JOURNAL OF RESEARCH IN SCIENCE, MATHEMATICS AND TECHNOLOGY EDUCATION
DOI: 10.31756/jrsmte.321

Impact of Fathom on Statistical Reasoning among Upper Secondary Students

Nanteni Ganesan

Kwan Eu Leong

University of Malaya, Malaysia

Abstract: The teaching and learning of statistical reasoning is becoming challenging due to the change in the perspective emphasizing on the deeper understanding rather than basic statistics computations. As suggested by researchers, implementing technologies able to develop student interest in the topics leads to deeper understanding. Hence, this study used dynamic software, Fathom for teaching statistical reasoning. The purpose of this study is to examine the statistical reasoning understanding among upper secondary students after using dynamic software, Fathom. The sample consists of seventy-two students randomly assigned to control and experimental groups. The experimental group underwent an intervention where they learnt statistical reasoning using Fathom while the control group learnt statistical reasoning using traditional learning method not involving Fathom. Statistical Reasoning Assessment (SRA) was used in this study as the instrument for measuring statistical reasoning. The research hypothesis data were analyzed using MANCOVA test. The findings showed a significant difference across four statistical reasoning constructs namely Describing Data, Organizing Data, Representing Data and Analyzing and Interpreting Data between students in the control and experimental groups. Furthermore, the results of the analysis emphasized that the students who learned statistical reasoning using Fathom performed better than students in the control group. In brief, the upper secondary students' statistical reasoning enhanced after implementing Fathom.

Keywords: *Data; Fathom; Graphs; Statistics Constructs; Statistical Reasoning; Technology*.

Introduction

Nowadays, a lot of information can be expressed in the form of accumulated data. The data are transformed into different graphical representations for example bar graphs or histograms (Meletiou & Lee, 2002). Direct visualization of numerical data involved in production, pollution, alimony and so forth become a concern in today's society (Bruno & Espinel, 2009). Understanding statistical graphs with the ability to read and interpret graphs is part of learning in every community. Math-related learning statistics focused on data distribution. Good understanding of data distribution in graphs requires mastery of basic concepts of statistical learning. This is because statistics is a subset of mathematics that specifically teaches us how to collect, organize the data, how to analyze the data and teach us ways to interpret and to present the data analysis to others (Sorge & Schau, 2002). Moreover, statistics as a subject is also important for future career and job performance. It can help enhance one's analytical skills, thus increasing job marketability.

According to Hall and Heyde (2014), the data analysis and interpretations done by statisticians can help to explain a complex business environment and measure company performance. Companies also can keep track of their business inventories, employee attendance and their productivity daily hence enabling better management decisions. Besides, a company can use statistical data to compete with other firms by analyzing the data from both companies. By gathering customer feedback for statistical analysis, companies can also improve product or customer service satisfaction. Eventually, learning statistics is not based on computations only but reasoning and thinking skills are major components in describing, evaluating and analyzing data sets. This study helps student decision-making in career and for the future through learning statistical reasoning.

Statistical reasoning is the method people use to reason through quantifiable contemplations and understanding factual data. Statistical reasoning could include associating one idea to another, for example mean and boxplot. Reasoning implies understanding and having the capacity to clarify factual procedures, and having the capacity to interpret measurable results (Garfield, 2002). Statistical reasoning is seen as reasoning the mental representations and associations that students have with statistical ideas. During the 1990s there was great demand for basic tutoring to acknowledge statistics education focused on thinking and reasoning. One of the widespread disagreements developed was that traditional techniques of teaching statistics never lead students to reason statistically.

At the point when students utilize the information to make inferences, they are utilizing their thinking on the grounds, as indicated by Galotti (2008), that thinking is the psychological procedure that changes given data to achieve conclusions. Students should figure out how to utilize their thinking when they make determinations from the information. Using insights, students have to figure out how to present conclusions from the information in some structure and graphically. Furthermore, they should figure out how to analyze the information. They will utilize the data implanted in these representations and synopses to make inference about the information.

The statistics teaching at secondary school level deserves attention. Through an interesting approach, lessons can be made meaningful with active discussion; it is expected that learning will be more meaningful through statistical reasoning. According to Accrombessy (2006), in this era, the rapid development of computer science and new information and communication technologies, it is useful to introduce the concept of instructional software for teaching statistical reasoning. Therefore, students can change their behavior toward statistics and participate actively in the course.

Literature Review

Many studies show that students encounter problems in describing, analyzing and interpreting the histogram. DelMas, Garfield, and Ooms (2005) noted that, first, the students are confused by the terms "horizontal" and "vertical" which led to difficulties in plotting or interpreting data. Second, students are confused between histograms and bar graphs leading to confusion over how an observation on the vertical axis reflects the values of variables. Finally, students are unable to read information from histograms correctly, leading to wrong responses when answering questions about the value of a specific frequency.

Many statistical ideas and rules are complex, which create difficulties for students in understanding the terms and procedures. Many students are struggling with basic mathematics (such as fractions, decimals and algebraic formulas), and this interferes with learning the related statistical content.

The context in numerous statistical problems may mislead the students, causing them to rely on their experiences and often-faulty perceptions to produce an answer, rather than select an appropriate statistical procedure. Students link statistics with mathematics and expect the focus to be on numbers, computations, formulas and one right answer. They are uncomfortable with the messiness of data; the different possible interpretations based on different assumptions, and the extensive use of data interpretation and communication skills.

Heavily lecturer-based instructions that focus on computations and discrete methods fail to teach statistical reasoning adequately hence leaving students unprepared for today's data-driven world. The statistical reform movement argued that it is essential that teachers of statistics focus on teaching underlying process or reasoning skills (GAISE, 2005). Unfortunately, the teaching emphasizes more on computational techniques; students are unable to see the big picture or develop reasoning skills. They are unable to perform well in statistical reasoning because of several other factors such as lack of attention, lazy to think, memory overload or inability or misconception in statistics (Saldanha & Thompson, 2002). According to Pratt and Ainley (2008), reasoning in statistics has become a focus of research. This is due to problems occurring when students are unable to make interpretations of statistical output. Several studies have given evidence supporting that students misinterpret and are unable to give reasons for the result or answers obtained from data.

Students need to develop an understanding of the underlying processes involved in statistics and learn to ask questions that challenge their reasoning about the processes. According to Galotti (2008), when students use data to make conclusions, they are using their reasoning skills. Reasoning is the cognitive process that transforms given information in order to reach conclusions. Students need to learn to use their reasoning when they draw conclusions from the data. The dynamic educational software tool, Fathom, enables teachers and students to use, modify, and develop embedded micro worlds for themselves. In such micro worlds, the students can rapidly represent the information in an assortment of diagrams, including bar chart, function plots, scatter plots, histograms and use interactive explorative features (Andreas, 2014). Using the software, they will learn to represent the data in tabular form and graphically; then they learn to summarize the data. Students use the information embedded in these representations and summaries to draw conclusion about the data using their reasoning skills when they draw conclusions from the data. Moreover, Fathom provides students with the tools to build simulations that explain concepts of probability and statistics. In Fathom, students can plot values and functions on top of bivariate information and shift them powerfully with sliders to demonstrate the impacts of variables. This allows students to develop an understanding of abstract concepts and the interrelationships between concepts and enhance their reasoning (Friel, 2007; Garfield, Chance, & Snell, 2000). Teacher and students can focus on interpretation of outcomes and understanding concepts rather than on spending more hours in solving basic computational problems.

This is supported by Meletiou and Stylianou (2003) who developed a course which has its main component a technological tool, Fathom. This study was designed to investigate the effects of a technology-based course on student understanding of graphical representations of data. Specifically, they examined how technology affected students' perception of data presented graphically and their approaches to problems involving a strong graphical element. The findings showed that technology integration in the classroom brought about important changes in students' ways of learning statistics. Moreover, using Fathom as a learning tool had increased students' interest in actively pursuing problems involving a difficult graphical element.

Jones et al. (2000) mentioned four constructs in their studies: describing data, organizing data, representing data, and analyzing and interpreting data which provides specific descriptors of students' reasoning at each level. Their model is used in this study as guidance to prepare and evaluate statistical reasoning assessment. The models of development in statistical reasoning can be helpful in assessing and monitoring students' performances over time, as well as in evaluating the effectiveness of classroom instruction. By assessing and observing changes in students' reasoning according to the model, researchers are able to identify weaknesses in the teaching methodology and technology, and after refinement and changes, then reassessed the students' reasoning. This cycle of assessment and refinement has great potential in evaluating the pedagogical effectiveness of teaching methodology and use of Fathom.

Purpose of the study

The purpose of this study is to examine the usage of Fathom in mathematics teaching and learning, particularly in the secondary school syllabus, which focus on Form Four Statistics topic. In order to achieve that, this study's aim is to:

- Investigate if this learning method has led to any differences in Describing Data, Organizing Data, Representing Data and Analyzing and Interpreting Data.

Methodology

Design of the study

To address the research question, the researcher used a quantitative research approach. Quantitative study is more suitable compared to qualitative because of the measurements of the variables are through statistical inference. A quasi-experimental non-equivalent pretest and posttest design study was carried out for eight weeks using statistical activities based on Fathom software to improve the performance and effectiveness of statistical reasoning among Form Four (16 years of age) students. This design enables the researcher to compare intact groups when random assignment is not possible (Campbell & Stanley, 1963). A quasi-experimental design is deemed the strongest research design when a true experimental design is not possible. Quasi-experimental design allowed the researcher to use existing groups; thus it is more convenient than a true experimental design. According to Chua (2012), quasi-experimental design is commonly used to assess the effectiveness of a program if the respondents cannot be distributed randomly.

Sampling

The sample of study, selected through convenience sampling, consisted of seventy-two Form Four students from a secondary school in Selangor state in Malaysia. Thirty-four students were in the experimental group (taught using Fathom); the remaining thirty-eight students were in the control group. Since the entire sample had mixed ability students therefore the researcher randomly assigned samples into two groups.

Instrument

The statistical reasoning instrument used in this study, namely pre-test and post-test, used the same questions in both tests. The pre-test and post-test from both groups are the same to obtain meaningful results. The topic covered in these instruments is descriptive statistics, which involves the measures of central tendency and graph distribution. The tasks in the instrument based on technology align with statistical reasoning. The study's instrument was adapted from "Developing Statistical Reasoning Assessment Instrument for High School Students in Descriptive Statistics" (Chan, Ismail, & Sumintono, 2016). The original assessment was designed based on the initial statistical reasoning framework to evaluate students' statistical reasoning levels across the four constructs developed by Jones et al. (2000), which fulfill the needs of this study instrument.

Jones et al. (2000) and Mooney (2002) mentioned four constructs in their studies: a) describing data, b) organizing data, c) representing data, and d) analyzing and interpreting data. These four constructs were then adopted from the work of Chan et al. (2016). The framework designed by the researchers could identify students' understanding in statistical reasoning. Many students face difficulties in reading, analyzing and interpreting data. Besides that, students have misconceptions about measures of central tendency. This continues with difficulties to represent tasks into mathematical graphs, which leads to incapability to compare graphs and distributions (Ciancetta, 2007; Clark, Kraut, Mathews, & Wimbish, 2007; Cooper & Shore, 2008; Mevarech, 1983; Pollatsek, Lima, & Well, 1981). Therefore, the framework designed by the researchers to overcome this tendency is more suitable as the framework for this study as it helps in the data collection process to answer the research questions. The framework developed by the researcher consists of sub-processes for each construct.

Initially, the topics of descriptive statistics covered in the original assessment tool were measures of central tendency and measures of variability. The statistical reasoning assessment used in this study was slightly modified and eliminated items to fit the Fathom-based instructions and to be aligned with the syllabus. Items such as measuring of variability were eliminated because Form Four level focuses on measures of central tendency, histogram and frequency polygons. Besides that, the instrument scoring was based on a rubric adapted from the Developing Statistical Reasoning Assessment Instrument for High School Students in Descriptive Statistics (Chan et al., 2016). Rubric of selected items from original assessment was modified to fit the Fathom-based answers.

Reliability and Validity

The instrument used to assess effectiveness of Fathom based instruction in enhancing students' statistical reasoning had undergone evaluation in pilot study before it was administered. The pilot study was intended to investigate any weakness in the research design. It was conducted under the same condition using similar respondents and the same instrument planned for the study. The pilot study was also intended to test how well the design can be applied in the field, to find errors in the data collection instrument and to locate errors in interpretation of the data collected. This pilot study was conducted at the secondary schools in Selangor (not the place of the actual research but with a similar background to the sample of the actual study). Internal consistency reliability was used to check the instrument reliability as this study involved only one instrument to be administered to all the respondents. In terms of practicality of the instrument, when pilot study is conducted, the respondents are asked to comment on the wording, timing and their understanding of the items. Some 30 subjects in the sample (N = 30) participated in the Statistical Reasoning Assessment. The instrument obtained Cronbach alpha coefficient of .82, indicating good internal consistency.

After the pilot test, inter-rater reliability was evaluated with two different raters to examine the consistency and precision of the statistical reasoning assessment rubric. Inter-rater reliability helps to identify the rubric of the instrument considered relatively subjective and precise scoring. The Pearson correlation was used in this study to measure how consistent raters were in marking the Statistical Reasoning Assessment. Correlation coefficients were used in this study, as it is nearly perfect to measure the association between two independent raters. One of the raters was a school mathematics teacher with 10 years of experience teaching mathematics in secondary school. Results of the Pearson correlation coefficient indicated a very strong and positive correlation between the two raters scoring, r (30) = .981, $p < .01$. This indicated strong positive consistency between both raters in scoring which make the rubric of the statistical reasoning assessment reliable.

Implementing Fathom in Statistical Reasoning Classroom

The teacher introduced dynamic software, Fathom, to the experimental group students at the beginning of the study. They were taught how to use the software using Fathom tutorial and manual guidance before being taught the statistical reasoning lessons. The researcher or teacher had prepared a number of instructional activities implementing Fathom as a tool of teaching and learning statistical reasoning. Students were encouraged to use Fathom to build graphs and analyze data from graphs that have been built. Time was provided for students to involve in discussions and presentation of their answer, to make inferences and conclusions. After solving statistical questions using software, students were required to apply the knowledge learnt during the lessons in class activities.

Data Collection and Analysis

In the beginning, students from the control and treatment group were given the pre-test. Pre-test was conducted to ensure that two groups were equal in understanding of statistics. Students were instructed to show all the steps involved in their solutions. After the instructional activities done with control (without intervention) and experimental (Fathom-based intervention) groups, students from both groups were given posttest to measure the differences between their scores. In addition, they were also told that the tests that they sat for in this study would not affect their own school’s test score. All the questions in the tests were subjective based on the statistical reasoning.

The quantitative data were analyzed with the Statistical Packages for the Social Sciences Personal Computer (SPSS). For answering the research question, MANCOVA was used to analyze whether the learning method has led to any differences between the four constructs. MANCOVA was chosen because it involves the use of covariance that serves as a measure. Another reason is MANCOVA is to control the control variable (covariate) which is a factor that does not want to be studied but it affects the dependent variables. This allows seeing the exact effect of independent variables on dependent variables without unwanted interference.

Results

Is there any significant difference in Describing Data, Organizing Data, Representing Data and Analyzing and Interpreting Data between Form Four students in the control and experimental groups?

To answer research question two, Multivariate analysis of covariance (MANCOVA) statistics was used. A few assumptions of MANCOVA statistics analysis need to be met before running the test. Figure 1 shows the results of testing assumptions indicated that there were no univariate or multivariate outliers as assessed by boxplot.

Moreover, the data were normally distributed for statistical reasoning constructs namely, Describing Data, Organizing Data, Representing Data and Analyzing and Interpreting Data for both groups. Table 2 shows the Shapiro-Wilk test ($p > .05$) for checking the assumption of normality.

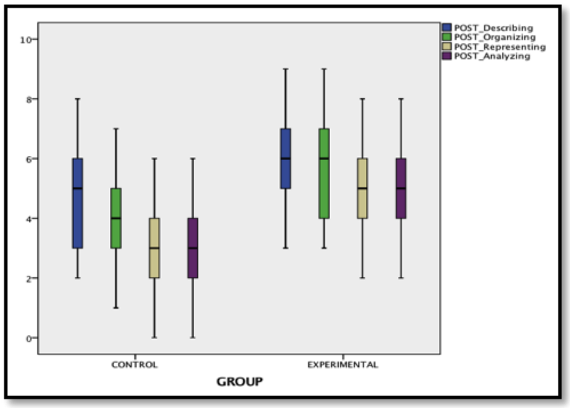

Figure 1. Assumptions of no univariate.

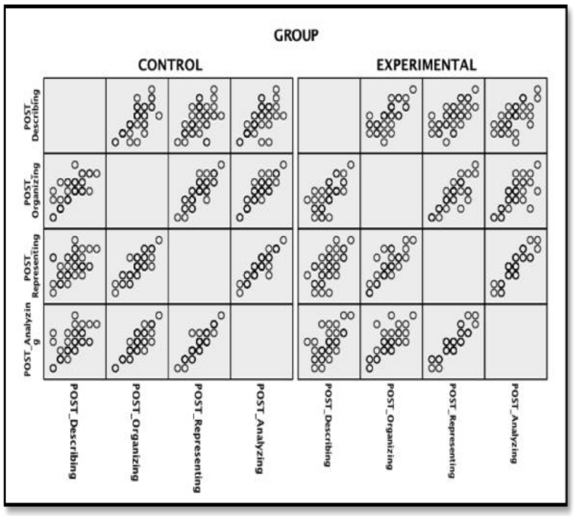

Figure 2. Linearity between DV and Covariate for constructs of IV

Table 2

Assumption of Multivariate Normality

		Shapiro-Wilk		
	Group	Statistics	df	*p*
Posttest Total Average Describing Data	Control	.952	38	.104
	Experimental	.955	34	.171
Posttest Total Average Organizing Data	Control	.954	38	.121
	Experimental	.951	34	.136
Posttest Total Average Representing Data	Control	.954	38	.119
	Experimental	.954	34	.166
Posttest Total Average Analyzing and Interpreting Data	Control	.956	38	.138
	Experimental	.954	34	.157

Table 3 shows there was homogeneity of variance matrices, as assessed by Box's M test (*M* = 14.08, *F* = 1.32, *p* = .212).

Table 3

Box's Test of Equality of Covariance Matrices

Box's M	14.080
F	1.321
df1	10
df2	22750.965
P	.212

Table 4 shows that the assumption of homogeneity of variance-covariance was met, as assessed by Levene's test of homogeneity of variances for the four constructs namely; Describing Data (*F* = 1.103, *p* = .297), Organizing Data (*F* = .245, *p* = .662), followed by Representing Data (*F* = .320, *p* = .574) and lastly Analyzing and Interpreting Data (*F* = .002, *p* = .967).

Table 4

Assumption of Homogeneity of Variances

	F	*df1*	*df2*	*p*
Posttest Total Average Describing Data	1.103	1	70	.297
Posttest Total Average Organizing Data	.245	1	70	.662
Posttest Total Average Representing Data	.320	1	70	.574
Posttest Total Average Analyzing and Interpreting Data	1.02	1	70	.967

The shape of the scatterplot shows linearity of variable was oval-shaped hence the relationships between variables were linear as shown in Figure 2.

The next assumption homogeneity of regression slope was met since the interaction term while controlling the pretest score was not statistically significant for all the four constructs as illustrated in table 5.

Table 5

Homogeneity of Regression Slope Tests of Between-Subjects Effects

Source	Dependent Variable	*df*	*F*	*P*
Group*Pre_Describing*Pre_Organizing*Pre_Representing*Pre_Anlayzing	Posttest Total Average Describing Data	2	.87	.43
	Posttest Total Average Organizing Data	2	3.81	.30
	Posttest Total Average Representing Data	2	1.39	.26
	Posttest Total Average Analyzing and Interpreting Data	2	1.81	.18
Error	Posttest Total Average Describing Data	64		
	Posttest Total Average Organizing Data	64		
	Posttest Total Average Representing Data	64		
	Posttest Total Average Analyzing and Interpreting Data	64		

The first construct Describing Data *F* (2,64) = .87, *p* = .43 was not statistically significant, Organizing Data *F* (2,64) = 3.81 was not statistically significant, *p* = .30, Representing Data *F* (2,64) = 1.39 was not statistically significant, *p* = .26 and lastly Analyzing and Interpreting Data *F* (2,64) = 1.81, *p* = .18 was not statistically significant.

The result of adjusted means of post-test score for four constructs is presented in Table 6.

Table 6

Adjusted Mean of Posttest Scores for Each Statistical Reasoning Constructs in Control and Experimental Groups.

Dependent Variable	Group	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Posttest Describing	Control	4.331[a]	.165	4.001	4.661
	Experimental	5.277[a]	.178	4.921	5.633
Posttest Organizing	Control	4.766[a]	.182	4.403	5.129
	Experimental	5.644[a]	.196	5.252	6.036
Posttest Representing	Control	4.536[a]	.113	4.311	4.762
	Experimental	5.401[a]	.122	5.157	5.644
Posttest Analyzing	Control	4.144[a]	.146	3.853	4.436
	Experimental	4.809[a]	.157	4.495	5.124

a. Covariates appearing in the model are evaluated at the following values: PRE_Describing = 4.58, PRE_Organizing = 3.61, PRE_Representing = 3.08, PRE_Analyzing = 1.36.

The post-test score of Describing Data while controlling pretest for control group (M = 4.33, SE = .165, 95% CI [4.01, 4.66]) was different than for the experimental group (M = 5.28, SE = .178, 95% CI [4.92, 5.63]). The adjusted mean of posttest score of Organizing Data was different between control (M = 5.28, SE = .178, 95% CI [4.92, 5.63]) and experimental (M = 5.28, SE = .178, 95% CI [4.92, 5.63]) groups. The posttest score of Representing Data for control (M = 5.28, SE = .178, 95% CI [4.92, 5.63]) and experimental (M = 5.28, SE = .178, 95% CI [4.92, 5.63]) was also different; and also, the total average score of posttest of Analyzing and Interpreting Data was different between control (M = 5.28, SE = .178, 95% CI [4.92, 5.63]) and experimental (M = 5.28, SE = .178, 95% CI [4.92, 5.63]) groups.

These differences are visualized by the generated plots of estimated marginal means of posttest scores in terms of Describing Data, Organizing Data, Representing Data and Analyzing and Interpreting Data as shown in figure 3.

The multivariate tests in table 7 show that the differences between the control and experimental groups on the statistical reasoning constructs was statistically significant, *F* (4, 63) = 7.00, $p < .005$; Wilks' Λ = .60, with large effect size and observed power (partial $\eta 2$ = .60, observed power = 1). Thus, this study rejects the null hypothesis that the mean of the posttest score of Describing Data, Organizing & Reducing Data, Representing Data and Analyzing and Interpreting Data of Form Four students was not significantly different between experimental and control groups after controlling for pretest scores.

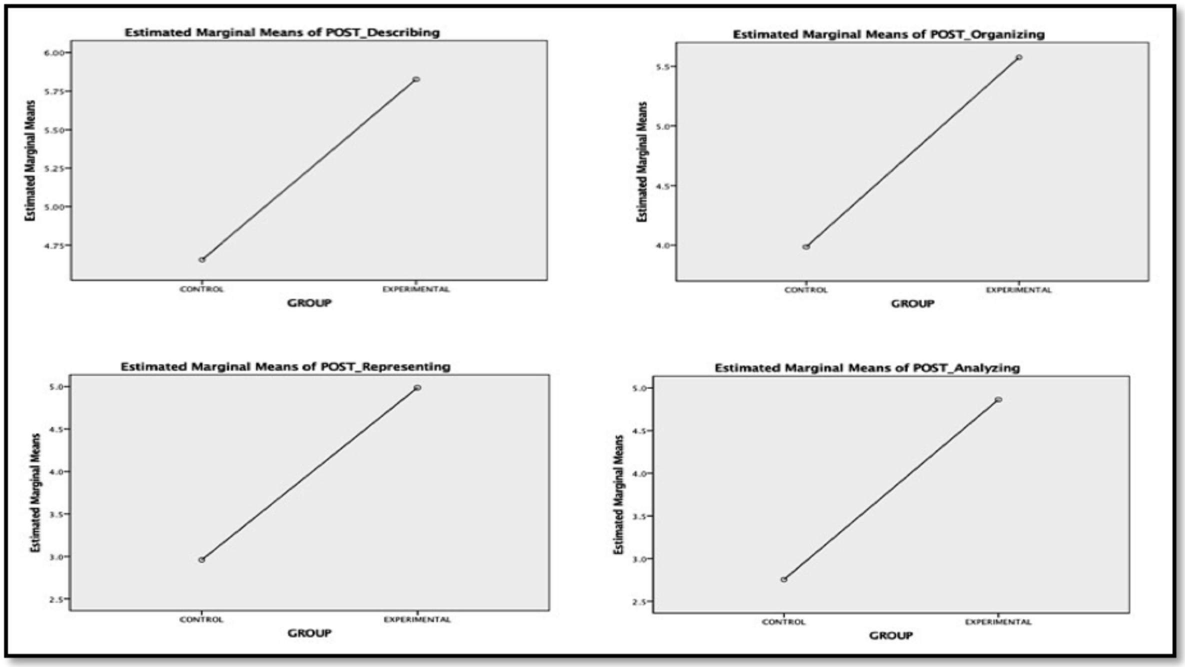

Figure 3. Estimated Marginal Means of Posttest Scores for Each Construct

Table 7

Adjusted Mean of Posttest Scores for Each Statistical Reasoning Constructs in Control and Experimental Groups Multivariate Tests

	Value	*F*	*Hypothesis df*	Error df	*p*	Partial Eta Squared	Observed Power
Pillai's Trace	.308	15.177[b]	4.000	63.00	.000	.596	1.000
Wilks' Lambda	.692	7.004[b]	4.000	63.00	.000	.596	1.000
Hotelling's Trace	.445	7.004[b]	4.000	63.00	.000	.596	1.000
Roy's Largest Root	.445	11.211[b]	4.000	63.00	.000	.596	1.000

Table 8 shows the Tests of Between-Subject Effects as well as the follow-up Univariate Anova illustrated that the Describing Data posttest score F (1,66) = 9.37, $p <$ .0005, partial $\eta 2$ = .88; Organizing Data posttest score F (1,66) = 16.23, $p <$ .0005, partial $\eta 2$ = .80; Representing Data posttest score F (1,66) = 25.69, $p <$ .0005, partial $\eta 2$ = .72 and Analyzing and Interpreting posttest score F (1,66) = 28.33, $p <$ .0005, partial $\eta 2$ = .70 were statistically significantl different between the control and experimental groups after controlling the pretest scores.

Table 8

Univariate Tests

Dependent Variable		SS	df	Mean Square	*F*	*p*	Partial Eta Squared	Observed Power
PostTest Describing	Contract	146.024	1	20.738	9.373	.003	.876	1.000
	Error	20.738	66	2.212				
PostTest Organizing	Contract	155.133	1	38.159	16.234	.000	.803	1.000
	Error	38.159	66	2.351				
PostTest Representing	Contract	159.761	1	62.180	25.688	.000	.719	1.000
	Error	62.180	66	2.421				
PostTest Analyzing	Contract	156.703	1	67.266	28.331	.000	.699	1.000
	Error	67.266	66	2.374				

Post hoc analysis was not performed as this study has two groups only. The data provided sufficient evidence to conclude that there is a significant difference in terms of Describing Data, Organizing Data, Representing Data and Analyzing and Interpreting Data between Form Four students in the control and experimental groups while controlling for pretest scores.

Discussion

Based on the findings of this study, Fathom software can promote good teaching and learning outcome in statistical reasoning. This study analyzed the difference in statistical reasoning constructs, namely describing data, organizing data, representing data and analyzing and interpreting data between form four students in the experimental group when controlling for pre-test. The use of Fathom increased student scores in the experimental group compared to score of students in control group. By employing this teaching approach students learn how to understand statistical concepts by developing their reasoning skills in describing, organizing, representing and interpreting data. This study's activities and lessons developed the first construct, Describing Data, by asking students "why" questions to make them think and reason why such event or situation took place. Since Fathom helped to visualize the raw data in tables and graphs, students had the opportunity to focus on the meaning behind the data and graphs. Moreover, students are required to collect, organize data by themselves and discuss among peers and their teacher. Based on feedback students are able to understand that possibilities exist for different interpretations of the same data; there is no one correct answer in statistical reasoning contrary to what is projected in the traditional statistics classroom. Besides that, with help from Fathom, teachers can show precisely the arrangement of data in frequency tables and the measure of central tendency located (mean, median and mode). Fathom can show the outcome immediately and does not require more time to construct different graphs at one time especially during teaching. Therefore, students actively engaged in representing data differently. Their understanding on selecting the correct graph for the data will deepen as they are able to explain the reasons behind their choice. Students are able to analyze and interpret data better after using Fathom since the software allows them to learn by experimenting and gain experience with hands-on activities. This encouraged students to actively participate without fear of making mistakes.

Recent studies have proposed that technology-based learning with well-planned lessons will help students learn statistical concepts (Ben-Zvi, Gravemeijer, & Ainley, 2018; Brahier, 2016; Eichler & Zapata-Cardona, 2016). The findings supported Loveland and Schneiter (2014) who stated that both constructivist methodology and technology play a significant role in enhancing statistical reasoning ability and statistical concepts by providing students with access to view and design simulations. This study agreed with the previous studies, since it also exposed that implementing technology aligned with a statistical reasoning learning environment could have a positive effect on students' statistical reasoning. The findings supported Chance and Rossman (2006); Lane and Peres (2006); Mills (2004) that technology plays a significant role in enhancing students' statistical reasoning ability and statistical concepts.

Conclusion

In this study, the Fathom-based method was carried out to investigate its effect on Form Four students' statistical reasoning across four constructs namely describing data, organizing data, representing data and analyzing and interpreting data. The students showed a remarkable improvement in these skills through this new intervention.

From the theoretical aspect, the findings of this study are in congruence with constructivism learning theory. This suggested that if teachers perceive students as "active learning seekers" in the learning process, they are able to synthesize information to construct knowledge and understanding from prior knowledge. Furthermore, the results of this study are also in accordance with Garfield and Ben-Zvi (2008) whereby a statistical reasoning learning environment equipped with instruments and tools encouraged students to make and test inferences using data, involve in discussions and be able to explain ideas.

This study has provided an alternative approach for enhancing students' statistical reasoning skills. Therefore, mathematics educators highly recommended using Fathom to teach statistical reasoning and mathematics in a meaningful way. This could be coupled with research to establish better findings to conclusively ascertain whether Fathom does actually have an effect on learning of broader statistics concepts and on different levels of students.

References

Accrombessy, F. (2006). *An evaluation study of the process of reform of statistics teaching at the secondary level in Benin: Assessment and perspectives.* Paper presented at the Seventh International Conference on Teaching Statistics.

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of Statistics Learning Environments. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in Statistics Education* (pp. 473-502). Cham: Springer International Publishing.

Brahier, D. J. (2016). *Teaching secondary and middle school mathematics.* New York, NY: Routledge.

Bruno, A., & Espinel, M. C. (2009). Construction and evaluation of histograms in teacher training. *International Journal of Mathematical Education in Science and Technology, 40*(4), 473-493. doi: 10.1080/00207390902759584

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. *Handbook of research on teaching* (pp. 171-246). Chicago, IL: Rand McNally.

Chan, S. W., Ismail, Z., & Sumintono, B. (2016). A Framework for Assessing High School Students' Statistical Reasoning. *PloS One, 11*(11), e0163846.

Chance, B., & Rossman, A. (2006). *Using simulation to teach and learn statistics.* Paper presented at the Seventh International Conference on Teaching Statistics.

Chua, Y. P. (2012). Effects of computer-based testing on test performance and testing motivation. *Computers in Human Behavior, 28*(5), 1580-1586.

Ciancetta, M. A. (2007). *Statistics students reasoning when comparing distributions of data.* Citeseer.

Clark, J., Kraut, G., Mathews, D., & Wimbish, J. (2007). The fundamental theorem of statistics: Classifying student understanding of basic statistical concepts. Unpublished manuscript. Retrieved from

http://www1. hollins. edu/faculty/clarkjm/stat2c. pdf

Cooper, L., & Shore, F. (2008). Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. *Journal of Statistics Education, 16*(2), 1-13.

DelMas, Garfield, J., & Ooms, A. (2005). *Using assessment items to study students' difficulty reading and interpreting graphical representations of distributions.* Paper presented at the Fourth Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4).

Eichler, A., & Zapata-Cardona, L. (2016). Empirical research in statistics education. In *Empirical Research in Statistics Education* (pp. 1-37). Springer.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *A curriculum framework for K-12 statistics education*. GAISE report, American Statistical Association.

Gal, I., & Garfield, J. B. (1997). *The assessment challenge in statistics education* (Vol. 12). IOS Press.

Galotti, K. M. (2008). *Cognitive psychology in and out of the laboratory* (5th ed.). Thousand Oaks, CA: SAGE.

Garfield, Joan. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*(3), 58-69.

Garfield, Joan, & Ben-Zvi, D. (2008). *Developing students' statistical reasoning*. Springer.

Hall, & Heyde, C. C. (2014). *Martingale limit theory and its application.* Academic Press.

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

Lane, D. M., & Peres, S. C. (2006). *Interactive simulations in the teaching of statistics: Promise and pitfalls.* Paper presented at the Seventh International Conference on Teaching Statistics.

Lane-Getaz, S. J. (2006). What is statistical thinking, and how is it developed. *Thinking and reasoning about data and chance: Sixty-eighth NCTM Yearbook* (pp. 273-289). Reston, VA: NCTM.

Loveland, J., & Schneiter, K. (2014). *Teaching Statistics with Lectures or Activities: A Comparative Study.* Paper presented at the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Arizona, USA.

Meletiou, M., & Lee, C. (2002). *Student understanding of histograms: A stumbling stone to the development of intuitions about variation*. Paper presented at the The Sixth International Conference on Teaching Statistics (ICOTS6). https://www.stat.auckland.ac.nz/~iase/publications/1/10_19_me.pdf

Meletiou, M., & Stylianou, D. A. (2003). Graphical representation of data: The effect of the use of dynamical statistics technological tool. Retrieved from https://pdfs.semanticscholar.org/3463/fc7cbe9424dd8b4608e4466b4904e50c3052.pdf

Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics, 14*(4), 415-429. doi: 10.1007/bf00368237

Mills, J. D. (2004). Students' attitudes towards Statistics: Implications for the future *College Student Journal, 38*(3).

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

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics, 12*(2), 191-204. doi: 10.1007/bf00305621

Pratt, D., & Ainley, J. (2008). Introducing the special issue on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 3-4.

Saldanha, L., & Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, *51*(3), 257-270.

Sorge, C., & Schau, C. (2002). *Impact of engineering students' attitudes on achievement in statistics: A structural model.* Paper presented at the annual meeting of the American Educational Research Association. New Orleans.

Corresponding Author Contact Information:

Author name: Kwan Eu Leong

Department: Mathematics and Science Education

Faculty: Education

University, Country: University of Malaya, Malaysia

Please Cite: Ganesan, N., & Leong, K.E. (2020). Impact of Fathom on Statistical Reasoning among Upper Secondary Students. *Journal of Research in Science, Mathematics and Technology Education*, 3(1), 35-50. Doi: 10.31756/jrsmte.321

Received: November 03, 2019 ▪ Accepted: April 21, 2020