them on the board without showing the trials and errors behind the solution, leaving the students clueless as they were refrained from the thinking processes required in statistics in analyzing and interpreting the data and furthermore in arriving at results (Gattuso & Ottaviani, 2011).

Garfield et al. (2015) highlighted that statistical thinking had been promoted by statisticians in business, industry, and academia for the past three decades as the important ability for students to understand and learn statistics. In the 1990s, some outstanding statisticians had voiced out to call upon the statistical thinking's development among students (Moore, 1990). The American Statistical Association (1991) supported this move by stating the necessity for high school students to engage in exploring data, learn the techniques used for making formal inference, scheduling their lessons and studies, and analyzing on how to utilize statistics in the society. The American Association for the Advancement of Science (1993) also agreed with the move by emphasizing that graduated students ought to be able to analyze studies critically using proper statistical techniques. As a result, there was a big change in the documents of statistics education through the wide shift of focus onto statistical thinking and its incorporation to ensure students were enhanced and with this important ability.

Within the past decades of research studies that focused on teaching and learning of statistics, the research base of high school and secondary school statistics started to grow where there were researches which focused on developing models that explained students' statistical thinking (Groth, 2003). Various statisticians had followed the lead and started focusing their research on statistical thinking (Ben-Zvi & Garfield, 2004).

The research of statistical thinking had started to grow in numbers where statistical thinking research were conducted on various areas related to students, teachers, curriculum, etc. Garfield et al. (2015) agreed that more emphasis had been placed on the research on students at elementary and secondary school levels on the learning of key statistical ideas such as center, variability, and distribution to understand their statistical thinking. Among the highlighted agendas for the emphasis of statistical thinking were: more data and concepts' usage, less focus on theory and recipes, and encourage active learning (Pfannkuch, 2008). Cobb et al. (1991) mentioned that research-based frameworks that described student's thinking were important for reforming curriculum and instruction. Jones et al. (2001) utilized a research-based statistical thinking framework to encourage instruction at school level and managed to successfully showed the effectiveness of supplying teachers with knowledge about student' thinking in learning statistics.

Aside from students developing statistical thinking, teachers also should have the knowledge on the processes of students' thinking (Even and Tirosh, 2002). Current emphasis in secondary school was on teaching students the concepts of mean, median, mode, and construction of graphs rather than teaching them on how to think with data (Friel, O'Connor, & Mamer, 2006). Pfannkuch (2008) agreed by stating that the use of statistics and statistical thinking were not observed in classrooms, where students were restricted in exploring statistics to analyze data. Furthermore, the idea and knowledge to develop students' comprehension of statistical ideas and statistical thinking were not fully understood by researchers and teachers. The school curriculum had just started to incorporate ideas of statistical thinking and thus teachers ought to enhance their knowledge of statistical thinking to integrate them into the lessons for the students. Garfield et al. (2015) stated that although statistical thinking had been related to the goal of statistics instruction, there was a lot things unknown about the best way to teach and assess statistical thinking yet. It was without doubt that teachers played a vital role to demonstrate and facilitate teaching approach that will bring upon awareness on the purpose of statistics to the students. Teachers were emphasized to show abilities in having contextual knowledge about the situation to know where the data originated from and how the data will be interrogated, collected, and measured (Pfannkuch, 2008).

Statistical thinking had been a goal of statistics education for quite a long time, but there was no consensus about what statistical thinking is. Statisticians and textbook writers had come out with various definitions for statistical thinking, with no consensus to date (Garfield et.al, 2015).

The "statistical thinking" term was utilized in various articles and journals and came in various definitions as well. The usage of the term tended to overlap with the other two terms which are "statistical reasoning" and "statistical literacy" (Chance, 2002). Due to the lack of ways to identify the students' development of the ability to think, several statisticians had attempted to outline the meaning of statistical thinking.

Statistical thinking was referred to as the method where statisticians solve statistical problems by thinking about data and variability (Garfield & Ben-Zvi, 2008). Snee (1990) defined the ability as process of thoughts which identified the existence of variation around its presence in everything done, and every tasks were interrelated chain processes, and recognizing, organizing, computing, limiting, decreasing dispersion provided room for growth. Moore (1990) explained that the important elements included the dispersion of data processes, architecture of data production, measures, and interpretation of variation.

Garfield, delMas, and Chance (2003) defined statistical thinking as comprehension of processes and ideas behind statistical investigations and how they are being conducted. Statistical thinkers were able to understand context of problem to investigate and draw conclusions once they had collected and analyzed the data. The understanding of methods involved along with appropriate data analysis allowed the evaluation of statistical study's results to infer the representation of sample to the populations.

Chance (2002) highlighted that these definitions of statistical thinking suggested that there was a more global view of the statistical process, including comprehending the variability and statistical process as whole, which should be instilled in students. It was assumed that statistical thinking was only developed by statistician through various practice, experience, and working with other statisticians. Lately, there had been more and more calls for statistical thinking development to be instilled in novices such as students in the mental habits and problem solving skills required for thinking statistically.

Statistical methods were rarely being applied and only being used in limited situations. Most of the time the problem appeared in a rather direct questions such as "constructing the histogram to show the characteristics of the data", "perform a t-test to show the statistically different means". This type of approach only enabled students to have limited narrow view of the statistical application and not utilizing any problem solving or thinking abilities (Chance, 2002).

The students should be encouraged to view statistical process as whole by assigning them to assignments like projects where the students were required to come out with plans to formulate data collection, executing the data collection, data analysis, and data interpretation. The students were free to choose and decide whichever techniques they deemed appropriate to yearn the best solution for the tasks assigned. Aside from letting the students run wild with their imaginations for the projects, it was also best to provide feedback so that the students were able to learn from experience and apply the knowledge to other problems (Chance, 2002). Aizikovitsh, Clarke, and Kuntze (2014) stated that tasks assigned by teachers can serve as context to determine how students think before and after instruction, and hence should be focused on ensure that the tasks able to direct students in specific content's aspects and specify the information processing.

In terms of statistical thinking framework or processes, Ben-Zvi and Friedlaner (1997) highlighted patterns of statistical thinking in four modes: (a) Mode 0: Uncritical thinking, (b) Mode 1: Purposeful usage of a representation, (c) Mode 2: Purposeful approach of multiple representations: developing metacognitive abilities, (d) Mode 3: Creative thinking.

Wild and Pfannkuch (1999) developed a statistical thinking framework that consisted of a four-dimensional framework: (a) Dimension One: The Investigative Cycle, (b) Dimension Two: Thinking Types, (c) Dimension Three: The Interrogative Cycle, (d) Dimension Four: Dispositions and came to the conclusion of statistical thinking having the definition of "what a statistician does" and to have characteristics of instantly question and probe into involved issues and data.

Langrall and Mooney (2002) developed and validated a statistical thinking framework that involved four processes mainly describe data, data organization and reduction, data representation, and data analyzation and interpretation along where each process have four levels with rubric and descriptors to serve as guideline. Langrall and Mooney's statistical thinking framework will be adapted and used in this study to assist in the assigning of scores based on the descriptors of each level in the framework.

Standard-Based Curriculum for Secondary Schools (KSSM) was implemented in phases throughout all the secondary schools in Malaysia which encompassed the Content Standards, Learning Standards, evaluation and pedagogical approach. The Standard Document of Curriculum and Assessment suggested that teachers to use statistical enquiry approach (Ministry of Education Malaysia, 2018). The statistical enquiry approach highlighted was similar to the statistical investigation cycle by Langrall and Mooney's statistical thinking framework for developing statistical thinking.

2.4 Gender

Gender was defined as the variations in a male or female's scholastic achievement involving psychological and socio-cultural dimensions (Ewumi, 2012). The literature on gender and the differences were portraying more on the disadvantages of women back then. However, this statement was not so valid today, as more research showing that girls and women had more advantages in certain aspects of education (Buchmann, Diprete, & McDaniel, 2008).

Research on gender and its differences focused more on the academic performance of secondary school populations compared to young children. Academic performance had been measured all this while by either standardized tests and other assessments or school grades, where element of academic performance and abilities of the students were the indicators on how specific gender of male and female performed academically. (Buchmann et al., 2008).

The research on gender were overwhelming in the education area. A comprehensive analysis of 1,600 research in the achievement, personality, and social relations areas were identified on gender differences (Maccoby & Jacklin, 1974). Despite the wide amount of literature in this particular area, there were still disagreements in terms of gender differences, even in mathematics performance as well (Leahy & Guo, 2001). There were various arguments on whether males were better than females in terms of measures of achievement, or whether this gender differences in test scores were not significant anymore. Hyde et al. (1990) argued that gender differences in test scores had been narrowed down for the recent decades. Hedges and Nowell (1995) agreed with the statement by stating that the gender differences in test scores had been stable for the past 30 years.

The research on gender in terms of their learning of mathematics and statistics had been the highlight for quite some time. Carr and Davis (2001) mentioned that there was evidence of gender differences as early as the first grade in the learning of mathematics, where it can be observed through comprehensive analysis of the learning processes of children. Fennema et al. (1998) agreed by stating that girls in the first grade up to second grade were prone to concrete solution strategies while boys were prone to abstract solution strategies from a longitudinal study conducted and observed. Klein, Ali-Japha, and Hakak-Benizri (2010) stated that students of male and female in secondary school level showed differences in mathematics performance. Liu and Wilson (2009) agreed and added on that gender differences can be observed from the large scaled PISA mathematics assessments that assessed students of 15 years old, where more male students were performing better than female in that assessments.

In terms of mathematics achievements, there were studies that showed evidence of gender differences in related to spatial skills and verbal skills. Gallagher et al. (2000) stated that men were outperforming women with great difference on spatial skills related questions in an analysis of tests and slight difference on verbal skills related problems. Boys and girls tend to choose different type of strategies when it came to solving mathematical problems. Geary et al. (2000) agreed by stating that boys tend to prefer spatial imagery method of solution compared to verbal computation. Grimm (2008) added on by stating that girls shown higher frequency in the level of reading in the early elementary school and mathematics' progress compared to boys. These findings supported the fact that there was high association of test scores in terms of arithmetic skills and verbal memory from the analysis of the achievement of students on the arithmetic subset of the Wechsler scales (Karzmark, 2009).

There were various researches that studied on the gender differences in terms of statistics, which focused on identifying which group of gender, either male or female was superior. Chiesi and Primi (2015) stated that there was a gender difference when it came to students' attitude in learning statistics. It was highlighted that female might have less mathematical aptitude than male, where they showed less confidence and more negative attitude toward statistics. However, they were also research that showed that females were better than males in terms of learning statistics (Mahmud & Zainol, 2008). Gallagher and Kaufman (2005) agreed that boys and girls nowadays were equally demanding in terms of mathematics lessons in school, with noticeable observation that girls obtained better grades compared to boys. There were also research that revealed no gender difference of male and female in terms of statistics learning (Judi, Ashaari, Mohamed, & Wook, 2011). Spelke (2005) agreed that the gender differences in terms of cognition was so small which led to the conclusion that males and females had same aptitude for mathematics.

Aside from the highlight of the gender differences in the mathematical learning, there were various research on gender on their learning in terms of other education area. There were researches on gender's development that showed that it was affected by adolescents' perception and impersonation, where their act of good gender behavior was to be rewarded and bad gender behavior to be punished (Ewumi, 2012). However, the research on gender differences in terms of cognitive processes and abilities were lacking. A theory proposed to explain the traits existed in the two groups of gender suggested that male representing cognitive domain while female representing emotional domain (Hall & Lucas, 1976 in Klein, 2004).

Spelke (2005) summarized the patterns identified in cognitive differences of gender by stating that girls and women were more likely to be good at tests of verbal fluency, arithmetic calculation, and memory for the spatial locations of objects. On the other hand, boys and men were more likely to be good at texts of verbal analogies, mathematical word problems, and memory in terms of environment's geometric configuration.

In searching explanations on gender differences, Huber (2008) highlighted that biological differences can be a factor as well aside from social and economic factors. Biological differences in gender might contributed even a small part to the differences identified in male and female. However, this biological explanation is quick to be labeled as sexist although it is not necessary needed for a 'smarter sex' with 'better biology' to make conclusion that any cognitive ability to be linked to biological origins due to the fact that gender differences in cognitive tasks were quite well established (Halpern et al., 2005).

Hyde and Mezulis (2001) regarded that cognitive variances between males and females overlapped each other. For instance, Hyde and Else-Quest (2012) highlighted the overlaps in distribution of male and female's scores in mathematics and visuospatial tasks where boys achieved better performance than girls in mathematics and science but girls were better in overall grades and reading in comparison to boys. Tach and Farkas (2006) agreed by stating that girls had better skills in reading as early as kindergarten compared to boys and this observation still existed even up to elementary school level where boys continued to face problems with reading. Trzesniewski et al. (2006) highlighted that boys' reading difficulties were linked to other disadvantage such as antisocial behavior, whereby there was a relationship between those two in affecting each other.

Boys were known to have issues of reading difficulties, anti-social behavior, attention disorder, delayed and stuttering conversations (Rutter et al., 2004). Gender differences were trivial and neglect able when it came to anti-social behavior that originated from social relationships context, but boys still at risk of anti-social behavior that was neuro-development in origin (Moffitt et al., 2001). Girls were also better than boys in the aspect that they were putting more effort compared to boys in completing tasks and demonstrated eagerness to learn (Downey & Vogt Yuan, 2005). Moreover, girls in their adolescent stage showed higher levels of attentiveness and organizational skills, self-discipline, leadership qualities and interest, where all these traits promoted their academic success (Rosenbaum, 2001). There was a huge gap in boys and girls in terms of their abilities back then. Nowadays, boys and girls were more competitive in enrolling mathematics classes in high schools, though girls tend to obtain better grades than boys (Gallagher & Kaufman, 2005). Nonetheless, there were still instances where boys still obtaining better scores in certain aspects of domains compared to girls (Buchmann et al., 2008).

There were questions on gender in how it was affected by gender stereotypes in ways such as perceptions and socialization among the boys and girls. Students tend to perceive their own abilities due to stereotypes. For instance, due to stereotypes of women performing poorer on mathematics tests, women tend to experience anxiety when taking standardized mathematics tests which resulted in poor test performance portrayed from them (Steele, 1997). Macher et al. (2012) agreed by mentioning that male students were more favorable in terms of gender effects, as they had less statistics anxiety.

There were also debates on whether teachers favor a certain gender over the other. Research conducted based on classroom observation back in the 1990s were talking about how teachers favor boys over girls. It was said that recently it was the other way round where schools favor girls over boys (Sommers, 2000). There was no conclusive evidence in whether gender of teachers was the cause for the gender differences, where several studies identified no difference in males' performance when being taught by male teachers in comparison to female teachers (Sokal et al., 2007). Dee (2006) argued that there was indeed gender difference in terms of achievement when boy and girl students were taught by different gender of teachers. The girls' achievement was raised while the boys' achievement was lowered when taught by a female teacher subjects of science, social studies, and English. Nonetheless, it was unclear whether the result was caused by gender bias in teaching or that students tend to learn better from teachers of the same gender (Buchmann et al., 2008).

There was evidence that showed that gender gap was linked to family background. Before the mid-1960s, educated parents were not showing much favoritism on their sons where daughters were still able to have unity with them. However, that was not the case for less educated parents where they showed favoritism towards their sons compared to their daughters, hence the gender gap identified for college completion was seen the largest in favoring male among the families with less educated parents. After the mid-1960s, favoritism on sons showed a decline and for some cases there were even reversed situations where less educated parents or single mother showing more favoritism for daughters compared to sons (Buchmann & DiPrete, 2006).

Parents' socio-economic status also somehow influenced the gender gaps. There were several evidence suggesting that gender gaps in academic scores were noticeable among low income students (Hinshaw 1992). For instance, Entwisle et al. (2007) identified that girls were better in reading compared to boys in fifth grade, when both coming from low income families; while boys and girls coming from middle and high income families had same reading scores. Buchmann et al. (2008) highlighted that some research showed differences in gender of the child due to parental involvement, where parents played a role in the educational performance and attainment of children. Stevenson and Baker (1987) agreed that parents do affect the gender differences as parents had more involvement with daughters for home related activities and had more involvement with sons for school related activities. However, parents' involvement with sons declined as their sons grew older but not the case for daughters as their involvement remained the same. On the other hand, Muller (1998) disagreed by saying that parents' involvement with their children were not gender specific and did not contribute to the gender difference about mathematics.

Entwisle et al. (2007) added on that the large gender gap in terms on reading score among low income students of first and fifth grade were due to parent' lower reading expectations of boys. Gender differences were also identified to be caused by certain parenting styles was authoritative lacking (Mandara, 2006). However, Buchmann et al. (2008) argued that parenting styles and expectations were not justified to be deemed as the cause for gender differences, but only as one of the result. Parenting did affect children's gender in a way such as role modelling, where boys looked at fathers as an example while girls looked to their mothers as example to develop aspirations in terms of education and occupation.

2.5 Socio-Economic Status

Manstead (2018) viewed socio-economic status as the economic position and educational fulfillment of an individual. He added on that people perceive the terminology of socio-economic status rather than social class nowadays when it came to the topic of identities. This can be seen where people had reasonable sense of their status economically and educationally in comparison to others, and that the traditional boundaries existed between social classes had become less apparent. As such, there were more literatures that focused on indexing income and educational fulfillment for socio-economic status, rather than indexing means and production for social class.

Socio-economic status was often being related to students' academic achievement. Perry and McConney (2010) highlighted that the socio-economic status of students was strongly being associated with educational achievement locally, nationally, or internationally as per measured by the standardized assessment systems, and that this statement was well established in various educational literature of research. Sirin (2005) agreed by stating that the academic achievement of students was associated with the socio-economic profile of a school as shown in many international studies conducted.

Various past research had demonstrated the importance of parental education and family resources on the fulfillment of education of an individual. A family background with its resources such as financial capital, social capital, role models' access, information and adviser, attitudes of individuals, and previous academic performance had influence towards children's upbringing and were the determinant for educational inequalities. (Buchmann et al., 2008).

Benner, Boyle, and Sadler (2016) agreed by stating that parental educational involvement played a major role in students' academic achievement and performance in primary and secondary school. Hill and Tyson (2009) agreed by stating that students' academic outcomes had been linked to such involvement which was parents' active participation in the children's educational lives. It was also highlighted that parental educational involvement came in three different aspects which were home-based involvement, school-based involvement, and academic socialization.

First of all, home-based involvement consisted of the involvement of parents at home environment on their children such as monitoring and checking children's progress on their homework, and conduct educational activities from time to time to promote learning and enrichment. Second, school-based involvement consisted of the involvement of parents in school related activities such as participating parent-teacher conferences to promote communication with the teachers and other parents, volunteering to help out in the school for instance helping out at the canteen or the bookstore, and involved in school organizations. Third, academic socialization consisted of parents' involvement in terms of promoting their children's future educational and occupational development, in which parents passed on indirect messages to the school to voice out their educational expectations for their children, and thus allowed them to provide their children with the necessary tools for independence and educational success (Hill & Tyson, 2009). Woolley and Grogan-Kaylor (2006) agreed with the aspects of parental educational involvement and stated that students' academic success was affected by parental educational involvement, in which the involvement contributed to students' better performance on tests, higher scores and grades in examinations.

The importance of home environment on students' scholastic achievement was notable for its role in promoting big impact on students' psychological, emotional, social and economic state (Ewumi, 2012). Parents acted a major role in providing a home that stimulated the mental processes of students. Fafunwa (1998) highlighted that children's development of intellectual skills was a task ought to be performed by home. Intellectual skills can be developed if parents pay attention to children's education through providing academic needs and needed moral support. Parents with low income often viewed the education of children as the responsibility of teachers while middle and upper income parents tend to view education as a joint responsibility of parents and teachers (Ewumi, 2012). Nonetheless, some lower income family parents did make up to it by putting it more effort in engaging enrichment activities that can stimulate their children's progress and development and improve their cognitive abilities and social skills (Laueau, 2003).

Englund et al. (2004) identified that children who were successful academically in the early stages in school had higher expectations from parents in comparison to children who were not successful academically, in which these differences can be seen throughout elementary school. Parental educational involvement certainly gave a big impact on affecting the children's academic performance. Carolan and Wasserman (2015) agreed by adding on the statement that parents' expectation towards children were influenced by their socio-economic status, in which lower income parents and less educated parents were less likely to put high educational expectations on their children in comparison to higher socio-economic status parents. Low socio-economic status parents also involved less on their children's well-being in terms of their education and performance in school (Cheadle & Amato, 2011).

Similarly, children's prior achievement levels encouraged their parents to engage in activities involving their children in terms of educational involvement to bring forth benefits to their children without bias from parents' side. Despite the fact that there were sociodemographic markers to restrict the involvement of parents in terms of educational involvement activities due to parents' education levels, children's past performance may influence parents to take part in guiding their children in order to not hinder the children's future potential (Benner et al., 2016).

In addition, there were evidence that the relationships between parental educational involvement and children's academic performances might be influenced by various aspects of family socio-economic status. It was more distinguished to see stronger link for lower socio-economic status children compared to higher socio-economic status children in terms of parents' home-based involvement and academic socialization along with engagement with children (Benner et al., 2016).

McConney and Perry (2010) highlighted that socio-economic status may vary for different students, schools, and nations, in which past studies had shown evidence of students and their school socio-economic status strongly linked to their performances which varied across countries. It was highlighted that it was due to diverse schooling characteristics based on context existed in every nation. For instance, it was observed that there was stronger association between school socio-economic status and student achievement in countries which had higher concentration of high socio-economic status in smaller number of schools. Furthermore, it can be observed as well that there was stronger association between school socio-economic status and student achievement in countries that had higher resources and funding to their schools.

Low socio-economic status areas of schools tend to have lower achievement students, lower school graduates, and lower higher education students due to the reason that these areas were lacking in resources and expertise in comparison to higher socio-economic status areas of schools (Ewumi, 2012). Spring (2017) supported the claim by mentioning that schools in higher socio-economic status areas focused on enhancing the thinking abilities of students while schools in lower socio-economic status areas focused on rote-learning of students.

Schools were supposed to be a place with suitable learning environment that enable students to utilize their abilities and effort to pave their success in achievement. However, that was not the case as there was existence of inequalities in socio-economic status or social class that promoted advantages for class background (Manstead, 2018). However, Bourdieu and Passeron (1990) argued that social inequalities were reproduced by the school system by encouraging norms and values that were more suitable to middle class socio-economic status children, which extended to aid them to outperform their working-class peers. Stephens, Markus, and Philips (2014) further highlighted that children who grew up from middle class socio-economic status family needed not to worry about the family having financial issue in comparison with low class socio-economic status family.

Manstead (2018) highlighted that a person's socio-economic status reflected their upbringing on how they grew up in what sorts of environment can have a lasting influence on their thinking. Stephens, Markus, and Philips (2014) agreed by stating that social class or socio-economic status could result to patterns of thinking, in which they shape the self-concept through home, school, and work.

2.6 Past Research on Statistical Thinking

Scranton (2013) had conducted a research to examine statistical thinking of middle school students in a technological studying environment using TinkerPlots software program comparing to the traditional paper and pencil environment. TinkerPlots was a construction tool software program that allowed middle school students to conduct data analysis in a technological environment by using the tool to construct own graphs and manipulate them by organizing data into various type of representations without requirement of complicated statistics knowledge. The types of critical, creative, and statistical thinking of the students in both environments were characterized. The research questions highlighted in that study was: (i) When the traditional tools of paper and pencil was used, what was the critical and creative thinking of the middle school students in comparison to when using Tinkerplots software program? (ii) When the traditional tools of paper and pencil was used, what was the statistical thinking of the middle school students in comparison to when using Tinkerplots software program? Qualitative research design was utilized featuring multiple case studies. The population consisted of middle school students in the grade 6 and 7 from various socio-economic status backgrounds. Two different grades were selected to cover the thought processes of wider range. The participants selected for the case-study analysis were twelve students with six from each grade. Pilot study was conducted for one day per week consecutively in order to develop the tasks for the study and, observe the interaction from students with peers and the software, and determine the suitable time to be allocated for the students to use the software. The instrumentation which were the tasks consisted of questions that required the TinkerPlots software method and questions that required paper and pencil method for the solution, in which both tasks required the students to comprehend, organize, represent, analyze, and interpret data.

The focus of the tasks was to observe the critical, creative, and statistical thinking of the students when solving the tasks assigned. The findings revealed that students who utilized the paper and pencil task spent more time in creating representations than interpreting or analyzing the representations constructed. Students also did not identify the causal relationships of the data. On the other hand, most students in the TinkerPlots task constructed several complete representations that were easily modified and represented the data along with added context of both concretization and causal relationships to the data. The findings indicated that the statistical thinking of students were enhanced when they utilized Tinkerplots software.

Le (2017) conducted a research to evaluate the statistical thinking of students who signed up for introduction course statistics in hope of answering the research questions by revealing components of students' statistical thinking developed. An assessment of statistical thinking, also known as Modeling to Elicit Statistical Thinking (MODEST) was developed along with the test blueprint. Modifications were made to MODEST based on the feedback collected by reviewer. The assessment was then pilot testing to two cohorts of students, where one cohort consisted of senior, undergraduate students who major in statistics, while the other cohort consisted of students who enrolled in introductory statistics during a short semester. The responses from both cohorts of students were checked to determine whether statistical thinking was elicited from the items in MODEST assessment. The final revision was then administered online twice to 88 students who enrolled in introductory statistics course in the following year's semester, one at the early of the semester while the other one at the end of the semester. Both responses' from first administration and second administration were compared to answer the study's research questions. A test blueprint was created for the purpose of identifying the statistical thinking's characteristics for MODEST to measure, and the also for analyzing of the data collected in this study. The test blueprint consisted of four components of statistical thinking which were general problem-solving characteristics, statistical problem-solving processes, statistical problem-solving cognitive processes, and individual dispositions. In addition, the items in the assessment were developed based on the rubric blueprint. The obtained data were then analyzed by assigning points based on the components of statistical thinking, with 0 point given to unanswered, missing responses, or irrelevant responses. An overall score of statistical thinking were obtained based on several method of scoring. Aside from that, confidence intervals were computed using bootstrap percentile method to understand students' statistical thinking development. The findings indicated that students who entered the course had moderate statistical thinking's amount and gradually developed them throughout the course

Koparan and Güven (2013) carried out a study to analyze the statistical thinking levels of primary school pupils using a statistical thinking model and examine the differences among their statistical thinking levels. The sample of the study were 90 primary school pupils who were in their 6^{th}, 7^{th}, and 8^{th} grades. The study utilized developmental research method as the research design. Questions of open-ended and multiple choice were used by adapting questions from research literatures and getting feedbacks of experts. Students' responses were then analyzed according to the statistical thinking model to determine the levels of students. The result revealed the students' thinking differences between levels using qualitative data. Students are categorized in the first level in the higher levels of the statistical thinking model.

Pfannkuch and Rubick (2002) conducted a study to explore the statistical thinking of students with given data. Twelve-year old students were selected as the participants for this study to explore on their building up of recognition and understanding of relationships in a set of data. The students were given materials such as calculators and graph paper in the pencil and paper environment provided during an interview which was conducted for an hour. They were asked and prompted to justify their actions while thinking aloud during the interview session. In addition, the students were pushed to further think and reason on their given responses. The findings indicated no consistency in students' reasoning from the interview sessions conducted. The statistical thinking of the students across three representations of cards, tables, and graphs were analyzed and identified from the context of transnumeration, consideration of variation, reasoning with statistical models, and integrating statistical with contextual. The findings revealed several issues to be addressed for identifying the process of meanings construction from data of the students. The highlighted issues were previous contextual and statistical knowledge, thinking at higher level than constructed representations, representing and constructing actively, local and global thinking's connection, and the statistical thinking dialogue's changes across the representations.

Perry et al. (2000) conducted a study to determine the statistical thinking of elementary school students. The population for this study consisted of three groups of 20 students each. Four students were purposefully selected each from grades of 1 to 5. Students' selection was also based on teacher assessment and students' mathematic achievement, in which one student from both lower and upper quartile, and two students from the middle quartile of each grade level. The participants have variation in their age due to arraignments of different starting school timing of the students. The instrument utilized for this study was the Statistical Thinking Protocol which consisted of seven open-ended questions that meant to identify students' statistical thinking in terms of describing the data, organizing and reducing data, representing data, and analyzing and interpreting data. While the session was conducted, the research audio taped the students' responses for transcription, and collect students' drawings and graphs. A coding rubric was adapted from previous study to enable the researchers to code each question according to its construct and the portrayed students' level of thinking. The findings revealed that students' thinking levels across four constructs were generally increasing or remained constant along with the increase in their grade levels. Analysis of the statistical thinking of the students showed that students had similar level of thinking across at least three of the four constructs, with higher median on the first three levels of statistical thinking levels.

Sun and Buys (2010) conducted a study to evaluate a model for postgraduate level statistics course in university that was designed to develop the statistical reasoning and thinking of students. The model utilized for the study was called "Statistical Thinking Learning Environment" (STLE). The study utilized a cross-sectional research design with mixed method approach to evaluate the effectiveness of the STLE in developing the statistical thinking of the students. The sample of the study consisted of 10 students for the qualitative study and 90 students for the quantitative study. The students selected as the sample were undertaking either honors or postgraduate coursework program and were at the end of their statistics course study. Five students from previous semester and another five students from current semester were selected to participate in the interviews. Individual interviews were conducted to identify the opinions of students on the implementation process of STLE in statistical teaching. On the other hand, 90 students were approached and invited individually to complete questionnaire sent to them by hand, or by online via Survey Monkey website. The questionnaire consisted of survey to collect data on STLE approach and its relationship with the critical thinking and statistical thinking of the students. Out of the 90 students' responses, 34 useable responses were obtained. The students were of various nationality where 60% were Australian, and 40% were from other countries such as Taiwan, China, Indonesia, Malaysia, India, etc. The gender ratio of the samples was approximately equal. The STLE model was developed for the intention to enhance students' comprehension of statistics, and their ability of statistical thinking. In order to obtain qualitative data, semi-structured interview questions were utilized to obtained responses on the STLE's effectiveness and its relevance to students' critical and statistical thinking. In order to obtain quantitative data, questionnaire consisted of 20 items that measured STLE along four scales: critical reflection teaching method, research based course content, problem based course design, and methods for inspiring the learning of students. The items consisted of six-point Likert scale for the students to rate them. The items were validated using Cronbach alpha level of reliability which indicated strong reliability. A factor analysis was conducted and revealed high level of variance of the 20 questions. In addition, a descriptive analysis and inter-scale correlational analysis was conducted on the four STLE scales. A regression analysis was conducted as well to analyze the relationship between STLE and each of the variables in student learning enhancement, critical and statistical thinking, and social interaction's facilitation. The scores for items for each scale were averaged to find out the respondents' scale scores. The findings indicated that STLE had influence in the increase of students' statistical thinking and enhancement of their learning. It was concluded that STLE model was effective in ensuring that students were engaged in inference testing of data, and actively explain and discuss reasoning statistically.

Langrall and Mooney (2002) conducted a study to develop a framework to characterize the statistical thinking of middle school pupils. The Middle School Students' Statistical Thinking (M3ST) framework was developed and validated based on the observation and analyses of students' thinking and from past literatures. The M3ST framework consisted of four statistical processes mainly describing data, organizing and reducing data, representing data, and analyzing and interpreting data. Descriptors for each of the sub-process of the framework are developed based on a protocol to evaluate students' thinking. Clinical interviews were conducted based on protocol to assist in the development of the framework and its four levels of processes to collect responses to be categorized in four categories mainly idiosyncratic, transitional, quantitative, or analytical.

Groth (2006) done a study to explore the statistical thinking of students. The sample of the study were fifteen students aging from 14 to 19. Clinical interview with duration of approximately two to three hours for each students was conducted to collect responses from students using several statistical thinking tasks. The statistical thinking tasks covered a few concepts which were the law of large numbers, linear data transformations, and averages. The interview question for the concept of law of large numbers were meant to uncover students' comprehension of the law of large numbers' foundation. The first part of the tasks questioned students' observation in the difference between the mean of two random samples with different size. The second part of the tasks were meant to identify whether students' understanding on the difference between the mean for smaller random sample in comparison with larger random sample. The interview showed that students were oblivious in the mean difference for two samples although the students had completed a year of statistics course in college level. Next, the interview questions on linear data transformations meant to identify students' thinking on the center and spread of data after the data had gone through transformations. The interview session showed that students had misconception about data transformation in terms of the center and spread of the data. Interview tasks were also conducted on determine student's comprehension of the concept of averages. An item was designed to show 20 annual incomes, with one huge income as an outlier, and the mean income not portraying the typical income of the median income. The findings revealed that students seemed to only believe that mean was the only method to calculate the typical income and some students totally ignored the outlier of huge income. In conclusion, the findings indicated that there was inconsistency in students' intuitive understandings on the concepts in statistics such as the big numbers' law. Attention and assistance were needed to be given by teachers to students on how to design and evaluate appropriateness of experimental and non-experimental studies other than relying on statistics course and textbooks.

The key findings of past research highlighted the topic of interest on identifying the statistical thinking abilities of students of different age and education levels from elementary up to postgraduate. The findings from past research shown consistency in highlighting the issue faced in measuring the statistical thinking abilities of the students. Students were identified to still have misconception in mathematical concepts and display lower to average ability of statistical thinking abilities. The unavailability of the statistical thinking lessons and activities and the suitable instrument to measure the statistical thinking ability were deemed as gap. Past researches on statistical thinking showed that most students do not develop high level of statistical thinking as they are not focused on interpreting and analyzing. Students' intuitive understandings on the concepts in statistics are lacking. There was also gap of study where there was a lack of research on statistical thinking conducted in Asia and Malaysia.

2.7 Past Research on Gender

The study of gender was prioritized and focused mainly on education and psychological recently where the focus was on gender aspects and thinking in the last two decades**.** Penner and Paret (2008) carried out a study to identify the effect of gender differences on mathematics achievement of young children. The sample selected for the study was of a large size featuring a national wide personification sample of kindergarten students of class 1998-99 in public or private school in United States. The collection of data featuring longitudinal method was done at six points in time as those children advanced from kindergarten to fifth grade. The gender differences were examined using quantile regression models. The students' mathematics achievement was measured using uniform mathematics test featuring cognitive abilities. The findings revealed that boys initially perform well in the top distributions but worse in the bottom distributions in comparison to girls. The mathematics achievement of boys surpassed girls throughout the distributions by third grade. The result indicated the impact of some independent variables on the gender differences such as the students' socio-economic status and ethnicity. It was found out that the male students performing better than female students came from families with high socio-economic status backgrounds as well, proposing that gender dynamics in socio-economic status families have large influence on the segregating the gender in terms of science occupations. Gender differences were also observed for kindergartners in terms of race, where large male advantages were portrayed from Asians families while large female advantages were portrayed from Latino families in the top of the distribution. The quantile regression model revealed similar result, showing emergence of gender gap as early as kindergarten and will affect students' position in the education system up to pave way for achievement gaps found in high school.

Klein et al. (2010) carried out a study to identify the effect of gender on mathematics achievements, verbal, and spatial. Eighty kindergarten children (40 boys and 40 girls) aged 5 to 6 years old from urban central region of Israel were selected as the sample for the study. Most of the kindergartners had parents with qualifications of high school graduates and coming from middle socio-economic status neighborhoods. KeyMath test was utilized to determine the mathematical thinking of the children. The test was adapted to feature content of rote and rational counting and was administered individually with time allocated for test duration of approximately 30 to 40 minutes. The KeyMath test consisted of 14 subtests for instance addition, subtraction, multiplication, division, etc were divided into three major areas. The responses from the test were coded as correct or incorrect. Aside from the mathematics achievement test, verbal tests and spatial tests were conducted as well to evaluate the kindergartner's verbal ability and spatial skills. In terms of verbal tests, three tests were included which were the vocabulary test, the picture vocabulary test, and the auditory association test. These test were adapted from validated past studies with reliable instruments. On the other hand, the test of spatial skills utilized measures of colored matrices and mental rotation test, where these test required geometrical shapes' identification without spatial reasoning required for the problems. An additional mathematical communication between the teacher and children were conducted as well to include a play that featured two tasks of mathematical problem solving. In the play scenario located in the kindergarten's canteen area, kindergartners were required to purchase some items with a 20-dollar cartoon note, where the seller announced the price of each items and the children were required to compute the amount needed to be paid. The children were given the wrong change intentionally as well to see whether they are able to realize that. The mathematical communication between the teacher and children was evaluated using a scale from 0 to 4 (0 as no mathematical communication and 4 as the different mathematical communication aspects of situation displayed by the children). The measure of mathematical thinking, verbal, and spatial ability of all boys and girls were compared using analysis of variance (ANOVA). The findings indicated no effect of gender on mathematical achievements, or between their oral and spatial skills. Despite that, boys' spatial analysis and girls' oral skills affected the mathematics performance which indicated their use of different processes to solve mathematical problems.

Piaw (2014) carried out a study to identify whether gender and brain thinking style, had effects on creative thinking ability of Form Six pupils in Malaysia. Past studies had shown the relationship between brain thinking style and creative thinking ability, although the results were not consistent. In order to provide evidence for that relationship between creative thinking ability and two factors mainly gender and brain thinking style, the study was conducted utilized a non-experimental design. The data was collected using three paper-pencil psychological tests which were the Torrance Tests of Creative Thinking, the Styles of Learning and Thinking test, and the Watson-Glaser Critical Thinking Appraisal. The sample for the study was randomly selected from the population of 2491 students in the state of Selangor in Malaysia and constituted as 216 Form Six students. The sample consisted of 75 male and 141 female students, in which 79 students had left brain style, 118 students with right brain style, and 19 students with whole brain style of thinking. The sample size for the study was identified using the power analysis method. The three psychological tests were administered to the students with allocated time of 20, 30, and 45 minutes for each of them. The Torrance Tests of Creative Thinking consisted of three drawing activities to examine students' creative thinking ability in terms of originality, fluency, elaboration, abstractness of title and premature closure's resistance; while the Styles of Learning and Thinking test consisted of 28 multiple choice questions on brain hemispheric to examine the students' brain thinking style on whether they were using left or right brain or whole brain. The Watson-Glaser Critical Thinking Appraisal consisted of 80 multiple choice questions to examine the critical thinking ability of students. Before conduct the actual test administration, a pilot study was done to determine the three tests' reliabilities. Cronbach's alpha internal consistency of reliabilities shown strong reliability for the three tests. One-way Analysis of Variance (ANOVA) was utilized to analyze the main effect of gender and brain thinking style on creative thinking ability. Multivariate analysis of covariance (MANCOVA) was utilized to analyze the interaction effect of gender and brain thinking style on creative thinking ability. Results of MANCOVA test indicated correlation between gender and creative thinking ability; and significant correlation between right brain thinking and learning style with all the five components identified in creative thinking ability. Although test results revealed that those two variables were the factors that affected the ability of creative thinking, there was no significant interaction of the two independent variables on ability of creative thinking.

Baltaci, Yildi and Özcakir (2016) carried out a descriptive study to determine the relation between gender, metacognitive variances, styles of learning, and mathematics scores of students using a relational screening model. The study utilized descriptive research design and was conducted using relational screening model to identify the existence of relationship between the variables using statistical analysis. Sample of 330 fifth grade students between 10 to 11 years old were selected using purposeful sampling method of maximum variation sampling. The instrument utilized for the study was the “Metacognitive Awareness Scale for Children” and “Learning Styles Scale” adapted from past studies. The “Metacognitive Awareness Scale for Children” instrument consisted of 12 items with three-point Likert scale with maximum point of 36 points, while the “Learning Styles Scale” instruments consisted of 94 true-false items. Data analysis using the Statistical Package for Social Science (SPSS) for normality and homogeneity tests was used to analyze the data collected from both of the instruments. Skewness and kurtosis values was analyzed as one of requirement of parametric for the normality test. Independent samples t-test and ANOVA were conducted as well for data analysis. The findings indicated no relation between gender and styles of learning but found relation between mathematics scores and styles of learning; mathematics scores and metacognitive awareness; gender and metacognitive awareness; and styles of learning and metacognitive awareness. This result showed the significance to education in terms in how teachers, parents, and administrators can understand the metacognitive awareness and learning styles of students.

Martin et al. (2017) carried out a research to identify whether gender and experience had effects on statistical reasoning. This study comprised of two hundred and one undergraduate and graduate who were proficient in English as the sample. The undergraduates were selected based on invitation through standard psychology department participant pool, while the graduates were selected based on invitation through email which was sent out to the entire email list of the graduate students. The sample for the study had different levels of experience in statistics from 0 up to 4 or more statistics course taken. Graduate students were included due to them having completed at least two course in statistics in order to allow hypothesis testing to relate to experience. Questionnaire consisted of items with 5-point Likert scale was to be filled by the participants as well before the assignment of the main task. The instrument used was a 20-item Statistical Reasoning Assessment and some other individual measures of thinking dispositions and measures of cognitive ability. Before the assignment of the main task instrument, participants were required to fill three self-report questionnaires with allocated time of 30 minutes. After that, participants were given 60 minutes of allocated time for the five main pencil and paper tasks. Descriptive analysis using 2(gender) x 3(experience) between-subjects' ANOVA was conducted to analyze the data. The result shown that males have better performance than females overall. Furthermore, result also revealed the direct and indirect effect of gender on statistical reasoning through its influence on thinking dispositions.

Bart, Hokanson, Sahin, and Abdelsamea (2015) carried out a study to identify the gender's effect upon 8th and 11th grade students' creative thinking abilities. The sample selected for this study consisted of a large number of students from independent public school in the rural district. The participants for the study were 996 students in their 8th grade, in which 503 students were boys and 493 students were girls; and 748 students in their 11th grade, in which 407 students were boys and 341 students were girls. Instruments consisting a Torrance test and questionnaire were administered to the sample. Parental consent was obtained by the researchers and the school district for the study. The teachers of the school district were trained by the researchers on the procedure for instruments' administration during the workshops. The responses were completed and scored by the publisher of the Torrance test. Data analysis using one-way ANOVA was selected to analyze the effect of gender. The result shown that 8th grade female students achieved better scores in comparison to 8th grade male students in the creative thinking subtests. The result also shown that 11th grade female students achieved higher scores in comparison to 11th grade male students on three subtests of creativity. Result also revealed no differences in both male and female's subtest for fluency as they have the same performance. The findings indicated that females matured early compared to males and suggested that the gap between both genders could close when males started to mature. It was highlighted of the need to integrate methods and activities that facilitate creativity in the curriculum to develop the creative thinking abilities of both males and females. Aside for curriculum, teachers were ought to prepare more problem solving tasks that promote higher order creative thinking abilities.

The key findings of past research highlighted that topic of interest on examining gender and students' mathematics achievement and other factors such as thinking. Past researches on gender revealed mix results of effect of gender on thinking and performance. There were findings that showed male's mathematics achievement higher than females. On the other hand, there were finding that indicated higher scores in female students compared to male students in creative thinking subtests as well. There were findings that show no gender differences in mathematical achievements too. The research on gender's effect on statistical thinking is inadequate which deemed as the gap.

2.8 Past Research on Socio-Economic Status

Jankowska and Karwowski (2019) conduct a study to find out the role played by parents' involvement and their socio-economic status in affecting children's creative thinking development. The sample of the study consisted of 75 public primary school students and their parents. Interviews along with test were conducted four times in 5-month intervals to collect the data necessary to measure children's creative thinking. Family's socio-economic status were identified as well. The result indicated that the latent growth changes in children's creative thinking varied across each other. Parents' involvement and their socio-economic status affected the initial level of creative thinking. It is identified that children coming from higher socio-economic status families often have higher creative thinking initially, yet the higher socio-economic status did not trigger more intense creativity growth.

Norfadillah, Hutagalung, Nor, and Isa (2017) performed a study to examine the effect of preschoolers' socio-economic status on their cognitive abilities level. The sample consisted of preschoolers from the area of Klang Valley, Malaysia. The socio-economic status of preschoolers was measured by parents' education level, fathers' income and fathers' occupation. McCharty Scales of Children Abilities test battery was translated and adapted to be used for the study. The result shown significant differences in the cognitive abilities of children by fathers' education. However, result shown no statistically significant difference of father's income and occupation, and mother's education in affecting children's cognitive abilities.

Rinn (2013) carried out a research for identifying correlation among socio-economic status, fathers' level of education, mother's level of education, and social support from family. The population of the study was 12,000 students who took Science, Technology, Engineering, and Mathematics (STEM) major courses from a research university in United States. The sample selected was 499 students which consisted of 373 female and 126 male students. The sample indicated their parents' education level, with 33 had mothers not graduated from high school while 32 had fathers not graduated from high school, 76 had mothers graduated with high school diploma while 64 had fathers graduated with high school diploma, 114 had mothers enrolled in college while 83 had father enrolled in college, 184 had mothers graduated from college while 178 had father graduated from college, 23 had mothers with training after graduated while 24 had fathers with training after graduated, and 69 had mothers graduated with Master's degree or PHD while 114 had fathers graduated with Master's degree or PHD. The sample indicated their family financial background as well in which 219 students were from low socio-economic status families, 261 students were from middle socio-economic status families, and 104 students were from high socio-economic status families. Data were collected from the samples using demographic questionnaire Self-Description Questionnaire III adapted from past validated study. The Self-Description Questionnaire III consisted of 136 items and was meant to measure the students' self-concepts by assessing four academic areas which were mathematics, verbal, general academic, and problem solving; and eight nonacademic areas which were physical appearance, physical ability, relations with same and opposite sex, relations with parents, spiritual religion, reliability, and emotional stability. The result shown that only the mathematical self-concept of female students was affected by their mother's education level where male students were not affected. On the other hand, no correlation was identified between fathers' level of education and students' mathematical self-concept.

Gustafsson et al. (2016) conducted a study to examine school characteristics that affect the correlation between grade 8 students' socio-economic status and mathematics achievement. 50 countries' grade 8 pupils who had previously participated TIMSS in 2011 were selected as the sample. TIMSS 2011 was reanalyzed using two level random slopes modelling to identify the effect of some variables pertaining the socio-economic status of the schools upon the mathematics achievement of pupils. Result indicated that socio-economic status variable had the largest effect on affecting mathematics achievement throughout various schools.

Caro (2009) conducted a study to find out the scholastic achievement gap in Canadian students' socio-economic status from child to teenager. The study used four-time point longitudinal design. The sample for the study involved 22831 children aged 0 to 11 in cycle 1 (year 1994 to 1995), with remaining numbers of 16903 children for cycle 2 (year 1996 to 1997), 16718 children for cycle 3 (year 1998 to 1999), and 15632 children for cycle 4 (year 2000 to 2001). Data such as children's information, families, and other socio-economic data were gathered. The mathematics achievement of the students was measured using the "Mathematics Computation Test" from the standardized Canadian achievement test that consisted of 15 questions related to addition, subtraction, multiplication, and division on whole numbers, decimals, fractions, percentages, negatives, and exponents. An interviewer was in charge in administering the questions while reading and recording the responses from students on answer sheet. A four-time point longitudinal design with regression techniques was applied for this study. In addition, hierarchical linear models and panel data models were used to gauge the gap's trajectory. Hierarchical linear models were used to gauge the students from low to high socio-economic status on their growth trajectories. Panel data models were used to include the scoring probability at or above ceiling value in the algorithm of model estimation when producing each level's gap of points' estimates. The result revealed stable gap in the period of time between 7 to 11 years of age and widen from age 11 to 15 years at quick rate. It was noticed that female students had better performance than male students. The students' mathematics growth rate was positive and significant statistically despite that the relationship was not constant between age and mathematics achievement of the students. On the other hand, the relationship between the students' socio-economic status and age was strong and significant statistically which suggesting that a widening and increasing gap between students from higher socio-economic status and students from low socio-economic status in terms of their mathematics achievement.

Ford (2013) had performed a study to identify the effect of socio-economic status on scholastic achievement of culturally diverse pupils. The study implemented mixed method research design to collect quantitative data of 207 middle school students' Criterion-Referenced Competency Tests (CRCT) assessments result and qualitative data of interview responses from 4 school pupils, 4 parents, and 4 middle school educators. The quantitative data was collected using the students' 2012 CRCT data in Reading, English, Social Studies, Science, and Mathematics and the data was analyzed using IBM PAW Statistics 18. The qualitative data was collected using semi-structured interviews which were then transcribed and coded. Purposive methods of sampling were used to identify the suitable students with equal mix of race and academic achievement for the study and identify teachers based on their teaching experience. The participants for the interview was consented before conducting the interview which lasted 30 minutes. The demographic information of the participants was collected as well to identify the relationship between the students' socio-economic status and their academic achievement. Descriptive analysis was used to identify the means, standard deviation for the students' demographic data. Multivariate Analysis of Covariance (MANCOVA) was used to analyze the difference in the academic achievement of the students from low socio-economic status with covariation to teachers' teaching experience and qualifications. Same analysis was used to analyze the difference in the academic achievement of the students from high socio-economic status with covariation to teachers' teaching experience and qualifications. The gender and grade level variables were then analyzed to determine whether they had influence in affecting the academic achievement of students from low to high socio-economic status. The result from quantitative data analysis shown no correlation between students' socio-economic status and achievement. It was also indicated that female students were performing better than male students in English Language, Social Studies, Reading, and Mathematics. The qualitative data revealed four themes which were communication, motivation, cultural awareness, and teaching and learning supports. In terms of communication, the highlights were on the teacher's role and parents' support in preparing a classroom suitable for students from different background to promote learning interaction. In terms of motivation, the highlights were on teacher's interest and duties along with building relationship with parents to ensure the school's success and good learning place for students. In terms of cultural awareness, the highlights were on the students' background knowledge, belonging and awareness to build knowledge and make connections. In terms of teaching and learning support, the highlights were teachers' background knowledge and efforts in providing good learning activities in classroom by instilling interactions and further ensure the success of school. It was identified that teachers and parents played major influential role in affecting students' academic achievement with their communication and involvement.

The key findings of past research highlighted the topic of interest in the student's socio-economic status background and how it influenced their academic achievement and thinking. Past research on socio-economic status revealed findings that show consistency in showing that socio-economic status played a part in affecting the students' academic achievement and thinking. Parents' socio-economic status in terms of their income and education reflected in what environment they were able to provide to their children. It was also found that parental involvement played a vital role in affecting the upbringing of the children as it was noticed the involvement put children above their peers.

Past researches on socio-economic status shown that it was a strong determinant in affecting students' growth in terms of thinking and performances. Children staying their parents with high socio-economic status often have better scholastic achievements in comparison with children staying with families of low socio-economic status. Parents' education and income affect somehow in students' socio-economic growth. The research on the socio-economic status' effects on thinking abilities were lacking as most of the research were looking at other variables other than thinking. The lack of researches on socio-economic status' effects on students' statistical thinking abilities was deemed as the gap of study.

2.9 Conceptual Framework

The conceptual framework (Figure 2.1) was constructed according to literature reviews of the theoretical and research framework of this study. This conceptual framework portrayed the effects of gender and socio-economic status on the pupils' statistical thinking abilities. In this study, conceptual framework acted as a research model to determine whether socio-economic status and gender can affect the statistical thinking abilities of students.

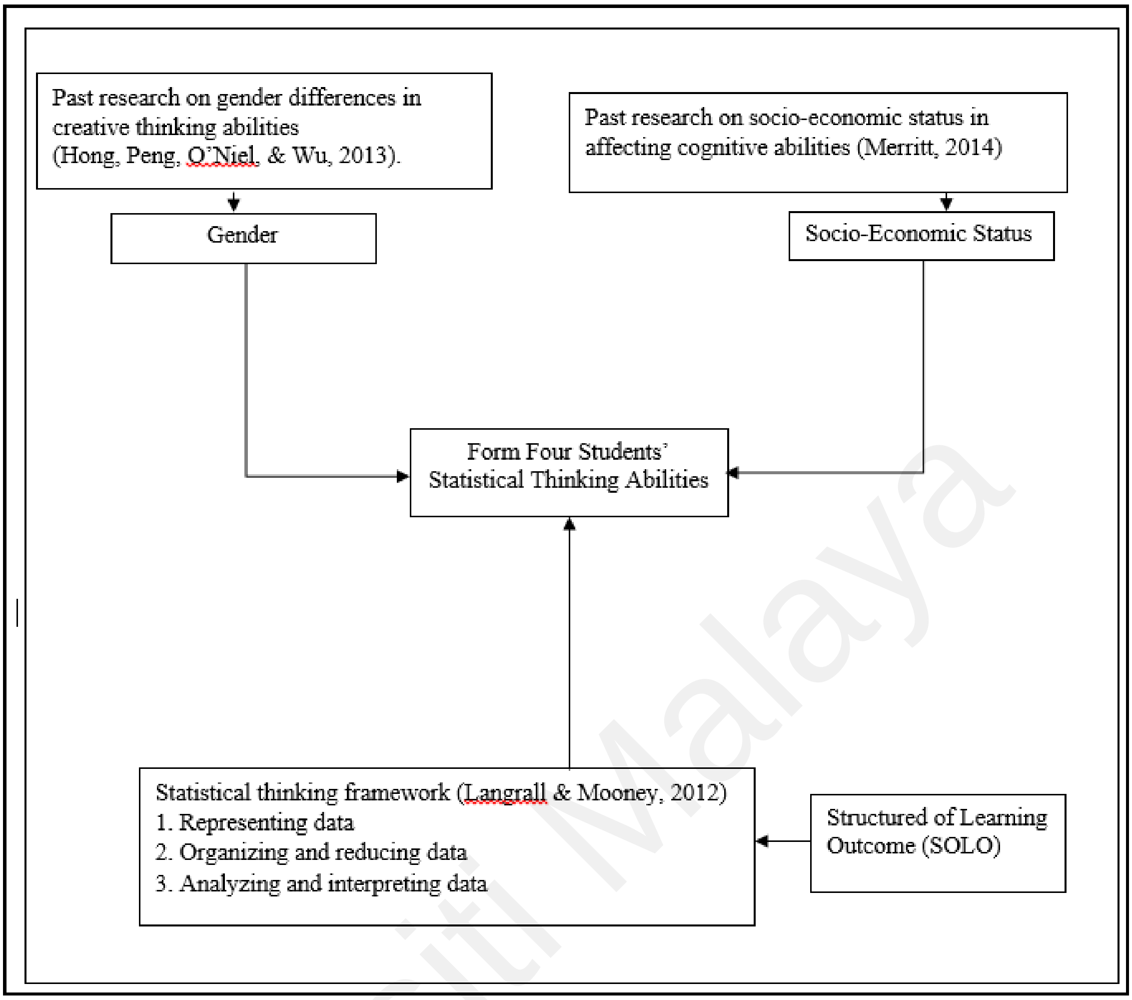


Figure 2.1. Conceptual Framework

The study of gender differences regarding the ability differences of males and females had caught attention of educational researchers which become intriguing variables for many studies (Kusumaningsih & Herman, 2018). For instance, Kousoulas and Mega (2009) conducted study to identify gender's effect on Greek primary school students' divergent thinking. Hong, Peng, O' Neil, and Wu (2013) conducted study upon effect of gender differences on students' creative thinking abilities and found interesting findings on the different creative thinking development between males and females. Upadhayay and Guragain (2014) identified that males and females have different cognitive abilities in terms of their awareness, memory, thought, visual and spatial processing, and executable functions.

There were also literatures on gender in learning statistics and mathematics. Garfield (2003) reported about a research done by Liu (1998) which shown that the existence of gender where males outperform females in their ability to prevent misconceptions. Research also revealed that gender affect mathematics achievement and this differences can be even seen in students from first grade (Rathbun et al., 2004). All the past literatures led to this study in determining the gender variable's effect on statistical thinking abilities of pupils.

There were various studies that associated students' socio-economic status and cognitive abilities. Merritt (2014) claimed that the cognitive incitement of students at home was affected by family's socio-economic status especially the lower income families. Parents' level of education related somehow in the students' quality of environment where higher education and training provided richer learning environment and better caregiving (Clarke-Stewart, Vendell, Burchinal, O'Brien, & McCartney, 2002).

Merritt (2014) stated that children of low socio-economic status often have risk on not developing necessary cognitive abilities for higher mathematics achievement due to various common low socio-economic status environmental factors. Downer and Pianta (2006) supported the claim by mentioning that pupils of lower socio-economic status were prone to lower mathematics achievement. The evidence led to the conduct and the start of this research to determine the effects of socio-economic status on pupils' statistical thinking abilities.

Numerous researches had stressed upon the development of higher order thinking in students in education area. It was in fact the central point of education goal where the main focus was teaching people to become good thinkers to solve problems (Gagné, 1980). In the statistics field, statistical thinking was a type of higher order thinking construct which was important for students to develop it (Le, 2017). The emphasis was stronger as statistics educators called for a change of pedagogy and content in statistics courses to prepare the students better in real world (Moore, 1997).

The statistics thinking framework developed and validated by Langrall and Mooney (2002) provided a useful cognitive model that portrayed the thinking of students. The model allowed better planning and establishing of the instruction and mathematics curriculum. The statistical thinking framework was adapted for this study and served as guideline for the identification of students' statistical thinking levels.

2.10 Summary

This literature review was classified into four major sections. The first section described on the statistical thinking framework theory. The second section described the key concepts which was statistical thinking ability, gender, and socio-economic status. The third section discussed on the past literatures on statistical thinking, gender, and socio-economic status. There were various studies that explained about the effect and correlation pertaining to the variables but there was a gap related to the inadequate of studies researching on effects of the gender variable and socio-economic status variable towards statistical thinking abilities which led to the conduct of this study to ascertain whether the variables were factors that influence students' statistical thinking abilities. Lastly, the fourth section showed the conceptual framework.

CHAPTER 3

METHODOLOGY

3.1 Introduction

Chapter three discussed the methodology for this study. There were nine main sections mainly research methodology, research design, population and sample, instrument, rubric, instruments' validity and reliability, procedure for data collection, procedure for data analysis, and ethics.

3.2 Research Design

This study utilized quantitative research methodology. Quantitative research methodology is defined by Creswell & Creswell (2017) as the method that involved instruments to obtain quantitative data for analysis purposes using statistical procedures. It was also deemed as the method to identify the relation between measureable variables in order to assess objective theories. Wiersma and Jurs (2008) supported the claim by stating that it was a plan to execute research through using numbers to describe phenomena, including participants' selection, data collection procedure, and data analysis procedure (Wiersma & Jurs, 2008). Quantitative research methodology was selected instead of qualitative research methodology as the study intended to use precise hypotheses to be backed up by supports or rejection from data collection methods (Johnson & Christensen, 2008). Data collection would utilize instruments to obtain information and proceed for analysis using statistical procedures and hypothesis testing. The study did not intend to build meaning of phenomenon from the opinions of participants as focused in the approach of qualitative research methodology.

The research designs for this study were survey descriptive design, causal comparative design, and factorial design. Survey descriptive design contained the characteristics of describing aspects and identify trends such as attitudes, opinions, behaviors, or characteristic of a population from the information collected through asking questions from a group of people taking part in the population (Creswell, 2012). Survey descriptive design carried the purpose of determining how population members discerning themselves on various different variables. In the case for this study, it was suitable to identify members of population to be classified according to two group of variables of gender and socio-economic status.

The survey design had strengths in its ability to be conducted for large amount of participants using analyzable questions (Fraenkel, Wallen, & Hyun, 2011). Besides that, the administration can be done quickly and reach geographically dispersed population. Generalization of sample to population for the result of the study can be done using this design (Creswell & Creswell, 2017). The survey design had its weakness as well. The information collected might subjected to response bias where the samples provided inaccurate information responses. Besides, survey did not control for variables that explained the independent and dependent variables' relationship (Cresswell, 2012).

The selection of survey descriptive design for this study was due to the need to gather information for the two groups of gender variable and socio-economic status variable. The information was needed for the answering this study's research questions to identify the effects of both variables on statistical thinking abilities of the students. Furthermore, survey descriptive design was chosen due to the fact that this study involved the use of instruments where the instruments were a necessity to be administered to the sample using survey research procedure to collect the data required for this study (Fraenkel et al., 2011). Cross-sectional technique was used in this study to a group of samples at just one point in time.

On the other hand, causal-comparative design was the approach of researchers to determine the possible factors which might influence individual groups. One or a few more categorical independent variables were studied on to see the effects on one or a few more quantitative dependent variables (Johnson & Christensen, 2008). This design had strength in its ability to allow the establishment of cause and effect by the researcher to see how dependent variable was being affected by independent variables (Williams, 2007). On the other hand, causal-comparative design had weakness as well. Researcher was not able to establish that the factor that was currently studied on was the only cause to the effect of the dependent variable.

Causal-comparative design was selected to interpret the remaining research questions for the identification of differences in students' statistical thinking abilities by gender and socio-economic status. This was in contrast to experimental study where the difference between or among groups were created instead by researcher to compare performance to identify the effects of the created difference. The observed difference in these two designs was the manipulation of the group difference variable where variable such as gender cannot be manipulated in comparison to variable such as teaching style that was able to be manipulated in experimental research design (Fraenkel et al., 2011). Moreover, this study was not going to establish cause and effect relationship and just to identify possible causes. Therefore, this study did not plan to manipulate any independent variables and thus the variables chosen for this study cannot be manipulated. As a result, causal-comparative design was selected instead of experimental design for this study.

3.3 Population and Sample

This section described the population of the study and the utilized process in the selection of sample. This study was planned to cover the population of all Form Four students who attended government secondary school and using the secondary schools' integrated curriculum for Mathematics subject prepared by the Ministry of Education (MOE) in Shah Alam area, Selangor, Malaysia in the year of 2019. The population of all government secondary schools in Shah Alam area consisted of a total of 29 secondary schools, where the number of students estimated to be at 5800, based on 8 classes in each school and 25 students from each class.

The sample for this study was selected based on random probabilistic sampling using two-stage cluster sampling technique. Probabilistic sampling was chosen for this study to ensure the equal and independent selection chance for each and every populations' member as random sample and thus enable the researcher to generalize the result to the population (Cresswell, 2002). Cluster random sampling was selected as it allowed selection of sample more conveniently due to difficulty in identifying the population (Cresswell, 2002). In this case, the number of students from the population of all secondary schools in Selangor are large and not easily identified. The advantages of cluster random sampling were also seen in its characteristics of easy implementation in schools and less time consuming.

The population of every government secondary schools in the area of Shah Alam, Selangor consisted of a total of 29 secondary schools. Two-stage cluster sampling method was used to cluster samples by schools. There were 29 schools, therefore there were 29 clusters. The first stage involving the random selection of 4 clusters from the 29 clusters. The second stage involved selection of the number of Form Four classes in the selected clusters were deemed as the second level clusters.

There were 8 second-level clusters from each of the selected 4 clusters. The 8 second-level clusters were then assigned numbers to randomly select 2 second-level clusters. The selected 2 second-level clusters from each clusters were constituted as the sample.

This study intended to collect data for the two group of variables which were the gender group and socio-economic status group. The gender variable had two groups which were male and female, while socio-economic status variable had three groups which were low, middle, and high socio-economic status. 180 Form Four students were constituted as the sample for this study. There were 90 male and 90 female students. 65 students were in low socio-economic status group, 98 students were in middle socio-economic status group, and 17 students were in high socio-economic status group. Gall et al. (2007) stated that there should be a minimum of 15 participants from each group for causal-comparative research to be conducted. The number of samples in each group for this study were sufficient.

3.4 Instruments

The research instruments in this study involved three components mainly the demographics questionnaire, Statistical Thinking Abilities Test (STAT) and a rubric for scoring the STAT. The demographics questionnaire consisted of a questionnaire to collect demographics information of students regarding their gender and socio-economic status. STAT was adapted in order to identify the statistical thinking abilities of students from their responses provided in answering the items. All of the items were either adapted or adopted from several different items from previous studies to be included in the STAT. Last but not least, the third component of the research instrument was a rubric to score the STAT.

3.4.1 Demographics Questionnaire

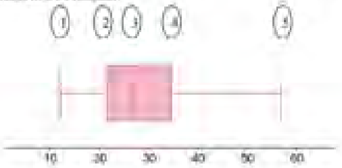
The demographics questionnaire consisted of items to collect information for the two groups of independent variables. The two groups of independent variables were mainly gender group and socio-economic status group. The gender group was categorized into male and female. The group of socio-economic status was categorized into low, middle, and high and was determined based on monthly household income of the students which was obtainable from students' registry information from their homeroom teacher and parents' education from questionnaire.

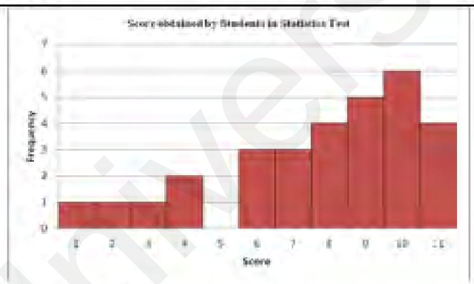
3.4.2 Statistical Thinking Abilities Test

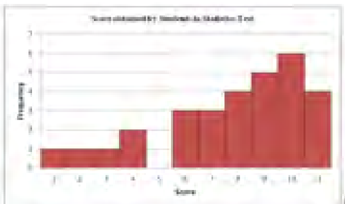
The Statistical Thinking Abilities Test (STAT) consisted of four main items and sub-items that sum up to 28 items altogether. The items were modified from items published in journals from several sources. The first three items were adapted from Chan and Ismail (2014) while the fourth item is adapted from Jones et al. (2000). Necessary changes were made through the adaptation of the items to match the sub-dimensions of the levels in the rubric. The mathematical content of the items was based on the topic of Statistics in the Form Four Mathematics curriculum prepared by the Ministry of Education, Malaysia. The complete table with adapted or adopted items and modified items of STAT along with justifications and remarks is shown in the Table 3.1 below.

Table 3.1.

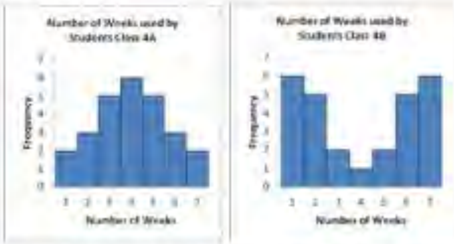
Adapted/Adopted Items and Modified Items of STAT

<table>
<tr><th>Adapted/Adopted Items</th><th>Revised Items</th></tr>
<tr><td colspan="2">Task 1(source: Shiau, W. C., & Ismail, Z. (2014). A technology-based statistical reasoning assessment tool in descriptive statistics for secondary school students. TOJET: The Turkish Online Journal of Educational Technology, 13(1).)</td></tr>
<tr><td>
The data below indicate the amount of protein (in grams) for various fast food sandwiches (Source: The Doctor's Pocket Calorie, Fat, and Carbohydrate Counter, 2002).
<table>
<tr><td>23</td><td>30</td><td>20</td><td>27</td><td>44</td><td>26</td><td>35</td><td>20</td><td>29</td><td>29</td></tr>
<tr><td>25</td><td>15</td><td>18</td><td>27</td><td>19</td><td>22</td><td>12</td><td>26</td><td>34</td><td>15</td></tr>
<tr><td>27</td><td>35</td><td>26</td><td>43</td><td>35</td><td>14</td><td>24</td><td>12</td><td>23</td><td>31</td></tr>
<tr><td>40</td><td>35</td><td>38</td><td>57</td><td>22</td><td>42</td><td>24</td><td>21</td><td>27</td><td>33</td></tr>
</table>
1) What are the highest and lowest amount of protein (in grams) for various fast food sandwiches?
2) Organize the data into GeoGebra spreadsheet.
3) Construct a frequency polygon using GeoGebra spreadsheet.
4) Record the values of the mean, median and standard deviation from the computer.
5) Describe the distribution of the graph in terms of its shape, center and variability.
6) Represent the data in another way.
7) Write the name of the feature at each of the labels on the five-number summary of the box plot and record the values from the computer.


<table>
<tr><th>No</th><th>Five-number summary</th><th>Value</th></tr>
<tr><td>1</td><td></td><td></td></tr>
<tr><td>2</td><td></td><td></td></tr>
<tr><td>3</td><td></td><td></td></tr>
<tr><td>4</td><td></td><td></td></tr>
<tr><td>5</td><td></td><td></td></tr>
</table>
8) Describe how the box plot is related to its matching histogram.
9) Which graph do you think represents the data better, the histogram or the box plot? Explain why.
</td><td>
The data below indicate the amount of protein (in grams) for various fast food sandwiches.
<table>
<tr><td>23</td><td>30</td><td>20</td><td>27</td><td>44</td></tr>
<tr><td>25</td><td>15</td><td>18</td><td>27</td><td>19</td></tr>
<tr><td>27</td><td>35</td><td>26</td><td>43</td><td>35</td></tr>
<tr><td>40</td><td>35</td><td>38</td><td>57</td><td>22</td></tr>
</table>
1) What is the mean for ungrouped and grouped data? Are they the same? Explain why.
2) What do you notice about the mean and median? Are they the same?
3) Construct a histogram to represent these data.
4) Describe the distribution of the graph in terms of its shape, center, and dispersion.
5) Represent the data in another way using frequency polygon.
6) Describe the measure of dispersion of the frequency polygon.
7) Describe how the frequency polygon is related to its matching histogram.
8) Which graph do you think represents the data better, the frequency polygon or the histogram? Explain why.
</td></tr>
<tr><td colspan="2">Task 2</td></tr>
<tr><td>


1) What does this graph tell you?
2) What is the mean of the graph? Explain how.
3) What is the mode of the graph? Explain how.
4) What is the median of the graph? Explain how.
5) Draw the graph using GeoGebra dynamic worksheet by dragging the red circle. Tick the check box of Show histogram, Show mean and Show median.
6) Compare your answer in Question 2, and 4 with the values shown on the computer. If the answers are different, explain why.
7) What is the range of the graph? Explain how.
8) What is the interquartile range of the graph? Explain how.
9) What is the standard deviation of the graph? Explain how.
10) Tick the check box of Show IQR and Show Std Dev.
11) Compare your answer in Question 8 and 9 with the values shown on the computer. If the answers are different, explain why.
12) Describe the distribution of the graph with respect to its shape, center and variability.
Another set of new scores obtained by students from a different class are as follows:
13) Drag the red circle to draw the new histogram.
14) Record the values of mean, median, interquartile range, and standard deviation from the computer.
Two students who each obtained a score of 1 are added to the graph.
15) Record the values of mean, median, interquartile range, and standard deviation from the computer.
16) Compare the results in question 15 with question 14. What do you observe? Explain why.
17) Which measures of center is the most suitable to be used to represent the score obtained by students? Explain why.
18) Which measures of spread is the most suitable to be used to represent the score obtained by students? Explain why.
</td><td>
The following graph shows the score obtained by a group of students in a statistics test.

1) What is the mean, mode and median of the graph?
2) Compare your answer for mean and median. If the answers are different, explain why.
3) What is the range and interquartile range of the graph?
4) Compare your answer in interquartile range and range. If the answers are different, explain why.
5) Describe the distribution of the graph with respect to its shape, center and variability.
6) Construct a cumulative frequency graph. Which graphs represent the data better?
Two students who each obtained a score of 1 are added to the graph.
7) Compare the old and new graph. What do you observe about the mean? Explain why.
8) Which measures of center is the most suitable to be used to represent the score obtained by students in the new graph? Explain why.
9) Which measures of spread is the most suitable to be used to represent the score obtained by students in the new graph? Explain why.
</td></tr>
</table>

Task 3


The following graphs illustrate the number of weeks used by the students from class 4A and 4B to finish reading a storybook.

1) What are the highest and lowest number of weeks used by the students from class 4A to finish reading a storybook?
2) What are the highest and lowest number of weeks used by the students from class 4B to finish reading a storybook?
3) Predict which class has the lower standard deviation. Explain why.
4) Drag the red circle on the GeoGebra dynamic worksheet to create the histograms for Class 4A and Class 4B. Tick the check box of Show Std Dev.
5) Compare the answer you predicted in Question 3 to the value shown on the computer. If the answers are different, explain why.
6) Are there any similarities or differences between the two graphs produced on the computer? Explain.

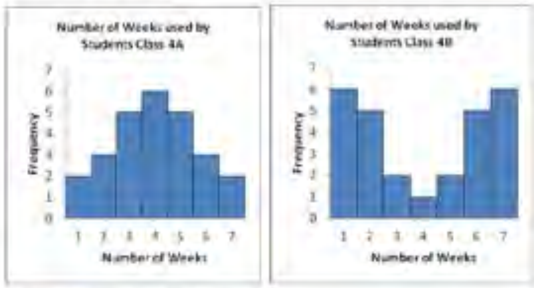
The teacher did a survey of the number of weeks used by the students from class 4A and 4B to finish reading a storybook during the school holidays. The following data indicated the results of the survey.

Week	1	2	3	4	5	6	7
Class 4A	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Class 4B	0	[illegible]	5	6	4	[illegible]	0

7) Predict which class has the larger standard deviation. Explain why.
8) Drag the red circle on the GeoGebra dynamic worksheet to create the histograms. Tick the check box of Show Std Dev.
9) Compare the answer you predicted in Question 7 to the value shown on the computer. If the answers are different, explain why.
10) Are there any similarities or differences between the two graphs produced on the computer? Explain.

The graphs below show the number of weeks used by the students from class 8C and 8D to finish reading a storybook.

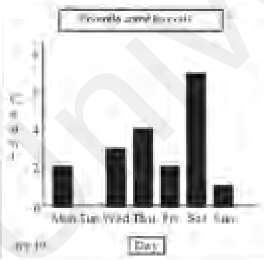
The following graphs illustrate the number of weeks used by the students from class 4A and 4B to finish reading a storybook.



1) Predict the mean for class 4A and 4B.
2) What are the mean and median for class 4A? What do you notice about them?
3) Construct a line graph to represent class 4B.
4) Which graph represents the data better?
5) Compare the answer you predicted in Question 1 to the value calculated using calculator. If the answers are different, explain.

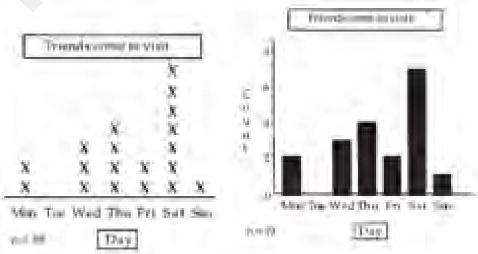
Task 4 (source: Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical thinking and learning*, 2(4), 269-307.)

D1 What does this picture tell you? (Follow-up probe) How many friends came to visit each day?
A2 What can't you tell from this picture?
A3 Which day had the highest number of visitors? Which day had the lowest number of visitors?
A4 How many friends came to visit Sam during this week?
C1 About how many friends came to visit Sam each day?
C2 What is the average number of friends that came over to visit Sam each day?
A5 About how many friends would you expect to come to Sam's place every week during the summer vacation?
A6 How many friends would you expect to visit Sam during a 4-week month?



D1 What does this picture tell you?
D2 How are the two pictures alike? Do you think the pictures represent the same data?
C3 Which one of these pictures would be more useful?
E1 Can you draw either of these graphs in another way?

Sam had some friends visiting him each day during one week in the summer. The number of friends and the days they visited were displayed like this:



1) How are the two pictures alike? Do you think the pictures represent the same data?
2) Which one of these pictures would be more useful? Explain why.
3) Describe the dispersion of the data presented in the graphs.
4) Represent either of the graphs in another way.
5) Identify the mean and mode of the graph.
6) Compare the mean and median. Explain.

3.4.3 Rubric

The Middle School Students' Statistical Thinking (M3ST) framework developed by Langrall and Mooney (2002) was adapted as the rubric for this study to evaluate the students' statistical thinking abilities based on their responses. The rubric was used to assign scores in terms of one, two, three, or four to students' responses according to the four levels of each process. The original M3ST framework contained four processes mainly: describing data, organizing and reducing data, representing data, and analyzing and interpreting data. The four processes were then matched across four cognitive levels of SOLO with descriptors. The original framework is modified and adapted to be used for this study. The final rubric consisted of three processes after removal of describing data.

The first component of the rubric is ability to organize and reduce data. It was adapted from the original framework as shown in the screen capture in Figure 3.1. The changes made were the removal of the "grouping data" subprocess and changes to the descriptors of the second and third subprocesses in Table 3.2.

Organizing and Reducing Data **In general, students performing or explaining questions, tasks, or activities which involve organizing and reducing data will . . .**	**Level 1 - Idiosyncratic** **O.1.1** Not attempt to group data. **O.2.1** Not be able to describe data in terms of representativeness or typicalness. **O.3.1** Not be able to describe the spread of the data in terms representative of the spread.	**Level 2 - Transitional** **O.1.2** Group data but not in a summative form. **O.2.2** Describe the typicalness of data using invented measures that are partially valid. **O.3.2** Describe the spread of the data using invented measures that are partially valid.
	Level 3 - Quantitative **O.1.3** Group data in a summative form or group data by creating new categories or clusters. **O.2.3** Describe the typicalness of data using a measure of center from a flawed procedure or a valid and correct invented measure. **O.3.3** Describe spread of data using a measure from a flawed procedure or a valid and correct invented measure.	**Level 4 - Analytical** **O.1.4** Group data in a summative form by creating new categories or clusters. **O.2.4** Describe data using a valid and correct measure of center. **O.3.4** Describe spread of data using valid and correct measure.

Figure 3.1. Screen Capture of M3ST Framework, Organizing and Reducing Data

Table 3.2

First component of the Rubric, Ability to Organize and Reduce Data

Construct	Level 1: Idiosyncratic (1)	Level 2: Transitional (2)	Level 3: Quantitative (3)	Level 4: Analytical (4)
Ability to organize and reduce data				
O.1 Describe data using measures of center	O.1.1 Unable to reduce data using measures of center	O.1.2 Reduces data using measures of center through invented measures with major errors	O.1.3 Reduces data using measures of center through invented measures with minor errors	O.1.4 Reduce data using measures of center correctly
O.2 Describe data using measures of spread	O.2.1 Unable to reduce data using measures of spread	O.2.2 Reduces data using measures of spread through invented measures with major errors	O.2.3 Reduces data using measures of spread through invented measures with minor errors	O.2.4 Reduce data using measures of spread correctly

The second component of rubric is ability to represent data. It was adapted from the original framework as shown in the screen capture in Figure 3.2. The revised rubric is demonstrated in Table 3.3.

Representing Data **In general, student students performing or explaining questions, tasks or activities which involve representing data will . . :**	**Level 1 - Idiosyncratic** **R.1.1** Be unable to construct a display or constructs a display for that is both incomplete and unrepresentative of the data. **R.2.1** Evaluate the effectiveness of data display based on irrelevant features or reasons.	**Level 2 - Transitional** **R.1.2** Construct a display that is partially complete and representative of the data or complete and unrepresentative of the data. **R.2.2** Evaluate the effectiveness of data display based on relevant display features.
	Level 3 - Quantitative **R.1.3** Construct a complete and representative display. The display may have a few minor flaws. **R.2.3** Evaluate the effectiveness of a data display based on relevant display features with some reference to the context the data is presented.	**Level 4 - Analytical** **R.1.4** Construct a complete, representative and appropriate display. **R.2.4** Evaluate the effectiveness of a data display based on relevant display features and the context the data is presented.

Figure 3.2. Screen Capture of M3ST framework, Representing Data

Table 3.3

Second Component of the Rubric, Ability to Represent Data

Construct	Level 1: Idiosyncratic (1)	Level 2: Transitional (2)	Level 3: Quantitative (3)	Level 4: Analytical (4)
Ability to represent data				
R.1 Constructing a data display	**R.1.1** Unable to construct a display	**R.1.2** Construct a display that is partially complete or correct	**R.1.3** Construct a complete and representative display. The display may have a few minor flaws.	**R.1.4** Construct a complete, representative and appropriate display.
R.2 Evaluating the effectiveness of data displays	**R.2.1** Unable to evaluate the effectiveness of data display	**R.2.2** Evaluate the effectiveness of data display based on irrelevant display features or reasons	**R.2.3** Evaluate the effectiveness of a data display based on relevant display features	**R.2.4** Evaluate the effectiveness of a data display based on relevant display features with reference to the context the data is presented.

The third component of the rubric is ability to analyze and interpret data. It was adapted from the original framework as shown in the screen capture in Figure 3.3. The changes made were the removal of the "inferencing from a given data set or display" and "using comparable reasoning" subprocesses as seen in Table 3.4.

Analyzing and Interpreting Data **In general, student performing or explaining questions, tasks or activities which involve analyzing and interpreting data will . . .**	**Level 1 - Idiosyncratic** **A.1.1** Make no or incorrect comparisons within data displays or data sets. **A.2.1** Make no or incorrect comparisons between data displays or data sets. **A.3.1** Make inferences that are not based on the data or inferences are based on irrelevant issues. **A.4.1** Not use relative thinking.	**Level 2 - Transitional** **A.1.2** Make a single correct comparison or a set of partially correct comparisons within or between data displays or data sets. **A.2.2** Make a single correct comparison or a set of partially correct comparisons between data displays or data sets. **A.3.2** Make inferences that are partially based on the data. Some inferences may be only partially reasonable. **A.4.2** Use relative thinking qualitatively.
	Level 3 - Quantitative **A.1.3** Make local or global comparisons within data displays or data sets. **A.2.3** Make local or global comparisons between data displays or data sets. **A.3.3** Make inferences that are primarily based on the data. Some inferences may be only partially reasonable. **A.4.3** Uses relative thinking quantitatively but not in a reasonable manner.	**Level 4 - Analytical** **A.1.4** Make local and global comparisons within data displays or data sets. **A.2.4** Make local and global comparisons between data displays or data sets. **A.3.4** Make reasonable inferences based on data and the context. **A.4.4** Uses relative thinking quantitatively in a reasonable manner.

Figure 3.3. Screen Capture of M3ST framework, Analyzing and Interpreting Data

Table 3.4

Third Component of the Rubric, Ability to Analyze and Interpret Data

Construct	Level 1: Idiosyncratic (1)	Level 2: Transitional (2)	Level 3: Quantitative (3)	Level 4: Analytical (4)
Ability to analyze and interpret data				
A.1 Making comparisons within data sets or data displays	**A.1.1** Make no or incorrect comparisons within data displays or data sets.	**A.1.2** Make a single correct comparison or a set of partially correct comparisons within or between data displays or data sets.	**A.1.3** Make local or global comparisons within data displays or data sets.	**A.1.4** Make local and global comparisons within data displays or data sets.
A.2 Making comparisons between data sets or data displays	**A.2.1** Make no or incorrect comparisons between data displays or data sets.	**A.2.2** Make a single correct comparison or a set of partially correct comparisons between data displays or data sets.	**A.2.3** Make local or global comparisons between data displays or data sets.	**A.2.4** Make local and global comparisons between data displays or data sets.

The scores' weightage was decided and assigned based on descriptors of the rubric. Points from one up to four could be obtained, where the points obtained representing the four levels of statistical thinking processes. For example, score of one point representing level 1 of the statistical thinking process; score of two points representing level 2 of the statistical thinking process, score of three points representing level 3 of the statistical thinking process, and score of four points representing level 4 of the statistical thinking process. The adapted rubric from Langrall and Mooney (2002) had been validated by them in their study. The scoring validation of the rubric descriptor scale was done by an experienced lecturer in statistics and an experienced mathematics teacher.

3.5 Validity and Reliability of the Instrument

A pilot study was conducted on 30 students from a school not involved with sample before the actual study. The participants of the pilot study were 30 Form Four students who would not be participating in the actual study. The students had pre-requisite knowledge regarding the statistics content required to answer the instruments. A pre-test was conducted through the administering of instruments and supervised by the researcher. A post-test was later carried out two weeks after the pre-test was done. The instruments were collected and scored according to the rubric. Feedbacks were taken into account into adjusting sufficient time for answering the instruments. The result was utilized to enact the instruments' validity and reliability.

Validity referred to evidence support for the inferences made by researchers basing on the collected data using a specific instrument where the inferences were appropriate, correct, meaningful, and useful (Fraenkel et al., 2011). There were three main types of validity which were the face-related validity, content validity, and the criterion-related validity. This study used the rubric developed by Langrall and Mooney (2002). According to Langrall and Mooney (2002), the rubric had high content and criterion-related validity. This study also utilized the face-related evidence of validity where a few panel of experts in the area of expertise validated the STAT. A Mathematics teacher whom had 20 years of experiences in teaching secondary school Mathematics and an expert who have Masters in statistics were consulted to evaluate the appropriateness of the items in the STAT. The rubric along with the STAT were given to them for them to evaluate the items' content and whether the items can be assigned scores based on the rubric. The experts also checked the items in terms of its difficulty, and whether the items matched the mathematics syllabus learnt by the students.

There were a few feedbacks provided. First of all, the experts commented on the instruments' items difficulties. The items were covered by the syllabus of Form Four's statistics. However, the difficulty of the items was rather high, where only students with good comprehension of statistics were expected to answer the items correctly. Next, the experts highlighted on the items' ordering sequence, where certain sub items were suggested to be rearranged on the ordering to fit the sequence of events to strengthen the link between the items. The time allocation was highlighted as well for the time needed to answer the items. Overall, they were satisfied with the items and rubric.

The suggestions from the experts were considered and necessary amendments were made accordingly. Regarding the difficulty of the items, the difficulty was adjusted and tuned slightly without affecting too much of the items' structure. The justification was that the study meant to gauge Malaysia Form Four students' statistical thinking levels. Therefore, it was inevitable to have a certain standard of difficulty on the items that were able to measure students' statistical thinking abilities levels. Moreover, the items were adapted from other statistical thinking studies that had validated and administered them to their secondary school students, which proven to bring results of identifying students' level of statistical thinking abilities. Next, the review on the items' ordering sequence were spot on, and was changed as per feedbacks. The time allocation was adjusted as well to 2 hours to allow sufficient time for the students to answer.

Reliability referred to the degree of consistency of obtained measured scores for every administration of instruments (Creswell & Creswell, 2017). The test-retest method was selected which utilized administration of instruments to the same sample at two different times with sufficient time interval. The scores correlate at a positive and high level such at .6 if they are reliable (Creswell & Creswell, 2017). The administering of the STAT instrument was administered twice to 30 students not involved in the sample of studies with an elapsed time interval of two weeks. The reliability coefficient was calculated using Pearson correlation. Test-retest results indicated that the items for sub-process describe data using measures of center were significantly stable (Pearson's r = .95). Table 3.5 showed the test-retest reliability's result.

Table 3.5:

Test-Retest Reliability Result for Sub-process Describe Data using Measures of Center

		PreO1	PostO1
PreO1	Pearson Correlation	1	.95
	Sig. (2-tailed)		.000
	N	30	30
PostO1	Pearson Correlation	.95	1
	Sig. (2-tailed)	.000	
	N	30	30

Test-retest results showed that the items for sub-process describe data using measures of spread were significantly stable (Pearson's r = .86). Table 3.6 presented the test-retest reliability's result.

Table 3.6

Test-Retest Reliability Result for Sub-process Describe Data using Measures of Spread

		PreO2	PostO2
PreO2	Pearson Correlation	1	.86
	Sig. (2-tailed)		.000
	N	30	30
PostO2	Pearson Correlation	.86	1
	Sig. (2-tailed)	.000	
	N	30	30

Test-retest results showed that the items for sub-process constructing a data display were significantly stable (Pearson's r = .96). Table 3.7 presented the test-retest reliability's result.

Table 3.7

Test-Retest Reliability Result for Sub-process Constructing a Data Display

		PreR1	PostR1
PreR1	Pearson Correlation	1	.96
	Sig. (2-tailed)		.000
	N	30	30
PostR1	Pearson Correlation	.96	1
	Sig. (2-tailed)	.000	
	N	30	30

Test-retest results showed that the items for sub-process evaluate the effectiveness of data displays were significantly stable (Pearson's r = .96). Table 3.8 presented the test-retest reliability's result.

Table 3.8:

Test-Retest Reliability Result for Sub-process Evaluate the Effectiveness of Data Displays

		PreR2	PostR2
PreR2	Pearson Correlation	1	.96
	Sig. (2-tailed)		.000
	N	30	30
PostR2	Pearson Correlation	.96	1
	Sig. (2-tailed)	.000	
	N	30	30

Test-retest results showed that the items for sub-process making comparisons within data sets or data displays were significantly stable (Pearson's r = .97). Table 3.9 presented the test-retest reliability's result.

Table 3.9

Test-Retest Reliability Result for Sub-process Making Comparisons within Data Sets or Data Displays

		PreA1	PostA1
PreA1	Pearson Correlation	1	.97
	Sig. (2-tailed)		.000
	N	30	30
PostA1	Pearson Correlation	.97	1
	Sig. (2-tailed)	.000	
	N	30	30

Test-retest results showed that the items for sub-process making comparisons between data sets or data displays were significantly stable (Pearson's r = .94). Table 3.10 presented the test-retest reliability's result.

Table 3.10

Test-Retest Reliability Result for Sub-process Making Comparisons between Data Sets or Data Displays

		PreA2	PostA2
PreA2	Pearson Correlation	1	.94
	Sig. (2-tailed)		.000
	N	30	30
PostA2	Pearson Correlation	.94	1
	Sig. (2-tailed)	.000	
	N	30	30

Last but not least, the inter-rater method was done based on the scores rated by the two experts. After administering the pilot test, the Statistics expert as judge 1 and an experienced Mathematics expert as judge 2 independently scored the responses of students according to the questions. The degree of agreement between the first rater and the second rater was calculated with Intraclass Correlation Coefficient (ICC). The rubric was used by both raters in rating students' responses. Table 3.11 showed the results of the inter-rater reliability of the STAT using rubric.

Table 3.11

Inter-Rater Reliability Results of ICC

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		LL	UL	Value	*df1*	*df2*	*P*
Single Measures	.93[a]	.85	.97	32.528	29	29	.000
Average Measures	.97[c]	.92	.98	32.528	29	29	.000

The two rater's scoring of the STAT using the rubric showed high reliability degree. The average measure ICC was .97 with a 95% confidence interval from .92 to .98. It was concluded that the inter-rater reliability results showed that the rubric scores correlated 96% of the time.

Item analysis procedure was carried out to analyze the item difficulty index and item discrimination index of the Statistical Thinking Abilities Test instruments. As the instruments was partial credit, the item difficulty index was calculated as the total sum of scores obtained by the students in a certain question divided by the multiplication of maximum credit assigned to that question with the total frequency count of responses in that question. On the other hand, the discrimination index was calculated by using the sum of scores obtained by top 30% students subtracting the sum of scores obtained by bottom 30% students in a certain question divided by the multiplication of maximum credit assigned to that question with the total frequency count of responses in that question.

Item difficulty index ranging from 0 to 1, where items with index ranging below 0.3 were considered too difficult, and items with difficulty index ranging above 0.8 were too easy. Items with difficulty index ranging between .3 and .8 were considered good and acceptable (McCowan & McCowan, 1999). The item difficulty index for four questions in the instruments were above 0.4 which were good and acceptable.

Item discrimination index ranging from 0 to 1, indicating discrimination index of .4 or more for good items; discrimination index of .3 to .39 for reasonably good items; discrimination index of .2 to .29 for marginal items; and discrimination index of less than .2 for bad items (Ebel & Frisbie, 1991). The item discriminated the students better the higher the value of item discrimination index. Table 3.12 showed the obtained item difficulty index and the item discrimination index for pilot test of Statistical Thinking Ability Test instruments to 30 students. The four items in the instruments had discrimination index of items ranging between .5 and .7. They discriminated the students fairly well.

Table 3.12

Item Difficulty Index and Item Discrimination Index of Instruments

Students \ Questions	Q1	Q2	Q3	Q4	Total
1	18	16	10	8	52
2	16	11	4	5	36
3	15	13	9	8	45
4	16	13	8	8	45
5	13	12	8	7	40
6	12	10	8	7	37
7	12	10	6	8	36
8	11	8	5	6	30
9	13	10	9	8	40
10	10	12	5	3	30
11	11	8	4	3	26
12	10	7	2	3	22
13	9	7	3	4	23
14	9	8	3	1	21
15	8	5	5	3	21
16	6	5	5	3	19
17	6	5	0	3	14
18	5	3	3	3	14
19	6	3	3	2	14
20	6	5	2	2	15
21	2	3	1	2	8
22	2	4	2	2	10
23	3	4	3	1	11
24	5	3	2	1	11
25	5	5	1	1	12
26	3	3	2	2	10
27	2	2	1	0	5
28	4	5	2	1	12
29	3	2	2	0	7
30	3	2	3	3	11
max	18	16	10	8	52
Item difficulty index	0.45185	0.425	0.40333	0.45	0.43397
Item discrimination index	0.57778	0.5125	0.53	0.6875	

3.6 Data Collection Procedure

This study was conducted to Form Four students in government secondary school located in Shah Alam area in the year of 2019 after obtaining permission from the schools. Data were collected using demographic questionnaires and STAT. The instruments were administered to four schools where 180 students' responses to the instruments were collected. Students were briefed about the instructions and supervised throughout the allocated time to finish the test. They were required to write their answers in the answer sheet of the bilingual instruments. The test took approximately two hours. The duration of time was allocated based on pilot test and feedbacks from experts.

3.7 Data Analysis Procedure

The data analysis procedure consisted of scoring the Statistical Thinking Abilities Test (STAT) using the rubric and analyzing the scores. First of all, the data collected from the students' responses were scored through matching the students' responses with the rubric by the researcher. The weightage of the scores was decided and assigned based on descriptors of the rubric. Scores obtained in terms of one up to four points indicating the four levels in the statistical thinking framework. The median for the scores obtained was calculated to show the students' level of statistical thinking and level for each sub process.

The scores were assessed later to answer the research questions. The statistical analysis of the study was conducted using Statistical Package for the Social Sciences (SPSS). The analysis was done using a few different hypothesis tests. A table presenting a list of research questions and its matching statistical test analysis were presented as seen in table 3.13 below. The test of assumptions required to run the hypotheses tests for each research question were examined before the carrying out the analysis.

Table 3.13

Data Analysis Method for Each Research Questions

Research Questions	Statistical Analysis
RQ1: What are the Form Four pupils' statistical thinking abilities?	Descriptive
RQ2: Is there any significant difference in the statistical thinking abilities of Form Four pupils by gender?	Mann-Whitney U test
RQ3: Is there any significant difference in the statistical thinking abilities of Form Four pupils by socio-economic status?	Kruskal Wallis H test

The first research question was evaluated based on descriptive statistics. The scores obtained from the STAT instrument in identifying students' statistical thinking abilities were analyzed to compute the frequency and percentage. Bar chart was used to portray the frequency of scores obtained by 180 students in each of the statistical thinking abilities processes to show students' level of statistical thinking abilities.

The second research question was analyzed using Mann-Whitney U test to identify the difference between statistical thinking abilities by two gender groups mainly male and female. There were several assumptions to be met for using this test analysis. The first assumption was that the measurement should be done at ordinal or continuous level for dependent variable. In this study, dependent variable of statistical thinking abilities was measured at four ordinal levels, thus fulfilled the first assumption. The second assumption was the categorization of independent variables into two categorical and independent groups. In this study, the independent variables of gender with two groups of male and female met the criterion. The third assumption was to have independence of observations between groups. The samples in each group for this study were different. The fourth assumption was for the two variables to be not normally distributed. In this study, the obtained data for statistical thinking abilities levels by gender were not normally distributed. The fifth assumption was that the shape of both distributions to be the same right skewed shape, as the data were not normally distributed.

The third research question was analyzed using Kruskal Wallis H test to identify the differences of three socio-economic status groups mainly low, middle, and high socio-economic status by statistical thinking abilities. There were several assumptions to be met for usage of this test analysis which were similar to the assumptions of Mann Whitey U test. The first assumption stated the measures to be done at ordinal or continuous level for dependent variable. Statistical thinking ability was the dependent variable and measured at four ordinal levels, thus fulfilled the first assumption. The second assumption indicating requirement of three categorical and independent groups for the independent variable. The socio-economic status variable consisted of three categorical and independent groups which were low, middle, and high. Thus, the second assumption was fulfilled. The third assumption was to have independence of observations between groups. The samples in each group for this study were different. The fourth assumption was for the two variables to be not normally distributed. In this study, the obtained data for scores of statistical thinking abilities levels by socio-economic status were not normally distributed. The fifth assumption was for that the distributions for both groups to have the same variability, as the data were not normally distributed.

3.8 Ethics

Before conducting the data collection through the administration of instruments, the necessary approvals were required. First of all, the application for the approval to conduct research was applied to the Ministry of Education, Malaysia. Upon getting the letter of approval from the ministry, the application for approval to conduct research was then applied to the Selangor State Education Department. After getting the approval letter from both Ministry of Education, Malaysia and Selangor State Education Department, an appointment with principal of the schools in area of Shah Alam was arranged to obtain permission for data collection. Once the permission from school was granted, the instruments were administered to selected classrooms of Form Four students of the selected schools. Students were briefed of their privacy's protection before the administration of the instruments. Students' confidential information such as their names and household income would not be revealed. The names of students were coded and assigned with random numbers when entering the data collected from the instrument into SPSS. This information was stored separately from the data in hard copy format in the researcher office's locked cabinet.

3.9 Summary

All secondary schools Form Four students studying in the area of Shah Alam were constituted as the population. The sample comprised of 180 pupils. Survey descriptive and causal-comparative design were used. The instrument consisted of demographic questionnaires, Statistical Thinking Abilities Test (STAT), and rubric. The STAT was developed by adaptation of instruments from previous research to collect students' responses. The rubric was adopted from the statistical thinking framework of Langrall and Mooney (2002). The instrument was pilot-tested by the researcher before the actual study to establish the validity and reliability of STAT. Result of inter-rater reliability showed that the rubric scores correlated 96% of the time. Test-retest reliability results revealed that the STAT scores were significantly stable. The scores of pre and post tests were analyzed descriptively and inferentially to answer three research questions using multiple statistical tests namely descriptive, Mann-Whitney

U test, and Kruskal Wallis H test. The results of this study were presented in the subsequent chapter.

CHAPTER 4

FINDINGS

4.1 Introduction

This chapter presented the descriptive and inferential analysis used in this study and its results. The sample's demographic data obtained from questionnaire was highlighted. Research question one was evaluated using descriptive statistics while research question two was evaluated using Mann Whitney U test. In addition, research question three was analyzed using Kruskal Wallis H test. Before conducting hypotheses-testing on the research questions, the assumptions required were tested. The hypotheses were then examined at 5% level of significance.

4.2 Demographic Data

Demographic data was collected from 180 samples using a demographic questionnaire. The gender of samples was categorized into male and female, where 90 students were of male and another 90 students were of female as seen in Table 4.1.

Table 4.1

Descriptive Statistics for 180 Samples' Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	90	50.0	50.0	50.0
	Female	90	50.0	50.0	100.0
	Total	180	100.0	100.0	

Figure 4.1 presented the bar chart for male and female's number of students.

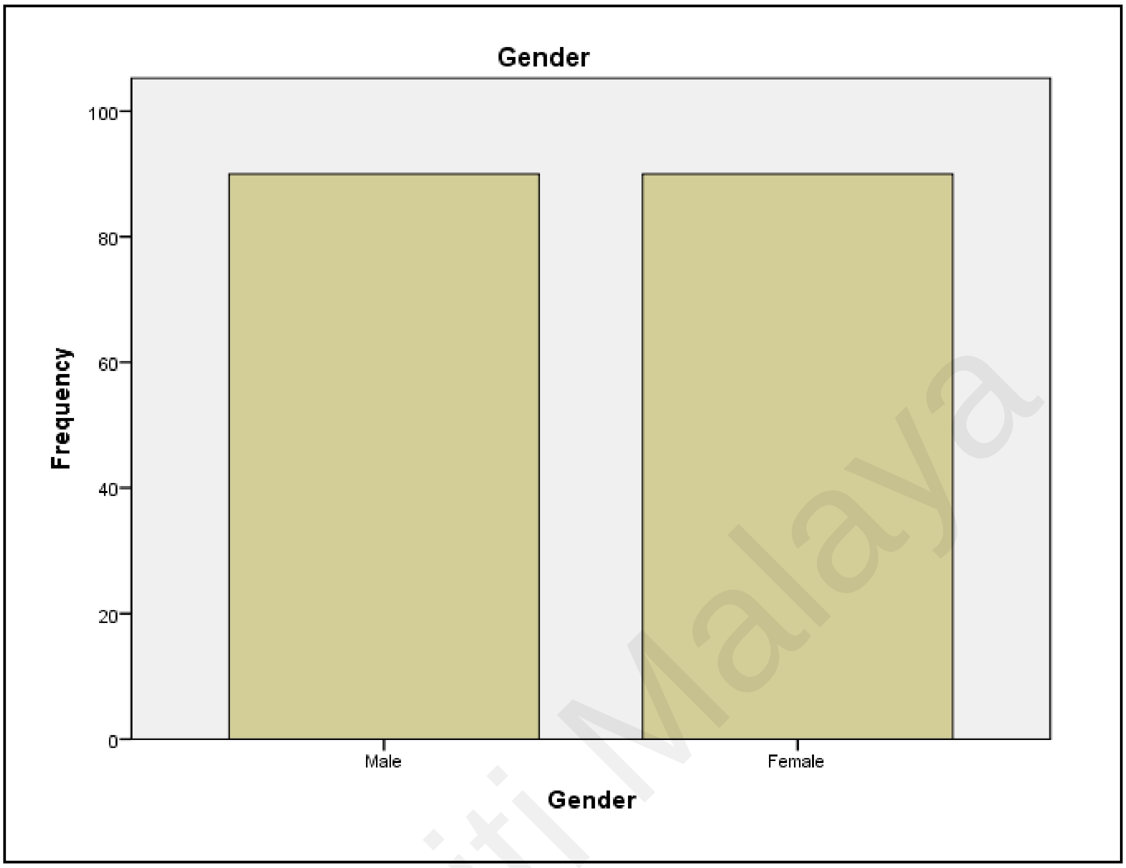


Figure 4.1. Bar Chart for Frequency of Male and Female Students

The socio-economic status of samples was categorized into low, middle, and high. The data collected from 180 samples shown that 65 students (36.1%) were at low socio-economic status, 98 students (54.1%) were at middle socio-economic status, and 17 students (9.4%) were at high socio-economic status, as shown in Table 4.2.

Table 4.2

Descriptive Statistics for 180 Samples' Socio-Economic Status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low	65	36.1	36.1	36.1
	Middle	98	54.4	54.4	90.6
	High	17	9.4	9.4	100.0
	Total	180	100.0	100.0	

Figure 4.2 presented the bar chart for number of students in three socio-economic groups.

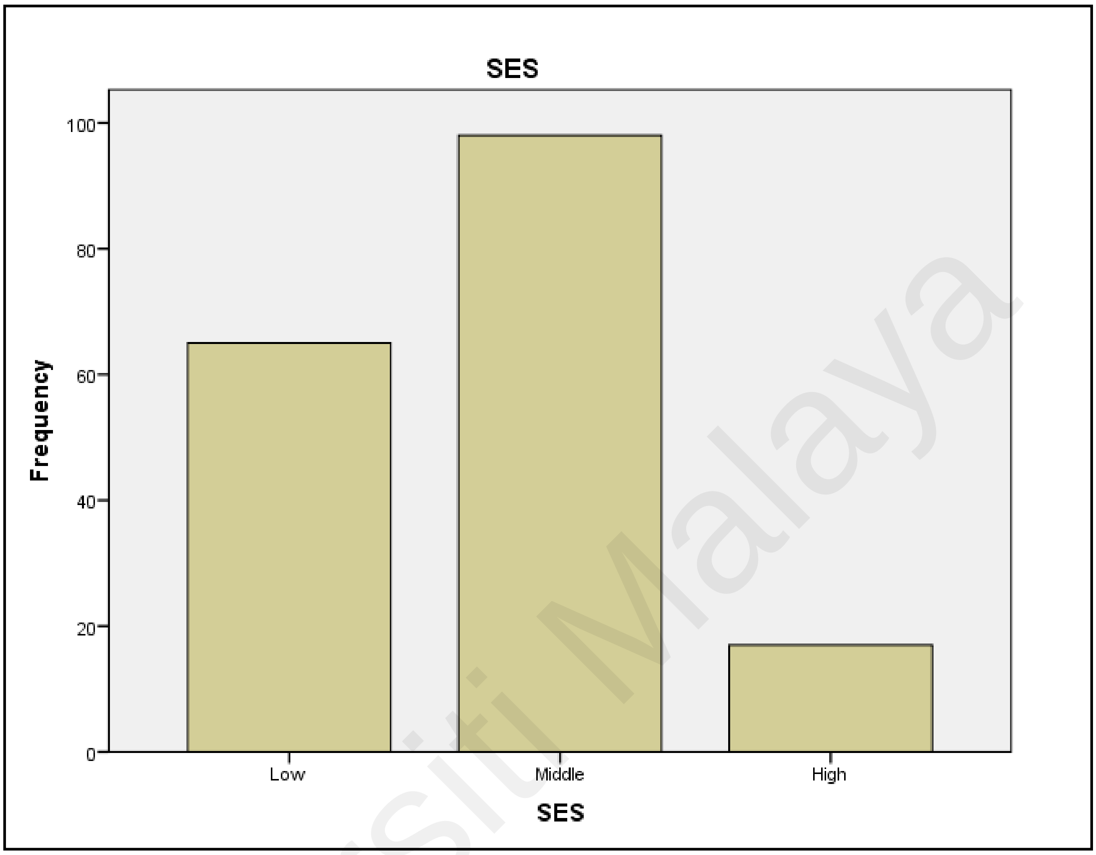


Figure 4.2. Bar Chart for Frequency of Low, Middle, and High Socio-Economic Status Groups

Out of 65 students from the low socio-economic status group, male students consisted of 30 students while female students consisted of 35 students. Meanwhile, out of 98 students from the middle socio-economic status group, 47 students were male, and 51 students were female. Lastly, 17 students from the high socio-economic status group consisted of 13 male students and 4 female students. Table 4.3 shown the number of students in each of the group for the 180 samples.

Table 4.3

Descriptive Statistics for 180 Samples' Gender and Socio-Economic Status

		Male		Female	
		n	%	n	%
Socio-economic status	Low	30	33%	35	39%
	Middle	47	52%	51	57%
	High	13	14%	4	4%
Total		90	100%	90	100%

4.3 Reporting of Findings for Research Question One

Descriptive statistics was utilized to answer the first research question of the study. The data obtained from the samples was used to compute the frequency and percentage. The students' scores were categorized according to four levels mainly: Level 1 (Idiosyncratic), Level 2 (Transitional), Level 3 (Quantitative), and Level 4 (Analytical). 76 students (42.2%) were at Level 1 statistical thinking ability, while 83 students (46.1%) were at Level 2 statistical thinking Ability. 18 students (10%) were at Level 3 statistical thinking ability. Lastly, 3 students (1.7%) were at Level 4 statistical thinking ability. Most of the students were at Level 2 of the statistical thinking abilities, shown in Table 4.4.

Table 4.4

Descriptive Statistics for 180 Students' Statistical Thinking Abilities Levels

	Level	Frequency	Percent
Valid	1	76	42.2
	2	83	46.1
	3	18	10.0
	4	3	1.7
	Total	180	100.0

Figure 4.3 presented the bar chart for frequency of 180 students' level of statistical thinking abilities.

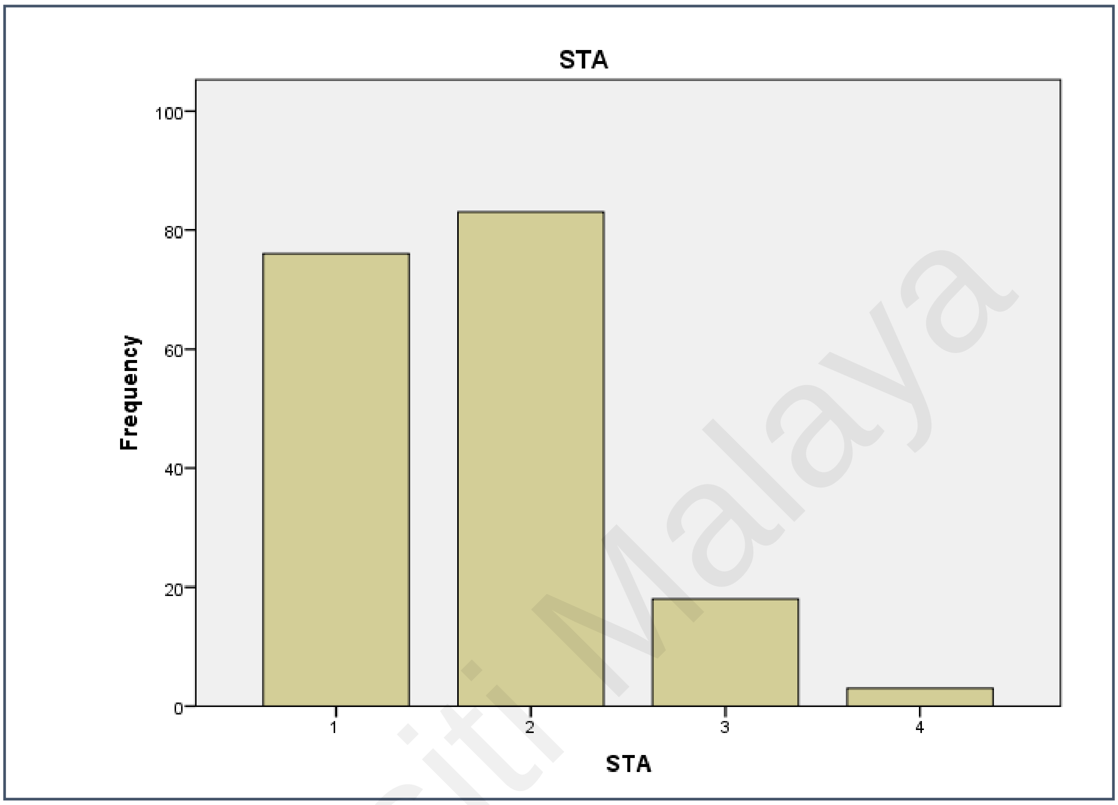


Figure 4.3. Bar Chart for Frequency of 180 Students' Levels of Statistical Thinking Abilities

The three processes of statistical thinking abilities were discussed as well according to the four levels. Table 4.5 showed the four levels of statistical thinking ability of students in the first process Ability to Organize and Reduce Data. 77 students (42.8%) were at Level 1 of the first process Ability to Organize and Reduce Data. 83 students (46.1%) were at Level 2. 18 students (10%) were at Level 3. Lastly, 2 students (1.1%) were at Level 4.

Table 4.5

Descriptive Statistics for 180 Students' Levels in Ability to Organize and Reduce Data

	Level	Frequency	Percent
Valid	1	77	42.8
	2	83	46.1
	3	18	10.0
	4	2	1.1
	Total	180	100.0

Figure 4.4 presented the bar chart for count of 180 students' level in Organize and Reduce Data process.

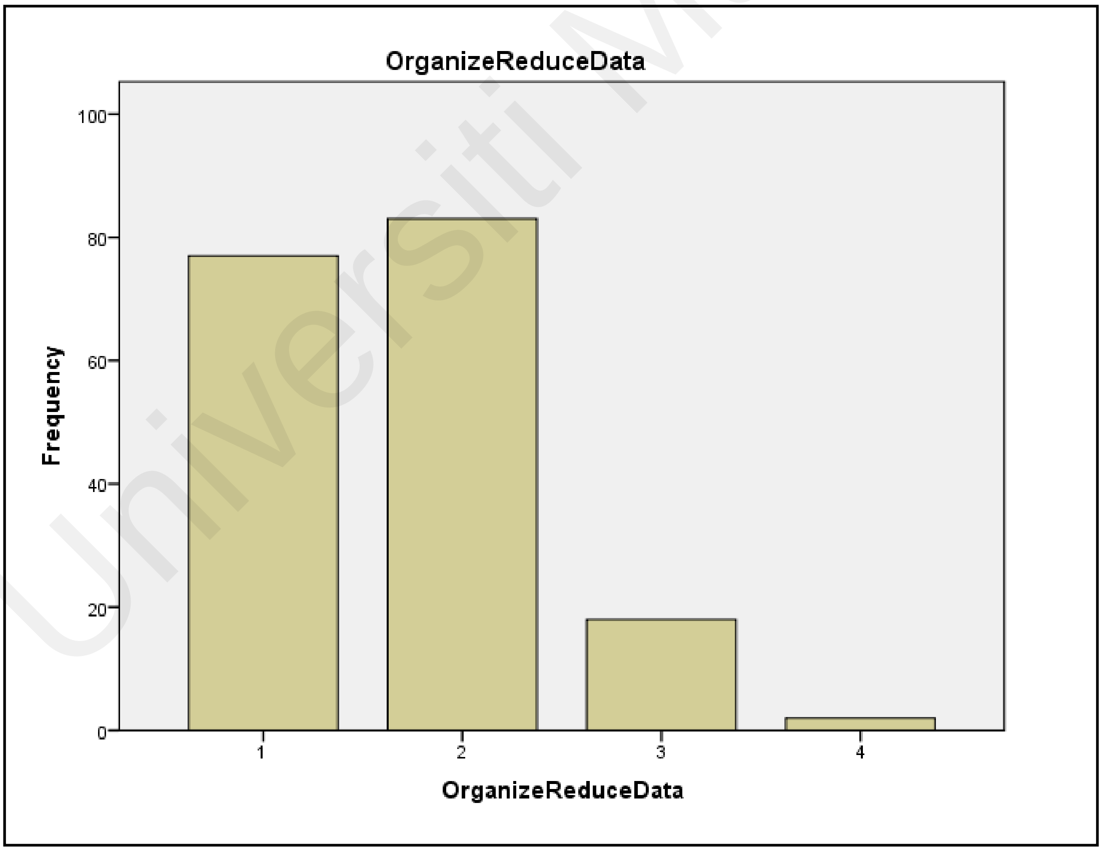


Figure 4.4. Bar Chart for Frequency of 180 Students' Levels in Organize and Reduce Data

Table 4.6 showed the four levels of statistical thinking ability of students in the second process Ability to Represent Data. 80 students (44.4%) were at Level 1 of the second process Ability to Represent Data. 74 students (41.1%) were at Level 2. 16 students (8.9%) were at Level 3. Lastly, 10 students (5.6%) were at Level 4.

Table 4.6

Descriptive Statistics for 180 Students' Levels in Ability to Represent Data

		Frequency	Percent
Valid	1	80	44.4
	2	74	41.1
	3	16	8.9
	4	10	5.6
	Total	180	100.0

Figure 4.5 presented the bar chart for count of 180 students' level in Represent Data process.

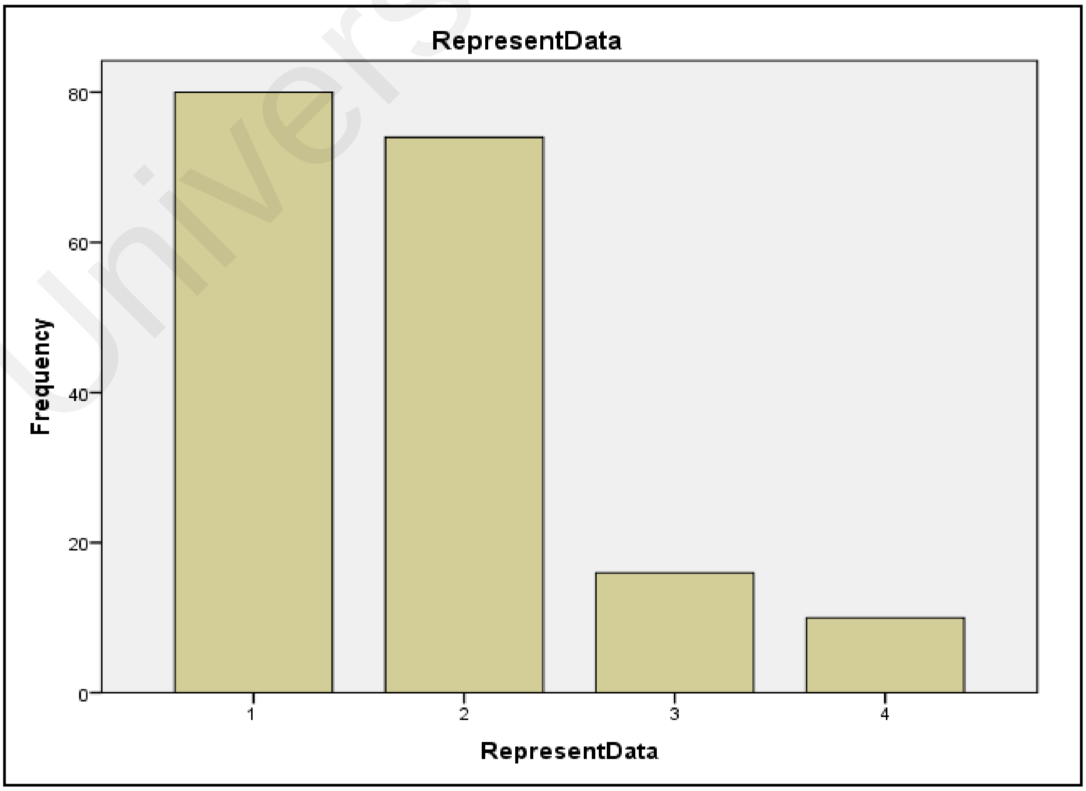


Figure 4.5. Bar Chart for Frequency of 180 Students' Levels in Represent Data

Table 4.7 showed four levels of statistical thinking ability of students in third process Ability to Analyze and Interpret Data. 80 students (44.4%) were at Level 1 of the third process Ability to Analyze and Interpret Data. 78 students (43.3%) were at Level 2. 12 students (6.7%) were at Level 3. Last, 10 students (5.6%) were at Level 4.

Table 4.7

Descriptive Statistics for 180 students' Levels in Ability to Analyze and Interpret Data

		Frequency	Percent
Valid	1	80	44.4
	2	78	43.3
	3	12	6.7
	4	10	5.6
	Total	180	100.0

Figure 4.6 presented the bar chart for count of 180 students' level in Analyze and Interpret Data process.

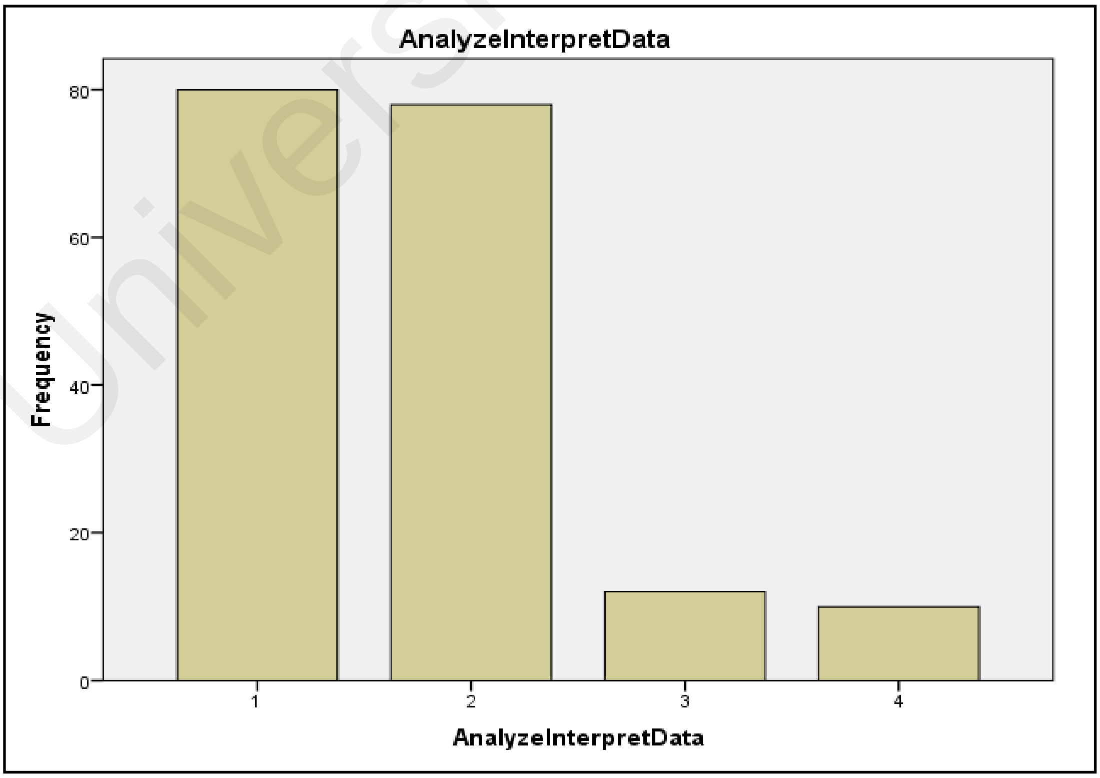


Figure 4.6. Bar Chart for Frequency of 180 Students' Levels in Analyze and Interpret Data

4.4 Reporting of Findings for Research Question Two

A Mann-Whitney U test was conducted to find out if statistical significance difference happened between statistical thinking abilities by gender. Mann-Whitney U test was selected due to the fact that the measured dependent variable was ordinal level. Before conducting analysis using Mann-Whitney U test, several assumptions have to be fulfilled.

The first assumption of dependent variables to be measured at the ordinal level was met. The dependent variable which is the statistical thinking abilities was measured using four levels according to the framework, where the four levels are ordinal. The second assumption of two categorical group of independent variables are met. Gender variable consists of male and female groups. The third assumption of the test was that the groups were of independence of observations. The samples participated in each group were different and independent of each other.

The result revealed that overall male students have slightly higher statistical thinking abilities level than female students, as portrayed in Table 4.8. The distribution of for both gender shown right skewed shape based on the mean ranks.

Table 4.8

Mean Rank Table for Gender

	Gender	N	Mean Rank	Sum of Ranks
STA	Male	90	95.61	8605.00
	Female	90	85.39	7685.00
	Total	180		

Table 4.9 presented result of the Mann-Whitney U test analysis which shown that the male pupils' level of statistical thinking abilities was higher than female students (U = 3590, p = .148). However, as the p value is more than .05, it can be

interpreted that there was no significant difference of levels of statistical thinking abilities by gender.

Table 4.9

Mann-Whitney U Test of Levels of Statistical Thinking Abilities by Gender

	STA
Mann-Whitney U	3590.000
Asymp. Sig. (2-tailed)	.148

a. Grouping Variable: Gender

4.5 Reporting of Findings for Research Question Three

A Kruskal-Wallis H test was carried out to find out whether a statistical significance difference existed between median of statistical thinking abilities by socio-economic status. Kruskal-Wallis H test was selected due to the fact that the measured dependent variable was at ordinal level. Before conducting analysis using Kruskal-Wallis H test, several assumptions have to be fulfilled.

The first assumption of dependent variables to be measured at the ordinal level was met. The dependent variable which is the statistical thinking abilities was measured using four levels according to the framework, where the four levels are ordinal. The second assumption of two or more categorical group of independent variables are met. Socio-economic status variable consists of three groups which are low, middle, and high. The third assumption of the test was that the groups were of independence of observations. The samples participated in each group were different and independent of each other.

The result revealed that overall students from three group of socio-economic status were at level 2 of statistical thinking abilities. A Kruskal-Wallis H test indicated a statistically significant difference in statistical thinking abilities of the socio-economic status groups, $\chi2(2) = 34.591$, $p = .001$, as seen in Table 4.10. The result shown that the three group of socio-economic status have the same shape based on the different mean ranks obtained by the three groups.

Table 4.10

Kruskal Wallis Test of Levels of Statistical Thinking Abilities by Socio-Economic Status

Test Statistics[a,b]	STA
Chi-Square	34.591
df	2
Asymp. Sig.	.001

a. Kruskal Wallis Test
b. Grouping variable: SES
STAT = Statistical Thinking Abilities

The mean rank statistical thinking abilities were shown in Table 4.11 of 60.07 for low socio-economic status, 101.06 for middle socio-economic status, and 145.97 for high socio-economic status.

Table 4.11

Mean Rank Table for Socio-Economic Status

	SES	N	Mean Rank
STA	Low	65	60.07
	Middle	98	101.06
	High	17	145.97
	Total	180	

4.6 Summary

A section on demographic data revealed the information about the gender and group and socio-economic status group of samples in frequency and percentages. Descriptive statistics was used to answer the first research question and the results revealing majority students at level 1 and 2 of statistical thinking abilities. Mann-Whitney U test was used to answer the second research question and result showed no significant difference of levels of statistical thinking abilities by gender. Kruskal-Wallis H test was used to investigate the third research question and revealed significant difference of statistical thinking abilities' level by socio-economic status.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Introduction

This study was conducted to determine the effects of gender and socio-economic status on Form Four pupils' statistical thinking abilities. This chapter presented findings with attempts to obtain answers for the research questions. Data was collected and analyzed statistically using descriptive statistics, Mann-Whitney U test, and Kruskal-Wallis H test. This chapter presented summary of findings, thorough findings' discussion with conclusions, implications of the study, and recommendations for further studies.

5.2 Summary of Findings

5.2.1 Research Question One

What are the statistical thinking abilities of Form Four pupils?

The first research question was analyzed using descriptive statistics. The results showed that most students were at level 1 and level 2 of statistical thinking abilities.

In terms of the three processes of statistical thinking abilities, 77 students (42.8%) were at Level 1, 83 students (46.1%) were at Level 2, 18 students (10%) were at Level 3, and 2 students (1.1%) were at Level 4 of the first process Ability to Organize and Reduce Data. Meanwhile, 80 students (44.4%) were at Level 1, 74 students (41.1%) were at Level 2, 16 students (8.9%) were at Level 3, and 10 students (5.6%) were at Level 4 of the second process Ability to Represent Data. Lastly, 80 students (44.4%) were at Level 1, 78 students (43.3%) were at Level 2, 12 students (6.7%) were at Level 3, and 10 students (5.6%) were at Level 4 of the third process Ability to Analyze and Interpret Data.

5.2.2 Research Question Two

Is there any significant difference in the statistical thinking abilities of Form Four pupils by gender?

The second research question was analyzed using Mann-Whitney U test. The result showed that male students' level of statistical thinking abilities was higher than female students (U = 3590, p = .148). Based on the results, data failed to provide sufficient evidence to make conclusion that there is significant difference of statistical thinking abilities by gender.

5.2.3 Research Question Three

Is there any significant difference in the statistical thinking abilities of Form Four pupils by socio-economic status?

The third research question was analyzed using Kruskal-Wallis H test. The result shown that there was statistically significant difference in statistical thinking abilities of the socio-economic status groups, $\chi 2(2)$ = 34.591, p = .001, with mean rank statistical thinking abilities of 60.07 for low socio-economic status, 101.06 for middle socio-economic status, and 145.97 for high socio-economic status.

5.3 Discussion

In this section, the discussions of the results in chapter four is presented. The discussion is divided into two parts and two major findings basing upon the two objectives of the study. The first part discussed the effect of gender on Form Four pupils' statistical thinking abilities, while the second part discussed the effect of socio-economic status on Form Four pupils' statistical thinking abilities.

5.3.1 Effect of Gender on Statistical Thinking Abilities of Form Four Pupils

There were several studies that prompted the research into gender effect. Martin et al. (2017) studied on the effect of gender on statistical reasoning, where gender was found to show statistically significant direct effect in influencing statistical reasoning directly, and indirectly through influence on thinking dispositions. In their studies, Martin et al. (2017) reported that males have higher statistical reasoning in comparison with females. Piaw (2014) conducted study to identify the effects of gender group and brain thinking style group on Form Six students' creative thinking in Malaysia. Results shown that the gender variable and thinking style variable were factors that affected creative thinking significantly.

Despite that, the findings from current study indicated no evidence to show effect of gender on statistical thinking abilities. The findings were in line with the reported results in a study conducted by (Matud, Rodgriguez, and Grande, 2007), where gender had no significant effect on creative thinking. The results shown that despite men showing slight superiority than female on the measures, the differences were slight and disappeared when women had high educational level.

The question come to the reason of the absence of effect of gender on statistical thinking abilities. Ewumi (2012) highlighted that gender's development was affected by adolescents' perception and impersonation, and that good gender behavior to be rewarded and bad gender behavior to be punished. Furthermore, a proposed theory explaining the traits existed in the two groups of gender suggested that male representing cognitive domain while female representing emotional domain (Hall & Lucas, 1976 in Klein, 2004). Hyde and Else-Quest (2012) highlighted the overlaps in the cognitive differences of male and females in terms of mathematics and visuospatial tasks.

Hyde (2005) conducted comprehensive meta-analyses on gender differences for many cognitive tasks and identified close-to-zero differences in mathematical abilities for gender. Lindberg et al. (2010) agreed by stating that many studies conducted in 1990 to 2007 shown no differences in mathematics performance by gender.

Sample size might play a role as well in its effect size power in affecting the significance of the gender factor. As the study utilized moderate sample size for gender, it might be inadequate to see the effects on statistical thinking abilities of pupils.

Students from both gender were exposed to the same curriculum and pedagogy in schools. Based on that, both gender students were expected to have the same thinking abilities. This is in line with the result of the study that show no effect of gender variable on affecting statistical thinking abilities of pupils.

The results conducted revealed higher statistical thinking abilities of male students compared to female students. The result was in line with the view of Steele (1997) and Macher et al. (2012) that female students tend to have statistics anxiety and thus resulted in portrayal of lower statistical thinking abilities by them in comparison to male students.

As there was a lack of gender effect research on statistical thinking abilities, and with available literatures only on gender focusing on other thinking abilities, it was uncertain whether gender was a factor that affect student's statistical thinking abilities. This led to the line of thought that gender itself might not be a factor that showed effect on statistical thinking abilities. The differences in statistical thinking abilities cannot be assigned to gender. Thus, gender did not have main effect on students' statistical thinking abilities.

5.3.2 Effect of Socio-economic Status on Statistical Thinking Abilities of Form Four Pupils

There were several studies that prompted the study on effect of socio-economic status variable. Jankowska and Karwowski (2019) conducted a study to determine family socio-economic status' effect on the development of children's creative thinking. The result indicated significant differences in children's creative thinking by family's socio-economic status. It was identified that higher socio-economic status families had children with higher creative thinking, and vice versa. Norfadillah et al. (2017) studied on socio-economic status' effect on preschooler's mental abilities in Klang Valley area, also revealed result of significant differences in children's cognitive abilities by parents' education. Children had higher cognitive abilities with parents that had higher education. The results from these studies were consistent with the findings from current study that indicated effect of socio-economic status on Form Four pupils' statistical thinking abilities.

The effect of socio-economic status variable on statistical thinking abilities was observed at three different groups mainly the low socio-economic status group, middle socio-economic status group, and high socio-economic status group. The effect was supported by the claim of (Blums, Belsky, Grimm, and Chen, 2017) that a family's socio-economic status affected students' cognitive ability. Hackman, Farah, & Meaniey (2010) highlighted that students' development of brain structures and cognitive functions were affected by socio-economic status. Parents coming from higher socio-economic status group were prone to be committed with engaging with children activities that triggered intellectual ideas (Conger & Donnellan, 2007). These parents knew the importance of children's development of intellectual abilities and prioritized children's education and intellectual abilities' development at young age. Meanwhile, lower socio-economic status family's children tend to be deprived of these privileges and educational materials, resulting in delays in children's cognitive abilities and often prevailed compared to other peers (Fasig, 2000). Low socio-economic status school areas did not have adequate expertise and funding to cover the inequalities imposed by low socio-economic status home environment (Ewumi, 2012).

The results conducted revealed highest statistical thinking abilities from high socio-economic status group, and lowest statistical thinking abilities from the low socio-economic status group. This result was in line with the study conducted by Jankowska and Karwowski (2019) that higher socio-economic status family with parental involvement had big impact on the development of the students' statistical thinking abilities.

There were many studies that related socio-economic status as the main factor in the educational researches. There were sufficient evidences to conclude the socio-economic status' effects on pupils' statistical thinking abilities. As such, differences in statistical thinking abilities can be assigned to socio-economic status variable.

5.4 Implication of the Study

This study has important implications for improving pupils' statistical thinking abilities. The result revealed that gender was not a factor that affected students' statistical thinking abilities. Meanwhile, results shown that socio-economic status variable had effect on pupils' statistical thinking abilities. The findings provide implications to teachers, curriculum planners, and researchers.

5.4.1 Implication to Teachers

The rubric for the statistical thinking abilities framework from this study can be utilized by school teachers to assign scores to pupils' responses and understand better the level of pupils' statistical thinking abilities according to the three processes of statistical thinking abilities, mainly data organization and reduction ability, data representation ability, and data analyzation and interpretation ability. Teachers with knowledge of students' thinking levels will have idea on how to guide students to progress and attain next desired level of thinking. Teachers can focus on the abilities of students solely on the respective processes and improve student's statistical thinking abilities from there.

The statistical thinking processes can be deemed as the students' learning outcome in their lesson plan. The descriptors will allow the teachers to gauge the levels of students' statistical thinking in certain components. For instance, students should be able to create a complete, representative and appropriate display. Students should also be able to assess whether the displayed data was effective by judging upon the features. Teachers are able to design appropriate and effective instruction for the students according to the sub processes of learning outcomes.

This study also alerts teachers that the area of statistics covered in school's Mathematics curriculum is inadequate as it is only a small subtopic throughout the whole Form Four of Mathematics curriculum syllabus. Teachers may need to incorporate additional instruction in their lesson plan to involve real time situation problems and expose the students with some basic statistical software for statistical data analysis which are suitable for their syllabus.

By implying that gender is not an effect on statistical thinking abilities, teachers may eliminate that factor and focus on teaching methods or instructions that does not involve gender-based activities. Teachers can monitor and identify students coming from different background of socio-economic to assist the development of their statistical thinking abilities appropriately. Teachers can communicate with students' parents as well to understand the students' background to identify the issues faced by the students to solve them.

5.4.2 Implication to Researchers

The findings of this study will have great implication to researchers. For researchers who plan to conduct research in statistical thinking abilities, they are able to further pursue in depth study on the statistical thinking abilities by looking into other factors which might affect the students' statistical thinking abilities and the interaction effect among them.

As researchers need proper measurement data in order to understand and conduct their research, the instruments in this study can serve as a guideline to be adapted or adopted by researchers to utilize them in measuring statistical thinking abilities of students. With a statistical thinking framework presented in this study, researchers are able to improve on the current instruments to incorporate questions that could elicit students' process of thoughts more effectively and further include the observed responses into the existing framework.

5.4.3 Implication to Curriculum Planners

The findings of this study has implication to curriculum developers and textbook writers. Curriculum developers are able to understand statistical thinking abilities and its processes better and incorporate them into the mathematics curriculum to suit students of all abilities with the availability of instruments and rubric for scoring from this study. Curriculum developers and textbook writers can gain insight from the study to make improvement to the current curriculum and include activities in the form of modules or instructional materials to serve as guideline for the teachers.

5.4.4 Implication to Teaching Educators

Teaching educators also benefits from the implication of this study in a way that they could gauge the ideas and operations behind statistical thinking and apply the knowledge to monitor student's levels of statistical thinking abilities to ensure that students develop thinking abilities according to their own pace so that they would not prevail from their peers. Educators could also take note regarding the effect of socio-economic status variable on statistical thinking abilities from the findings of this study and plan the necessary pedagogy to be implemented in lessons on students from various socio-economic status backgrounds.

5.5 Contribution of this study

This study has provided contribution to the theory. As the theory for this study was adapted using the statistical thinking framework by Mooney and Langrall (2002), the rubric and instruments were able to assign scores to the students' responses to identify the levels of statistical thinking of the students.

5.6 Recommendations for Future Research

Based upon the findings and the implications, there is a need to carry out further research in the future. Further research could be done in the future for addressing the limitation and delimitations to overcome them.

Referring to the limitation, the study is delimited to only the population of every secondary school pupil in the Shah Alam area, Selangor. Thus, to make the finding more generalize to bigger population, further study is recommended to expand the scope to include other locations and other states such as Perak, Kelantan, Penang.

Referring to the limitation, the study is delimited to only Form Four syllabus of statistics that cover the measure of central tendency, and graphical representation. Further study should include other areas of statistics.

Other level of education could be considered as well besides secondary school level education. The current study is only limited to the subpopulation of Form Four students in Shah Alam area of schools. Further study should include other Forms.

Referring to the limitation, the study is delimited to only survey and causal-comparative research design. Further study should include other research designs such as experimental research design. Experimental research design could be done to include an approach different from the traditional to elicit students' statistical thinking abilities.

The current study only incorporates three processes of statistical thinking abilities. Future researchers could investigate statistical processes or sub-processes that were left out from the framework to further improve the current framework.

This study has revealed result of no effect of gender on statistical thinking abilities with no significance difference of students' statistical thinking abilities by gender. Future researchers could attempt to address this, by using larger sample size and increase the effect power.

5.7 Conclusion

Findings from past researches have shown gaps in students' statistical thinking in comprehending statistical concepts. It is necessary to identify students' levels of statistical thinking to address the gaps which prompted the current study that intended to measure students' levels of statistical thinking abilities, and furthermore the factors that have effect on it.

This study intended on identifying the effects of gender variable and socio-economic status variable on Form Four students' statistical thinking abilities. The findings have shown evidence that show significant effects of socio-economic status on students' statistical thinking abilities.

This study hopes to be a catalyst that prompts future researchers and mathematics educators to apply the findings to various different area of school statistics to improve the statistical thinking abilities levels of the students.

References

Aizikovitsh-Udi, E., Clarke, D., & Kuntze, S. (2014). Hybrid tasks: Promoting statistical thinking and critical thinking through the sema mathematical activities.

Alordiah, C. O., Akpadaka, G., & Oviogbodu, C. O. (2015). The Influence of Gender, School Location and Socio-Economic Status on Students' Academic Achievement in Mathematics. *Journal of Education and Practice, 6*(17), 130-136.

Association, A. S. (1991). Guidelines for the teaching of statistics. *Alexandria, VA: Author*.

Baer, J., & Kaufman, J. C. (2008). Gender differences in creativity. *The Journal of Creative Behavior, 42*(2), 75-105.

Baltaci, S., Yildiz, A., & Özcakir, B. (2016). The relationship between metacognitive awareness levels, learning styles, genders and mathematics grades of fifth graders. *Journal of Education and Learning, 5*(4), 78.

Bart, W. M., Hokanson, B., Sahin, I., & Abdelsamea, M. A. (2015). An investigation of the gender differences in creative thinking abilities among 8th and 11th grade students. *Thinking Skills and Creativity, 17*, 17-24.

Bécares, L., & Priest, N. (2015). Understanding the influence of race/ethnicity, gender, and class on inequalities in academic and non-academic outcomes among eighth-grade students: findings from an intersectionality approach. *PloS one, 10*(10), e0141363.

Benner, A. D., Boyle, A. E., & Sadler, S. (2016). Parental involvement and adolescents' educational success: The roles of prior achievement and socioeconomic status. *Journal of youth and adolescence, 45*(6), 1053-1064.

Ben-Zvi, D., & Friedlander, A. (1997). Statistical thinking in a technological environment. *Research on the role of technology in teaching and learning statistics*, 45-55.

Ben-Zvi, D., & Garfield, J. B. (2004). *The challenge of developing statistical literacy, reasoning and thinking*: Kluwer academic publishers.

Biggs, J. (1991). Multimodal learning and the quality of intelligent behavior. *Intelligence: Reconceptualization and measurement*, 57-76.

Biggs, J. B. (2011). *Teaching for quality learning at university: What the student does*: McGraw-hill education (UK).

Biggs, J. B., & Collis, K. F. (1982). *Evaluation the quality of learning: the SOLO taxonomy (structure of the observed learning outcome)*: Academic Press.

Blums, A., Belsky, J., Grimm, K., & Chen, Z. (2017). Building links between early socioeconomic status, cognitive ability, and math and science achievement. *Journal of Cognition and Development, 18*(1), 16-40.

Bourdieu, P., & Passeron, J.-C. (1990). *Reproduction in education, society and culture* (Vol. 4): Sage.

Bower, G. H., & Hilgard, E. R. (1981). *Theories of learning*: Prentice-Hall.

Buchmann, C., & DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American sociological review, 71*(4), 515-541.

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annu. Rev. Sociol, 34*, 319-337.

Caro, D. H. (2009). Socio-economic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education, 32*(3), 558-590.

Carolan, B. V., & Wasserman, S. J. (2015). Does parenting style matter? Concerted cultivation, educational expectations, and the transmission of educational advantage. *Sociological Perspectives, 58*(2), 168-186.

Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary educational psychology, 26*(3), 330-347.

Castles, I., Tayt, J., & O'Donnell, G. (1992). *Surviving statistics: A user's guide to the basics*: Australian Bureau of Statistics.

Chan, S. W., & Ismail, Z. (2014). A Technology-Based Statistical Reasoning Assessment Tool in Descriptive Statistics for Secondary School Students. *Turkish Online Journal of Educational Technology-TOJET, 13*(1), 29-46.

Chance, B., delMas, R., & Garfield, J. (2003). *The Web-Based ARTIST: Assessment resource tools for improving statistical thinking project.* Paper presented at the Invited paper, Joint Statistical Meetings, San Francisco. Retrieved from https://apps3. cehd. umn. edu/artist/articles/AERA_2003. pdf.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education, 10*(3).

Cheadle, J. E., & Amato, P. R. (2011). A quantitative assessment of Lareau’s qualitative conclusions about class, race, and parenting. *Journal of family Issues, 32*(5), 679-706.

Chiesi, F., & Primi, C. (2015). *Gender differences in attitudes toward statistics: Is there a case for a confidence gap?*

Clarke-Stewart, K. A., Vandell, D. L., Burchinal, M., O’Brien, M., & McCartney, K. (2002). Do regulable features of child-care homes affect children’s development? *Early childhood research quarterly, 17*(1), 52-86.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in mathematics Education*, 3-29.

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annu. Rev. Psychol., 58*, 175-199.

Council, N. R., & Committee, M. L. S. (2001). *Adding it up: Helping children learn mathematics*: National Academies Press.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative*: Prentice Hall Upper Saddle River, NJ.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in mathematics Education*, 382-393.

Dee, T. S. (2006). The why chromosome: How a teacher's gender affects boys and girls. *Education Next, 6*(4), 68-76.

Ding, C. (2006). Using regression mixture analysis in educational research. *Practical Assessment Research and Evaluation, 11*(11), 1-11.

Downer, J. T., & Pianta, R. C. (2006). Academic and cognitive functioning in first grade: Associations with earlier home and child care predictors and with concurrent home and classroom experiences. *School Psychology Review, 35*(1), 11.

Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *Sociological quarterly, 46*(2), 299-321.

Duerr, S. R. (2012). Relationships between Gender, Socioeconomic Status, Math Attitudes, and Math Achievement: An International Investigation. *ProQuest LLC*.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement*: Prentice Hall.

Edinyang, S., Ubi, I., Usang, E. E. P., & Adalikwu, R. A. (2013). Gender, socio-economic status, teacher qualification and their interaction on students' retention ability in Social Studies in Akwa Ibom State Nigeria. *International Knowledge Sharing Platform, 2*, 35-40.

Englund, M. M., Luckner, A. E., Whaley, G. J., & Egeland, B. (2004). Children's achievement in early elementary school: Longitudinal effects of parental involvement, expectations, and quality of assistance. *Journal of educational psychology, 96*(4), 723.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education, 80*(2), 114-138.

Even, R., & Tirosh, D. (2002). Teacher knowledge and understanding of students' mathematical learning. *Handbook of international research in mathematics education*, 219-240.

Ewumi, A. M. (2012). Gender and socio-economic status as correlates of students' academic achievement in senior secondary schools. *European Scientific Journal, 8*(4), 23-36.

Fafunwa, A. B. (1998). *Memoirs of a Nigerian minister of education*: Macmillan Nigeria Publishers, Limited.

Fasig, L. G. (2000). Toddlers' understanding of ownership: Implications for self-concept development. *Social Development, 9*(3), 370-382.

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational researcher, 27*(5), 6-11.

Ford, Y. P. (2013). The relationship between socio-economic status and the academic achievement of culturally diverse students.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2011). *How to design and evaluate research in education*: New York: McGraw-Hill Humanities/Social Sciences/Languages.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education*, 124-158.

Friel, S. N., O'Connor, W., & Mamer, J. D. (2006). More than "Meanmedianmode" and a bar graph: What's needed to have a statistical conversation. *Thinking and reasoning with data and chance*, 117-137.

Gagne, R. M. (1965). The Conditions of Learning. New York: Holt, Rinehart and Winston. *Inc., l970*.

Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely, M., & Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. *Journal of experimental child psychology, 75*(3), 165-190.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*(3).

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*: Springer Science & Business Media.

Garfield, J., & Ben-Zvi, D. (2018). The Growing Importance of Statistics.

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in mathematics, 88*(3), 327-342.

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

Gattuso, L., & Ottaviani, M. G. (2011). Complementing mathematical thinking and statistical thinking in school mathematics *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 121-132): Springer.

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of experimental child psychology, 77*(4), 337-353.

Giagazoglou, P. (2013). The interaction effect of gender and socioeconomic status on development of preschool-aged children in Greece. *Infants & Young Children, 26*(2), 177-186.

Goni, U., & Bello, S. (2016). Parental Socio-Economic Status, Self-Concept and Gender Differences on Students' Academic Performance in Borno State Colleges of Education: Implications for Counselling. *Journal of Education and Practice, 7*(14), 21-27.

Greene, J. (2014). *Memory, Thinking and Language (PLE: Memory): Topics in Cognitive Psychology*: Psychology Press.

Grimm, K. J. (2008). Longitudinal associations between reading and mathematics achievement. *Developmental neuropsychology, 33*(3), 410-426.

Groth, R. E. (2003). *Development of a high school statistical thinking framework.* Illinois state university.

Groth, R. E. (2006). An exploration of students' statistical thinking. *Teaching Statistics, 28*(1), 17-21.

Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2016). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*.

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature reviews neuroscience, 11*(9), 651.

Halpern, D. F., Wai, J., & Saw, A. (2005). A Psychobiosocial Model. *Gender differences in mathematics*, 48.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*(5220), 41-45.

Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: a meta-analytic assessment of the strategies that promote achievement. *Developmental psychology, 45*(3), 740.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: causal relationships and underlying mechanisms. *Psychological bulletin, 111*(1), 127.

Hong, E., Peng, Y., O'Neil Jr, H. F., & Wu, J. (2013). Domain-General and Domain-Specific Creative-Thinking Tests: Effects of Gender and Item Content on Test Performance. *The Journal of Creative Behavior, 47*(2), 89-105.

Huber, J. (2008). Prefatory. *Annu. Rev. Sociol, 34*.

Hyde, J. S., & Else-Quest, N. (2012). *Half the human experience* (Vol. 1): Nelson Education.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of women quarterly, 14*(3), 299-324.

Hyde, J. S., & Mezulis, A. H. (2001). Gender difference research: Issues and Critique. *Encyclopedia of gender*, 551-560.

Jankowska, D. M., & Karwowski, M. (2019). Family factors and development of creative thinking. *Personality and Individual Differences, 142*, 202-206.

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*: Sage.

Jones, G. A., Langrall, C. W., Thornton, C. A., Mooney, E. S., Wares, A., Jones, M. R., . . . Nisbet, S. (2001). Using students' statistical thinking to inform instruction. *The Journal of Mathematical Behavior, 20*(1), 109-144.

Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical thinking and learning, 2*(4), 269-307.

Judi, H. M., Ashaari, N. S., Mohamed, H., & Wook, T. M. T. (2011). Students profile based on attitude towards statistics. *Procedia-Social and Behavioral Sciences, 18*, 266-272.

Kaboub, F. (2008). Positivist paradigm. *Encyclopaedia of Counselling, 2*, 343.

Karzmark, P. (2009). The effect of cognitive, personality, and background factors on the WAIS-III arithmetic subtest. *Applied Neuropsychology, 16*(1), 49-53.

Kawashima, N., & Shiomi, K. (2007). Factors of the thinking disposition of Japanese high school students. *Social Behavior and Personality: an international journal, 35*(2), 187-194.

Klein, J. (2004). Who is most responsible for gender differences in scholastic achievements: pupils or teachers? *Educational Research, 46*(2), 183-193.

Klein, P. S., Adi-Japha, E., & Hakak-Benizri, S. (2010). Mathematical thinking of kindergarten boys and girls: Similar achievement, different contributing processes. *Educational Studies in mathematics, 73*(3), 233-246.

Kohan-Mass, J. (2016). Understanding Gender Differences in Thinking Styles of Gifted Children. *Roeper Review, 38*(3), 185-198.

Koparan, T., & Güven, B. (2013). A study on the differentiation levels of middle school students' statistical thinking. *Elementary Education Online, 12*(1), 158-178.

Kousoulas, F., & Mega, G. (2009). Students' divergent thinking and teachers' ratings of creativity: Does gender play a role? *The Journal of Creative Behavior, 43*(3), 209-222.

Kusumaningsih, W., & Herman, T. (2018). *Gender differences in algebraic thinking ability to solve mathematics problems.* Paper presented at the Journal of Physics: Conference Series.

Langrall, C. W., & Mooney, E. S. (2002). *The development of a framework characterizing middle school students' statistical thinking.* Paper presented at the Sixth International Conference on Teaching Statistics (ICOTS6), Cape Town.

Le, L. (2017). Assessing the Development of Students' Statistical Thinking: An Exploratory Study.

Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. *Social forces, 80*(2), 713-732.

Leiker, V. C. (1993). *The relationship between an integrated learning system, reading, and mathematics achievement, higher order thinking skills and certain demographic variables: a study conducted in two school districts.* Baylor University.

Liu, H.-J. C. (1998). *A cross-cultural study of sex differences in statistical reasoning for college students in Taiwan and the United States*: University of Minnesota.

Liu, O. L., & Wilson, M. (2009). Gender differences in large-scale math assessments: PISA trend 2000 and 2003. *Applied Measurement in Education, 22*(2), 164-184.

Macher, D., Paechter, M., Papousek, I., & Ruggeri, K. (2012). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European journal of psychology of education, 27*(4), 483-498.

Mahmud, Z., & Zainol, M. (2008). Examining postgraduate students' perceived competency in statistical data analysis and their attitudes toward statistics. *International Journal of Education and Information Technologies, 2*(1), 79-86.

Malaysia, D. o. S. (2017). Report of household income and basic amenities survey 2016.

Malaysia, M. o. E. (2013). Malaysia education blueprint 2013–2025: Ministry of Education Malaysia Putrajaya.

Malaysia, M. o. E. (2018). Standard Document for Curriculum and Assessment for Form Four Mathematics.

Mandara, J. (2006). The Impact of Family Functioning on African American Males' Academic Achievement: A Review and Clarification of the Empirical Literature. *Teachers College Record*.

Manstead, A. S. (2018). The psychology of social class: How socioeconomic status impacts thought, feelings, and behaviour. *British Journal of Social Psychology, 57*(2), 267-291.

Martin, N., Hughes, J., & Fugelsang, J. (2017). The Roles of Experience, Gender, and Individual Differences in Statistical Reasoning. *Statistics Education Research Journal, 16*(2).

Massey, A., & Miller, S. J. (2006). Tests of hypotheses using statistics. *Mathematics Department, Brown University, Providence, RI, 2912*, 1-32.

Matlin, M. W. (1989). Cognition. New York: Holt, Rinehart and Winston: Inc.

Matud, M. P., Rodríguez, C., & Grande, J. (2007). Gender differences in creative thinking. *Personality and Individual Differences, 43*(5), 1137-1147.

McConney, A., & Perry, L. B. (2010). Socioeconomic status, self-efficacy, and mathematics achievement in Australia: A secondary analysis. *Educational Research for Policy and Practice, 9*(2), 77-91.

McCowan, R. J., & McCowan, S. C. (1999). Item Analysis for Criterion-Referenced Tests. *Online Submission*.

Merritt, A. R. (2014). *Building Links Between Early Socioeconomic Status, Cognitive Ability and STEM Achievement*: University of California, Davis.

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*: Cambridge University Press.

Mooney, E. S. (2002). A framework for characterizing middle school students' statistical thinking. *Mathematical thinking and learning, 4*(1), 23-63.

Moore, D. S. (1990). Uncertainty. *On the shoulders of giants: New approaches to numeracy*, 95-137.

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International statistical review, 65*(2), 123-137.

Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association, 93*(444), 1253-1259.

Muller, C. (1998). Gender differences in parental involvement and adolescents' mathematics achievement. *Sociology of Education*, 336-356.

Norfadillah, W., Hutagalung, F., Nor, M. M., & Isa, Z. M. (2017). A Study of Socioeconomic Status (SES) on Cognitive Abilities Among Preschoolers in Klang Valley, Malaysia. *Advanced Science Letters, 23*(3), 2141-2144.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37*(1), 239-253.

Perry, B., Putt, I., Jones, G. A., Thornton, C. A., Langrall, C. W., & Mooney, E. S. (2000). *Elementary school students' statistical thinking: An international perspective.* Paper presented at the PME CONFERENCE.

Perry, L., & McConney, A. (2010). School socio-economic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal of Education, 54*(1), 72-85.

Pfannkuch, M. (2008). Training teachers to develop statistical thinking. *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study, 18*.

Pfannkuch, M., & Horring, J. (2004). Developing statistical thinking in a secondary school: A collaborative curriculum development. *Curricular development in statistics education: International Association for Statistical Education*, 204-218.

Pfannkuch, M., & Rubick, A. (2002). An exploration of students' statistical thinking with given data. *Statistics Education Research Journal, 1*(2), 4-21.

Piaget, J. (1983). Piaget's theory in P. Mussen (ed)"Handbook of child psychology", Vol. 1: New York: Wiley.

Piaw, C. Y. (2014). Effects of Gender and Thinking Style on Student's Creative Thinking Ability. *Procedia-Social and Behavioral Sciences, 116*, 5135-5139.

Project, Staff, A. A. f. t. A. o. S., American Association for the Advancement of Science, W., DC., Project, A. A. f. t. A. o. S., & Science, A. A. f. t. A. o. (1993). *Benchmarks for science literacy*: Oxford University Press.

Rathbun, A., & West, J. (2004). From kindergarten through third grade: Children's beginning school experiences: Mathematica Policy Research.

Reese, H. W., Lee, L.-J., Cohen, S. H., & Puckett Jr, J. M. (2001). Effects of intellectual variables, age, and gender on divergent thinking in adulthood. *International Journal of Behavioral Development, 25*(6), 491-500.

Rinn, A. (2013). Family context predictors of math self-concept among undergraduate STEM majors: An analysis of gender differences. *Journal of the Scholarship of Teaching and Learning*, 116-132.

Rosenbaum, J. E. (2001). *Beyond College for all: Career paths for the forgotten half*: Russell Sage Foundation.

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., . . . Carroll, J. (2004). Sex differences in developmental reading disability: new findings from 4 epidemiological studies. *Jama, 291*(16), 2007-2012.

Scheaffer, R. L., & Stasny, E. A. (2004). The state of undergraduate education in statistics: A report from the CBMS 2000. *The American Statistician, 58*(4), 265-271.

Schwartz, S., & Whitin, D. (2006). Graphing with four-year-olds. Exploring the possibilities through staff development. *Thinking and reasoning with data and chance Sixty eight NCTM Yearbook*, 5-16.

Scranton, M. A. (2013). Examining middle school students' statistical thinking while working in a technological environment.

Shmueli, G. (2010). To explain or to predict? *Statistical science, 25*(3), 289-310.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research, 75*(3), 417-453.

Snee, R. D. (1990). Statistical thinking and its contribution to total quality. *The American Statistician, 44*(2), 116-121.

Sokal, L., Katz, H., Chaszewski, L., & Wojcik, C. (2007). Good-bye, Mr. Chips: Male teacher shortages and boys' reading achievement. *Sex roles, 56*(9-10), 651-659.

Sommers, C. (2000). The war against boys: How misguided feminism is harming our young men. New York: touchstone.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *American Psychologist, 60*(9), 950.

Spring, J. (2017). *American education*: Routledge.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613.

Stephens, N. M., Markus, H. R., & Phillips, L. T. (2014). Social class culture cycles: How three gateway contexts shape selves and fuel inequality. *Annual Review of Psychology, 65*, 611-634.

Stevenson, D. L., & Baker, D. P. (1987). The family-school relation and the child's school performance. *Child development*, 1348-1357.

Sun, J., & Buys, N. (2010). *Developing postgraduate students' statistical thinking in university: Evaluation of a Statistical Thinking Learning Environment Model.* Paper presented at the EDULEARN10 Proceedings.

Tach, L. M., & Farkas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research, 35*(4), 1048-1079.

Taylor, S., & Yu, D. (2009). The importance of socio-economic status in determining educational achievement in South Africa. *Unpublished working paper (Economics). Stellenbosch: Stellenbosch University*.

Thompson, A. D., Simonson, M. R., & Hargrave, C. P. (1992). *Educational technology: A review of the research*: Association for educational communications and technology.

Trzesniewski, K. H., Moffitt, T. E., Caspi, A., Taylor, A., & Maughan, B. (2006). Revisiting the association between reading achievement and antisocial behavior: New evidence of an environmental explanation from a twin study. *Child development, 77*(1), 72-88.

Upadhayay, N., & GUraGaiN, S. (2014). Comparison of cognitive functions between male and female medical students: a pilot study. *Journal of clinical and diagnostic research: JCDR, 8*(6), BC12.

Usiskin, Z., & Hall, K. (2015). The relationships between statistics and other subjects in the K-12 curriculum. *Chance, 28*(3), 4-18.

Watson, J. M. (1997). Assessing statistical thinking using the media. *The assessment challenge in statistics education, 12*, 107-121.

Wiersma, W., & Jurs, S. G. (2005). Research methods in education: An introduction.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International statistical review, 67*(3), 223-248.

Williams, C. (2007). Research methods. *Journal of Business & Economics Research (JBER), 5*(3).

Woolley, M. E., & Grogan-Kaylor, A. (2006). Protective family factors in the context of neighborhood: Promoting positive school outcomes. *Family relations, 55*(1), 93-104.