IMPACT OF STATISTICAL REASONING LEARNING ENVIRONMENT MODEL USING TINKERPLOTS ON YEAR FIVE STUDENTS' STATISTICAL REASONING

PAVETHIRA SELORAJI

FACULTY OF EDUCATION

UNIVERSITY OF MALAYA

KUALA LUMPUR

2018

IMPACT OF STATISTICAL REASONING LEARNING ENVIRONMENT MODEL USING TINKERPLOTS ON YEAR FIVE STUDENTS' STATISTICAL REASONING

PAVETHIRA SELORAJI

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF MATHEMATICS EDUCATION

WITH INFORMATON COMMUCATION AND TECHNOLOGY

FACULTY OF EDUCATION

UNIVERSITY OF MALAYA

KUALA LUMPUR

2018

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Pavethira a/p Seloraji

Matric No: PGG 140003

Name of Degree: Master of Mathematics Education with Information Communication and Technology

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Impact of Statistical Reasoning Learning Environment Model Using Tinkerplots on Year Five Students' Statistical Reasoning

Field of Study: Mathematics Education with ICT

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

Students in primary are familiar with handling data yet there are some still struggling with reasoning. The purpose of this study is to examine the impact of *TinkerPlots* on statistical reasoning among Year Five pupils. The research utilised the quasi experimental research design. Two intact classrooms were selected using convenience sampling. The sample in the study consists of 46 Year Five students. There was one classroom as the control group while the other was the experimental group. The experimental group went through interventions using *TinkerPlots* in a Statistical Reasoning Learning Environment (SRLE) class for a month where else the control group through the traditional method. The research conducted a pre- test and post-test for both the groups. A paired samples t-test was conducted to check the improvement between pre-test and post-test after traditional teaching approach. There was a significant improvement in the scores for control group (M = -3.04, SD = 2.98), t (22) =-4.90, p<.0005. A paired samples t-test was also conducted to determine the significant improvement between pre-test and post-test in using *TinkerPlots* on students' statistical reasoning skills. There was a significant improvement in the scores for the experimental group (M = -5.74, SD = 2.09), t (22) =-13.15, p<.0005. One-Way ANCOVA is conducted to determine if the mean post-test score different based on the teaching approach using *TinkerPlots* while controlling for pre-test score. The results of the One-Way ANCOVA suggested a statistically significant effect of the covariate, pre-test score, on the dependent variable, post-test score F (1,43) = 43.37, p<.05. One-Way MANOVA test was conducted to determine if there is any significant difference in the constructs of statistical reasoning between control and experimental group. The

results of the One-Way MANOVA suggested a statistically significant effect on both Representing Data $F (1, 43) = 20.10$, $p<.05$ and Analysing and Interpreting Data $F (1, 43) = 34.87$ $p<.05$. Thus, this study has important contributions for improving students' statistical reasoning skills. The results of this study has proven that teaching and learning in a SRLE class using *TinkerPlots* software could improve students' statistical reasoning skills.

IMPAK PENGGUNAAN TINKERPLOTS TERHADAP PENAAKULAN STATISTIK DALAM KALANGAN MURID TAHUN LIMA

ABSTRAK

Murid-murid di sekolah rendah berakrab dengan mengendalikan data. Namun ada yang masih bergelut dengan penaakulan statistik. Tujuan kajian ini adalah untuk mengkaji kesan TinkerPlots terhadap penaakulan statistik dalam kalangan murid Tahun Lima. Kajian ini menggunakan reka bentuk penyelidikan kuasi eksperimen. Dua buah kelas telah dipilih menggunakan persampelan mudah. Sampel dalam kajian ini terdiri daripada 46 orang murid Tahun Lima. Terdapat satu kelas sebagai kumpulan kawalan manakala satu lagi kumpulan eksperimen. Kumpulan eksperimen melalui intervensi dengan menggunakan *TinkerPlots* dalam kelas Statistical Reasoning Learning Environment (SRLE) selama sebulan manakala kumpulan kawalan melalui kaedah tradisional. Penyelidikan dijalankan secara pra ujian dan pasca ujian untuk kedua-dua kumpulan. Ujian t-sampel yang telah dilakukan telah dijalankan untuk menentukan peningkatan ketara antara pra ujian dan pasca ujian menggunakan *TinkerPlots* mengenai kemahiran penaakulan statistik murid. Terdapat peningkatan yang signifikan dalam skor untuk kumpulan eksperimen ($M = -5.74$, $SD = 2.09$), $t (22) = -13.15$, $p < .0005$. ANCOVA dijalankan untuk menentukan sama ada skor pasca ujian berkenaan, min berbeza berdasarkan pendekatan pengajaran menggunakan *TinkerPlots* sementara mengawal untuk skor pra ujian. Keputusan ANCOVA mencadangkan kesan yang ketara secara statistik terhadap skor kovarian, skor pra ujian, kepada pembolehubah yang bersandar, skor pasca ujian $F (1,43) = 43.37$, $p <.05$. Ujian MANOVA dijalankan untuk menentukan sama ada terdapat perbezaan

yang ketara dalam konstruk penaakulan statistik di antara kumpulan kawalan dan eksperimen. Keputusan MANOVA mencadangkan kesan yang ketara secara statistik pada kedua-dua konstruk iaitu Mewakili Data $F(1, 43) = 20.10$, $p < .05$ dan Menganalisis dan Mentafsir Data $F(1, 43) = 34.87$, $p < .05$. Oleh yang demikian, kajian ini mempunyai sumbangan penting untuk meningkatkan kemahiran penaakulan statistik pelajar. Hasil kajian ini telah membuktikan bahawa pengajaran dan pembelajaran dalam kelas SRLE menggunakan perisian *TinkerPlots* dapat meningkatkan kemahiran penalaran statistik pelajar.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my God to guide me through this journey, to show peace in my mind during the stressful period; correct path to follow and face any challenges upon completing my dissertation.

Secondly, a huge thank you to my supervisor, Dr Leong Kwan Eu. His confidence in me to pass through each and every stepping stone in my learning journey of my research, encouraged me to continue my work confidently. Moreover, he helped to find and suggested a lot of resources for me to read which could help for my dissertation. He also recommended many workshops and trainings organised by or held in Universiti Malaya which helped me with my dissertation writing. Adding to that, he never hesitated to set appointment and give me feedback for my writing.

Thirdly, I would like to thank my mother, Sunthari a/p Karuppiah who understood my status of pursuing my studies, took care of my responsibilities sometimes to not to give me extra chores. I also would like to thank my sister, Thesikga a/p Seloraji who helped me to analyse the data I have collected.

I'm hugely indebted to my best friend, Seethalakshmi a/p Suppiah, to be the strongest backbone upon completing my dissertation. I'm grateful to her who was there to listen whenever I was really down during this journey and gave motivation to continue and pushed me through this. Her endless support and trust in me motivated me to complete my work.

I would like to express my gratitude to my former Year Leader, Mr Ian Bartlett as well who initiated feedback on the instruments created for this research. His sharing of knowledge helped me to improve my instrument for this study.

I am also thankful to Ms Louise Shepard, the former Deputy Principal who supported to conduct the research in the school and helped me to check my language on the instruments set for this research.

Finally, I would like to thank my close friend, Rubini Radza who understood my entire journey and had strong believe in me.

CONTENTS Page

Chapter 2 Literature Review

Chapter 3 Research Methodology

Chapter 4 Findings

Chapter 5 Conclusions, Implications and Recommendations

LIST OF TABLES

Table

LIST OF CHART

Chart

LIST OF DIAGRAM

Diagram

LIST OF GRAPHS

Graph

LIST OF APPENDICES

CHAPTER ONE

INTRODUCTION

1.1 Background of Research

Technology is a prominent tool in learning Mathematics. Integrating technology in teaching and learning Mathematics, give more chances and potential to find out many new ways of learning and understanding the concept in Mathematics. It is essential that teachers and students have regular access to technologies that support and advance mathematical sense making, reasoning, problem solving, and communication. Effective teachers optimize the potential of technology to develop students' understanding, stimulate their interest, and increase their proficiency in mathematics. When teachers use technology strategically, they can provide greater access to mathematics for all students (NCTM, 2000).

Statistics is a discipline on its own rather than a part of mathematics. Statistical problems can be solved using statistics knowledge and it's not the same as solving mathematical problems using knowledge.

Statistics education is playing an important role in mathematics because the students are exposed to real world situations and have to make decisions wisely based on the interpretation made and be able to reason for the choice. Garfield and Ben-Zvi (2008) claimed that statistics education is also an emerging field that grew out of different disciplines and is currently establishing itself as a unique field of study. The two main disciplines from which statistics education grew are statistics and mathematics education. It followed by the point that, statistics education has been the focus for researchers in many disciplines, perhaps because statistical reasoning is used in many disciplines and provides so many interesting issues and challenges (Garfield & Ben-Zvi, 2009). The National Council of Teachers of Mathematics (NCTM), first in *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989) and then later in *Principles and Standards for School Mathematics* (NCTM, 2000), emphasized the importance of statistics education as a part of the Data Analysis and Probability content standard (NCTM, 2016). It lets students to formulate questions and collect, organize, and display relevant data to answer these questions. Additionally, it emphasizes learning appropriate statistical methods to analyze data, making inferences and predictions based on data, and understanding and using the basic concepts of probability (NCTM, 2016).

The fact that quantitative information is everywhere and numerical data are increasingly presented with the intention of adding credibility to advertisements, arguments, or advice. Most would also agree that being able to provide good evidence-based arguments and to be able to critically evaluate data-based claims are important skills that all citizens should have, and therefore, that all students should learn as part of their education (Watson, 2006).

Statistical reasoning is something that the students predict or assume based on the data obtained. They are then able to come to a conclusion by stating the hidden reason behind it. Statistical reasoning can accompany deductive and inductive reasoning in inquiry situations where hypotheses are formulated and tested for experimentals or surveys designed to answer specific questions (Lavigne & Lajoie, 2007). Statistical reasoning is described as "the way people reason with statistical ideas and make sense of statistical information. It involves making interpretations based on sets of data, or statistical summaries of data. Students need to be able to combine ideas about data and chance, which leads to making inferences and interpreting statistical results (p. 101)" (Garfield & Chance, 2000).

Statistical reasoning involves making interpretations based on sets of data, representations of data, or statistical summaries of data. Much of statistical reasoning combines ideas about data and chance, which leads to making inferences and interpreting statistical results. Underlying this reasoning is a conceptual understanding of important ideas, such as distribution, centre, spread, association, uncertainty, randomness, and sampling (Garfield, 2003).

Many people think of mathematics and statistics as the same thing, and therefore, confuse statistical reasoning with mathematical reasoning (Garfield & Gal, 1999). Today's leading statistical educators view these disciplines and types of reasoning as quite distinct. Gal and Garfield (1997), "distinguish between the two disciplines in the following ways:

- In statistics, data are viewed as numbers with a context. The context motivates procedures and is the source of meaning and basis for interpretation of results of such activities.
- The indeterminacy or "messiness" of data distinguishes statistical investigations from the more precise, finite nature characterizing mathematical explorations.
- Mathematical concepts and procedures are used as part of the solution of statistical problems. However, the need for accurate application of computations is rapidly being replaced by the need for selective, thoughtful, and accurate use of technological tools and increasingly more sophisticated software programs.
- Many statistical problems do not have a single mathematical solution, but instead, start with a question and result in an opinion supported by certain findings and assumptions.

> These answers need to be evaluated in terms of quality of reasoning, adequacy of methods employed, and nature of data and evidence used (p.207)".

Using a collaborative classroom research model that implemented activities and gathered data in three different institutions, (delMas, Garfield, & Chance, 1999) studied the development of reasoning about sampling distributions, using a simulation program and research-based activities. They found that student performance on a specially designed post-test, to assess students' reasoning about sampling distributions, improved as the activity was changed to imbed assessments within the activity. The Sampling Distributions program, developed by delMas, allows students to interact with the concept visually, in a dynamic, interactive environment. Lunsford, Rowell, and Goodson-Espy (2006) replicated this study in a different type of undergraduate course and found similar results.

Lane and Tang (2000) compared the effectiveness of simulations for teaching statistical concepts to the effectiveness of a textbook; while Aberson, Berger, Healy, Kyle, and Romero (2000) studied the impact of a Web-based, interactive tutorial used to present the sampling distribution of the mean on student learning. In a study of students' reasoning about the standard deviation, delMas (2005) had students manipulate a specially designed software tool to create histograms with the highest or lowest possible standard deviation, given a set of fixed bars. He identified some common ways students understand and misunderstand the standard deviation, such as thinking of "spread" as spreading butter, being evenly distributed in a graph. He also found that students had difficulty reasoning about bars in a histogram having density, in that they represent several points on a particular interval on a graph.

Another topic of interest to statistics educators has been the use of online instruction either in a Web-based course or “hybrid/blended” course, in which a significant amount of the course learning activity has been moved online, making it possible to reduce the amount of time spent in the classroom. For example, Utts (2003) and Ward (2004) found no differences in course performance for students in a hybrid versus a traditional course, and concluded that hybrid courses were not resulting in decreased student performance, although Utts noted lower evaluations by students in the hybrid courses. However, no significant differences in course performance do not imply that there were no real differences in student outcomes for the compare instructional methods.

Keeler and Steinhorst (1995), Giraud (1997), and Magel (1998) investigated different methods of cooperative learning in teaching statistics at their institutions, and found generally positive results. Keeler and Steinhorst (1995) found that when students worked in pairs, the final grades were higher and more students stayed in the course than in previous semester. Giraud (1997) found that using cooperative groups in class to work on assignments led to higher test grades than students in a lecture class. Magel (1998) found that implementing cooperative groups in a large lecture class also led to improved test scores compared to grades from a previous semester that did not use group work.

Meletiou and Lee (2002) organized their curricula along a Project-Activities-Cooperative Learning-Exercises model emphasizing statistical thinking and reasoning and an orientation toward investigating conjectures and discovery of results using data. Students were assessed on their understanding at the beginning and end of the course. Increased understanding was observed on tasks requiring statistical reasoning such as deducing whether a set of data could have been drawn at random from a particular population.

One of the major areas of current interest is the role technological tools (such as computers, graphing calculators, software, and Internet) can play in helping students develop statistical literacy and reasoning. Research on simulation training indicates that even a well-designed simulation is unlikely to be an effective teaching tool unless students' interaction with it is carefully structured (Lane & Peres, 2006). Simulations, however, can play a significant role in enhancing students' ability to study random processes and statistical concepts (Lane & Peres, 2006; Lane & Tang, 2000; Mills, 2004).

Clements (2000) provides a summary and a rationale for moving beyond mundane exercises to higher order learning experiences. This view is shared by Shaffer and Kaput (1999) in suggesting that technology offers to potential to find new ways to learn mathematics. Suggestions of context linked investigations to enhance beginning inference through explorations with TinkerPlots using real data are appearing in the professional literature (Watson, 2008; Watson & Wright, 2008). Classroom research based on the innovations introduced within TinkerPlots has further illustrated new ways to learn statistics. Given the growing interest and the availability of body-sensing technologies, we see an opportunity for the educational technology community to explore new forms of teaching and learning that involve this class of tools (Lee & Thomas, 2011).

Some of the most informative studies were not designed to focus on the use of technology, but on larger teaching experimentals that combined innovative instructional activities and technological tools to promote student reasoning about a particular topic, such as distribution (e.g., Bakker, 2004a, Cobb, 1999; Cobb & McClain, 2004). These studies focused on the use of a set of Minitools, applications created to help students move along a learning trajectory. Similarly, the studies of Makar and Confrey (2005) and Rubin, Hammerman, and Konold (2006) explore teachers' knowledge and reasoning as they use innovative software (Fathom or TinkerPlots).

The types of research studies that explore technology in statistics education can be grouped into three categories:

1. Development, use, and study of particular tools (e.g., the creation and use of Fathom software – Biehler, 2003; Minitools – Cobb & Moore, 1997).
2. How use of particular tools help develop students' reasoning (e.g., use of Sampling SIM software to develop reasoning about sampling distributions – Chance, delMas & Garfield, 2004).
3. Comparison of tools (e.g., comparing ActivStats, CyberStats, and MM∗Stat multimedia – Alldredge & Som, 2002; Symanzik & Vukasinovic, 2002, 2003, 2006).

Research on the role of technology in teaching and learning statistics has been increasing over the last decade. In 1996, a special International Association for Statistical Education (IASE) Roundtable was convened in Granada, Spain to discuss the current state of research on the role of technology in statistics education at that time. While much of the work reported at the roundtable (Garfield & Burrill, 1997) was on the development of new tools to help students learn statistics, there was a clear call for more research on appropriate ways to use these tools to promote student learning. It was suggested that a new research agenda was needed to identify appropriate methodologies for future studies on this topic as well as to explore new ways to use technology in studying this topic (Hawkins, 1997). Given the changes in technology in the past decade, ideas about both of these aspects of technology are still emerging. In this section, we highlight some of the more recent research questions being explored and the types of studies involved, particularly with respect to developing students' statistical reasoning. Ben-Zvi describes how technological tools are now being designed to support statistics learning in the following ways (2000, p. 128):

1. Students' active construction of knowledge, by "doing" and "seeing" statistics.

2. Opportunities for students to reflect on observed phenomena.

3. The development of students' metacognitive capabilities, that is, knowledge about their own learning and thought processes, self-regulation, and control.

In addition, technological tools can bring exciting curricula based on real-world problems into the classroom; provide scaffolds and tools to enhance learning; and give students and teachers more opportunities for feedback, reflection, and revision (Bransford, Brown & Cocking, 2000).

1.2 Statement of the Problem

Most students in primary are familiar with handling data yet there are some still struggling with reasoning. There are several critical issues discussed in the research area, statistical reasoning.

Firstly, the critical issue where the students in Malaysia did not perform well in TIMMS and PISA. Of the TIMSS 2011 eighth grade participants that also participated in 2007 and have comparable data, there were both participants with increases and participants with decreases in average mathematics achievement over the period. Six countries (Hungary, Jordan, Malaysia, Sweden, Syria, and Thailand) had lower achievement. Among the countries with an overall decrease in mathematics achievement, only Jordan and Malaysia had decreases in all four content domains which are number, algebra, geometry, and data and chance. In the year of 2011, difference from overall Mathematics Score Malaysia scored 11 in Number which the subscale score is significantly higher than overall mathematics score. In other content domains Malaysia scored -10 in Algebra, -8 in geometry and -11 in data and chance which indicates the subscale scored significantly lower than overall mathematics score. The results shows the issue of the students did not perform well in data which leads the researcher to do research on the topic data.

On the other hand, among the six countries with an overall decrease, Jordan and Malaysia had lower achievement in all three cognitive domains which are knowing, applying and reasoning. In TIMSS 2011, difference from overall Mathematics score in knowing is 4 which the subscale score is significantly higher than overall mathematics score; applying is -1 and reasoning is -14 which the subscale score is significantly lower than overall mathematics score (Mullis, Martin, Foy, & Arora, 2012). The results shows that the students did not perform well in reasoning questions as well. Thus, the researcher would have interest to focus on reasoning.

PISA has some expected competencies, whereby, Level 6 and Level 5 are in the category of Advanced. There are only 3.1% students who achieved Level 6. To achieve Level 6, students need to focus on conceptualising, generalising and utilising information based on their modelling of complex problem situations; link different information sources and representations and flexibly translate between them; capable of advanced mathematical thinking and reasoning and provide accurate interpretations of their findings. On the other hand, about 12.7% students have achieved Level 5, where to achieve the particular level, the students need to use broad, well-developed thinking and reasoning skills, appropriately linked representations, symbolic and formal characterisations, and insight pertaining to these situations and communicate their interpretations and reasoning. (*Malaysia Education Blueprint 2013-2025*, 2012).

Students that achieve Level 4 or Level 3 are categorized under intermediate. 31.6% students had reached Level 4 and 56% of students have achieved Level 3. To achieve Level 4, the students need to focus on working effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions; select and integrate different representations; including symbolic representations and linking them directly to aspects of real-world situations. However, to achieve Level 3, the students need to execute clearly described procedures; select and apply simple problem-solving strategies; interpret and use representations based on different information sources and reason directly from them and develop short communications reporting their interpretations, results and reasoning. (*Malaysia Education Blueprint 2013-2025*, 2012).

In conclusion, the PISA findings show there are less statistical reasoning skill is showed by the students in advanced and intermediate level. As a result, the researcher wants to study students' statistical reasoning in this study.

The second critical issue is that primary school students are facing difficulties in making statistical reasoning. Burrill and Camden (2005) propose that "students seem to be mastering statistical procedures and vocabulary but are not able to use statistical reasoning in a meaningful way" and that "an over-emphasis in school syllabi on answering questions rather than posing them, and making decisions based only on data displays produces an approach based on absoluteness of data that stifles the development of statistical thinking" (p. 4).

Chan and Ismail (2012) point out many students find difficulties in understanding the concepts and this often lead to misconceptions in statistical reasoning, especially in descriptive statistics and inferential statistics. Such situation will definitely impede students' learning of statistics and curtail their enthusiasm if this problem is not overcome.

Moreover, statistical inference is a topic that is typically not taught until university, because students traditionally find it very difficult (Abelson 1995; Cobb and Moore 1997; Erickson 2006; Garfield and Ahlgren 1988). However, researchers have highlighted the opportunities for building pre-tertiary students' understandings of inferential reasoning at a much younger age by focusing on informal aspects of statistical inference (Makar and Ben-Zvi 2011; Pratt and Ainley 2008).

Another problem is that statistics education, until now, has shied away from informal inference and has not developed a language with shared meanings, nor a shared understanding of how to talk about graphs. Whether the research is focussing on students' cognition by using innovative technology such as Fathom (Key Curriculum Press Technologies, 2000) or using students' own products the problem of communicating and articulating the meaning of the statistical representations remains difficult.

Friel, Curcio, and Bright (2001) consider that research is needed on understanding what it is about the nature of reasoning that makes comparing data sets such a challenging task. Furthermore, they believe that graph comprehension involves an interplay between visual shapes, visual decoding, judgement, and context.

The researcher also specified that while students have a fairly good "out of school" understanding of the concept of sample, they have difficulty making the transition to the formal, statistical meaning of this term and the related connotations and conceptualizations (Watson & Moritz, 2000).

Despite the increase in statistics instruction at all educational levels, historically the discipline and methods of statistics have been viewed by many students as a difficult topic that is unpleasant to learn (Garfield & Ben-Zvi, 2008). Moreover, one type of statistical reasoning that is difficult for students is to make conclusions about differences between groups. It is well documented that students who know how to compute the arithmetic mean and median are mostly not inclined to use such measures when comparing groups (Konold & Higgins, 2002). In addition to, at primary (elementary) school level statistics is often reduced to frequency counts and bar graphs, with rules for calculating mean and range added later. Indeed, the local curriculum itself does not give strong and specific emphasis to interpreting, reading, critiquing, and questioning data (Victorian Curriculum and Assessment, 2005).

Overall, third critical issue in this study discusses, that there is lack of knowledge and skills in using ICT. This is due to, teachers who do not have enough time to integrate technology in class; training given to the teachers based on software learning is teacher-centred and lack of participation on exploring the software; lack of comfort in using new technologies and lack of administrative support. Thus, the researcher is inquisitive to study the effectiveness of using technology in teaching statistical reasoning.

In summary, some problems have been identified. Students are unable to recognise the essentials of statistical reasoning; and also tend to misinterpret the concept in statistical reasoning. Students find complicating in making statistical meaning.

Thirdly, the critical issue is that lack of technology introduced in primary. There are a few factors that affect the usage of information technology in the education field. Factors like personality, attitude and environment are known to have positive relation with ICT usage in the classroom. On the other hand, low level of knowledge and skills coupled with limited sources were known to be the deterrent factor for successful ICT usage in the classroom (Razak & Embi., 2001).

Indeed two factors have been identified as the main factors in the application of technology in the teaching and learning of mathematics. The first factor is the teachers' perception that the use of technology is not able to help in the teaching and learning of mathematics. This was further worsened by the fact that teachers always claim that they do not have sufficient time to prepare for ICT integrated lessons (Zakaria, Daud, & Nordin, 2007).

In the present teachers' professional development courses, there are hands-on activities but this was not supported by relevant modules or manuals for the facilitators and the course participants. The activities conducted in those courses are teacher-centred and in most situations, courses are conducted using softcopy materials supplied by vendors. The approach was rather ineffective in the learning of a particular software which normally requires active participation from the participants (Arshad, Yaacob, Yusof, & Latih, 2000).

Perhaps more than other fields, mathematics as a subject is thought to have benefited and established a stronger intrinsic link with the development of computers in recent times but four issues were identified as critical to proper and effective use of computer technologies in the mathematics classroom. Top among them is computer attitude, followed by software selection, a proper utilization direction, and Web-based professional development of mathematics teachers (Kadijevich, 2002). Yuen and Ma (2001) noted not unlike any other innovation, teachers initially resisted the use of computers in education. As a matter of fact, the term “computerphobia” and “computer anxiety” were coined and entered in the literature vocabulary due to teacher (not student) resistance to computer use.

It is followed by the lack of support for teachers. According to Ritchie (1996), schools are not yet effectively implementing instructional technologies in spite of the increase in the capacity of available educational technology. This study identified lack of administrative support as one of the most critical impediments for the integration of instructional technology. Administrative support is needed in order to provide funding for computer labs, consistent technical support for teachers, and on-going professional development for teachers to have the opportunity to learn new technologies and their use in classrooms. Even when the technology is in place and the technical support is available, teachers need much more support and professional development in learning how to implement a new pedagogy with technology since technology alone does not make for effective teaching. To maximize the benefits of technology for students, teachers need to spend time modifying what they will teach, how they will teach it, and how they will assess it using technology (U.S. Congress, Office of Technology Assessment, 1995).

Furthermore, there is lack of awareness of and comfort with new technologies. Probability and statistics are specialized subjects, and many schools may not have a faculty member whose expertise is in these areas. Since teachers’ schedules are very demanding, little time is available to learn about new technologies and their capabilities. Teachers who have learned statistics decades earlier may not be comfortable using the new tools and may not believe in the value of their use. In some cases, teachers may be able to attend conferences and hear about new technologies, but this is usually not enough time for them to appreciate the benefits of the technology and fully learn how to effectively use it in the classroom. Unless teachers are provided with long-term support for learning to use and implement technology, they are unlikely to use it in their classrooms. Internet-based communities of teachers are becoming an increasingly important tool for overcoming teacher's isolation and need for support (Levin & Waugh, 1998).

Current syllabus and teaching strategies of statistics (data handling) do not allow students to explore on the topic. Thus, the students do not remember on what they have learned and especially they are not able to transfer their knowledge to the real world situation. It leads the students to fail to make statistical reasoning. Saying that, the study would like to research on suitable teaching strategies based on the theory, "constructivism" which allows the students to explore and build their knowledge in order to develop their statistical reasoning skills.

In conclusion, past research studies show that there are some issues resulting with poor scoring in TIMSS and PISA, difficulties in making statistical reasoning among primary students and lack of technology used in primary. Thus, the interest to research and study on the impact of a technology tool in learning data handling. This is because, the study is predicting that there is an impact of using TinkerPlots software in statistical reasoning among primary school students. To sum up, the study would like to check whether the assumption made is true or false.

1.3 Theoretical Framework

The research study conducted is based on Statistics Reasoning Learning Environment (SRLE) model. SRLE, is known as an effective and positive statistics classroom, can be viewed as a learning environment for developing in students a deep and meaningful understanding of statistics and helping students develop their ability to think and reason statistically. The model is based on six principles of instructional design as below (Garfield & Ben-Zvi, 2008);

1. Focuses on developing central statistical ideas rather than on presenting set of tools and procedures.
2. Uses real and motivating data sets to engage students in making and testing conjectures.
3. Uses classroom activities to support the development of students' reasoning.
4. Integrates the use of appropriate technological tools that allow students to test their conjectures, explore and analyze data, and develop their statistical reasoning.
5. Promotes classroom discourse that includes statistical arguments and sustained exchanges that focus on significant statistical ideas.
6. Uses assessment to learn what students know and to monitor the development of their statistical learning as well as to evaluate instructional plans and progress.

The research is to enhance statistical reasoning using a technological tool. Thus, SRLE model is suitable to be used to direct the research because all the criteria from the model are to develop statistical reasoning among students.

The theory is focused based on the SRLE model. The research study uses the theory called constructivism. A recent theory of learning which has been widely accepted in education communities stems from earlier work by Jean Piaget, and has been labelled 'constructivism.' This theory describes learning as actively constructing one's own knowledge (Von Glasersfeld, 1987). Today, this is the guiding theory or much research and reform in mathematics and science education. Constructivists view students as bringing to the classroom their own ideas, material. Rather than 'receiving' material in class as it is given, students restructure the new information to fit into their own cognitive frameworks. In this manner, they actively and individually construct their own knowledge, rather than copying knowledge 'transmitted', 'delivered' or 'conveyed' to them. A related theory of teaching focuses on developing students' understanding, rather than on rote skill development, and views teaching as a way to provide opportunities for students to actively construct knowledge rather than having knowledge 'given' to them (Garfield, 1995). Constructing knowledge is focused in the study because the technology tool used in the research enhances the students' ability to link and build their knowledge with what they already know which magnifies students' statistical reasoning.

The implication of current theories of learning is that good instructional practice consists of designing learning environments that stimulate students to construct knowledge. This involves activities that provide students many opportunities to think reason and reflect on their learning, as well as discussing and reflecting with their peers. It does not mean that teachers should never tell students anything directly and instead should always allow them to construct knowledge for themselves. Rather, it means that learning is enhanced when teachers' pay attention to the knowledge and beliefs that learners bring to a learning task, use this knowledge as a starting point for new instruction and monitor students' changing conceptions as instruction proceeds (Garfield & Ben-Zvi, 2009).

Constructivism used in the study instead of other theories because the research is interested in encouraging students to construct knowledge and building new knowledge by connecting their new ideas on what they have learnt earlier. There are four characteristics of learning in constructivism: (a) an emphasis on understanding; (b) a focus on the processes of knowing (e.g., Piaget, 1978; Vygotsky, 1978); (c) the principle that people construct new knowledge and understandings based on what they already know and believe; and (d) the importance of helping people take control of their own learning, predict their performances on various tasks, and to monitor their current levels of mastery and understanding (metacognition, e.g., Brown, 1975; Flavell, 1976). The treatment for the study is planned based on SRLE model. The characteristics stated for constructivism shares same characteristics as SRLE model. Particularly for this research study in treatment, the students are required to develop their knowledge through activities. They need to test their predictions by exploring the data. To sum up, the characteristics support the research objective because they enhance statistical reasoning using and exploring technology tool.

The assumptions are important in this research. One of the importance is to have a smooth process of research. The sample chosen is a particular year group from a school. It makes the research conducted easier. Moreover, the assumptions also can narrow down the research. With a selected sample, the research is only focusing on statistical reasoning where appropriate results can be obtained based on the objectives. The assumptions also can easily collect the data, analysis the data and as well to interpret the data. Assumptions in your study are things that are somewhat out of your control, but if they disappear your study would become irrelevant. For example, if you are doing a study on the middle school music curriculum, there is an underlying assumption that music will continue to be important in the middle school program. If you are conducting a survey, you need to assume that people will answer truthfully. If you are choosing a sample, you need to assume that this sample is representative of the population you wish to make inferences to (Simon, 2011). According to the constructivist theory of learning, people learn by constructing knowledge, rather than by receiving knowledge. In the most general sense, the contemporary view of learning in accordance with the constructivist theory is that 'new knowledge and understandings are based on the existing knowledge and beliefs we already have' and are grounded in our experiences (Cobb 1994; Piaget 1978; Vygotsky 1978). Moreover, Statistical Reasoning Learning Environment' (SRLE) and is built on the constructivist theory of learning (Garfield & Ben-Zvi, 2009).

Following are the assumptions in the research;

1. Learning occurs by passively, but rationally, reflecting on stimuli in the environment and can be inferred as a consequence of change in behaviour among Year Five students.
2. Different individuals would exhibit the same behaviours if they were given the same stimuli and reinforcements.
3. Statistical reasoning is necessary knowledge to be developed by each student based on his or her own experience.
4. Skills consisting of empirical activities are undertaken by the Year Five students quickly and accurately.
5. Year five students achievement in statistical reasoning consists of statistical reasoning skills that are mastered by them.

6. Year five students achievement in statistical reasoning can be measured using established testing in particular.
7. Year Five students start learning statistics since Early Years.
8. Item sets in the statistical reasoning test is adequate representation of content of data handling topic of Year Five.

1.4 Purpose and Research Questions

The purpose of this study is to examine the impact of *TinkerPlots* on statistical reasoning among Year Five pupils. The objectives of the study includes:

1. To identify the improvement between pre-test and post-test of the control group Year Five students' statistical reasoning.
2. To identify the improvement between pre-test and post-test of the experimental group Year Five students' statistical reasoning.
3. To identify the difference in Year 5 students' statistical reasoning in post-test between experimental group and control group when controlling for pre-test.
4. To identify the difference in the four constructs of statistical reasoning.

In overall, the research gives importance to the following four research questions below;

1. Is there any significant improvement between pre-test and post-test of the control group Year Five students' statistical reasoning?
2. Is there any significant improvement between pre-test and post-test of the experimental group Year Five students' statistical reasoning?

3. Is there a significant difference in Year Five students' statistical reasoning in post-test between the control and experimental group when controlling for pre-test?
4. Is there any significant difference in the four constructs of statistical reasoning in Year Five students?

Based on the research questions, three research hypothesis are formed.

1. The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in control group after traditional teaching approach.
2. The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in experimental group using *TinkerPlots*.
3. The mean of post-test scores of Year 5 experimental group students is higher than the control group students when controlling the pre-test score.
4. There is a mean difference in Year 5 students' constructs of statistical reasoning in post-test based on using *TinkerPlots*.

1.5 Definition of Terms

There are several terms used in the research. The research comprises the terms such as statistics, statistical reasoning *TinkerPlots,* SRLE and traditional teaching.

Statistics. *Statistics* is concerned with scientific methods of collecting, organising, summarising, presenting, analysing and interpreting numerical data (Croxton & Cowden, 1968).

Statistical reasoning. *Statistical reasoning* is the use of statistical ideas and tools to summarize and draw assumptions besides making conclusions from the data (Lovett, 2001). In this study, statistical reasoning is measured by using the total scores in the pre-test and post-test.

TinkerPlots. *TinkerPlots* is a tool that helps students to make associations between different data handling stages, such as: collecting, organizing, formulating and testing hypotheses about the data (Konold et al., 2005).

Statistical Reasoning Learning Environment (SRLE). *SRLE* is an effective and positive statistics classroom which is student centred where the teacher facilitates developing of knowledge by guiding through discussions and activities (Garfield et al., 2008).

Traditional teaching. *Traditional teaching* puts the responsibility for teaching and learning mainly on the teacher and it is believed that if students are present in the lesson and listen to the teacher's explanations and examples, they will be able to use the knowledge. It is teacher centred where teacher delivers knowledge by telling and explaining (Boumová, 2008).

1.6 Limitations and Delimitations

There are several limitations and delimitations in the research. There are four limitations which consist of sampling, theory, research design and method of data collection where else the three delimitations are on the topics of mathematics, sample and research question.

Convenience sampling is used in the study. The sample is selected based on the researcher's judgment. The study does not give an equal chance of being selected for all the year 5 students from the school. Also, it is very unlikely that a representative sample is being chosen. Thus, it is less generalised to the population.

The theory that has been used in the research is constructivism. It only focuses on psychomotor domain and does not discuss about spiritual domain where it is synthesis of belief systems, social concepts and emotional stability.

The next limitation is based on the research design. The research used quantitative study which is quasi experimental study. The research is also not generalized statistically to a bigger population because of lack of random assignment. Moreover, this design will not take other factors into account on the impact of TinkerPlots in students' statistical reasoning. For example, other factors may influence the study such as students' extra maths class or any other class that develop their critical thinking; level of students' interest in maths subject and could be the extent of students' exposure to the real world.

Moreover, the method of data collection is one of the limitations. The research collected the data based on pre-test and post-test where results can be bias. Other than that, the data collected is also responded from students is subjective. Adding up to that, the method used was unable to analyse the impact simultaneously. Therefore, the method of this study is limited testing effect and historical effect.

One of the delimitations of the research is the topic of mathematics. This research particularly focuses on statistical reasoning from the topic statistics using *TinkerPlots*. The research does not concentrate on other topics, the software is also not used widely for other topics from mathematics like probability.

The sample used also is one of the delimitations. The sample is a class of Year Five pupils from a particular international school in Federal Territory. The sample size is restricted to 50 students where each class or group consist of 25 students. The research is not conducted with other Year Five pupils in the same school. In addition, the research is also not conducted in other international schools.

The delimitations can be tackled with some steps. One of the ways is by choosing the sample very carefully where it can represent the overall population. Moreover, the researcher can also have further researches in reasoning different topics in future.

1.7 Significance of the Study

There are several groups that could benefit from the finding of the research who are mathematics text book writers, mathematics teachers and mathematics curriculum esearchers. Mathematics text book writers hopefully get an idea of inserting *TinkerPlots* in teaching and learning in the text book where it may help the teachers and students to know on how to use the software in learning mathematics as well as encourage them to explore more in the topic to prove or test the conjectures.

The research also hopefully gives benefits to mathematics teachers to implement *TinkerPlots* in teaching mathematics. Teachers expected to get an idea about using *TinkerPlot*s in teaching mathematics in data handling. In addition, teachers are also assumed to explore and familiarise in *TinkerPlots* in order to guide the students in developing their statistical reasoning skills by using *TinkerPlots.* They are assumed to benefit by planning their lesson plans which include *TinkerPlots* in teaching where it makes students interested in learning the topic; exploring the topic and generating their own findings from the data collected.

Moreover, hopefully the mathematics curriculum researchers can get further ideas to conduct their research topic. It is expected to help broaden the particular research area where it may be very useful for the students and teachers in learning mathematics using ICT. Hopefully in future, learning mathematics using ICT will be a norm in educational organisations where students will not easily get bored with mathematics lesson and clearly understand the concept. On the other hand, ICT also should play a role to develop the students' reasoning skills.

1.8 Summary

This chapter consists of background of research, statement of problem, theoretical framework purpose and research questions, definition of terms, limitations and delimitations and significance of study. A research problem is chosen and explained. The justification for the research problem is explained. Then, the explanations are given to theoretical framework, purpose of the research and research questions. Finally, definition of terms is given, limitations and delimitations are described and significance of study discussed. Based on Chapter One, the research moves forward to literature review in Chapter Two, research methodology in Chapter Three, the findings in Chapter Four, and discussion, conclusions and implications of the study in Chapter Five. Next, all references listed under the heading References, while the support materials are attached thereto under Appendix.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Chapter Two discusses the main points which are namely review of previous studies related to statistics education, statistical reasoning and technological tools to learn statistics. The second part of this chapter contains the difficulties of learning statistics. The third part will be the teaching and learning of statistics. It is followed by the next section which discusses the teaching and learning statistics using technology. The fifth part of this chapter discusses the research literature on the definition and meaning of statistical reasoning.

2.2 Constructivism Theory

Although there are different versions of constructivism, the basic idea is that people learn by constructing knowledge, rather than by receiving knowledge. In the constructivist theory, the meaning of "knowing" has shifted from being able to remember and repeat information to being able to find and use it (Simon, 1995). In the most general sense, the contemporary view of learning in accordance with the constructivist theory is that new knowledge and understandings are based on the existing knowledge and beliefs we already have and are grounded in our experiences (e.g., Cobb, 1994; Piaget, 1978; Vygotsky, 1978). We learn by doing. And when we learn, our previous knowledge does not go away; it is integrated with the new knowledge. One implication for teaching is that teachers need to pay attention to the incomplete understandings, the false beliefs, and the naıve renditions of concepts that learners bring with them to a given subject. Teachers then need to build on these ideas in ways that help each student achieve a more mature understanding. If students' initial ideas and beliefs are ignored, the understandings that they develop can be very different from what the teacher intends (Bransford et al.,2000).

Constructivism emphasizes the interaction of individuals and situations in the acquisition and improvement of abilities and knowledge (Cobb & Bowers, 1999). Constructivism is different in comparison with conditioning theories that highlight the effect of the environment on the individual as well as with information processing theories that place the locus of learning within the mind with little attention to the context in which it occurs. It has common assumption with social cognitive theory stating that individuals, behaviors, and environments interact in mutual fashion (Bandura, 1986, 1997). A key assumption of constructivism is that individuals are active learners and build knowledge for themselves and in order to understand material appropriately, the learners need to discover the basic principles (Geary, 1995). Constructivists vary in how much they attribute this process wholly to learners. Some believe that mental structures come to reflect reality, whereas others (radical constructivists) believe that the individual's mental world is the only reality (Schunk, 1996). Constructivists also vary in the extent to which they attribute the construction of knowledge to social interactions with teachers, peers, parents, and others (Bredo, 1997). Another constructivist assumption is that teachers should not teach in the traditional sense of delivering instruction to a group of students. Rather, they should structure situations such that learners become actively involved with content through manipulation of materials and social interaction (Schunk, 1996). Activities include observing phenomena, collecting data, generating and testing hypotheses, and working collaboratively with others. Students are taught to be self- dependent and engaged actively in their learning by determining goals, observing and gauging progress, and going beyond basic necessities by discovering interests (Bruning, Schraw, Norby, & Ronning, 2004; Geary, 1995).

2.3 Difficulties of Learning Statistics

According to Garfield and Ben-Zvi (2008), firstly many statistical ideas and rules are complex, difficult, and or counterintuitive. It is therefore difficult to motivate students to engage in the hard work of learning statistics. Secondly, many students have difficulty with the underlying mathematics (such as fractions, decimals, proportional reasoning, and algebraic formulas) and that interferes with learning the related statistical concepts. A third reason is that the context in many statistical problems may mislead the students, causing them to rely on their experiences and often faulty intuitions to produce an answer, rather than select an appropriate statistical procedure and rely on data-based evidence. Finally, students equate statistics with mathematics and expect the focus to be on numbers, computations, formulas, and only one right answer. They are uncomfortable with the messiness of data, the ideas of randomness and chance, the different possible interpretations based on different assumptions, and the extensive use of writing, collaboration and communication skills. This is also true of many mathematics teachers who find themselves teaching statistics.

Besides that, Garfield and Ben-Zvi (2008), mentioned that many studies reveal that it is difficult to determine the impact of a particular teaching method or instruction tool on students' learning in a course due to limitations in study design or assessments used. While teachers would like research studies to convince them that a particular teaching method or instructional tool leads to significantly improved student outcomes, that kind of evidence is not actually available in the research literature. Moreover, the research studies on attitudes and anxiety suggest that there are few strong (or large) predictors of how well students do in a statistics course, that there is little change in attitudes from beginning to end of a first course in statistics (and sometimes negative changes) and that difficulties in students' reasoning and poor attitudes are fairly widespread. The evidence does not show that if students are good in mathematics or have good attitudes, they will be likely to succeed in statistics, which is contrary to many teachers' beliefs. Instead, students who may not be strong in mathematics may work hard, enjoy the subject matter, and do very well in an introductory statistics course.

2.4 Teaching and Learning Statistics in Primary and Secondary Schools

According to (Drews, Dudgeoun, Lawton & Surtees, 2007), in Key Stage 1, Processing, Representing and Interpreting data is found within (Number) in the National Curriculum for Mathematics in England (DfEE, 1999). In addition, the National Numeracy Starategy Framework (DfEE, 1999) identifies that Key Stage 1 children should be developing an understanding for organising and using data within the strand of solving problems. Handling Data forms Attainment Target 4 of the National Curriculum for Key Stage 2. During Key Stage 2 this includes interpreting data in tables, charts and graphs, including those generated by a computer. By end of Key Stage 2, children will be able to calculate the range and average for a set of data.

As a further matter, Lehrer, Kim and Jones (2011) emphasized that students often learn procedures for measuring, but rarely do they grapple with the foundational conceptual problem of generating and validating coordination between a measure and the phenomenon being measured. Coordinating measures with phenomenon involves developing an appreciation of the objects and relations in each as well as establishing their mutual correspondence. The research supported students' developing conceptions of statistics by positioning them to design measures of center and of variability for distributions that they had generated through repeated measure of a length. After students invented and explored the viability of their measures individually, they participated in a public (whole-class conversation) forum featuring justification and reflection about the viability of their designed measures. We illustrate how individual invention enticed students to attend to, and to make explicit, characteristics of distribution not initially noticed or known only tacitly. Conceptions of statistics and of relevant characteristics of distribution were further expanded as students justified and argued about the utility and prospective generalization of particular inventions. Teachers supported student learning by highlighting prospective relations between characteristics of measures and characteristics of distribution as they emerged during the course of activity in each setting.

In the past three decades, statisticians in business, industry, and academia have promoted statistical thinking as an important outcome for students learning statistics. In what has since become one of the most influential documents in statistics education (Garfield, Le, Zieffler, & Ben-Zvi, 2014). Likewise Moore (1998), has recommended that students gain multiple experiences with the messy process of data collection and exploration, discussions of how existing data are produced, experiences which ask them to select appropriate statistical summaries and draw evidence-based conclusions. Not to mention this, according to Cobb and Moore (1997), if students equate statistics with mathematics, they may expect the focus to be on a single correct outcome. However, statistics offers distinctive and powerful ways of reasoning that are distinctly different from mathematical reasoning.

By the same token, The National Council of Teachers of Mathematics (NCTM) gives importance to data analysis. Grades 3–5 Expectations: In grades 3–5 all students should design investigations to address a question and consider how data-collection methods affect the nature of the data set; collect data using observations, surveys, and experimentals; represent data using tables and graphs such as line plots, bar graphs, and line graphs; recognize the differences in representing categorical and numerical data.

Adding to that NCTM (2016), also stated to select and use appropriate statistical methods to analyze data. Grades 3–5 Expectations: In grades 3–5 all students should describe the shape and important features of a set of data and compare related data sets, with an emphasis on how the data are distributed; use measures of center, focusing on the median, and understand what each does and does not indicate about the data set; compare different representations of the same data and evaluate how well each representation shows important aspects of the data.

Furthermore, NCTM (2016) gives importance to develop and evaluate inferences and predictions that are based on data. Grades 3–5 Expectations: In grades 3–5 all students should propose and justify conclusions and predictions that are based on data and design studies to further investigate the conclusions or predictions.

The importance for data analysis; select and use appropriate statistical methods to analyse data; to develop and evaluate inferences and predictions that are based data for other grades like Pre K-2, Grades 6-8 and Grades 9-12 are added in Appendix A.

2.5 Teaching and Learning Statistics with Technology

Many research studies over the past several decades, however, indicate that most students and adults do not think statistically about important issues that affect their lives (Ben-Zvi & Garfield, 2004). Biehler, Ben-Zvi, Bakker and Makar (2012), stated in parallel to these developments, statistical "packages" such as SPSS (http://www.spss.com) and BMDP (http://www.statistical-solutions-software.com) were developed for supporting the statistical practitioner. For many decades these two tools were characterized as a "black box" with a collection of statistical methods, where the user analyzed the statistical problem, selected the appropriate method (predominantly numerical), and obtained the corresponding results. However, neither interactive working styles nor statistical graphs were very much supported with these packages at that time.

On the other hand, Olive and Makar (2010) explained technologies for learning statistics should mirror the theory and practice of professional statistics packages to keep the gap between learning statistics and using statistical methods professionally as small as possible. Another perspective is to use technology to improve the learning of statistics. The focus in this perspective is on other affordances of technology, such as making statistics visual, interactive and dynamic, focusing on concepts rather than computations, and offering the opportunity to experimental with data to make it engaging for students.

In the early 1990s, ProbSim and DataScope were developed by Cliff Konold's team for doing probability simulations and data analysis, respectively, that were easy to learn and simple enough for students. The drawback was that they did not support conventional statistical experimentals or the creation of new methods (Ernie, 1996). In addition to, Chance, Ben-Zvi, Garfield and Medina (2007) discussed many technological tools are available for statistics instruction. Choosing technology or a combination of technologies that is most appropriate for the student learning goals, could involve a complex set of considerations and decisions about how to best choose and use these tools, how often to use them, and for what purposes and activities. In a like manner, Biehler et al. (2012) acknowledged that there are many types of technological tools and resources to support the learning and teaching of statistics.

These include: (a) statistical software packages, (b) spreadsheets, (c) applets/stand-alone applications, (d) graphing calculators, (e) multimedia materials, (f) data repositories, and (f) educational software. The goal of this section is to provide an overview of these types of tools and common examples of each, and to highlight the requirements for software from an educational perspective.

Technology in statistics class enables students to have adequate time to explore, analyze and interpret data. Another benefit is that information technology can assist students in understanding the abstract ideas of statistics. Students can display and visualize data in multiple representation form such as histograms and boxplots by using computer (Garfield & Ben-Zvi, 2008).

According to Rubin (2007), information technology has the potential to make the complex concepts and ideas of statistics more accessible.

2.5.1 Teaching and Learning Statistics with technology in Primary School

A group of statistics educators developed the vision that, to realize the potential of technology at school level, would require the creation of specific tools adapted to inexperienced students' needs that could also grow up with them as they gained expertise (Konold & Higgins, 2010). Biehler et al. (2012), soon after Fathom software, TinkerPlots (Konold & Miller, 2005; http://www.keypress.com) was developed by Cliff Konold's team as a kind of little sibling of Fathom for younger children.

2.5.2 Teaching and Learning Statistics with technology in Secondary School

New technological tools supported the careful checking of more complex assumptions of traditional procedures, for instance by residual analysis or graphical data exploration. More robust methods could therefore be developed and implemented by the support of technology (Biehler et al., 2012).

Moreover, Biehler et al. (2012) stated the new technological support allowed the user more easily to locate patterns in the association of two variables, select an appropriate functional model, and then check the residuals for deviation from the model. This is a much more challenging process of statistical reasoning than just applying the algorithm of least squares. Moreover, this more complex process is more adequate to solve real problems and thus technology indirectly contributed to the empowerment of statistics to solve such problems.

Further to that, Biehler et al. (2012) also said the software Fathom (Finzer, 2001; http://www.keypress.com) was later developed for secondary and tertiary statistical learning and realized and extended many of the envisioned features. Biehler et al. (2012) also mentioned statistical packages are computer programs designed for performing statistical analyses. Several packages are commonly used by statisticians, including SAS (http://www.sas.com), SPSS (http://www.spss.com), and Minitab (http://www.minitab.com). Although these packages were mainly designed for use by science and industry, they have evolved into statistics learning tools for students and are increasingly used in introductory statistics classes.

2.6 The Concept of Statistical Reasoning

According to Gal and Ginsburg (1994), statistical reasoning refers to the ability to understand and integrate statistical concepts and ideas in order to interpret data and make decisions based on a given context. In addition, positive stimulation of students' self and value beliefs about statistics has beneficial effects. These motivational factors influence the development of adequate statistical thinking during teaching and learning process, the structural application of the knowledge obtained in real life situations, and future interest in statistics.

Lavigne and Lajoie (2007) opinionated, statistical reasoning can accompany deductive and inductive reasoning in inquiry situations where hypotheses are formulated and tested for experimentals or surveys designed to answer specific questions. For example, one might have a hypothesis that being athletic changes one's resting heart rate. We can conduct an experimental and deduce from our theory that if we have matched groups and each group (athletic and couch potatoes) engages in a 15-min activity then the athletic group should have an average heart rate lower than the couch potatoes assuming that they are matched for age, diet, gender, etc. In addition, from this fact we can infer inductively that generally there is something about exercise that leads to changes in resting heart rate. To come to this conclusion, we must decide which statistical test is appropriate to the question, and this decision requires some reasoning about the nature of the tests themselves and the type of data involved. In other words, reasoning about data and reasoning about statistical measures are necessary. Thus, while deductive and inductive reasoning are involved, statistical reasoning is also required for making sense of the data upon which the general reasoning is based, and this sense-making is based on inferences made given one's knowledge of statistics.

2.6.1 The Measurement of Statistical Reasoning

According to Gundlach, Andrew, Richards, Nelson and Levesque-Bristol (2015) web-augmented traditional lecture, fully online, and flipped sections, all taught by the same instructor with the same course schedule, assignments, and exams in the same semester, were compared with regards to student attitudes; statistical reasoning;

performance on common exams, homework, and projects; and perceptions of the course and instructor. The Survey of Attitudes Toward Statistics-36 (SATS-36) instrument and eight questions from the Statistical Reasoning Assessment (SRA) were given both at the beginning and end of the semester to measure change. The students selected their own sections, but the students in the sections were similar demographically, with similar pre-course college grade point averages. The SATS-36 showed increases in affect, cognitive competence, and perceived easiness and decreases in value, interest, and effort from beginning to end of the semester for all sections. Only affect and perceived easiness showed any differences for section, with traditional higher than online on average for both. Results from the SRA questions showed an increase in correct statistical reasoning skills and decrease in misconceptions for all sections over the semester. Traditional students scored higher on average on all three exams, but there were no significant differences between sections on homework, the project, or on university evaluations of the course or instructor. Results are contextualized with prior educational research on course modalities, and proposals for future research are provided.

2.6.2 The Teaching and Learning of Statistical Reasoning

Gil, Ben-Zvi & Apel, (2008) presented an initial framework to assess creative praxis of primary school students involved in learning informal statistical inference in statistical inquiry settings. In building the suggested framework, the research adapt the three common characteristics of creativity in the mathematics education literature, namely, fluency, flexibility, and novelty, to the specifics of learning statistics. The research used this framework to capture creative praxis of three sixth grade students in a 60-min statistical inquiry episode. The episode analysis illustrates the strengths and limitations of the suggested framework. The research finally consider briefly research and practical issues in assessing and fostering creativity in statistics learning.

Gal (2002) argued that understanding, interpreting, and reacting to real-world messages that contain statistical elements go beyond simply learning statistical content. He suggested that these skills are built on an interaction between several knowledge bases and supporting dispositions. Statistical literacy skills must be activated together with statistical, mathematical and general world knowledge.

Ben-Zvi (2006) found that the growing samples task design combined with "what-if" questions not only helped students make sense of the data at hand, but also supported their informal inferential reasoning by observing aggregate features of distributions, identifying signals out of noise, accounting for the constraints of their inferences, and providing persuasive data-based arguments see also (Gil & Ben-Zvi, 2011).

The growing awareness of students to uncertainty and variation in data enabled students to gain a sense of the middle ground of "knowing something" about the population with some level of uncertainty and helped them develop a language to talk about the grey areas of this middle ground (Ben-Zvi, Bakker, Aridor, & Makar, 2012).

2.6.3 Statistical Reasoning Studies with Technology

According to Pratt and Noss (2010), in a series of studies on children's understanding about chance and distribution, these researchers examined students' emergent understanding of randomness while engaging with software especially designed to support their probabilistic reasoning. The students' understandings were then analyzed by comparing them to experts' meanings for randomness. Unlike the experts, the children shifted rapidly between four meanings for randomness. Moreover, their choices for these meanings seemed to be triggered by seemingly superficial (from the statistical point of view) aspects of the data. These findings helped the researchers re-design the software to better steer the micro-evolution of students' knowledge toward the focus of the expert's aspirations. Besides that, progress in the understandings of teaching and learning of statistical reasoning and the availability of high quality technological tools for learning statistics have enabled the relatively young field of statistics education to integrate and readily capitalize on these advances (Garfield & Ben-Zvi, 2008).

2.7 Statistical Reasoning Learning Environment (SRLE) in Teaching and Learning Statistics

The focus of this research is based on Statistics Reasoning Learning Environment (SRLE). A few research supports on why it's essential to follow the criteria of SRLE class. Bakker (2004b) suggests that asking students to make conjectures about possible samples of data pushes them to use conceptual tools to predict the distributions, which helps them develop reasoning about samples. In these processes, "what-if" questions prove to be particularly stimulating.

Furthermore, standards in the area of mathematical education have largely emphasized the importance of being engaged in inquiry-based learning. Inquiry is the process in which students solve problems, pose questions, construct solutions and explain their reasoning (e.g., (NCTM: National Council of Teachers of Mathematics, 2000, PISA: Programme for International Students Assessment, 2003). One obvious way to bring students into the processes of inquiry learning is by offering them environments and tasks that allow them to carry out the processes and help them build a personal knowledge that they can use and explain what they learn. Rapid advances in computer-based learning have facilitated the opportunities to empower inquiry learning (Gil et al., 2008b).

Several important implications for teacher education programmes emerge from this analysis. First, a holistic approach to the statistical process is needed in order for teachers to understand the importance of spending time assisting students with question formulation and data collection (this analysis found these process components to be underrepresented in the state standards). Second, teachers will need to be prepared to facilitate discussions with students around the expectations that promote statistical reasoning. That is, in many states (to varying degrees), statistics education has moved beyond calculating means and constructing graphs, and it is important that teachers know how to implement these new expectations. Finally, it seems important that teachers begin to see state expectations as a minimum requirement. That is, teachers working in states that expect students only to "do" the process components and that lack attention to statistical reasoning and/ or the statistical process should be encouraged to enhance their instruction to include these critical components of statistical literacy (Newton, Dietiker, & Horvath, 2011).

Besides, students need authentic, situated, and rich experiences in taking samples and learning how samples do and do not represent the population prior to formal higher studies of statistics. These experiences may include collecting data through surveys and experimentals, where they learn characteristics of good samples and reasons for bad samples (e.g., bias), and creating models using simulation tools (e.g., TinkerPlots) to study the relationship between sample and population. These experiences may help students develop a deeper understanding of sampling and (informal) inference, as they repeatedly deal with taking samples, repeated samples, and simulations (Ben-Zvi, Bakker, & Makar, 2015).

The impact of classroom contexts on the development of students' statistical reasoning and thinking abilities and on the improvement of their attitude and beliefs require further study. Even if further empirical study is first required on this issue, it is worthy pointing out the appropriateness of using personalized and specific items when measuring components of affect behaviour (Olani, Hoekstra, Harskamp, & Werf, 2010).

2.8 TinkerPlots in Teaching and Learning Statistics

Better adoption of statistical software among students is one of the important ways to improve the students' statistical knowledge and consequently to strengthen their positive attitude towards statistics (Bastürk, 2005).

There are a few reasons on why TinkerPlots software is chosen. TinkerPlots is a data analysis tool with simulation capabilities (since version 2.0) that has especially been designed for supporting young students' development of statistical reasoning (Grade 4 of primary school to middle-school students, students from the age of 9 onwards) (Biehler et al., 2012).

According to Konold and Miller (2005), TinkerPlots allows students systematically to build their understanding of statistical representations and concepts through exploring data. Also TinkerPlots has the tools serve for students to test their hypothesis, to construct different statistical representations, to see how different data affects their analysis and to see in what degree they need additional data to learn more about nature of the given real life phenomenon (Rubin, 2007).

According to Fitzallen (2007), here are six fundamental criteria of TinkerPlots that made it a popular teaching and learning software. First of all, it has user-friendly and easy to use interface. Secondly, it allows data to be represented in a variety of forms. Thirdly, it aids in translating between mathematical expression and natural language. Fourthly, it expands the memory when organizing or reorganizing data. Fifthly, it offers multiple entry points for abstract concepts, and finally it provides visual representations of activities.

Likewise, by using the dynamic software in statistics such as "TinkerPlots" the students are able to develop their understanding in statistics including grasp of basic concepts before they study advanced topics in Normal distribution. TinkerPlots empower students to use their ability to create graphical representation, which will enable them to develop their visualization skills, thinking skills, concepts and understanding (Khairiree & Kurusatian, 2009).

2.9 Previous Research Findings

Based on Garfield, Le, Zieffler and Ben-Zvi (2014), the research paper describes the importance of developing students' reasoning about samples and sampling variability as a foundation for statistical thinking. Research on expert–novice thinking as well as statistical thinking is reviewed and compared. A case is made that statistical thinking is a type of expert thinking, and as such, research comparing novice and expert thinking can inform the research on developing statistical thinking in students. It is also posited that developing students' informal inferential reasoning, akin to novice thinking, can help build the foundations of experts' statistical thinking. The research is related to this study because it encourages to do further research on students' reasoning.

Fitzallen and Watson (2010) also implemented a recent study using the TinkerPlots software to enhance statistical reasoning of primary five and six graders. They were able to construct plots and utilize plots to support their thinking on the given data. From these three studies, noticed that the TinkerPlots software is a dynamic graphing software package that is usually used in primary schools and middle schools. The research done by Fitzallen and Watson (2010) provides idea of using of a particular and suitable software which is TinkerPlots for this study, for the primary students.

A study carried out by Ben-Zvi (2006), is one that had left a positive impact concerning utilization of information technology in developing students' statistical reasoning ability. The study was carried out on primary five graders to increase their informal ideas of inference and argumentative skills using the TinkerPlots software. Results showed that the TinkerPlots software can support students' multiplicative reasoning, aggregate reasoning, recognition of the value of large samples, and variability justification. The study conducted by Ben-Zvi encourages this study as well because the study is on developing on students' statistical reasoning and using *TinkerPlots* software.

According to Heriques and Oliveira (2016), all students, even in the less successful groups, were able to demonstrate some aspect of informal statistical inference during the lessons using *TinkerPlots* software. Students drew their conclusions based on the data they had collected from their class and often used the data to make inferences about an unknown population (whole school). A few students used probabilistic language for describing their generalizations but even those who made it had difficulties in including references to levels of uncertainty. These results highlight the need of working on probabilistic language issues, helping students to evolve from a deterministic perspective of inference to include uncertainty in their statements.

Wei Chan, Zaleha and Sumintono (2015) conducted a research on the "Impact of Statistical Reasoning Learning Environment (SRLE)". Based on their findings, SRLE is a powerful instructional model that integrates six principles. By using this model, the students could engage actively in the classroom. The study also confirmed the effectiveness of this model since it improved the students' learning outcomes. It is also possible to perform studies, regardless of whether they are related or not related to this subject matter, on different grade levels to further confirm the effectiveness of SRLE. Besides, different types of educational software can be used as the technological tool as well.

2.10 Conceptual Framework

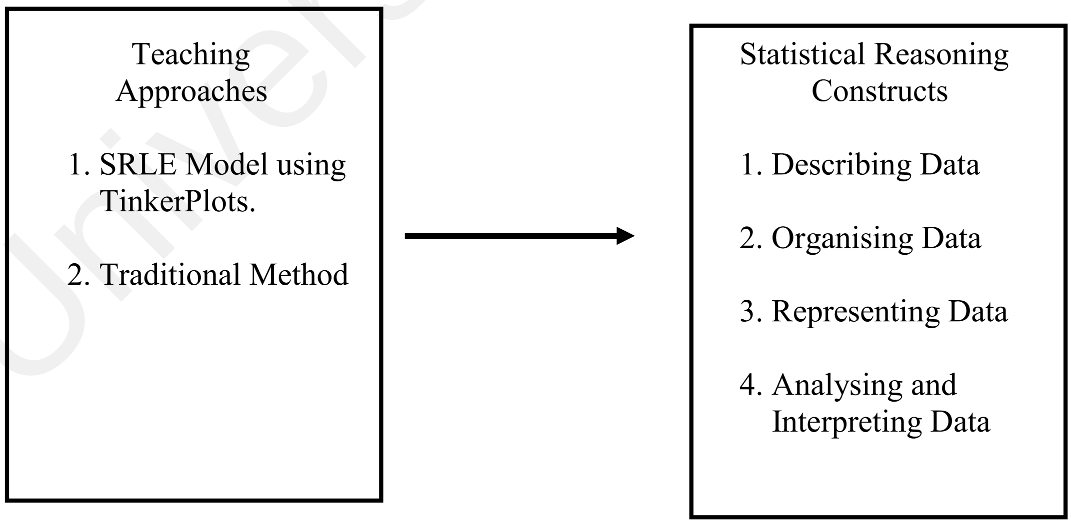


Chart 1: *Conceptual Framework for the study of statistical reasoning among Year 5 pupil using TinkerPlots*

The conceptual framework shown in Chart 1, was constructed based on the review of the literature on theoretical framework and research framework of this study. In order to build students' statistical reasoning skills, two approaches were used. One of the approaches is using interventions which was created based on the combination of SRLE model and TinkerPlots software. The interventions were created based on the theory called, "constructivism". The theory influences the study by contributing a few characteristics. The characteristics coincides with SRLE model as well as contributes towards students' statistical reasoning skills. The model is based on six principles of instructional design described by Cobb & McClain (2004) that on first glance may seem very similar to the six Guidelines for Assessment and Instruction in Statistics Education (GAISE, 2005a, 2005b) recommendations for teaching statistics. The design principles focuses on developing central statistical ideas rather than on presenting set of tools and procedures; uses real and motivating data sets to engage students in making and testing conjectures; uses classroom activities to support the development of students' reasoning; integrates the use of appropriate technological tools that allow students to test their conjectures, explore and analyze data, and develop their statistical reasoning; promotes classroom discourse that includes statistical arguments and sustained exchanges that focus on significant statistical ideas and uses assessment to learn what students know and to monitor the development of their statistical learning as well as to evaluate instructional plans and progress (Garfield & Ben-Zvi, 2008).

Moreover, utilizing TinkerPlots software in the study also linked with SRLE model where the role of technology to explore data, illustrate concepts, generate simulations, test conjectures and collaborate (Garfield & Ben-Zvi, 2008).

The role of technology suggested in SRLE model is parallel with constructivism theory because the students use the technology to integrate new knowledge based on their prior knowledge. The other approach used in the study is traditional method. In this study, traditional method is only used for control group where else the experimental used the interventions approach only.

The conceptual framework acted as the proposed research model of the study to examine whether the experimental group with interventions using TinkerPlots software and SRLE model would improve students' statistical reasoning based on statistical constructs such as describing data, organising data, representing data and interpreting data. On the other hand, the conceptual framework also examines whether the students improve in traditional approach. To conclude, the research hypothesized there is a difference in describing data, organising data, representing data and interpreting data among Year 5 students of this study.

2.11 Summary

In conclusion, scientific articles and previous studies provide some basic information to serve as a guide for the implementation of this study. Among the aspects that are supplied from the analysis and discussion in this chapter, is the technology used to teach and learn statistics education. Statistical reasoning studied in this research is knowledge built based on statistics education. Finally, it analyses the past research on statistics education from the perspective of teaching and learning statistics in primary and perspective of teaching and learning statistics in secondary that provides some basic information. Therefore, this study was carried out to obtain reasonable more information about statistical reasoning from the perspective of the children themselves.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

Chapter three is divided into ten parts which are introduction, research design, research hypotheses, population and sample, data collection method, instruments, validity of instruments, reliability of instruments, data analysis method and summary. First section which is the introduction part, summarises the sections covered in this chapter. In the second section, the contents for chapter three are underlined and the type of research design and justification of using the research design are explained. Research hypotheses are formed in the third section. It is followed by population, location, research sample, and sampling method explained in the fourth section. In the fifth section, the type of data and the method of collecting data are explained. Then, the explanation about the type, purpose and the contents of instrument are discussed in section six and in section seven and eight are about the validity of the instruments and reliability of instruments respectively, the usage of the findings are explained. Finally, in the ninth section the analysis of data is explained and the tenth section ends with the summary of chapter three.

3.2 Research Design

The purpose of this study was to investigate the impact of using *TinkerPlots* on students' statistical reasoning. In order to accomplish the purpose, the study answers the research questions as in chapter 1, the research carries out a quantitative method. Quantitative methods express the assumptions of a positivist paradigm which holds that behaviour can be explained through objective facts. Design and instrumentation persuade by showing how bias and error are eliminated (Firestone, 1987). There are a few reasons of choosing quantitative for the study. Quantitative research is based on a positivist philosophy which assumes that there are social facts with an objective reality apart from the beliefs of individuals. Moreover, quantitative research seeks to explain the causes of changes in social facts, primarily through objective measurement and quantitative analysis (Taylor, 1984).

This research carried out an experimental research. Experimental research is unique in two very important aspects. It is the only type of research that directly attempts to influence a particular variable, and when properly applied, it is the best type for testing hypotheses about cause-and-effect relationships. In an experimental study, researchers look at the effects of at least one independent variable on one or more dependent variables. The independent variable in experimental research is also frequently referred to as the experimental, or treatment, variable. The dependent variable, also known as the criterion, or outcome, variable, refers to the results or outcomes of the study. The major characteristic of experimental research that distinguishes it from all other types of research is that researchers manipulate the independent variable. They decide the nature of the treatment, whom it is to be applied, and to what extent. Independent variables frequently manipulated in educational research include methods of instruction, types of assignment, learning materials, rewards given to students, and types of questions asked by teachers. Dependent variables that are frequently studied include achievement, interest in a subject, attention span, motivation, and attitudes toward school. After treatment has been administered for an appropriate length of time, researchers observe or measure the groups receiving different treatments (by means of a posttest of some sort) to see if they differ. Another way of saying this is that researchers want to see whether the treatment made a difference. If the average scores of the groups on the posttest do differ and researchers cannot find any sensible alternative explanations for this difference, they can conclude that the treatment did have an effect and is likely the cause of difference. Experimental research, therefore enables researchers to go beyond description and prediction, beyond the identification of relationships, to at least a partial determination of what causes them (Jack, Norman, & Helen, 2015, pp.265-266).

Quasi experimental will be conducted in the research. The research design involves two groups which are an experimental and a control group. Both the groups are assigned to do a pre-test at the beginning. After the pre-test, the experimental group have interventions where else the control group will learn using traditional method. Eventually, both the groups will sit for a post-test. Diagram 1.2 shows the research design of the research that will be conducted.

Diagram 1.2

Research Design

Experimental Group	O1	X1	O2
Control Group	O1		O2

O1 represents the pre-test

O2 represents the post-test

X1 represents the students learning data handling using TinkerPlots Software in SRLE

The quasi experimental design was selected, because the analysis of students' reasoning skill using *TinkerPlots* can be done between groups which are with or without intervention or treatment for the experimental group only. For this research, only the experimental group is undergoing interventions using *TinkerPlots* in learning data handling. Moreover, the reason for choosing experimental research is to test the hypotheses in the research. A few hypotheses are formed in the research based on the research questions, so the most suitable design to conduct the research is experimental research.

Other experimental design, like true experimental, is not selected for the research because normally in true experimental the sample or participants are randomly assigned where else in quasi experimental design, the participants are not randomly chosen. According to (Springer, 2010), true experimental designs are based on random assignment of participants to an experimental group and at least one control and/ or comparison group. The essential ingredient of a true experimental design is that subjects are randomly assigned to treatment groups. However, quasi experimental designs do not include the use of random assignment (Jack et al., 2015). In addition to, quasi experimental is chosen compared to other experimental design because quasi experimental allows to have a few rival hypotheses since there is less control in it.

Quasi experimental designs are commonly employed in the evaluation of educational programs when random assignment is not possible or practical. Although quasi-experimental designs need to be used commonly, they are subject to numerous interpretation problems. The non-equivalent group, pre- test and post-test design partially eliminates a major limitation of the non-equivalent group, post-test only design. At the start of the study, the researcher empirically assesses the differences in the two groups. Therefore, if the researcher finds that one group performs better than the other on the post-tests/he can rule out initial differences (if the groups were in fact similar on the pre-test) and normal development (e.g. resulting from typical home literacy practices or other instruction) as explanations for the differences. Some problems still might result from students in the comparison group being incidentally exposed to the treatment condition, being more motivated than students in the other group, having more motivated or involved parents, etc. Additional problems may result from discovering that the two groups do differ on the pre-test measure. If groups differ at the onset of the study, any differences that occur in test scores at the conclusion are difficult to interpret (Gribbons & Herman, 1997).

The research can consider alternative explanations for any observed differences in outcome measures. If the treatment group outperforms the control group, consider a full range of plausible explanations in addition to the claim that the innovative practice is more effective. Program staff and participants can be very helpful in identifying these alternative explanations and evaluating the plausibility of each. Moreover, using multiple evaluation methods, evaluators should be careful in collecting the right kinds of information when using experimental frameworks. Measures must be aligned with the program's goals or objectives. Additionally, it is often much more powerful to employ multiple measures. Triangulating several lines of evidence or measures in answering specific evaluation questions about program outcomes increases the reliability and credibility of results (Gribbons & Herman, 1997).

3.3 Research Hypotheses

To answer the research questions formed in chapter 1, some hypotheses are developed and tabulated as in Table 3.1.

Table 3.1
Research Hypotheses

Research Questions	Research Hypotheses	
RQ1	Null hypothesis, H_0: The mean of post-test scores and the mean of the pre-test scores of Year 5 students in control group are not different.	H_0: $\mu_{pre\text{-}test} = \mu_{post\text{-}test}$
	Research Hypothesis, H_1: The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in control group.	H_1: $\mu_{post\text{-}test} > \mu_{pre\text{-}test}$
RQ2	Null hypothesis, H_0: The mean of post-test scores and the mean of the pre-test scores of Year 5 students in experimental group using *TinkerPlots* are not different.	H_0: $\mu_{pre\text{-}test} = \mu_{post\text{-}test}$
	Research Hypothesis, H_1: The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in experimental group using *TinkerPlots*.	H_1: $\mu_{post\text{-}test} > \mu_{pre\text{-}test}$
RQ 3	Null hypothesis, H_0: There is no mean difference of Year 5 students' statistical reasoning in the post test scores between the control and experimental group when controlling for pre-test.	H_0: $\mu_{control\ group} = \mu_{experimental\ group}$
	Research Hypothesis, H_1: 2. There is a mean difference in Year 5 students' statistical reasoning in post-test based on using *TinkerPlots*, controlling for pre-test.	H_1: $\mu_{experimental\ group} \neq \mu_{control\ group}$

RQ 4	Null hypothesis, H_0: There is no mean difference of Year 5 students' constructs of statistical reasoning in the post test scores on using *TinkerPlots*.	H_0: $\mu_{control\ group} = \mu_{experimental\ group}$
	Research Hypothesis, H_1: 2. There is a mean difference in Year 5 students' constructs of statistical reasoning in the post-test based on using *TinkerPlots*.	H_1: $\mu_{experimental\ group} \neq \mu_{control\ group}$

3.4 Population and Sample

The population of this study comprised Year Five students in international schools in the district of Selangor, Malaysia which focuses on UK curriculum.

Convenience sampling is chosen for the study. A convenience sample is a group of individuals who (conveniently) are available for study. Researchers tend to obtain samples from groups they have ready access to (Springer, 2010). The obvious advantage of the study is convenience (Jack et al., 2015). However, there a few disadvantages in the sampling. Firstly, the study can be biased. Secondly, the generalization can be made more plausible if data are presented to show that the sample is representative of the intended population on at least some relevant variables (Jack et al., 2015). The disadvantages was tackled because common problems are notified among international school students. On that purpose, international school students were chosen. The students represent other international school students who are in year 5 with similar characteristics.

One of the international schools was chosen from the district of Selangor. The school comprises UK curriculum. Two group of students were chosen from Year 5 classes. The group of students were chosen because statistical reasoning is mainly tested in Key Stage 2. Year 6 was not chosen because the interventions cannot be done due to the time limitation that they have, as they have to prepare themselves for checkpoint. Year 3 was also not chosen because of the new transition from Key Stage 1 to 2. Teachers require more time with Year 3 students to assist them to settle down in Key Stage 2.

The research was held at one of the branches of an international school in Selangor. The other branches are located in North (Penang), South (Johor) and East Malaysia (Miri). The school consists of four classes of year five. The research has used convenience sampling. Two intact classes from year five were chosen as a control group and an experimental group. There are several reasons to choosing the sample and the convenience sampling. One of the reasons is, it is easier to conduct the research where the reaseacher can plan on how to collect the data. The sampling is convenient because the researcher has intentions to achieve the objectives from the selected sample knowing their characteristics. It is also convenient to arrange the plan and conduct it. In addition, the sampling saves time as well as money. It is relatively fast and inexpensive. Moreover, the sampling is also helpful to collect useful data that would not be collected in probability sampling.

The research consists of 52 respondents from one of the year groups, from Key Stage 2, from the selected school. The students in the year group are mix ability group students. Majority of the students are Malaysian and the remaining are non-Malaysian from Korea, America, New Zealand, Singapore, Spain and Africa.

The ICT lab is set up with a desktop, LCD, projector, smart board and white board. The students sit on the floor (carpet time) where the teacher does the starter and main teaching. Students respond to the lesson verbally or by jotting down on the lap board to show to the teacher. The assistant teacher checks on students' feedback too.

Then, the students continue their tasks assigned for the day as planned. Each student has a desktop to learn data handling using *TinkerPlots*. The experimental group underwent the process where else the control group learnt data handling in the class room using traditional method.

The research conducted a treatment which is known as interventions. The interventions were planned based on the Statistical Reasoning Learning Environment (SRLE) class. Simultaneously, the interventions were planned based on the constructs of data handling. The learning objective for each lesson planned to liaise with SRLE class and constructs of data handling. At the same time, the learning objective for the lesson focuses on measuring the research objectives.

3.5 Data Collection

At first, two intact groups were respectively chosen from the year group to sit for the pre-test. After the pre-test, the experimental group had interventions for a month and a week using *TinkerPlots*. The interventions for experimental group was planned for 10 lessons as Table 3.2. The interventions was developed based on constructivism theory. The theory requires to have collaborative activities which build the students' understanding through investigating statistical problems and discuss their ideas and reasoning with other students. Learning facilitated by making and testing of conjectures (Garfield & Ben-Zvi, 2008). Thus, based on the criteria suggested in the theory, the interventions planned integrated with the theory to build students' statistical reasoning.

For the first week, the students were introduced to *TinkerPlots* by the teacher. The lesson was carried out for an hour. The following week, the teacher started relating to the topic, data handling with *TinkerPlots*. The students were given a project to conduct later. They had to collect data, analyse and interpret the data using *TinkerPlots* software.The students also did practice on the worksheets prepared by the teacher. The interventions took 40 minutes to explore data handling using *TinkerPlots*. 10 minutes was used to answer some questions prepared by the teacher and last 10 minutes was to discuss students' answers related to reasoning. The instructional activities for the interventions is attached in Appendix B.

Meanwhile, the control group was taught using traditional method. The lesson planning for the control group planned for 3 lessons, tabulated in Table 3.3. The detailed lesson plans added in Appendix C. Once the interventions were fully completed, the following week all the groups were given a post-test.

There were two qualified Mathematics teachers who were involved in the research to conduct the lessons for the students. The teacher who taught the control group has experience in teaching Mathematics for 25 years in varied international schools. On the other hand, another teacher who has experience in teaching Mathematics for 10 years in both the local and international schools taught the experimental group.

Table 3.2

Lesson Planning for Interventions

LESSON/ LEARNING OBJECTIVE	TEACHER ACTIVITIES	STUDENT ACTIVITIES
Lesson 1 & 2 Learning Objective: I can create the cards (data) in *TinkerPlots*	1. Teacher to introduce the *TinkerPlots* Software and the main functions of the software: Cards, Table, Plot, Slider, Text. 2. Teacher to ask students to insert the cards and add the values or insert table as shown in the slides prepared to continue the values of the attributes.	1. Students to add the attributes and values in *TinkerPlots* which are given in the PowerPoint slide. 2. Students to answer the questions on the paper (Intervention 1) attached in appendix.

Lesson 3 Learning Objective: I can draw the bar chart/ line graph	1. Teacher to give input on what types of graphs can be chosen for data representation. 2. Teacher to remind students on how to draw bar chart and line graphs by showing examples on Smart board.	1. Students will be given graph paper to plot the graphs (bar chart/ line graph). 2. Students to refer to their data analysis on *TinkerPlots* and draw the graphs. 3. Students answer the worksheet (Intervention 2) after drawing the graph.
Lesson 4&5 Learning Objective: I can analyse bar /line graph using data	1. Teacher to facilitate the children to insert the data in *TinkerPlots*. 2. Teacher shows a few ways of exploring the data.	1. Students collect data within their class about favorite subject and least favorite subject. 2. Students will start inserting the data they have collected in *TinkerPlot Software*. 3. Students to try out as teacher's example and explore new ways of analyzing the data. 4. Students compare with their Talk Partner and discuss the inference. 5. Students will do the worksheet (Intervention 3) based on the analysis.
Lesson 6&7 Learning Objective: I can collect data	1. Teacher to prepare questionnaire and remind students what types of data will be collected.	2. Students will be given a questionnaire about can drinks (Intervention 4) attached in appendix. 3. Students will survey students from other classes. They do the survey in the cafeteria by randomly choosing the participants. 4. Students insert the data collected in the *TinkerPlots Software*.

Lesson 8 Learning Objective: I can analyse and interpret data	1. Teacher gives a few examples of interpreting data from the analysis.	2. Students to analyze the data using the software as Figure 4. 3. Students interpret the data. 4. Students to answer the questions on the worksheet (Intervention 5) given.
Lesson 9 Learning Objective: I can create my own survey	1. Teacher encourage students to propose a few ideas to survey at home or school. 2. Teacher gets feedback from the students.	1. Students choose the area they would like to conduct research. 2. Students design their survey form. 3. Students make conjectures.
Lesson 10 Learning Objective: I can collect, analyse and interpret data	1. Teacher to remind students to check whether the conjectures made are the same or not. 2. Teacher to have a whole class discussion about the survey conducted by each student.	1. Students collect data for their topic of interest. 2. Students to key in the data collected in *TinkerPlots.* 3. Students analyze the data. 4. Students to create questions based on the analysis and Talk Partner to answer the questions. Students need to evaluate the answers given by the Talk Partner. 5. Students to answer the questions (Intervention 6) given by the teacher.

Table 3.3

Lesson Planning for Traditional Teaching Approach

LESSON/ LEARNING OBJECTIVE	TEACHER ACTIVITIES	STUDENT ACTIVITIES
Lesson 1 Learning Objective: I can draw the bar chart	3. Teacher to give input on what types of graphs can be chosen for data representation. 4. Teacher to remind students on how to draw bar chart by showing examples on Smart board. 5. Teacher assigns homework.	4. Students will be given graph paper to plot bar chart. 5. Students to do homework given by the teacher.
Lesson 2 Learning Objective: I can draw the line graph	1. Teacher to give input on what types of graphs can be chosen for data representation. 2. Teacher to remind students on how to draw line graph by showing examples on Smart board. 3. Teacher assigns homework.	1. Students will be given graph paper to plot line graph. 2. Students do homework given by the teacher.
Lesson 3 Learning Objective: I can analyse bar /line graph using data	3. Teacher teaches the way of reading the data of bar/line graph. 4. Teacher shows the examples of questions can be asked from the bar/line graph. 5. Teacher teaches the way to answer the questions.	6. Students will answer the questions prepared by the teacher.

3.6 Instruments

3.6.1 Pre-test and Post-Test

A few instruments are used in the study. One of the instruments used is pre-test or post-test to find out the students' statistical reasoning. The pre-test and post-test was developed based on the statistical reasoning Table 3.4. The framework adapted from (Jones, Thornton, Langrall, Mooney, Perry & Putt, 2000). Describing data involves accurate reading of raw data or data demonstrated in charts, tables, or graphs (Jones et al., 2000). It combines the reading of data from the studies of (Curcio, Artz, Gal, & Garfield, 1997). In the study of (Jones et al., 2000) four processes were put forth including reading data representations, demonstrating awareness of essential graphing conventions, identifying when different displays represent displays of the same data and assessing different in terms of describing data, Mooney (2002) identified the existence of four sub- processes, namely demonstrating consciousness of exhibited features, distinguishing similar data in various data depictions, assessing the efficacy of data depiction in data presentation, and recognizing components of data values.

Four Key constructs were built in the framework which are Describing Data, Organising Data, Representing Data and lastly Analysing and Interpreting Data. A few sub processes were built under the constructs. A1 is extracting data which has 2 items; A2 is recognising general features of graphical representation which has 2 items; A3 is showing awareness to the displayed attributes of graphical representation which has an item; B1 is categorising data and B2 is organizing data and each sub process has 2 questions. C1 demonstrating data, C2 is identifying the different representations of the data and C3 is judging the effectiveness of different representations; D1 is making comparisons within data, D2 is making comparisons between two data; D3 is making prediction or inference or interpretations and D4 is making conclusion.

The instrument followed the framework as it is, when adopting constructs. However, the sub processes is adjusted to suit this study. The framework did not include the sub process, “Reduce data using measure of centre” and “Reduce data using measure of spread” under the construct of Organising Data. This is because, both the sub processes were planned for secondary level. Both the sub processes cannot be tested for primary students because the syllabus is not covering measure of centre and measure of spread for them. Thus, the researcher had to cancel the sub processes. Moreover, the researcher added an extra sub process under the construct of Analysing and Interpreting Data which is “Make conclusion”. This sub process is added because the students are given opportunity to predict and test it. It will encourage them to make a conclusion in the end based on the assumption made.

These sub-processes consist of extracting and generating information from the data or graph; showing awareness of the displayed attributes of graphical representation; and recognising the general features of the graphical representation. For the first sub-process, the students have to extract and generate explicit information while reading the data displays. They ought to be aware of the displayed attributes of graphical representation, which is composed of graphical conventions (e.g., title and axis labels) related to the second sub-process. This sub-process is identical to the first sub-process of (Mooney, 2002). Furthermore, the third sub-process is new to the framework where students need to identify the general features of the graphical representation including shape, center, and spread. By integrating these three features together, students will recognize them as a whole entity rather than isolated concepts (Garfield & Ben-Zvi, 2007). The statistical reasoning test is attached in Appendix F.

Table 3.4

The framework of the pre-test and post test

Num	Constructs	Code	Items
1	Describing Data		
i)	Extract data	A1	a) What was Alfonso's highest recored temperature? (1mark) b) What day and time was the lowest recorded temperature? (1mark)
ii)	Recognise general features of graphical representation	A2	c) What does Alfonso's highest temperature explain to you? (1 mark) d) Why is it important the each piece of data should have a regular interval? (1mark)
	Show awareness to the displayed attributes of graphical representation	A3	e) Why the graph is not starting from 0? (1 mark)
2	Organising Data		
i)	Categorise Data	B1	a) What is the data measuring? (1 mark) b) Arrange the dinosaurs from shortest to tallest. (1 mark)
ii)	Organise data	B2	c) How did you get the answer? (2 marks)
3	Representing Data		
i)	Demonstrate Data	C1	a) Draw a suitable type of graph to represent the data given in the graph paper provided. (2 marks)

ii)	Identify the different representations	C2	b) Why is data normally represented in graphs? (1 mark)
iii)	Judges the effectiveness of different representations.	C3	c) Do you think the graph you have chosen is appropriate to represent the data? Support your answer with a reason. (2 marks)
4	Analysing and Interpreting Data		
i)	Make comparisons within data	D1	a) What is the difference in amount Leanne had on Thursday and Saturday in Week 1? Explain how you know. (2 marks)
ii)	Make comparisons between two data	D2	b) Compare the data in Week 1 and Week 2, how can you evaluate about the amount Leanne had? (2 marks)
iii)	Makes prediction/inference/interpretations	D3	c) If Leanne has RM6 on Friday in Week 2, predict the amount that she has in Week 3 on Friday. Support your prediction with a reason. (2 marks)
iv)	Make conclusion	D4	d) Make a conclusion from the four line bar charts given. (1 mark)

3.6.2 Rubric of Statistical Reasoning Test

A rubric of Statistical Reasoning Test (pre-test and post-test) is developed referring to the framework of pre-test and post-test. The mark is allocated according to each question from the constructs. If the students manage to answer the question correctly 1 mark will be given or else 0 mark is given.

Nonetheless, if the students can give more than one explanation, 2 marks will be given; if the students manage to give one explanation, they will only be given 1 mark. 0 mark is given if the answer does not match the rubric.

Each question measures a particular construct adopted for this study. Question 1 (a) and 1(b) measure the construct of 'Extract Data'; Question 1(c) and 1(d) measure the construct of 'Recognise general features of graphical representation; the next construct, ' Show awareness to the displayed attributes of graphical representation' measured by Question 1(e); Question 2(a) and 2(b) measure the construct, 'Categorise Data'; Question 2(c) measures the construct, 'Organise Data'; the construct, 'Demonstrate Data' is measured by Question 3(a); ' Identify the different representations' is measured by Question 3(b), ' Judges the effectiveness of different representations' is measured by Question 3(c); Question 4(a) measures 'Make comparisons within data'; Question 4(b) measures 'Make comparisons between two data'; Question 4(c) measures ' Make predictions/inference/interpretations' and Question 4(d) measures the construct, 'Make conclusion.

The main constructs which are Describing Data, Organising Data, Representing Data and Analysing and Interpreting Data are measuring the variables of statistical reasoning in this study. Thus, each main construct has a few questions to test the students' statistical reasoning. The students statistical reasoning is measured based on their scores in each construct.

The rubric was developed by 5 Mathematics experts from school. The possible answers were discussed by them. Then, the rubric was checked by the expert lecturer from Faculty of Education, Mathematics and Science Department.

3.7 Validity of the Instruments

Validation of the instruments in the research mainly focused on whether it measures the statistical reasoning among Year Five students. In addition to that, the focus was also on whether the instrument measures the suitability level of Year Five students. To accomplish the validity, the Pre-test and Post-test was given to the experts to evaluate the content. Four experts analysed the paper. Firstly, a teacher who has 10 years of experience in teaching both national and international curriculum, stated that Question 2(a) was unclear because of the vocabulary or terms used. Secondly, she asked to make Question 3(b) simple. Thirdly, she notified that there was a missing word which could mislead understanding, the question paper was corrected as per here suggestion.

Then, the paper was sent to be checked for Mathematics content of Year 5 to the Year Leader of Year 5. He is teaching Mathematics for 25 years. He checked the constructs of the paper and the items. A few ideas were adapted from him as Question 1(d) and 1(e). Furthermore, he said the amended Question 3(b) was clear.

After that, the paper was sent to the Deputy Head of Primary School, to check for any sentence structure error. She has experience in teaching and management of schools for about 15 years. She corrected the sentence structure of Question 1(c), 2(a), 4(c) and 4(d).

Finally, for content validity the paper was checked by a lecturer from Faculty of Education, Mathematics and Science Department. The validation rubric conducted by the experts of pre-test and post-test is attached in Appendix D.

The tasks of this technology-based statistical reasoning assessment tool had been validated by three experts, a crucial step which ensures that the items can evaluate the students' statistical reasoning level. The cooperation was carried out via electronic mail. The instrument was not validated concurrently by all experts, but was reviewed by one expert and amended accordingly before it was sent to the next expert. These three experts are lecturers from foreign universities that have published significantly influential works in the field of statistical reasoning. Expert A is an associate professor from the University of Minnesota, USA, with extensive experience in the field. He has taught statistics to university students for more than 20 years and has published countless papers about statistical reasoning in refereed journals, book chapters, and conference proceedings. Expert B is an associate professor from Illinois State University, USA, with years of teaching experience in statistics as well. He was actively involved in the development of models for statistical reasoning. Expert C is a senior lecturer from the University of New England, Australia who has numerous publications on statistical reasoning such as reasoning about sampling, reasoning about variation, informal inferential reasoning, and so forth. All experts contributed valuable views and suggestions to the constructed tasks other than helping to verify the accuracy of the English words used. Appropriate corrections were then made. Since this instrument is in dual language (English and Malay), two lecturers who are excellent in Malay helped to verify the language accuracy (Chan & Ismail, 2014a).

Internal validity and external validity are two sets of criteria that be used in evaluating the worthiness of an experimental design. Internal validity is the quality of and experimental design such that the results obtained can be attributed to the manipulation of the independent variable, whereas external validity is the quality of an experimental design such that the results can be generalized from the original sample and by extension, to the population from which the sample originated. Along with their seminal contribution to understanding and designing experimentals, Campbell and Stanley (1963) identified threats to both internal and external validity (Moutinho & Hutcheson, 2011).

Internal validity of this study was improved by controlling extraneous variables. All the students in the study fully participated. Thus, there is no missing values in both pre-test and post-test. Moreover, test-retest is done for this study. The Pearson correlation test results for the group of subjects (n=23), the test-retest correlation values for pre-test and post-test are 0.65 at the significance of $p<.05$. This means that the instrument is suitable for obtaining reliable data from other subjects who have the same characteristics as this group of research subjects.

External validity in the study was improved by conducting it in natural setting. There is no changes made in the environment of the study as it may change the outcome of the study. The environment of the study is referred to the place of the research conducted where the culture of the school was not amended to suite the research. Convenience sampling is used in the study instead of random sampling to select the participant to improve the internal validity. The results of the study cannot be generalised. However, the results can be represented and limited to the international primary schools with similar curriculum.

3.8 Reliability of Instruments

A pilot study has been conducted in the same school but using Year 5 class from different academic year (2014/2015). The students have left to the next year group. The sample was collected from the current academic year (2015/2016). The pre-test was piloted in the pilot study for validation and reliability purposes besides to check understanding of questions by students. The students were given instructions for the test for 5 minutes. The students were given half an hour to complete the paper which has 4 questions. Question 1 had 1(a), 1(b), 1(c), 1(d), 1(e), 1(f); Question 2 had 2(a), 2(b), 2(c); Question 3 had 3(a), 3(b), 3(c) and Question 4 has 4(a), 4(b), 4(c), 4(d). After 30 minutes, the papers were collected. 10 students out of 25 could not complete the paper within the given time. A few adjustments were considered to be made to the paper based on the students' feedback on the paper.

Firstly, the time for pre-test and post-test was increased for an extra 15 minutes because students complained that there was not enough of time to read questions and answer them. With the action taken, all the students will be able to answer all the questions within the given time. Secondly, Question 3(b), "Describe how the data relates to the graph you have chosen?" was stated that it was unclear for students. Thus, the question was changed to "Why is data normally represented in graphs?" The question was constructed based on the constructs as Table 3.3. The question was then checked by two experts and 10 students were randomly chosen from the same pilot test group to answer the question. All students understood what has been asked and almost all answered correctly. Thirdly, Cronbach's alpha reliability correlation value of the paper, .691 was improved to .703 by discarding Question 1 (f), "If Alfonso is at Antarctica, what could be his temperature? Discuss your answer". The initial SPSS output for the Cronbach's alpha added in Table 3.5 followed by Table 3.6 (Item Total

Statistics) to check the Cronbach's alpha if item deleted. The final SPSS output for the Cronbach’s alpha added in Table 3.7. Table 3.5, Table 3.6 and Table 3.7 are added in Appendix E. The paper was corrected by discarding Question 1 (f) after consulting the experts as well. The research could not discard any other items because it measures the statistical reasoning of the students which is the purpose of the research.

The validity and reliability of the technology-based statistical reasoning assessment tool had been measured by previous study too.

Two raters were involved in statistics; both of them are lecturers from local universities and are proficient in statistics and mathematics education. Rater A is an associate professor from Universiti Teknologi Malaysia and has 15 years of teaching experience in statistics and mathematics. The rater’s field of specialization is in advanced mathematical thinking and problem solving. Meanwhile, rater B is a senior lecturer from the same university who has extensive teaching experience in statistics and mathematics subjects as well. He was a lecturer in the Islamic Azad University, Iran, before joining the current university. The researcher tabulated the four constructs, sub-processes, and items before both raters were asked whether they agree or disagree. This was done by either giving a (√) or (X). Both raters were requested to judge the appropriateness of the items under the four constructs within a two week period before an in-depth discussion was held. Then, the percentage of agreement was calculated based on their judgment (Chan & Ismail, 2014a).

The three experts who validated the instrument had commented on the strengths and weaknesses of the instrument. Concerning instrument strength, expert A mentioned that there were some good items in this assessment tool. In addition, expert A also pointed out that it is acceptable to have both statistical literacy and statistical reasoning items in the instrument as some content is interconnected and sometimes statistical reasoning is the subset of statistical literacy (delMas, 2002). Expert B stated that there were two good questions to assess statistical reasoning, i.e., 'Describe the distribution of the graph with respect to its shape, center, and variability' and 'Which graph do you think represents the data better, the histogram or the box plot? Explain why.' Expert C found this instrument interesting and is looking forward to reading the published results.

The inter-rater reliability for this assessment tool is reasonably consistent. Since the instrument has strong validity and reliability, it is highly recommended that this instrument be used not only at the secondary school level, but also at the university level (Chan & Ismail, 2014b).

3.9 Data Analysis

The study consists of data procedure, showing on how to score the pre-test and post-test using the rubric prepared and analysing the score using inferential analysis.

After the pre-test was conducted, the examiner checked the students' responses and marked the answer based on the rubric prepared. The students were given appropriate scores based on their answers. The same process was taken for post-test as well. The rubric of the pre-test and post-test is attached in Appendix G.

Marks were collected for pre-test and post-test for each question and total marks was keyed in Statistical Package for Social Sciences (SPSS). The data was analysed based on the research questions. Research question 1 to 4 used inferential analysis.

Research Question 1: To answer research question 1, a dependent t-test or paired t-test was used in SPSS to compare the means in between pre-test and post-test this is to look whether there is any difference in the control group of the students' score after traditional teaching approach.

Paired-samples t-test was chosen for the first research question, fulfilled the assumptions. Firstly, it measured the students' scores using the same instrument or test-paper after some duration. Secondly, the dependent variable which is pre-test and post-test is interval scale. Thirdly, two classes were randomly chosen from the population of Year 5 international schools. At last, the data was also checked for normality using Shapiro-Wilks test because the sample is below 30.

Research Question 2: To answer research question 2, a dependent t-test or paired t-test was used in SPSS to compare the means in between pre-test and post-test. This is to look whether if there is any difference in experimental group of the students' score after having interventions using *TinkerPlots*.

Paired-samples t-test is suitable for research question 2, because the sample or individual was tested twice before and after a period of time using the same instrument or test paper which is pre-test and post-test. In addition to, the requirements for the test have been made by the questions. At first, the data used as the dependent variable in the research (the pre-test and post-test scores) is interval scale. In the second place, two classes were randomly chosen from the population of Year 5 international schools. Then, repeated measurements was taken using the same sample and test paper. Lastly, the data was also checked for normality using Shapiro-Wilks tests because the sample is below 30.

Research Question 3: To answer this research question the post-test scores in control and experimental group was also analysed using One –Way ANCOVA. The test was conducted in SPSS to find out if there is any difference among Year 5 control group and experimental group in terms of the students' score after the treatment of having interventions using *TinkerPlots* by controlling the pre-test score.

One –Way ANCOVA is chosen for the third research question, has met the assumptions and requirements required. Firstly, the sample size contains more than fifteen subjects in normal conditions, the study has 25 students which is enough to obtain an accurate result. Secondly, the dependent variable in the research which is pre-test and post-test scores of students known as interval scale. Thirdly, the research is conducted to determine if there is a statistical significant difference between two groups of data which are the experimental group and control group. For the questions, the data was also checked for normality using Shapiro-Wilks test. The reason of the test to be conducted is to check whether the data is normally distributed.

Research Question 4: To answer this research question the post-test scores in control and experimental group was also analysed using MANOVA. The test was conducted in SPSS to find out if there is any difference among Year 5 experimental group in terms of the students' scores in statistical reasoning constructs after the treatment of having interventions using *TinkerPlots*.

One-Way MANOVA was chosen for the fourth research question, has met the assumptions and requirements required. Firstly, research question 4 studies on two groups which are known as independent variables. The two independent variables were measured in interval scale. Secondly, there are four dependent variables which were studied in research question 3. The dependent variables are the constructs of the statistical reasoning. Thirdly, the sample size contains more than fifteen subjects in normal conditions, the study has 46 students which has adequate sample size to obtain an accurate result. For the research question, the data was checked for normality using Shapiro-Wilks test. The reason the test is conducted would be to check whether the data is normally distributed. Moreover, assumptions has been met for independence of observations. There is no relationship between the observations in each group where different sample or participants participated in each group. Finally, there is no significant outliers.

Table 3.8 shows the summary of the data analysis method of the research questions from 1 to 4.

Table 3.8
Data Analysis Method of Each Research Questions

Research Questions	Statistical Analysis
1. Is there any significant improvement between pre-test and post-test of the control group Year Five students' statistical reasoning?	One-tailed Paired T-Test
2. Is there any significant improvement between pre-test and post-test of the experimental group Year Five students' statistical reasoning?	One-tailed Paired T-Test
3. Is there a significant difference in Year Five students' statistical reasoning in post-test between the control and experimental control group when controlling for pre-test?	One –Way ANCOVA
4. Is there any significant different in the four constructs of statistical reasoning in Year Five students?	One- Way MANOVA

One of the problems with data analysis is the difficulties in getting sample of 30 for the study. Moreover, the researcher also overlooked the testing assumptions. The problems was tackled by ensuring of cleaning data and no outliers or missing data.

Based on the questions, scores were difficult to determine for certain answers. The rubric was overcome by the Maths experts.

3.9 Summary

This chapter consists of introduction, research design, population and sample, data collection, instrument, pilot study and data analysis. A research design for the research and the justifications are explained. Then, population and sample for the research are chosen and explained. Next, the methods for data collection are explained. Justifications for methods used for the research are given. It is followed by the choice of the instruments for the research which is explained together with justification made for the choice. Finally, the pilot study that was conducted is described followed by the data analysis. The discussion on the analysis will be carried out in the next chapter which is Chapter 4, Research Findings based on the research methodology in this chapter.

CHAPTER FOUR

FINDINGS

4.1 Introduction

This chapter analyses and discusses the analysis presenting the results of descriptive and inferential analysis which correspond to the four research questions. The first section explores research question one using descriptive and inferential statistics of control group before and after traditional teaching approach. Paired t-test was conducted to answer the research question. The second section, covers research question two which uses descriptive and inferential statistics of experimental group before and after interventions. Paired t-test was conducted as well to answer research question two. Assumptions for the test were conducted beforehand and discussed for both research question one and two. The third section of the chapter notifies descriptive and inferential statistics of post-test between experimental group and control group by controlling the pre-test scores. One-Way ANCOVA test was used to answer the research question. The last section, research question three also addresses descriptive statistics and inferential statistic of different in the constructs of statistical reasoning. One-Way MANOVA was used for research question three. Prior to the test, assumptions of One-Way ANCOVA and One-Way MANOVA were tested for both the research question three and four. All hypotheses were evaluated at the 5% level of significance.

4.2 Descriptive and Inferential Analysis of Pre-test and Post-test

Pre-test and post-test of this study was developed by the researcher of the study for 50 students and the researcher managed to gather information from 46 students from the test. The students from both groups (experimental group and control group) were asked to sit for the test before intervention. After intervention for experimental group, both the groups retook the test. The demographic information is added in Table 4.1.

Table 4.1
Demographic Information

Demographic Variables		Experimental Group		Control Group	
Gender	Male	13	56.5%	13	56.5%
	Female	10	43.5%	10	43.5%
Nationality	Domestic	15	65.2%	14	60.9%
	International	8	34.8%	9	39.1%

4.2.1 Results of Analysis for RQ 1

To derive research question 1, "Is there any significant difference between the mean score of the pre-test and post-test of the control group?" the researcher analysed the data using descriptive statistics by finding mean and standard deviation before and after intervention, as tabulated in Table 4.2. To satisfy the assumptions a paired t-test, a test of normality was conducted and tabulated in Table 4.3. After that, the data was also analysed in inferential statistics using paired samples t-test to test the research hypothesis stated in the study. The paired-samples t-test results is tabulated in Table 4.5. Table 4.4 shows the adjusted means of post-test of each group.

Table 4.2
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
PreTest	23	2.00	17.00	7.96	3.64
PostTest	23	6.00	17.00	11.00	3.36
Valid N (listwise)	23				

Table 4.2 shows the mean score of pre-test and post-test in control group after the traditional teaching approach. It shows the magnitude of the difference between the tests and can be seen which test has a higher mean. The post-test has higher mean (*M* = 11.00, *SD* = 3.36) than the pre-test (*M* = 7.96, *SD* = 3.64).

Table 4.3
Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
PreTest	.94	23	.21
PostTest	.94	23	.16

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

One of the requirements of the tests conducted in the study is normality, to check whether the score in the population is normally distributed. Thus, the study administered normality test. Table 4.3 shows the normality test table regulated for paired samples t-test.

The number of participants was below 50, hence, Shapiro-Wilk was used to check the normality. Based on Table 4.3, Shapiro-Wilk test tests whether the scores of the pre-test and post-test in control group are statistically significantly different from a normal distribution. The null hypothesis for the pre-test, the sample data of pre-test is not significantly different than normal population, fail to be rejected since the significance difference is .21 which is more than the alpha level .05. Therefore, the scores in pre-test are normally distributed.

On the other hand, the null hypothesis for the post-test, after the traditional teaching approach and normal distribution are not statistically significantly different from a normal distribution. Thus, the null hypothesis fail to be rejected since the significance difference is .16 which is more than the alpha level .05. In conclusion, it is presumed that the scores in the post-test is normally distributed.

Table 4.4
Paired Samples Test

		Paired Differences					t	df	Sig. (1-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PreTest - PostTest	-3.04	2.98	.62	-4.33	-1.76	-4.90	22	.00

A paired samples *t*-test was conducted to determine the significant improvement between pre-test and post-test after the traditional teaching approach on students' statistical reasoning skills. Based on Table 4.3, the null hypothesis is rejected. There was a significant improvement in the scores for control group (M = -3.04, SD = 2.98), t (22) =-4.90, p<.0005. The effect size was = 1.02. It was calculated based on the formula,

$$d = \frac{\overline{x}_t - \overline{x}_c}{s_{pooled}}$$

Key to symbols:
d = Cohen's size
x = mean (average of treatment or comparison conditions)
s = standard deviation
Subscripts: t refers to the treatment condition and c refers to the comparison condition (or control condition).

According to Cohen's (1988) interpretation, this is interpreted as a large effect. There is nearly one standard deviation unit of difference between the means of the pre-test and post-test scores. These results proves that the students in control group performed better in post-test after traditional teaching approach for data handling.

4.2.2 Results of Analysis for RQ 2

In order to answer research question 2, "Is there any significant difference between the mean score of the pre-test and post-test of the experimental group?", the researcher analysed the data using descriptive statistics by finding mean and standard deviation before and after intervention as tabulated in Table 4.5. To satisfy the assumptions of paired t-test, a test of normality was conducted and tabulated in Table 4.6. After that, the data was also analysed in inferential statistics using paired samples t-test to test the research hypothesis stated in the study. The paired-samples t-test results is tabulated in Table 4.8. Table 4.7 shows the adjusted means of post-test of each group. The data was also analysed and showed in Graph 4.1 to show clear and simple representation of comparison of control and experimental groups.

Table 4.5
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Before Intervention	23	6.00	16.00	9.57	2.52
After Intervention	23	11.00	20.00	15.30	2.48
Valid N (listwise)	23				

Table 4.5 shows the mean score of pre-test and post-test in experimental group using *TinkerPlots*. It shows the magnitude of the difference between the tests and can be seen which test has a higher mean. The post-test has higher mean (M = 15.30, SD = 2.48) than the pre-test (M = 9.57, SD = 2.52).

Table 4.6
Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
Before Intervention	.92	23	.06
After Intervention	.96	23	.42

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

One of the requirements of the tests conducted in the study is normality, to check whether the score in the population is normally distributed. Thus, the study administered normality test. Table 4.6 shows the normality test table regulated for paired samples t-test.

Since the number of participants was below 50, Shapiro-Wilk was used to check the normality. Based on Table 4.6, Shapiro-Wilk test tests whether the scores before and after interventions are statistically significantly different from a normal distribution. The null hypothesis for the before interventions and normal distribution are not statistically significantly different from a normal distribution. Thus, the null hypothesis fail to be rejected since the significance difference is .06 and presumed there is not statistically significant difference between before interventions and normal distribution. Therefore, the scores before interventions are normally distributed.

On the other hand, the null hypothesis for the after interventions, after interventions and normal distribution are not statistically significantly different from a normal distribution is fail to be rejected since the significance difference is .42. In conclusion, presumed that the scores after interventions is normally distributed as well.

Table 4.7

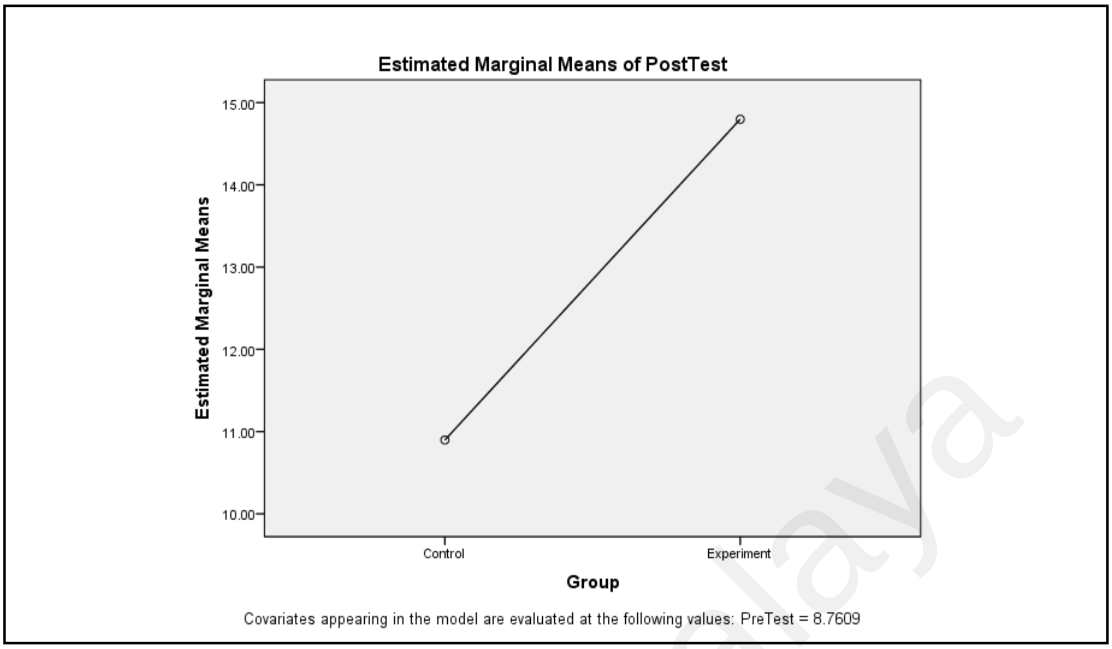
Adjusted Means of Posttest of Each Groups

Dependent Variable: PostTest

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Control	10.90[a]	.42	10.05	11.75
Experimental	14.80[a]	.42	13.95	15.65

a. Covariates appearing in the model are evaluated at the following values: PreTest = 8.7609.

Since the required assumptions were met the descriptive and inferential analyses on post-test scores were conducted. The adjusted means of post-test scores of experimental group and control group 10.90 (*SE* = .421, 95% CI [10.05, 11.75]) and 14.80 (*SE* = .42, 95% CI [13.95, 15.65]) respectively.



Graph 4.1: *Adjusted Means of Posttest Scores for Control and Experimental Groups*

Graph 4.1 illustrates the estimated marginal means of post-test scores where the adjusted mean of post-test score for Experimental group was higher than the adjusted mean of post-test score of Control group.

Table 4.8
Paired Samples Test

		Paired Differences					T	df	Sig. (1-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Before Intervention - After Intervention	-5.74	2.09	.44	-6.64	-4.83	-13.15	22	.00

A paired samples *t*-test was conducted to determine the significant improvement between pre-test and post-test in using *TinkerPlots* on students' statistical reasoning skills. Based on Table 4.8, the null hypothesis is rejected. There was a significant improvement in the scores for the experimental group (M = -5.74, SD = 2.09), t (22) =-13.15, p<.0005. The effect size was = 2.74. It was calculated based on the formula,

$$d = \frac{\bar{x}_t - \bar{x}_c}{s_{pooled}}$$

Key to symbols:
d = Cohen's size
x = mean (average of treatment or comparison conditions)
s = standard deviation
Subscripts: t refers to the treatment condition and c refers to the comparison condition (or control condition).

According to Cohen's (1988) interpretation, this is interpreted as a very large effect. There is nearly two standard deviation units of difference between the means of the pre-test and post-test scores. These results proves that the students in experimental group using *TinkerPlots* performed better in post-test after using *TinkerPlots* for data handling.

4.2.3 Results of Analysis for RQ 3

The next research question "Is there a significant difference in Year 5 students' statistical reasoning in post-test between the control and experimental group when controlling for pre-test?" was answered using descriptive statistics and inferential statistics.

One-Way ANCOVA test was conducted to determine if the mean post-test score is different based on the teaching approach using TinkerPlots while controlling for pre-test score and tabulated in Table 4.10. A few assumptions were tested in ANCOVA in this study which include a) independence of observations, b) homogeneity of variance, c) normality, d) linearity, e) independent of the covariate and the independent variable and f) homogeneity of regression slopes. The results of the normality test tabulated as Table 4.9 is homogeneity of variance; Table 4.10 and Table 4.11 are to show the normality; Table 4.12 is independence sample test and Table 4.13 is homogeneity of regression slope.

a) Independence of Observations

In examining the Graph 4.2 (attached in Appendix I) for evidence of independence, the scatterplot does suggest evidence of independence with relative randomness of points above and below the horizontal line at 0.

In examining the Graph 4.3 (attached in Appendix I) for evidence of independence, the scatterplot does suggest evidence of independence with relative randomness of points above and below the horizontal line at 0.

The independence of observations was met by random assignment of students to instructional method. This assumption was also confirmed by review of scatterplot of residuals against the levels of the independent variable. A random display of points around 0 provided further evidence that the assumption of independence was met. Adding up to that, there is no relationship between the observations in each group where different participants participated in each group. Thus, the assumptions has been met.

b) Homogeneity of Variance

Table 4.9

Levene's Test of Equality of Error Variances[a]

Dependent Variable: PostTest

F	df1	df2	Sig.
.11	1	44	.745

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + PreTest + Group

According to Levene's test in Table 4.9, the homogeneity of variance assumption is satisfied F (1, 44) = .11, p>.05. Thus, the null hypothesis is fail to be rejected. In conclusion, it indicates that the assumption of homogeneity of variance is met.

c) Normality

Table 4.10

Descriptives

		Statistic	Std. Error
Residual for Post Test	Skewness	-.27	.35
	Kurtosis	.36	.69

Table 4.11

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
Residual for PostTest	.99	46	.81

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The assumption of normality was tested and met via examination of the residuals. Based on Table 4.10 and Table 4.11, review of the S-W test for normality (SW= .99, df= 46, p= .81) and skewness (-.27) and kurtosis (.36) statistics suggested that normality was a reasonable assumption. The boxplot and histogram (attached in Appendix E) suggested a relatively normal distributional shape of the residuals. The Q-Q plot suggested normality was reasonable. In general, there is evidence that normality has been met.

d) Independence of the Covariate and Independent Variable

Table 4.12
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
PreTe-st	Equal variances assumed	1.29	.26	-1.74	44	.09	-1.61	.92	-3.47	.25
	Equal variances not assumed			-1.74	39.17	.09	-1.61	.92	-3.47	.25

Independence of the covariate and independent variable was met by random assignment of students to teaching approach. This assumption was also confirmed by an independent t test which examined the mean difference on the covariate (pre-test) by independent variable (teaching approach). The results were not statistically significant, $t(44)= -1.744$, $p<.005$, which further confirms evidence of independence of the covariate and independent variable. There was not a mean difference in pre-test based on teaching approach using *TinkerPlots*.

f) Homogeneity of Regression Slopes

Table 4.13
Tests of Between-Subjects Effects
Dependent Variable: PostTest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	525.29[a]	19	27.65	7.76	.000	.850
Intercept	4565.20	1	4565.20	1281.21	.000	.980
Group	134.82	1	134.82	37.84	.000	.593
PreTest	197.64	14	14.12	3.96	.001	.681
Group * PreTest	27.52	4	6.88	1.93	.135	.229
Error	92.64	26	3.56			
Total	8211.00	46				
Corrected Total	617.94	45				

a. R Squared = .850 (Adjusted R Squared = .741)

b. R Squared = .556 (Adjusted R Squared = .535)
c. Computed using alpha = .05

Homogeneity of regression slopes was suggested by similar regression lines evidenced in the scatter-plots of the dependent variables and covariates by group (reported earlier as evidence for linearity). This assumption was not confirmed by statistically significant interaction of pre-test by group, $F(4, 26) = 1.93$, $p < .005$.

Table 4.14
Tests of Between-Subjects Effects

Dependent Variable: PostTest

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	448.69[a]	2	224.34	56.10	.00	.73	113.10	1.00
Intercept	268.33	1	268.33	68.17	.00	.61	68.17	1.00
PreTest	171.10	1	171.10	43.47	.00	.50	43.47	1.00
Group	163.52	1	163.52	41.55	.00	.49	41.55	1.00
Error	169.25	43	3.94					
Total	8211.00	46						
Corrected Total	617.94	45						

a. R Squared = .726 (Adjusted R Squared = .713)

b. Computed using alpha = .05

The results of the One-Way ANCOVA suggests a statistically significant effect of the covariate, pre-test score, on the dependent variable, post-test score $F(1,43) = 43.47$, $p<.05$. More importantly, there is a statistically significant effect for teaching approach using *TinkerPlots* $F(1, 43) = 41.55$, $p<.05$, with a large effect size and strong power (partial $\eta^2_{group} = .491$, observed power = 1.000). The effect size using partial eta squared suggests that about 49.1% (of the variance in post-test scores can be accounted for teaching approach using *TinkerPlots* when controlling for pre-test score.

4.2.4 Results of Analysis for RQ 4

Research question 4, "Is there any significant different in the constructs of statistical reasoning between control and experimental group of Year Five students?" was answered by using descriptive statistics and inferential statistics analyses. The results of descriptive statistics is tabulated in Table 4.15. The data was also analysed and shows adjusted means of each construct in Table 4.16. Table 4.17 and Table 4.18 are tabulated to show the normality and followed by the Box's Test results tabulated in Table 4.19. Table 4.20 shows the Levene test and Table 4.21 shows the Multivariate test. Moreover, Graph 4.7, Graph 4.8, Graph 4.9 and Graph 4.10 show clear and simple representation of adjusted means of Describing Data, Organising Data, Representing Data and Analysing and Interpreting Data respectively. Correspondingly, One-Way MANOVA test conducted and tabulated in Table 4.22.

Table 4.15
Descriptive Statistics

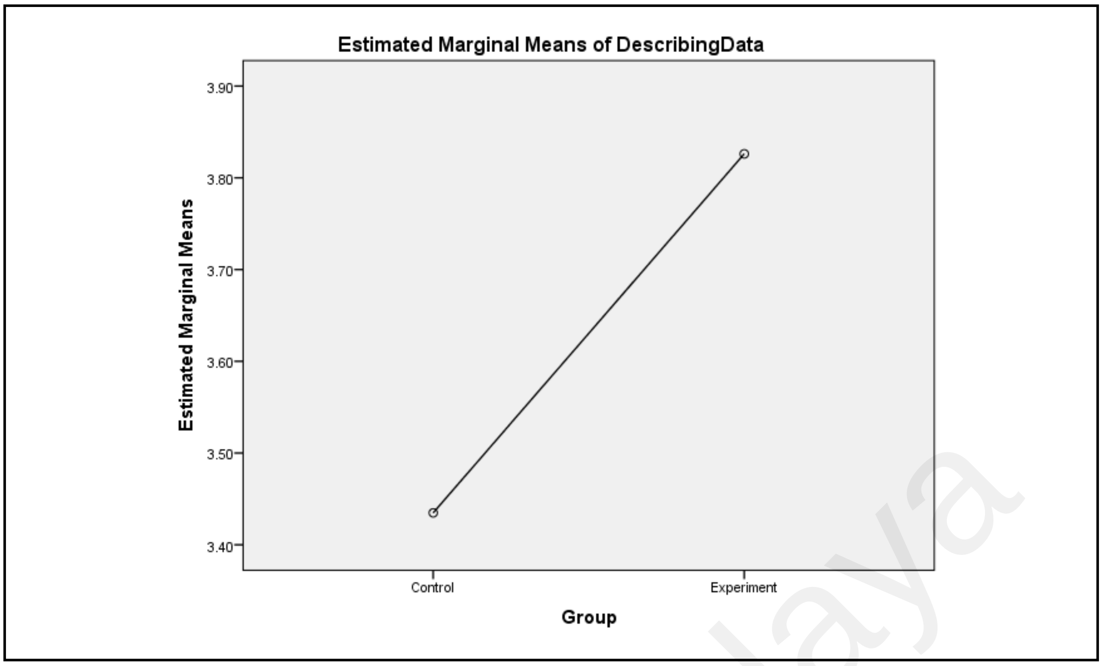
	Group	N	Std. Deviation	Mean
	Control	23	1.08	3.43
Describing Data	Experimental	23	.72	3.83
	Total	46	.93	3.63
	Control	23	1.03	2.61
Organising Data	Experimental	23	1.01	2.87
	Total	46	1.02	2.74
	Control	23	.83	2.17
Representing Data	Experimental	23	.88	3.30
	Total	46	1.02	2.74
	Control	23	1.72	2.65
Analysing and Interpreting Data	Experimental	23	1.29	5.30
	Total	46	2.01	3.98

Based on Table 4.15, descriptive statistics which include mean and standard deviation are reported for each constructs of statistical reasoning for control and experimental groups. The mean score Describing Data observed has higher mean in experimental group (*M*= 3.83, *SD*= .72) compared to control group (*M*= 3.43, *SD*= 1.08). It followed by the next construct, Organising Data has higher mean also in experimental group (*M*= 2.87, *SD*= 1.01) compared to control group (*M*= 2.61, *SD*= 1.03). The construct, Representing Data also shows higher mean in experimental group (*M*= 3.30, *SD*= .88) than control group (*M*= 2.17, *SD*= .83). Finally, Analysing and Interpreting Data shows a higher mean (*M*= 5.30, *SD*= 1.29) in experimental group compared to control group (*M*= 2.65, *SD*= 1.72).

Table 4.16
Grand Mean

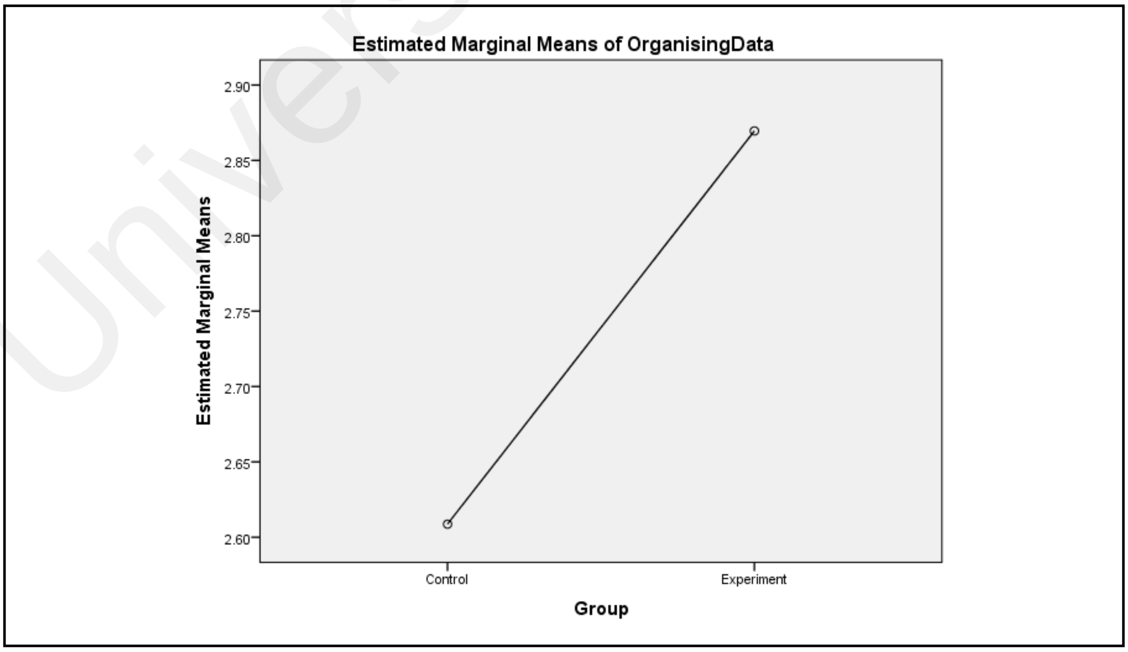
Dependent Variable	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Describing Data	3.63	.14	3.36	3.90
Organising Data	2.74	.15	2.44	3.04
Representing Data	2.74	.13	2.49	2.99
Analysing and Interpreting Data	3.98	.23	3.53	4.43

The adjusted means of each constructs of statistical reasoning are tabulated in Table 4.16, Describing Data 3.63 (*SE*= .14, 95% CI [3.36, 3.90]); Organising Data 2.74 (*SE*= .15, 95% CI [2.44, 3.04]); Representing Data 2.74 (*SE*= .13, 95% CI [2.49, 2.99]) and Analysing and Interpreting Data 3.98 (*SE*= .23, 95% CI [3.53, 4.43]).



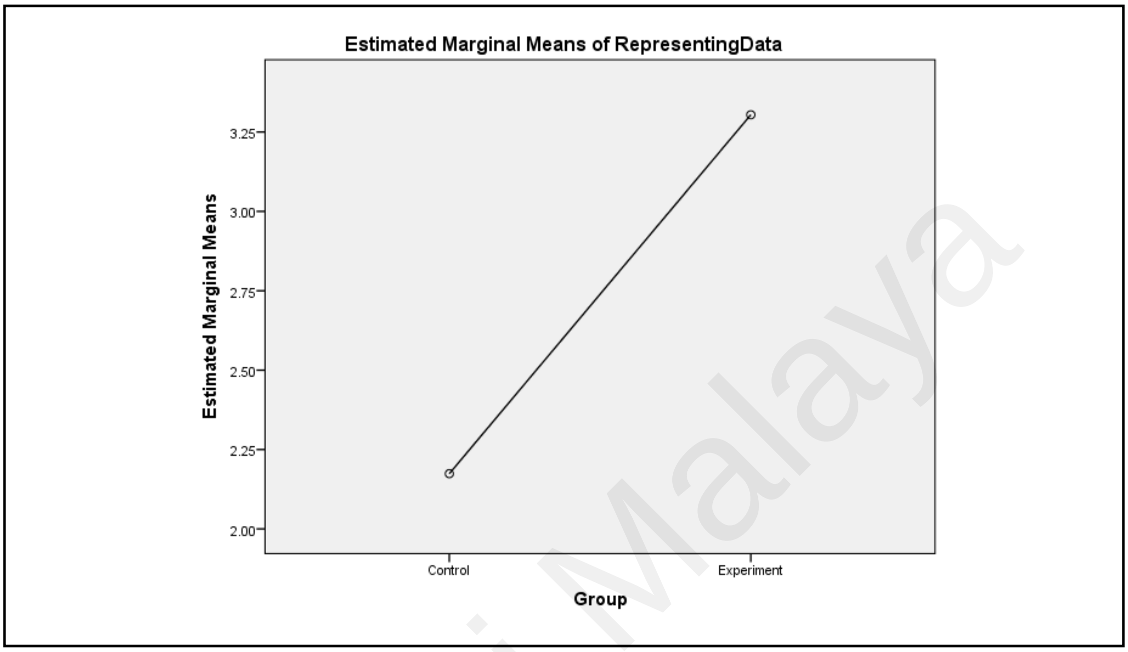
Graph 4.7: *Adjusted Means of statistical reasoning construct of Describing Data*

Graph 4.7 illustrates the estimated marginal means of experimental group in statistical reasoning constructs of Describing Data is higher than the adjusted means of control group.



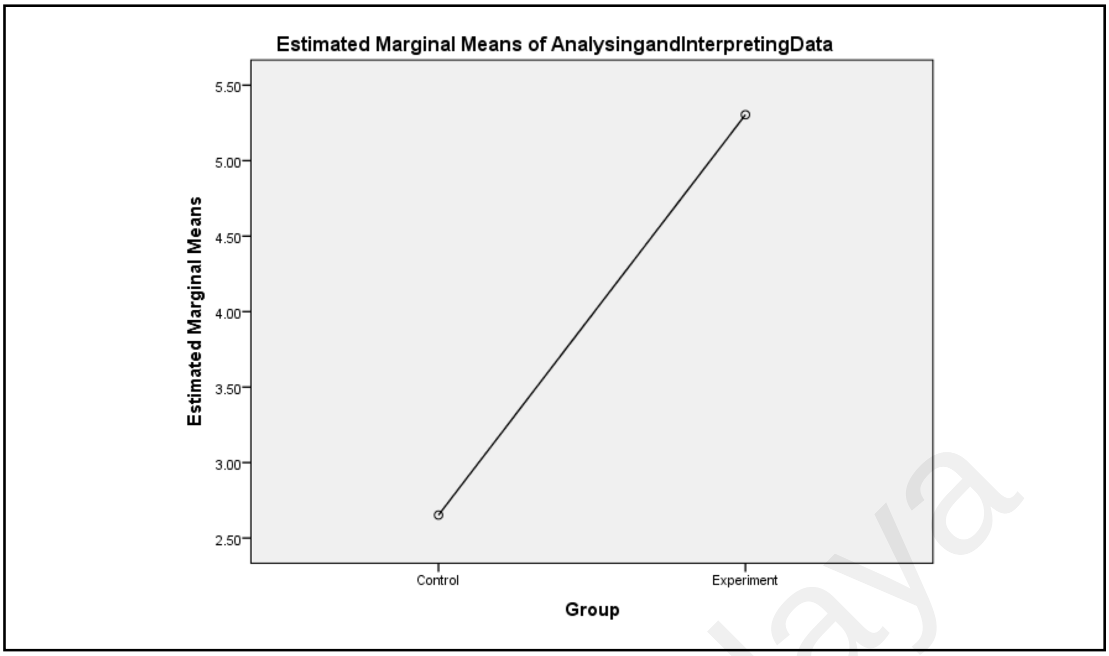
Graph 4.8: *Adjusted Means of statistical reasoning construct of Organising Data*

Graph 4.8 illustrates the estimated marginal means of experimental group in statistical reasoning constructs of Organising Data is higher than the adjusted means of control group.



Graph 4.9: *Adjusted Means of statistical reasoning construct of Representing Data*

Graph 4.9 illustrates the estimated marginal means of experimental group in statistical reasoning constructs of Representing Data is higher than the adjusted means of control group.



Graph 4.10: *Adjusted Means of statistical reasoning construct of Analysing and Interpreting Data*

Graph 4.10 illustrates the estimated marginal means of experimental group in statistical reasoning constructs of Analysing and Interpreting Data is higher than the adjusted means of control group.

A MANOVA test was conducted to determine if there is any significant difference in the constructs of statistical reasoning. A few assumptions were taken into consideration and tested which are a) independence of observations b) normality c) homogeneity of variance-covariance matrices.

a) Independence of observations

The independence of observations was met by random assignment of students to instructional method, where two different sample set of groups were chosen for control and experimental group respectively.

b) Normality

Table 4.17
Descriptives

		Statistics	Std Error
Describing Data	Skewness	-.27	.35
	Kurtosis	-.71	.69
Organising Data	Skewness	.03	.35
	Kurtosis	-1.33	.69
Representing Data	Skewness	-.36	.35
	Kurtosis	-.93	.69
Analysing and Interpreting Data	Skewness	-.21	.35
	Kurtosis	-.58	.69

Table 4.18
Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
Describing Data	.88	46	.000
Organising Data	.83	46	.000
Representing Data	.86	46	.000
Analysing and Interpreting Data	.94	46	.023

a. Lilliefors Significance Correction

The assumption of normality was tested and met via examination of the residuals. Based on Table 4.17 and Table 4.18, review of the S-W test for normality of Describing Data (SW= .88, df= 46, $p<.005$), skewness (-.27) and kurtosis (-.71) statistics suggested that normality was a reasonable assumption; review of the S-W test for normality of Organising Data (SW= .83, df= 46, $p<.005$), skewness (.03) and kurtosis (-1.33) statistics suggested that normality was a reasonable assumption; review of the S-W test for normality of Representing Data (SW= .86, df= 46, $p<.005$), skewness (-.36) and kurtosis (-.93) statistics suggested that normality was a reasonable assumption and review of the S-W test for normality of Analysing and Interpreting Data (SW= .94, df= 46, p>.005), skewness (-.21) and kurtosis (-.58) statistics suggested that normality was not a reasonable assumption.

Table 4.19
Box's Test of Equality of Covariance Matrices[a]

Box's M	10.28
F	.93
df1	10
df2	9255.78
Sig.	.51

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Group

Since normality test (Shapiro-Wilk) in Table 4.18 is violated, Box's M test was conducted to address the issue. Based on Box's M in Table 4.19, the results indicated that the equality of covariance was not statistically significant, F (10, 9255.78) = .93, p =.51. Thus, the null hypothesis is failed to be rejected so that the equality of covariance matrices has been met.

c) Homogeneity of variance-covariance matrices

Table 4.20
Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Describing Data	7.63	1	44	.01
Organising Data	.03	1	44	.86
Representing Data	.01	1	44	.95
Analysing and Interpreting Data	.81	1	44	.37

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Group

Based on Levene's test in Table 4.20, the homogeneity of variance assumption is satisfied in each and every construct. The construct Describing Data shows [F (1, 44) = .01] with an alpha level .05, p (.01) is significant. Thus, the null hypothesis is rejected. Moreover, the construct Organising Data [F (1, 44) = .86] shows the null hypothesis is rejected because it is not significant. Next construct, Representing Data [F (1, 44) = .95] also shows it is not significant, the null hypothesis is rejected. Lastly, Analysing and Interpreting Data [F (1, 44) = .37] shows not significant. Thus, the null hypothesis is rejected as well. In conclusion, the results indicate that the assumptions of homogeneity of variance is met.

Table 4.21
Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Group	*Wilks' Lambda*	.45	12.72[b]	4.00	41.00	.00	.55

a. Design: Intercept + Group
b. Exact statistic

Multivariate Tests indicated that the differences between the two groups on the combined dependent variables was statistically significant, F (4, 41.00) = 12.72, p <.05; Wilks' Lambda = .00, as table 4.21. This study at the 5% level of significance rejects the null hypothesis stating the mean of total average of the post-test score of Describing Data, Organising Data, Representing Data, and Analysing and Interpreting Data, the statistical reasoning constructs of Year 5 students was not different between the control group and experimental group after controlling for pre-test scores.

Table 4.22

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[e]
Group	Representing Data	14.70	1	14.70	20.10	.000	.314	20.10	.99
	Analysing and Interpreting Data	80.89	1	80.89	34.87	.000	.442	34.87	1.00
	Describing Data	36.96	44	.84					
	Organising Data	46.09	44	1.05					

a. R Squared = .045 (Adjusted R Squared = .024)

b. R Squared = .017 (Adjusted R Squared = -.006)

c. R Squared = .314 (Adjusted R Squared = .298)

d. R Squared = .442 (Adjusted R Squared = .429)

e. Computed using alpha = .05

One-Way multivariate analysis of variance (MANOVA) was conducted to test the hypothesis that there would be one or more mean differences between the statistical reasoning constructs (Analysing and Interpreting Data compared to Describing Data, Organising Data and Representing Data) and pre-test and post-test scores. The results of the One- Way MANOVA in Table 4.22 suggested a statistically significant effect on both Representing Data F (1, 43) = 20.10, p <.05, with a large effect size and strong power (partial $\eta 2$ = .31, observed power=1.00) and Analysing and Interpreting Data (F (1, 43) = 34.87 p <.05, with a large effect size and strong power (partial $\eta 2$ = .31, observed power=.99).

4.3 Summary

Based on the analysis in descriptive, experimental group students have scored better in post-test compared to the pre-test. The paired sample t-test also indicates there was a significant difference between the pre-test and post-test where students scored higher in post-test compared to pre-test. Next, the descriptive and inferential analysis of the second research question found out that the experimental group scored better than control group. The experimental group has higher mean than control group. Moreover, there was a significant difference in constructs of statistical reasoning which are Representing Data and Analysing and Interpreting Data. Further discussions, conclusions and implications will be carried out in the last chapter, Chapter 5.

CHAPTER FIVE

CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter discusses the important points of this research. First section of this chapter is about the summary of the study. The second section evaluates the summary of findings within the research. The third section discusses the implications based on the results analysed. The last section of the chapter is looks into recommendations for further research.

5.2 Summary of Study

The main purpose of this study is to examine the impact of *TinkerPlots* on statistical reasoning among Year Five pupils.

A quasi experiment was conducted in an international school in Selangor. 52 Year 5 students from a particular school was chosen as the sample using convenience sampling but 46 of them participated in the research. Then, the students were divided into two groups, control group and experimental group. Each group had 23 students. Both groups of students sat for the pre-test. Then, only the experimental group underwent interventions for a month and a week; where else control group studied data handling in traditional method. After the interventions, both the groups sat for post-test.

The pre-test and post-test was designed by adapting the framework from (Jones et al., 2000) The study hypothesized as following;

1) The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in control group after the traditional teaching approach.
2) The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in experimental group using TinkerPlots.
3) The mean of post-test scores of Year 5 experimental group students is higher than the control group students when controlling the pre-test score.
4) There is a mean difference in Year 5 students' constructs of statistical reasoning in post-test based on using *TinkerPlots*.

5.3 Summary of Findings

The study has answered the purpose of the research, the impact of using *TinkerPlots* in statistical reasoning by answering the research questions developed. Summary of results of hypotheses is tabulated in Table 5.1. Then, the summary of the research findings are answered in parallel to the research questions.

Table 5.1
Summary of results of hypotheses

	Hypotheses	T-value/ F-value	Results
H1	The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in control group after the traditional teaching approach.	-4.90*	Supported
H2	The mean of post-test scores is greater than the mean of the pre-test scores of Year 5 students in experimental group using *TinkerPlots*.	-13.146*	Supported

H3	There is a mean difference in Year 5 students' statistical reasoning in post-test based on using TinkerPlots, controlling for pre-test.	23.97*	Supported
H4	There is a mean difference in Year 5 students' constructs of statistical reasoning in post-test based on using TinkerPlots.	41.54*	Supported

Note: (p<.05)*

Question 1: Is there any significant improvement between pre-test and post-test of the control group Year Five students' statistical reasoning?

The research question is answered using descriptive statistics which are mean and standard deviation. The post-test has higher mean (M = 11.00, SD = 3.36) than the pre-test (M = 7.96, SD = 3.64). Correspondingly, inferential statistics is used. A paired t-test was conducted. The results shows that there was a significant difference in the scores for the control group (M = -3.04, SD = 2.98), t (22) =-4.90, p<.0005. The results obtained from the paired t-test indicated that there was an improvement in post-test compared to pre-test. Based on the obtained result, at the 5% level of significance, thus, the data provides sufficient evidence to conclude that the Year 5 students' statistical reasoning in control group improved significantly after the traditional teaching approach. The students could show improvement in their post-test, showing that, they have received the input in their class and applied it in the repeated test.

Question 2: Is there any significant improvement between pre-test and post-test of the experimental group Year Five students' statistical reasoning?

The research question is answered using descriptive statistics which are mean and standard deviation. The post-test has higher mean (M = 15.30, SD = 2.48) than the pre-test (M = 9.57, SD = 2.52). Correspondingly, inferential statistics is used. A paired t-test was conducted. The results shows that there was a significant difference in the scores for the experimental group (M = -5.57, SD = 2.09), t (22) =-13.15, p<.05. The results obtained from the paired t-test indicated that there was an improvement in post-test compared to pre-test. Based on the obtained result, at the 5% level of significance, thus, the data provide sufficient evidence to conclude that the Year 5 students' statistical reasoning in experimental group improved significantly after the interventions. The students managed to improve because they manage to construct knowledge based on the theory, "constructivism" used in the study. Moreover, they took their own responsibility of learning with teacher's guidance; predict their data as the characteristics of constructivism. Thus, the theory also plays an important role in influencing the study.

Question 3: Is there a significant difference in Year 5 students' statistical reasoning in post-test between the control and experimental control group when controlling for pre-test?

One-Way ANCOVA test was conducted for this research question. The results of the ANCOVA suggest a statistically significant effect of the covariate, pre-test score, on the dependent variable, post-test score $F_{pre\text{-}test}$ (1,43) = 43.47, p<.05. Also, there is a statistically significant effect for teaching approach using *TinkerPlots* F_{group} (1, 43) = 41.55, p<.05, with a large effect size and strong power (partial η^{2}_{group} = .491, observed power = 1.00). The results obtained from the ANCOVA test indicated that there was an improvement in experimental group compared to control group in post-test compared to pre-test controlling the pre-test. According to the obtained result, at the 5% level of significance, thus, the data provides sufficient evidence to conclude that the Year 5 students' statistical reasoning in experimental group improved significantly better than the students in control group after the interventions. The experimental group managed to perform better than control group because interventions created for the students in the group based on the theory, "constructivism". The learning was more student-centred and teacher's role was to facilitate the students in developing the knowledge through discussions and activities. Therefore, the students were able to build and link their knowledge on their own as the criteria of the theory. However, the control group used traditional method where teacher did the delivery of the knowledge and presented it. As a result, the students in the experimental group performed better than control group.

Question 4: Is there any significant different in the four constructs of statistical reasoning?

One-Way MANOVA test was conducted for this research question. $F(1, 43) = 20.10$, $p<.05$ is significant on Representing Data with a large effect size and strong power (partial $\eta 2 = .31$, observed power=1.00) and Analysing and Interpreting Data $F(1, 43) = 34.87$ $p<.05$, with a large effect size and strong power (partial $\eta 2 = .31$, observed power=.99). According to the results, at the 5% level of significance, thus the data provides sufficient evidence to conclude that the Year 5 students can enhance their statistical reasoning skills by using *TinkerPlots* software in a SRLE class. Students can perform better in statistical reasoning sub constructs such as demonstrating data; identify the different representations; judge the effectiveness of different representations; make comparisons within data; make comparisons between two data; make predictions/inference/interpretations and make conclusions. The students could perform better in mentioned sub constructs because the interventions planned was based on the theory, “constructivism” allowed them to explore the lessons and acquired the exposure in learning statistical reasoning and intensified statistical reasoning skills.

5.4 Discussions

The discussions contains two sections based on the objectives of the study. The first section discusses the impact of using TinkerPlots software in statistical reasoning, it is followed by the next section which discusses the impact of SRLE teaching approach.

5.4.1 Impact of using TinkerPlots Software in Statistical Reasoning

The difference between pre-test and post-test of the experimental group of Year Five students’ statistical reasoning was significant, suggested that using *TinkerPlot software* in data handling is effective in enhancing primary school students’ statistical reasoning compared to traditional approach. The findings of this study is consistent with a study carried out by (Ben-Zvi, 2006). The study is proven to give positive impact concerning utilization of information technology in developing students’ statistical reasoning ability. Results showed that the *TinkerPlots software* can support students’ multiplicative reasoning, aggregate reasoning, recognition of the value of large samples, and variability justification.

Past research by Pratt and Nose (2010) supports the findings too. He said in a series of studies on children's understanding about chance and distribution, these researchers examined students' emergent understanding of randomness while engaging with software especially designed to support their probabilistic reasoning. The students' understandings were then analyzed by comparing them to experts' meanings for randomness. Unlike the experts, the children shifted rapidly between four meanings for randomness. Moreover, their choices for these meanings seemed to be triggered by seemingly superficial (from the statistical point of view) aspects of the data.

Students easily get bored when they just have to learn in a classroom setting with books, whiteboard and teacher as the subject is delivered all the time. When they lose interest in a particular topic, there is a chance for them to find it difficult. They tend to learn more and discover things or knowledge on their own when they are given the opportunity to explore with their teacher's guidance. Thus, this study gives them chance to develop statistical reasoning.

Correspondingly, the results in this study also encourages the primary teachers to utilize *TinkerPlot software* while teaching and learning this particular topic, data handling. The practice of a technology tool in primary gives more learning possibilities for primary students to develop their knowledge. It provides more meaningful learning for students when they tend to explore the data on their own to make deductive meaning for what they are about to find.

In addition to, another study carried out by Fitzallen and Watson (2010) is similar to the findings of this research. They have implemented a study using the TinkerPlots software to enhance statistical reasoning of primary five and six graders.

The students were able to construct plots and utilize plots to support their thinking on the given data.

The study is also supported by Olive and Makar (2010), the focus in this perspective is on other affordances of technology, such as making statistics visual, interactive and dynamic, focusing on concepts rather than computations, and offering the opportunity to experiment with data to make it engaging for students.

The difference in Year 5 students' statistical reasoning in post-test based on using *TinkerPlot Software*, controlling the pre-test score was significant. The findings of the study is consistent with what was opinionated by (Ben-Zvi et al., 2015). They said, students need authentic, situated, and rich experiences in taking samples and learning how samples do and do not represent the population prior to formal higher studies of statistics. These experiences may include collecting data through surveys and experiments, where they learn characteristics of good samples and reasons for bad samples (e.g., bias), and creating models using simulation tools (e.g., TinkerPlots) to study the relationship between sample and population. These experiences may help students develop a deeper understanding of sampling and (informal) inference, as they repeatedly deal with taking samples, repeated samples, and simulations.

The finding is significantly different because students in experimental group had the chance to learn on their own pace which was student-centred. The students were enabled to do their survey to find out the subject they are interested in. Then, they had to analyse and interpret the data from the data collected. They were eager to check on their predictions. In other words, students were really motivated to learn and find out more when they were given task like investigation. Moreover, students preferred the visual studies in this particular topic. In conclusion, the use of a technology tool in primary does give an impact on learning.

On the other hand, the statement given in the previous research by Garfield and Ben-Zvi (2008) is proven in this study saying that, technology in statistics class enables students to have adequate time to explore, analyze and interpret data and information technology can assist students in understanding the abstract ideas of statistics. Students can display and visualize data in multiple representation form as well.

Lastly, the results is also consistent with Konold and Miller (2005), TinkerPlots allows students systematically to build their understanding of statistical representations and concepts through exploring data.

The standard of living and human evolution globally is upgrading day by day. The main platform for the changes is technology which provides them information at their fingertips giving them a chance to learn more. Looking at this perspective, technology is contributing a larger factor in learning. Students rely on technological tools as well to study. This research used TinkerPlot software and based on this research, the experimental group students used the technological tool to enhance their knowledge.

The research question shows positive results because students were excited to key in the data collected by them. They were keen to know the results from the analysis done in TinkerPlot software. They were not asked to do in a particular way but had to explore the software to understand about the data collected. Other than that, the students also shared about their analysis done in TinkerPlot with other classmates. So, the students find the topic easier and fun to learn which prompt them to make statistical reasoning.

5.4.2 Impact of SRLE Teaching Approach

The difference in Year 5 students' statistical reasoning in post-test using *TinkerPlot Software* in SRLE class, controlling the pre-test score was significant.

The findings of the research is parallel with Olani et al. (2010), stated the impact of classroom contexts on the development of students' statistical reasoning and thinking abilities and on the improvement of their attitude and beliefs require further study. Even if further empirical study is first required on this issue, it is worthy pointing out the appropriateness of using personalized and specific items when measuring components of affect behaviour.

The finding is consistent with what Gal (2002) suggested in the previous research. He mentioned understanding, interpreting, and reacting to real-world messages that contain statistical elements go beyond simply learning statistical content. He suggested that these skills are built on an interaction between several knowledge bases and supporting dispositions. Statistical literacy skills must be activated together with statistical, mathematical and general world knowledge. The students in the research tend to make statistical reasoning easily after expose to the real life situations.

The findings from this research also coincides with what Bakker (2004b) suggested, asking students to make conjectures about possible samples of data pushes them to use conceptual tools to predict the distributions, which helps them develop reasoning about samples. In these process, "what-if" questions prove to be particularly stimulating. Based on his statement, the students sampled in this research developed their statistical reasoning.

Moreover, the criteria used in the research is consistent with what was suggested by Newton et al. (2011), teachers have to understand the importance of spending time assisting students with question formulation and data collection. Furthermore, they need to be prepared to facilitate discussions with students around the expectations that promote statistical reasoning and begin to see and state expectations as a minimum requirement. That is, teachers that expect students only to "do" the process components and that lack attention to statistical reasoning and/or the statistical process should be encouraged to enhance their instruction to include these critical components of statistical literacy. Thus, the criteria used in the research enhanced students' statistical reasoning. It's proven in this research because there was a significant difference in Year 5 students' statistical reasoning in post-test between experimental group and control group by controlling the pre-test score. The experimental group students who underwent interventions in parallel with SRLE class did perform better than the students in control group.

Previous research also stated students in Malaysia did not perform well in TIMMS and PISA. According to Gil et al. (2008a), inquiry is the process in which students solve problems, pose questions, construct solutions and explain their reasoning (e.g., (NCTM: National Council of Teachers of Mathematics, 2000, PISA: Programme for International Students Assessment, 2003). They mentioned that one obvious way to bring students into the processes of inquiry learning is by offering them environments and tasks that allow them to carry out the processes and help them build a personal knowledge that they can use and explain what they learn. It was practiced in this research and findings show that the students could make statistical reasoning better after their interventions, because there was a significant difference between pre-test and post-test of the experimental group Year Five students' statistical reasoning.

The findings from this research stimulates the primary teachers on the pattern of teaching this topic. Teacher played an important role in this research. The teacher was mostly guiding the students in their research instead of delivering the entire process, predictions and answers. Throughout the interventions the teacher posed a lot of questions to make the students think and come out with their own answers. The teacher and the students had many discussions instead of traditional approach where teacher explains all the time and the students have to listen for the entire lesson. To sum up, when students are involved in discussions, they tend to concentrate on the topic of discussion and come up with a lot of great ideas. Later on, they have confidence in making statistical reasoning, and chances are lower in them having difficulty to come up with statistical reasoning.

Students will be excited and inspired to learn and find out on their own. The study suggested that learning SRLE class is impactful compared to traditional approach. They were asked to do deal with real data as one of the criteria of SRLE. The students made effort to spend time to look for the data and made them to go around to look and choose the correct sample. The students were enthusiastic to find information which was around them instead of referring to books. There are higher chances for students to perform better in TIMSS and PISA because inquiry based study is suggested to achieve good results in TIMMS and PISA.

5.5 Implications

5.5.1 Implications for Instructions

In this study, interventions were carried out for the experimental group students. Ten lessons were planned to coincide with Statistical Reasoning Learning Environment (SRLE) class which used *TinkerPlot* software. Results obtained, as explained in Chapter 4, students performed better after using *TinkerPlot* software. From the results obtained, some implications can be deduced for curriculum developers, teachers, mathematics educators, mathematics coordinators, students and mathematics curriculum researchers.

In relation to significance of the study in chapter 1, Mathematics book writer or curriculum developer will get an idea of including the part of technology which is using *TinkerPlot software* in the mathematics text book. As per results obtained in this study, using *TinkerPlot* software to learn data handling will enhance students' interest in learning. Thus, in the particular chapter of data handling, the curriculum developer can consider adding the way of using technology tool in this topic. Attaching the instructions of using *TinkerPlot* in data handling will guide the students to follow on their own and expand their understanding by exploring.

On the other hand, the curriculum developer can take into account of adding the instructions on how to use *TinkerPlot* software in data handling in the teacher's mathematics guidance book. Teacher may get possible ideas after referring to the book. They can get some knowledge on how to use the software and practice using it before implementing it in the class. So, the teachers will not be outdated if the part is included.

Moreover, teachers or mathematics educators will benefit from this study too. The study has proven that the students have performed better after using *TinkerPlot* software in a SRLE class. Thus, the teachers can implement teaching data handling using *TinkerPlot* software in a SRLE class. To improve themselves in teaching and of their professional development, teachers can consider using the teaching tool in their class. This is because many educational organizations give importance to teachers' development especially in exceling technology to implement in teaching and learning. In this case, the mathematics coordinator or mathematics leader in the school can consider providing training to teachers using TinkerPlot software. The training can help them for their professional development as well.

In addition to that, teachers also can use SRLE class model to teach data handling. The teaching method highly encourages students to be involved in the lesson. In like manner, it is proven in this study where students showed interest in learning in a SRLE class. The teacher's role becomes really important to encourage and facilitate students to develop their knowledge.

Other than that, learning data handling using TinkerPlot in a SRLE class will benefit the students. Firstly, the students will be motivated to use TinkerPlot to explore the data collected and test the conjectures made by them as well. Secondly, they tend to get involved in the lesson conducted where there will mainly be discussions as per SRLE criteria. Thirdly, the students are trained to explain their reasoning and justify the conclusions in the SRLE class.

Finally, this research will benefit the mathematics curriculum researchers as well. This study opens up some further ideas to do their research. This research gives them a thought to further their research to fill up the research gap.

5.5.2 Recommendations for Further Study

Based on the findings of this study and the implications, the research needs to be carried out further by integrating technology tool into teaching and learning mathematics in future. The recommendations for further research are made based on this research. Firstly, simple random sampling can be used in future research. This is to fulfil the requirement, the research population and sample means are normally distributed where it is considered in t-tests: one sample t-test, independent t-test, paired sample t-test and ANOVA tests. The requirement is needed to be taken into account because the sampling is chosen in a fair way and everyone is given equal chance for being selected and unbiased of representation of group. Thus, the results obtained from the sample described is robust as the sample is generalized to the population.

Secondly, the research can focus on sociocultural theory. The sociocultural perspective has a profound implication concerning the important role of enculturation processes in learning. Briefly stated, the process of enculturation refers to entering a community or a practice and picking up their points of view (Garfield & Ben-Zvi, 2008). The theory can focus on students' cognitive and cultural practice. This is especially the case with regard to statistical thinking, with its own culture, values, and belief systems, and habits of questioning, representing, concluding, and communicating (Garfield & Ben-Zvi, 2008).

Moreover, a mixed mode study can be considered in future. This study only focused on quantitative studies. It can be expanded with qualitative too next time. Researcher can look forward to conduct interviews to find out their understanding level of statistical reasoning, opinions and motivations. It helps to look into the problem faced by the students with data handling in detail.

In addition to that, this research studied students' statistical reasoning. In future, research can be conducted on data handling focusing on students' statistical concepts such as mean and median for upper primary students; variance, standard deviation and normal distribution for secondary school students.

The encouraging result of this study suggested that using TinkerPlot software could be a beneficial tool to learn data handling in primary schools. Saying that, this research only focused on two intact mixed ability groups from a primary international school. In future, further researches can be conducted in national primary schools as well such as national schools (sekolah kebangsaan) and national-type schools (sekolah jenis kebangsaan) which are Tamil and Chinese schools. Students from diverse background of the schools can give various results and this can be compared and explained further.

The research questions can be extended to find about students' attitudes towards learning statistical reasoning. The research only focused on the performance of students' statistical reasoning. The research did not find about students' attitudes of learning statistics (data handling) using TinkerPlots. The research can also find out students' attitudes towards learning statistics in a SRLE class in future.

5.6 Contributions of Study

This study has important contributions for improving students' statistical reasoning skills. The results of this study has proven that teaching and learning in a SRLE class using *TinkerPlots* software could improve students' statistical reasoning skills. Students are able to extract data, recognise general features of graphical representations, show awareness to the displayed attributes of graphical representation, categorise data, organise data, demonstrate data, identify the different representations, judges the effectiveness of different representations, make comparisons within data, make comparisons between two data, make predictions or inference interpretations and make conclusion. Adding up to that, students in experimental group could perform better than control group. They are exposed to the real world situation to explore and learn. Thus, the study can encourage teaching and learning statistics in a SRLE class using suitable technology tools in future. The positive results in this study can be a good example, for the teachers or educators, to implement in their statistics class which can produce more students who can make good statistical reasoning in future.

5.7 Conclusion of the Study

According to results obtained and comprehensive discussion, major findings emerged from this study. The independent variable of this study is the teaching approach which included traditional teaching and using TinkerPlots software to coincide with SRLE model. The Year 5 students from experimental group using teaching approach of SRLE model and TinkerPlots significantly improved. It shows that, the students show improvement in statistical reasoning constructs which are describing data, organising data, representing data, analysing and interpreting data. The students showed positive learning using TinkerPlots software. They managed to handle and explore original data collected; they could test the conjectures as well. The SRLE class conditions gave them the opportunity to improve their statistical reasoning skills. Thus, the students were able to solve reasoning based questions easily.

References

Abelson, R. (1995). Statistics as principled argument. *Hillsdale, NJ: Lawrence Erlbaum Associates*.

Aberson, C. L., Berger, D. E., Healy, M. R., Kyle, D., & Romero, V. L. (2000). Evaluation of an interactive tutorial for teaching the Central Limit Theorem. *Teaching of Psychology, x*(27), 289-291.

Alldredge, J. R., & Som, N. A. (2002). Comparison of multimedia educational materials used in an introductory statistical methods course. Proceedings of the sixth international conference on the teaching of statistics (ICOTS-6), Cape Town, South Africa.

Arshad, H., Yaacob, Y., Yusof, M. M., & Latih, R. (2000). *Implications of introducing Technology in Mathematics Education.* Paper presented at the Proceedings of the International Conference on Teaching and Learning.

Bakker, A. (2004a). *Design Research in statistics education: On symbolizing and computer tools (Doctoral Dissertation).* Utrecht, The Netherlands, CD Beta Press.

Bakker, A. (2004b). Reasoning about shape as a pattern i n variability *Statistics Education Research Journal, 3*(2), 64-83.

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

Bandura, A. (1997). *Self-efficacy: The exercise of control.*, New York: Freeman.

Bastürk, R. (2005). The effectiveness of computer-assisted instruction in teaching introductory statistics. *Journal of Educational Technology & Society, 8*, 170-178.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning, 2*(1-2), 127-155.

Ben-Zvi, D. (2006). *Scaffolding students' informal inference and argumentation.* Paper presented at the Proceedings of the Seventh International Conference on Teaching of Statistics, Salvador, Bahia, Brazil, 2–7 July, 2006. Voorburg, The Netherlands: International Statistical Institute.

Ben-Zvi, D., Bakker, A., Aridor, K., & Makar, K. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM Mathematics Education*, 13. doi:10.1007/s11858-012-0420-3

Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics, 88*(3), 291-303. doi:10.1007/s10649-015-9593-3

Ben-Zvi, D., & Garfield, J. (2004). The challenge of developing statistical literacy, reasoning, and thinking.

Biehler, R. (2003). Interrelated learning and working environments for supporting the use of computer tools in introductory classes.

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2012). Technology for Enhancing Statistical Reasoning at the School Level. 643-689. doi:10.1007/978-1-4614-4684-2_21

Bonawitz, E., Fischer, A., & Schulz, L. (2012). Teaching 3.5-Year-Olds to Revise Their Beliefs Given Ambiguous Evidence. *Journal of Cognition and Development, 13*(2), 266-280. doi:10.1080/15248372.2011.577701

Boumová, B. V. (2008). *Traditional vs. Modern Teaching Methods: Advantages and Disadvantages of Each.* Masaryk University.

Bransford, J., Brown, A. L., & Cocking, R. R. (2000). How people learn: Brain, mind, experience, and school.

Bredo, E. (1997). *The social construction of learning*. New York: Academic Press.

Brown, A. L. (1975). The development ofmemory: Knowing, knowing about knowing, and knowing how to know. *Advances in child development and behavior, 10*.

Bruning, R. H., Schraw, G. J., Norby, M. M., & Ronning, R. R. (2004). *Cognitive psychology and instruction* Upper Saddle River, NJ: Merrill/Prentice Hall.

Burrill, G., & Camden. (2005). *Curricular Development in Statistics Education, International Association for Statistical Education 2004 Roundtable.*

CA. (2000). Key Curriculum Press Technologies.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research on teaching*. Boston: Houghton Mifflin.

Chan, S. W., & Ismail, Z. (2012). The Role of Information Technology in Developing Students' Statistical Reasoning. *Procedia - Social and Behavioral Sciences, 46*, 3660-3664. doi:10.1016/j.sbspro.2012.06.123

Chan, S. W., & Ismail, Z. (2014a). Developing Statistical Reasoning Assessment Instrument for High School Students in Descriptive Statistics. *Procedia - Social and Behavioral Sciences, 116*, 4338-4343. doi:10.1016/j.sbspro.2014.01.943

Chan, S. W., & Ismail, Z. (2014b). A Technology- Based Statistical Reasoning Assessment Tool In Descriptive Statistics for Secondary School Students. *The Turkish Online Journal of Educational Technology, 13*(1), 18.

Chan, S. W., Ismai, Z., & Sumintono, B. (2015). The Impact of Statistical Reasoning Learning Environment : A Rasch Analysis. American Scientific Publishers, 21(5).

Chance, B. L., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. *The challenge of developing statistical literacy, reasoning, and thinking* 295-323.

Clements, D. H. (2000). From exercises and tasks to problems and projects: Unique contributions of computers to innovative mathematics education. *Journal of Mathematical Behavior, 19*, 9-47.

Cobb, G., & Moore, D. . (1997). Mathematics, statistics and teaching. *American Mathematical Monthly, 104*(9), 801-823.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*(7), 13-20.

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

Cobb, P., & Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. . *Educational Researcher, 28*(2), 4-15.

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. *The challenge of developing statistical literacy, reasoning, and thinking*, 375-396.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, N.J.: Lawrence Erlbaum.

Croxton. F.E, Cowden. D.J, & Klein. S. (1968). *Applied General Statistics* (Third Edition ed. Vol. 63). Prentice-Hall, Inc., Englewood Cliffs: Taylor & Francis.

Curcio, F. R., Artz, A. F. I., Gal, I., & Garfield, J. B. (1997). Assessing students' statistical problem solving behaviors in small-group setting. *The assessment challenge in statistics education*, 123-138.

delMas, R. C. (2002). Statistical literacy, reasoning, and learning: A commentary. *Journal of Statistics Education, 10*(3).

delMas, R. C. (2005). Activities to promote an understanding of variability.

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

DfEE. (1999). *The National Numeracy Strategy: Framework for Teaching Mathematics from Reception to Year 6*. London: DfEE.

Erickson, T. (2006, July, 2006). *Using simulation to learn about inference. In A.Rossman & B. Chance (Eds.), Working cooperatively in statistics education.* . Paper presented at the Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Bahia, Brazil.

Ernie, K. (1996). Technology reviews: DataScope and ProbSim *Mathematics Teacher, 89*, 359-360.

Finzer, W. (2001). Fathom dynamic statistics™ software. Key Curriculum Press.

Firestone, W. A. (1987). Meaning in Method: The Rhetoric of Quantitative and Qualitative Research. *Educational Researcher*.

Fitzallen, N. (2007). Evaluating data analysis software: The case of TinkerPlots. *Australian Primary Mathematics Classroom, 12*(1), 23-28.

Fitzallen, N., & Watson, J. (2010). *Developing Statistical Reasoning facilated by TinkerPlots*, Ljubljana, Slovenia.

Flavell, J. H. (1976). Metacognitive aspects of problem-solving. *The nature of intelligence, Hillsdale, NJ: Erlbaum*

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124-159.

GAISE. (2005a). Guidelines for assessment and instruction in statisctics education (GAISE) college report. *The American Statistical Education (ASA)*.

GAISE. (2005b). Guidelines for assessment and instruction in statisctics education (GAISE) report: A curriculum framework for PreK-12 statistics education *The American Statistical Education (ASA)*.

Gal. (2002). Developing statistical literacy: Towards implementing change. *Statistical Review, 70*(1), 46-51.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework Journal of Sttaistics Education,. *2*(2).

Gal, I., & Garfield, J. (1997). The Assessment Challenge in Statistics Education.

Garfield, J. (1995). How Students Learn Statistics. *InternufionnlSkUisfical Review, 63*(1), 25-34.

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

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.

Garfield, J., & Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning*: Springer.

Garfield, J., & Ben-Zvi, D. (2009). Helping Students Develop Statistical Reasoning: Implementing a Statistical Reasoning Learning Environment. *Teaching Statistics, 31*(3), 6.

Garfield, J., & Burrill, G. (1997). *Research on the role of technology in teaching and learning statistics.* Paper presented at the Proceedings of the IASE roundtable conference, University of Granada, Spain.

Garfield, J., & Chance, B. (2000). Assessment in Statistics Education: Issues and Challenges. *Mathematical Thinking and Learning*, 28.

Garfield, J., & Gal, I. (1999). *Teaching and assessing statistical reasoning*. Reston, VA: N.C.T.M.

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2014). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics, 88*(3), 327-342. doi:10.1007/s10649-014-9541-7

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

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50*, 24-37.

Gil, E., & Ben-Zvi, D. (2011). Explanations and context in the emergence of students' informal inferential reasoning. *Mathematical Thinking and Learning,, 13*(1-2), 87-108.

Gil, E., Ben-Zvi, D., & Apel, N. (2008). Creativity in Learning to Reason Informally About Statistical Inference in Primary School.

Giraud, G. (1997). Cooperative learning and statistics instruction. *Journal of Statistics Education, 5*(3).

Gribbons, B., & Herman, J. (1997). Practical Assessment, Research & Evaluation. *5*(14).

Gundlach, E., Andrew, K., Richards, R., Nelson, D., & Levesque-Bristol, C. (2015). A Comparison of Student Attitudes, Statistical Reasoning, Performance, and Perceptions for Web-augmented Traditional, Fully Online, and Flipped Sections of a Statistical Literacy Class. *Journal of Statistics Education, 23*(1), 33.

Hawkins, A. S. (1997). Myth-conceptions. *Research on the role of technology in teaching and learning statistics*, 1-14.

Henriques, A., & Oliveira, H. (2015). Students' informal inference when exploring a statistical investigation. Paper presented at the Ninth Congress of the European Society for Research in Mathematics Education, Prague, Czech Republic.

Jack, R. F., Norman, E. W., & Helen, H. H. (2015). *How to Design and Evaluate Research in Education* (Ninth Edition ed.). New York: McGraw Hill Education.

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

Kadijevich, D. J. (2002). *Four Critical Issues of Applying Educational Technology Standards to Professional Development of Mathematics Teachers*. Paper presented at the Proceedings of the 2nd International Conference on the Teaching of Mathematics at the undergraduate level, University of Crete.

Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education, 3*(2).

Khairiree, K., & Kurusatian, P. (2009). Enhancing Students' Understanding Statistics with TinkerPlots: Problem-Based Learning Approach. 10.

Konold, C., & Miller, C. D. (2005). TinkerPlots: Dynamic Data Exploration.

Lane, D. M., & Peres, S. C. (2006). Interactive simulations in the teaching of statistics: Promise and pitfalls.

Lane, D. M., & Tang, Z. (2000). Effectiveness of simulation training on transfer of statistical concepts. *Journal of Educational Computing Research, 22*(4), 383-396.

Lavigne, N. C., & Lajoie, S. P. (2007). Statistical reasoning of middle school children engaged in survey inquiry. *Contemporary Educational Psychology, 32*(4), 630-666. doi:10.1016/j.cedpsych.2006.09.001

Lee, V. R., & Thomas, J. M. (2011). Integrating physical activity data technologies into elementary school classrooms. *Educational Technology Research and Development, 59*(6), 865-884. doi:10.1007/s11423-011-9210-9

Levin, J., & Waugh, M. (1998). Teaching Teleapprenticeships: Frameworks for integrating technology into teacher education. *Interactive Learning Environments, 6*(1-2), 39-58.

Lovett, M. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. *Cognitive and instruction: Twenty-five years of progress*, 347-384.

Lunsford, M. L., Rowell, G. H., & Goodson-Espy, T. (2006). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. *Journal of Statistics Education, 14*(3).

Magel, R. C. (1998). Using cooperative learning in a large introductory statistics class. *Journal of Statistics Education, 6*(3).

Makar, K., & Ben-Zvi, D. (2011). The role of context in developing reasoning about informal statistical inference. *Special Issue of Mathematical Thinking and Learning, 13*(1&2).

Makar, K., & Confrey, J. (2005). Using distributions as statistical evidence in well-structured and ill-structured problems.

Malaysia Education Blueprint 2013-2025. (2012).

Meletiou, M., & Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistics Education Research Journal, 1*(2), 22-37.

Mills, J. D. (2004). Learning abstract statistics concepts using simulation. *Educational Research Quarterly, 28*(4), 18-33.

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

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

Moutinho, L., & Hutcheson, G. (2011). *The SAGE Dictionary of Quantitative Management Research*. London: SAGE Publications Ltd.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *Timss 2011 International Results in Mathematics* (pp. 520).

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*.

National Council of Teachers of Mathematics (2016). *Data Analysis and Probability, Principles and standards for school mathematics*.

Newton, J., Dietiker, L., & Horvath, A. (2011). *Teaching Statistics in School Mathematics- Challenges for Teaching and Teacher Education* (International Commission on Mathematical Instruction ed. Vol. 14).

Olani, A., Hoekstra, R., Harskamp, E., & Werf, G. v. d. (2010). Fostering Students' Statistical Reasoning, Self-Eficacy, and Attitudes : Findings From A Comprehensively Reformed Undergraduate Statistics Course. *ICOTS8*.

Olive, J., & Makar, K. (2010). Mathematical knowledge and practices resulting from access to digital technologies. *New York, NY: Springer*, 133-177. doi:10.1007/978-1-4419-0146-0_8 .

Piaget, J. (1978). Success and Understanding. *Cambridge, MA: Harvard University Press*.

Pratt, D., & Noss, R. . (2010). Designing for mathematical abstraction. *International Journal of Computers for Mathematical Learning, 15*(2), 81-97.

Pratt, D., & Ainley, J. (2008). Informal inferential reasoning. *Special Issue of the Statistics Education Research Journal, 7*(2), 3-4.

Razak, N. A., & Embi., M. A. (2001). *Meeting the technology challenge: The resistance factors in using information technology in technical schools.* Paper presented at the Proceedings of the International Conference on Technology and Vocational–Technical Education:Globalization and Future Trends.

Ritchie, D. (1996). The administrative role in integration of technology. *The administrative role in integration of technology, 80*(582), 42-52.

Rubin, A. (2007). Much Has Changed; Little Has Changed: Revisiting the Role of Technology in Statistics Education. *Technology Innovations in Statistics Education.*

Rubin, A., Hammerman, J. K., & Konold, C. (2006). Exploring informal inference with interactive visualization software.

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*, 359-382.

Shaffer, D. W., & Kaput, J. J. (1999). Mathematics and virtual culture: An evolutionary perspective on technology and mathematics education. *Educational Studies in Mathematics, 37*(2), 97-119.

Simon, M. (2011). Assumptions, Limitations and Delimitations. Retrieved from

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*(2), 114-145.

Springer, K. (2010). *Educational Research*. John Wiley&Sons,Inc: Jay O' Callaghan.

Symanzik, J., & Vukasinovic, N. (2002). *Teaching statistics with electronic textbooks.* Paper presented at the COMPSTAT 2002: Proceedings in computational statistics Heidelberg: Physica-Verlag.

Symanzik, J., & Vukasinovic, N. (2003). Comparative review of ActivStats, CyberStats, and MM*Stat. MSOR (Maths, Stats & Operational Research) Connections. *3, 1*, 37-42.

Symanzik, J., & Vukasinovic, N. (2006). Teaching an introductory statistics course with CyberStats, an electronic textbook. Journal of Statistics Education. *14, 1*.

Taylor, S. J., & Bogdan, R (1984). Qualitative research methods. The search for meanings (2nd edition).

U.S. Congress, Office of Technology Assessment (1995). *Teachers and technology: Making the connection. OTA-EHR-616. Washington DC: U.S. Government Printing Office.*

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician, 57*(2), 74-79.

Victorian Curriculum and Assessment Authority. (2005). *Victorian Essential Learning Standards-Discipline based Learning Strand: Mathematics.*

Von Glasersfeld, E. (1987). *Learning as a constructive activity. In Problem of representation in the teaching and learning of mathematics. C. Janvier*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Vygotsky, L. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Ward, B. (2004). The best of bothworlds: a hybrid statistics course. *Journal of Statistics Education, 12*(3), 74-79.

Watson, J. (2008). Eye colour and reaction time: An opportunity for critical statistical reasoning. *Australian Mathematics Teacher, 64*(3), 30-40.

Watson, J., & Wright, S. (2008). Building informal inference with TinkerPlots in a measurement context. *Australian Mathematics Teacher, 64*(4), 31-40.

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

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

Yuen, A. H. K., & Ma, W. K. (2001). *Teachers' Computer Attitudes: Factors influencing the Instructional Use of Computers.* Paper presented at the Proceedings of the International Conference on Computers in Education, Korea, 7p.

Zakaria, E., Daud, M. Y., & Nordin, N. M. (2007). Technology in Teaching and Learning Mathematics. *Trend in Teaching and Learning Mathematics*, 1-14.