EFFECTS OF GRAPHING CALCULATOR ON LEARNING INTRODUCTORY STATISTICS

Wei Wei
Mathematics Department, Metropolitan State University, Saint Paul, MN, United States
+Corresponding author: wei.wei@metrostate.edu

Katherine Johnson
Mathematics Department, Metropolitan State University, Saint Paul, MN, United States
katherine.johnson@metrostate.edu

ABSTRACT
Graphing calculators have been beneficial for teaching and learning statistics. Graphing calculators help students to visualize concepts, deal with real world data, and obtain accurate results promptly. However, heavily relying on the technology may hinder students' conceptual understanding and lead to misconceptions. In this study, we focus on the effect of using graphing calculator on students' performance of normal probability calculations and hypothesis testing, student' conceptual understanding of normal distribution and p-value, and students' retention on these concepts.
Key words: Graphing calculator, Hypothesis testing, Normal distribution, Introductory statistics curriculum

Introduction and Theoretical Perspectives

Graphing calculators have been widely used in statistics classrooms when the inferential statistics capabilities had been developed around the1990s (Dunham and Dick 1994, Garfield, 1995, Burrill, 1997). From then, graphing calculators became a portable and economical tool in learning statistics. The College Report of the American Statistical Association's Guidelines for Assessment and Instruction in Statistics Education (GAISE) project (2005) recommended the use of technology for developing concepts and analyzing data in a statistics curriculum, and the graphing calculator was one of the technologies recommended.

The benefits of using a graphing calculator in teaching statistics have been noted from several research studies. It was found to be easier to teach introductory statistics with a graphing calculator (Idris et al., 2003, Krishnan and Idris, 2013). Graphing calculators helped students to solve computational intensive problems, such as solving rate of change and derivatives, etc. (Roorda et al. 2016). With the use of a graphing calculator, instructors were able to spend more time on explanations of concepts and interpretations of analysis results, rather than focusing on complex computations (Garfield et al. 2000, Chance et al. 2007). The use of a graphing calculator allowed instructors to teach with real data rather than small made-up data and connect course materials to daily life (Lazari and Goel 2003; Chance et al. 2007). Graphing calculators offered the visual representations to help students better understand concepts (Sara et al., 2001, Kor and Lim, 2003, 2004, Graham et al., 2008). Griffith (1998, pp. 76) stated that "it may well be the case that technology does not help or hinder students in the memorization of facts, but technology does help students to develop conceptual understanding and problem solving abilities".

Specifically, there have been studies that quantitatively compared students' performance with and without using a graphing calculator in statistics courses. Tan (2012) compared students' performance on solving questions related to random variables, Poisson distribution, binomial distribution, and normal distribution with and without using a graphing calculator. He found that the graphing calculator approach significantly improved students' performance. Lazari and Goel (2003) compared two approaches: teaching introductory statistics using a TI-83 graphing calculator, versus, teaching the same course with the traditional method of not using a graphing calculator. The research showed that students' average final exam score was significantly higher when teaching with a TI-83 than when teaching without using a graphing calculator.

There are limitations to using graphing calculators. There have been issues of using graphing calculators with the partial view of graphs, not compiled with other statistics software files, and the graphical output with no labels and scales. Sometimes, these limitations can lead to students' misconceptions about graphs (Cavanagh and Mitchelmore, 2003, Chance et al. 2007). It has been reported that graphing calculator has some constraints within the classroom practice (Doerr and Zangor, 2000).

There are studies that had shown issues related to the effectiveness of using a graphing calculator on students' learning (Kharuddin and Ismail, 2017, Parrot and Leong, 2018). Collins and Mittag (2005) found that the use of a graphing calculator did not give students apparent advantages in their performance when solving statistical inference related questions in exams. Adam (1997) studied the effects of using a graphing calculator on students' understanding of functions and concluded that "They must have a basic understanding of the concept in order to understand the reasoning behind the operation of the graphing calculator. Otherwise, the student will see the graphing calculator as a machine for doing mathematics instead of a tool for learning."

Graphing calculators have the advantages for teaching and learning, but also have the disadvantages of prohibiting students' understanding of concepts if instructions were not given appropriately. In this study, we focus on the effects of using a graphing calculator on students' learning of normal probabilities and hypothesis testing. We investigated whether the use of graphing calculators in an introductory statistics course gave students an advantage in calculating normal probabilities and performing hypothesis tests, whether it helped students' conceptual understanding of normal transformation and understanding of one-sided and two-sided p-values, and whether it helped students' retention on understanding normal transformation and p-values, and performing hypothesis tests.

Methods

Study Design and Course Overview

The study was performed in a four-year university. Students who participated in the study were enrolled in four sections of an introductory statistics course. The students in the introductory statistics course were from a variety of majors, as this course is a required course for most of the majors. The class size of all sections were around 30 students. Students' consents were obtained at the beginning of the semester with a protocol approved by the Human Subject Review Board.

The four sections of the introductory statistics course in this study were taught by two instructors, both from the Mathematics Department. Each instructor taught two sections. Both instructors used the same textbook, *The Practice of* Statistics by Starnes et al. (2010), and similar teaching methodologies, including Microsoft PowerPoint slides, in-class group discussions, quizzes, one course project, and homework assignments from the textbook. Each instructor taught one section using a TI calculator (referred to as calculator section) and the other section without using a TI calculator (referred to as non-calculator section). Whether or not using TI calculator to teach was referred as pedagogy difference. The non-calculator sections were taught using the standard normal distribution table to find normal probabilities on a normal distribution, using formulas to calculate the test statistics and using graphing calculator to find p-values for a hypothesis test. The calculator sections were taught using TI calculator functions, *normalcdf* and *invnorm*, to calculate normal probabilities on a normal distribution, and using TI calculator functions, *1-PropZtest*, *2-PropZtest*, *T-Test*, *2-SampTtest*, and χ^2*-Test*, to perform hypothesis tests. The calculator and non-calculator sections had the same homework assignments, exams, and project from each instructor. But the two instructors had different homework assignments, exams and projects.

Assessments and Data Collection

The demographic survey was given at the beginning of the semester to investigate whether there was a difference in students' background, such as age, enrollment status, primary language and prior statistics classes taken between the calculator and non-calculator sections.

Two quizzes were created to test students' learning outcomes of the normal probabilities and hypothesis testing: quiz one was given after introducing the normal distribution; and quiz two was given after introducing the two-sample Z test. Three common questions from the final exam were used to investigate students' short-term retention of their knowledge of normal distribution, p-value and hypothesis testing. The structure of the data collection is shown in Figure 1. The two quizzes and final exam were given to the calculator sections and non-calculator sections in class during the same week.

The questions in the quizzes and final exams included both multiple choice questions and calculation questions. Quiz one included one multiple choice question testing students' conceptual understanding of the standard normal transformation, and two calculation questions testing students' ability to calculate normal probabilities and the endpoints under a normal distribution. Quiz two included one multiple choice question testing the conceptual understanding of difference between one-sided and two-sided p-values and a calculation question to evaluate students' ability to carry-out a hypothesis test. In the final exam, two multiple choice questions were given to test students' retention of their understanding of the normal transformation and one-tailed and two-tailed p-values, and one calculation question to test students' retention of performing a hypothesis test. The two multiple choice questions in the final exam were similar to the two multiple choice questions in quiz one and quiz two. The hypothesis testing question in quiz two was a two-sample proportion test, and the hypothesis testing question in the final exam was a one-sample mean test.

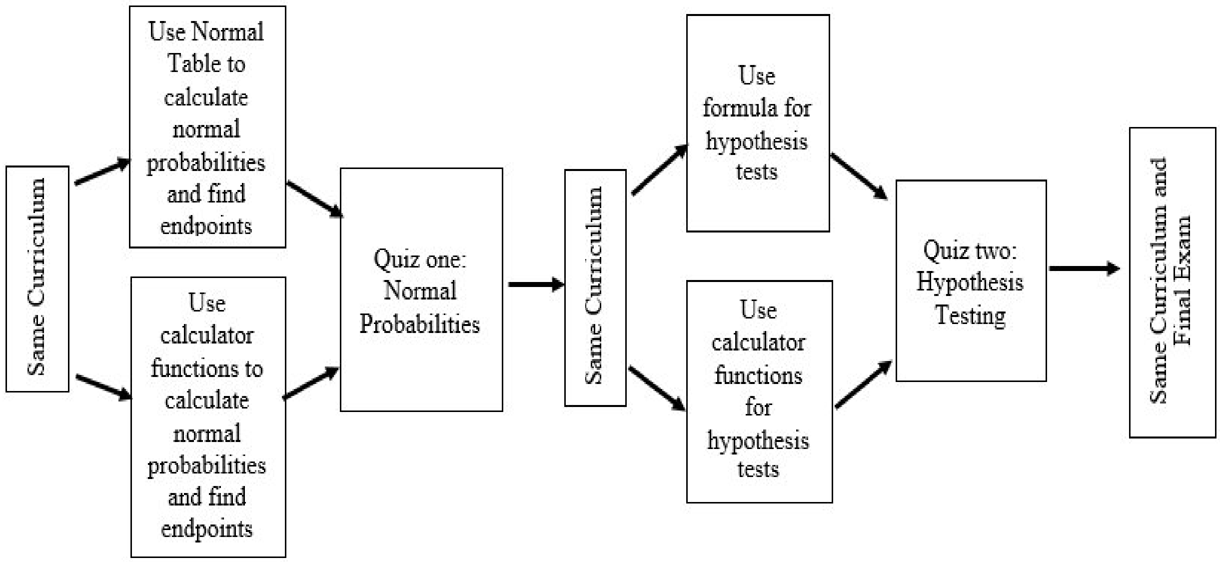

Figure 1. Assessment schedule of calculator and non-calculator sections.

Grading

All assessments were graded by their respective instructors according to a predetermined rubric. Each multiple choice question was worth one point and marked either correct or not correct (1 or 0 points). The calculation questions were worth different points according to the number of steps needed. Students' scores for each calculation question were converted into the percentage of possible points received for each question.

Statistical Analysis

The multiple choice questions and calculation questions were analyzed separately for each quiz and exam. A categorical analysis was used for multiple choice questions and a 2 (pedagogy) by 2 (instructor) analysis of variance (ANOVA) was used for calculation questions. Summary statistics were calculated for the calculation questions and the percent of correct responses were calculated for multiple choice questions.

For each multiple choice question, a Mantel-Haenszel procedure was performed. The correctness (1 or 0) was the dependent variable, pedagogy with two levels (calculator and non-calculator use) was the independent variable, and instructor with two levels (instructor one and two) was the strata. The Mantel-Haenszel procedure was used to determine if there was a significant association between the correctness and pedagogy with the consideration of instructor effects. The Mantel-Haenszel procedure contained two steps: the homogeneity of odds ratio test to detect whether there was an instructor effect, and the test of conditional independence to determine whether there was a significant correlation between the correctness and pedagogy.

For each calculation question, the dependent variable was the average percentage points and the independent variables were instructor with two levels (instructor one and two) and pedagogy with two levels (calculator and non-calculator). The 2 by 2 ANOVA was performed to detect the significant mean differences between the calculator and non-calculator sections and between the two instructors. All statistical analyses were carried out at a 0.05 significance level. All analyses were performed using JMP® Pro 9.0.2 and IBM Statistics SPSS 22.

Results

Analysis of Demographic Survey

Ninety-nine students agreed to participate in the study: 52 were in calculator sections and 47 were in non-calculator sections; 53 were taught by instructor one and 46 were taught by instructor two. Almost all of the students were between the ages of 18 and 40 years (87%); and 13% of the students were over 40 years. Almost three-fourths of the students (73%) were primarily English speakers. More than half (59%) were enrolled as full-time students. Nearly 20% of the students had taken a previous statistics class. As shown in Table 1, there is no difference in the students' demographics between the calculator and non-calculator sections.

	Instructor one		Instructor two	
	Calculator (n=27)	Non-calculator (n=26)	Calculator (n=25)	Non-calculator (n=21)
Age:				
18-40	21	22	24	19
41+	6	4	1	2
Primary Language:				
English	21	19	18	14
Other	6	7	7	7
Enrollment Status:				
Full-time	12	15	18	13
Part-time	15	11	7	8
Previous Statistics Class:	2	2	4	9

Table 1. Summary Statistics for Demographic Survey

Quiz One Results

The first question in quiz one tested students' conceptual understanding of a standard normal transformation. From the Mantel-Haenszel procedure, the homogeneity of odds ratio test revealed there was no significant instructor effect ($\chi^2(1, N=93) =0.198, p=0.66$), indicating the association between the correctness and pedagogy was not different between the two instructors. A significant association between the correctness and pedagogy was detected in the test of conditional independence ($\chi^2(1, N=93) =4.68, p=0.03$). Examination of the proportions of correctness showed that the calculator sections had 61.22% correct responses and the non-calculator sections had 36.36% correct responses. Figure 2 shows the percentage of correct responses by instructor for each pedagogy level.

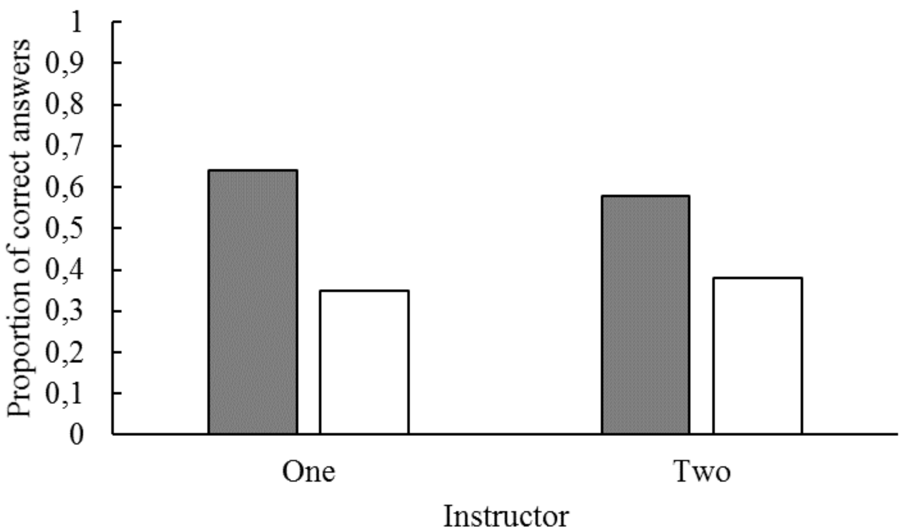

Figure 2. The proportion of correct answers for the multiple choice question testing normal transformation in quiz one. The grey bars represent the proportion of correct answers for calculator sections and the white bars represent that for non-calculator sections.

The two calculation questions in quiz one evaluated how to find the normal probabilities and how to find the endpoint for a given normal probability. The scores of the two questions were combined in the analysis. The 2 (pedagogy) by 2 (instructor) ANOVA showed that there was no instructor effects ($F (1, 89) =0.067, p=0.80$). The mean score of the calculator sections was significantly higher than that of the non-calculator sections ($F (1, 89) =6.94, p=0.0099$). There was no significant interaction between instructor and pedagogy ($F (1, 89) =0.067, p=0.80$). The mean score for the calculator sections was 82.65 (SD=27.28, CI = [74.82, 90.49]) and the mean score for the non-calculator sections was 66.67 (SD=30.71, CI = [57.33, 76.00]). Figure 3 shows the 95% confidence interval for mean score by instructor for each pedagogy level.

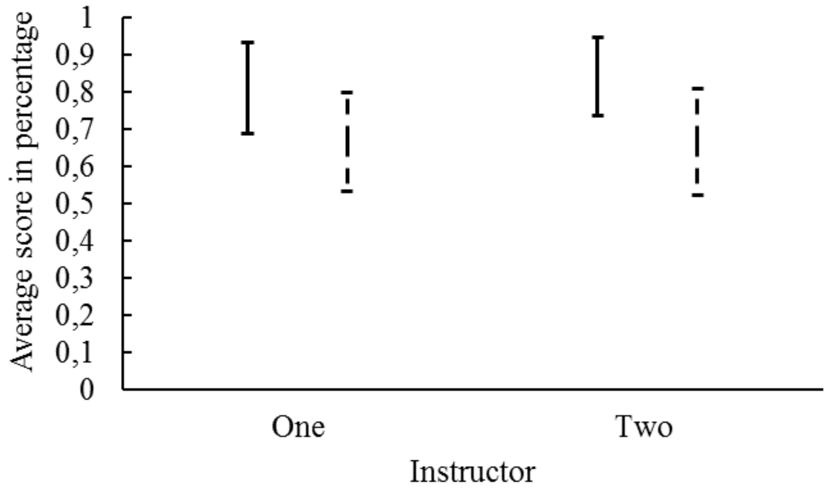

Figure 3. Average score of calculation questions testing normal probabilities and end points in quiz one. The solid bars represent the 95% confidence intervals for the calculator sections, and the dashed bars represent that for the non-calculator sections.

Quiz Two Results

The multiple choice question in quiz two tested students' conceptual understanding of one-sided and two-sided p-values. The homogeneity of odds ratio test in the Mantel-Haenszel procedure showed no instructor effect (χ^2 (1, N=88) =1.05, p=0.31), meaning the association between correctness and pedagogy was not significantly different between the two instructors. The test of conditional independence revealed a non-significant association between correctness and pedagogy (χ^2 (1, N=88) <0.001, p=0.99). The proportion of correct responses was 44.90% for the calculator sections and 43.59% for the non-calculator sections. Figure 4 shows the percentage of correct responses by instructor for each pedagogy level.

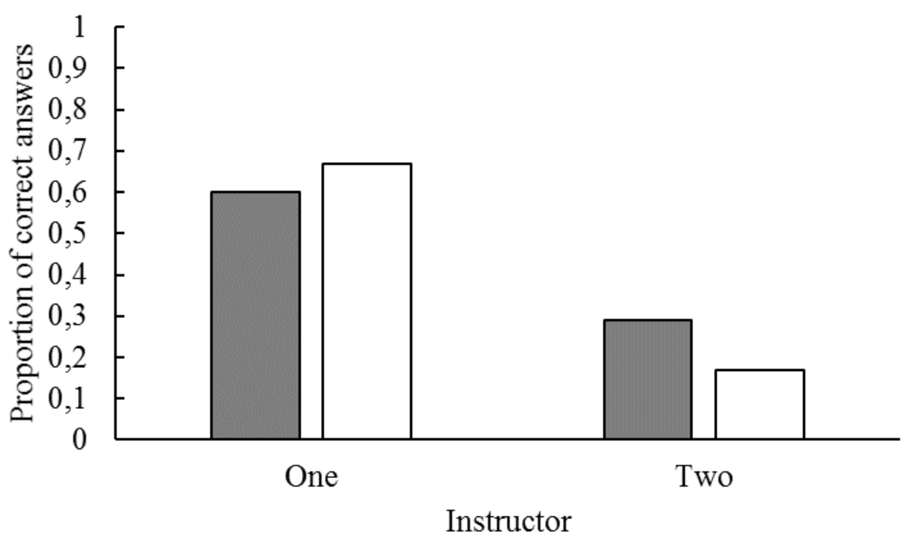

Figure 4. The proportion of correct answers for the multiple choice question in quiz two. The grey bars represent the proportion of correct answers for calculator sections and the white bars represent that for non-calculator section.

The calculation question in quiz two tested students' performance of a hypothesis test. The 2 (pedagogy) by 2 (instructor) ANOVA showed there was a significant effect from pedagogy (F (1,83)=10.49, p=0.0017), there was a significant effect from instructor (F(1,83)=7.53, p=0.0074), and there was a significant interaction between pedagogy and instructor (F(1,83)=9.83, p=0.0024). The mean score from instructor one's calculator section

(M=90.00, SD=18.84, CI= [82.22, 97.78]) was significantly higher than the mean score from instructor one's non-calculator section (M=57.81, SD=31.67, CI= [42.99, 72.63]). The mean score from instructor two's calculator section (M=88.02, SD=18.61, CI= [80.16, 95.88]) was not significantly higher than the mean score from instructor two's non-calculator section (M=87.5, SD=23.87, CI= [75.63, 99.37]). Figure 5 shows the 95% confidence interval for the mean score by instructor for each pedagogy level.

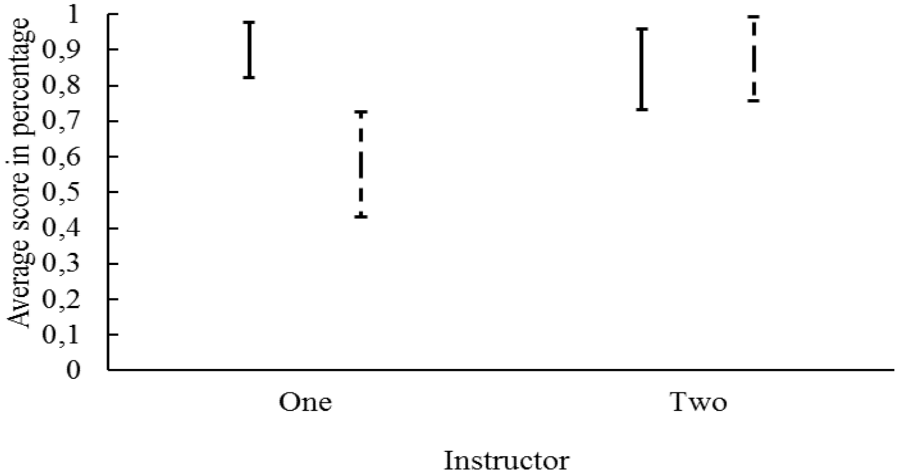

Figure 5. Average score of hypothesis testing question in quiz two. The solid bars represent the 95% confidence intervals for the calculator sections, and the dashed bars represent that for the non-calculator sections.

Final Exam Question Results

The first multiple choice question in the final exam tested students' retention of their understanding the normal transformation. The Mantel-Haenszel procedure revealed there was no significant instructor effect ($\chi^2(1, N=90) = 2.07, p=0.15$), and no pedagogy effects ($\chi^2(1, N=90) = 0.60, p=0.44$). That being said, there was no significant association between the pedagogy and proportion of correctness. The proportion of correct responses was 92% for the calculator sections and 85% for the non-calculator sections. Figure 6 shows the percentage of correct responses by instructor for each pedagogy level.

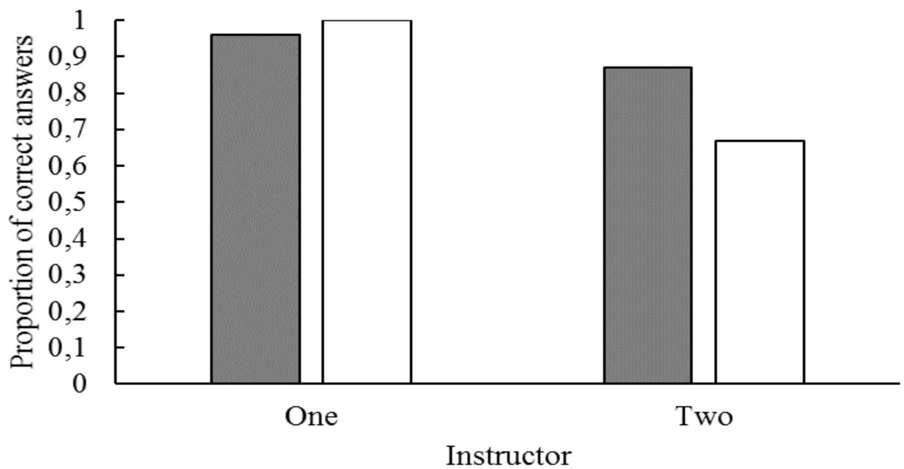

Figure 6. The proportion of correct answers for the normal transformation question in final exam. The grey bars represent the proportion of correct answers for calculator sections and the white bars represent that for non-calculator sections.

The second multiple choice question tested students' retention of their understanding of the one-sided and two-sided p-values. The homogeneity odds ratio test in the Mantel-Haenszel procedure revealed significant instructor effect ($\chi^2 (1, N=90) = 4.26, p=0.039$), meaning the association between correctness and pedagogy was significantly different between the two instructors. The test of conditional independence revealed a non-significant association between correctness and pedagogy (χ^2 (1, N=90) =2.15, p=0.14). The proportion of correct responses was 57% for the calculator sections and 64.5% for the non-calculator sections. Figure 7 illustrates the proportion of correct responses by instructor for each pedagogy level.

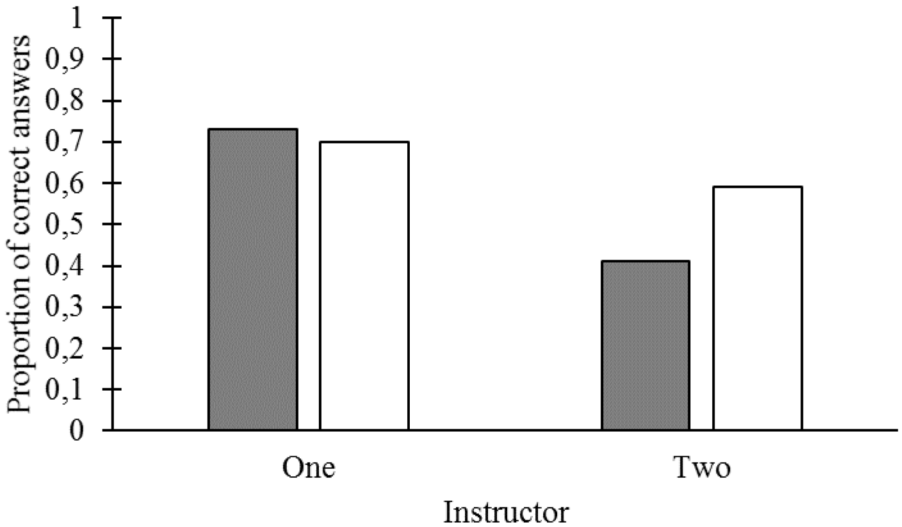

Figure 7. The proportion of correct answers by instructor and by pedagogy for the second multiple choice question in final exam. The grey bars represent the proportion of correct answers for calculator sections and the white bars represent that for non-calculator sections.

The analysis of the hypothesis testing question in the final exam showed that there was a significant difference between the calculator and non-calculator sections for instructor one, but no significant difference between the calculator and non-calculator sections for instructor two. The mean score was significantly higher for the calculator section (M=78.40, SD=19.02, CI= [70.87, 85.92]) than for the non-calculator section (M=62.39, SD=33.71, CI= [48.78, 76.01]) of instructor one. The mean score for the calculator section (M=70.22, SD=27.91, CI= [58.70, 81.74]) was slightly lower than that for the non-calculator section (M=72.78, SD=29.83, CI= [58.82, 86.74]) for instructor two. Figure 8 shows the 95% confidence interval for the mean score by instructor for each pedagogy level.

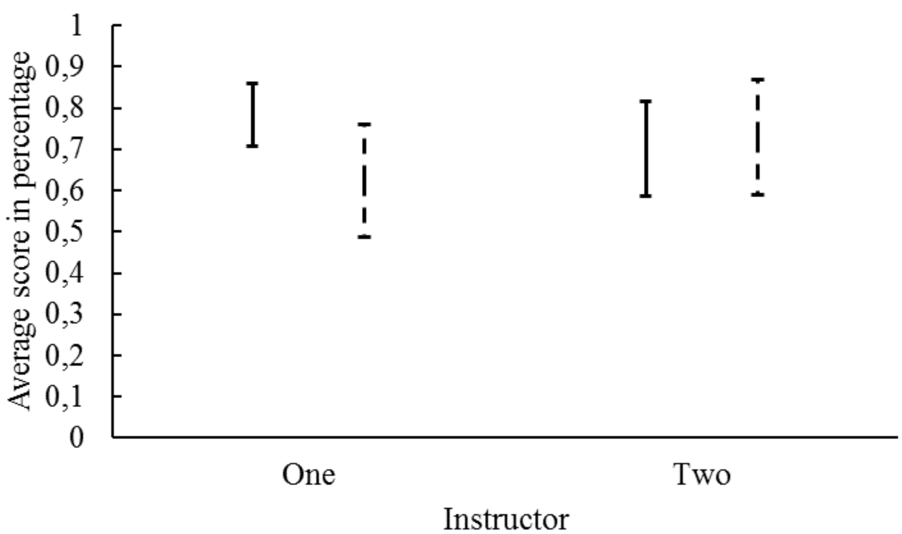

Figure 8. Average score of hypothesis testing question in final exam. The solid bars represent the 95% confidence intervals for the calculator sections, and the dashed bars represent that for the non-calculator sections.

Conclusion and Discussion

In this paper, we investigated whether using TI calculators in an introductory statistics course would help with students' performance for calculating normal probabilities and conducting hypothesis testing, and whether that would facilitate students' conceptual understanding of normal transformation and the difference between one-tailed and two-tailed p-values. Based on our analysis, the graphing calculator clearly helped students to calculate normal probabilities and perform hypothesis testing, although there were little discrepancies between the two instructors on the results of hypothesis testing. The instructor effects may due to the difference in the students' prior knowledge and the course schedule difference. For instructor one, the number of students who had previous statistics courses in calculator section was the same as that in non-calculator section. But for instructor two, the non-calculator section had twice as many students with previous statistics classes than the calculator section. For each week, instructor one's calculator section was scheduled after non-calculator section, while instructor two's calculator section was scheduled before non-calculator section. After teaching one round, the instructor may be more familiar with the materials, hence the learning of the later section may be more effective than the earlier section. That may create some advantages for the non-calculator section of instructor two. Our results showed that the use of TI calculators significantly helped with students' conceptual understanding of the normal transformation when that concept was first introduced, and significantly helped with students' performance of normal probability calculation and hypothesis testing. The study did not show any advantages or disadvantages of using graphing calculators on students' conceptual understanding of the one-tailed and two-tailed p-values.

Limitations

There was instructor effects for this study when testing students' performance of hypothesis testing and the understanding of p-value. The instructor effects may be due to the different between student populations, different prior knowledge, or different teaching methods.

References

Adams, T.L. (1997). Addressing students' difficulties with the concept of function: applying graphing calculators and a model of conceptual change. *Focus on Learning Problems in Mathematics*, 19, 43-51.

American statistics Association "College Report." Report from the Guidelines for Assessment and Instruction in Statistics Education (GAISE) project (2005). Available at http://www.amstat.org/asa/files/pdfs/GAISE/2005GaiseCollege_Full.pdf

Burrill, G. (1997). Graphing calculators and their potential for teaching and learning statistics. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics*. Voorburg, The Netherlands: International Statistics Institute.

Cavanagh, M., and Mitchelmore, M. (2003). Graphics calculators in the learning of mathematics: teacher understandings and classroom practices. *Mathematics Teacher Education and Development*, 2003 (5), 3-18.

Chance, B., Ben-Zvi, D., Garfield, J., and Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education*, 1(1).

Collins, L., and Mittag, K. (2005). Effect of calculator technology on student achievement in an introductory statistics course. *Statistics Education Research Journal*, 4 (1), 7-15. Available at http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(1)_Collins_Mittag.pdf

Doerr, H. M. and Zangor, R. Creating meaning for and with the graphing calculator. *Educational Studies in Mathematics* (2000) 41:143. https://doi.org/10.1023/A:1003905929557.

Dunham, P., and Dick, T. (1994). Research on graphing calculators. *The Mathematics Teacher*, 87, 440-445.

Garfield, J., Chance, B., and Snell, J.L. (2000). Technology in college statistics courses. In D. Holton et al. (Eds.), *The teaching and learning of mathematics at university levels: An ICMI study* (pp.357-370). Dordrecht, The Netherlands: Kluwer Academic Publishers. Available at http://www.dartmouth.edu/~chance/teaching_aids/books_articles/technology.html

Garfield, J. (1995). How students learn statistics. *International Statistical review*, 63 (1), 25-34.

Graham, E., Headlam, C., Sharp, J., and Watson, B. (2008). An investigation into whether student use of graphics calculators matches their teacher's expectations. *International Journal of Mathematical Education in Science and Technology*, 39 (2) 179-196.

Griffith, L. K. (1998). Impact of technology on pedagogy. In E.D. Laughbaum (Ed.), *Hand-held technology in mathematics and science education: A collection of papers* (pp. 76-77). Columbus, OH: The Ohio State University.

Idris, N., Tay, B.L., Goh, L.S., Mahfud, N., Ding, H.E., Aris, A., and Bakar, A.S.A. (2003). *A graphing calculator based instruction and its impact on the teaching and learning of mathematics*. Pulau Pinang, Malaysia: Universiti Sains Malaya Pulications.

Kharuddin, A. F., and Ismail, N. A. (2017). Graphing calculator exposure of mathematics learning in a partially technology incorporated environment. *EURASIA Journal of Mathematics Science and Technology Education.* 2017 13(6):2529-2537. DOI: 10.12973/eurasia.2017.01238a.

Krishnan, S., and Idris, N. (2013). The use of graphics calculator in a matriculation statistics classroom: a Malaysian perspective. *Technology Innovations in Statistics Education*, 7 (2). Available at http://escholarship.org/uc/item/10f0x4gf

Kor, L. K., and Lim, C. S. (2003). Learning statistics with graphics calculator: A case study. *Proceedings of 1st National Conference on Graphing Calculators* (pp. 18-26). Pulau Pinang, Malaysia: Universiti Sains Malaysia Publications.

Kor, L. K., and Lim, C. S. (2004). Learning statistics with graphics calculator: Student' viewpoints. *Integrating Technology in the Mathematical Sciences*, *USM Proceeding Series* (pp. 69-78). Pulau Pinang, Malaysia: Univerisiti Sains Malaysia Publications.

Lazari, A., and Goel S. (2003). Teaching statistics using graphics calculators versus traditional method. *Georgia Journal of Science*, 61(2), 84.

Parrot, M. A. S., and Leong, K. E. (2018). Impact of using graphing calculator in problem solving. *International Electronic Journal of Mathematics Education*, 13(3), 139-148. https://doi.org/10.12973/iejme/2704.

Roorda, G., Vos, P., and Drijvers, P. (2016). Solving rate of change tasks with a graphing calculator: a case study on instrumental genesis. *Digit Exp Math Educ* (2016) 2:228-252. Doi: 10.1007/s40751-016-0022-8.

Hennessy, S., Fung, P., and Scanlon, E. (2001). The role of the graphic calculator in mediating graphing activity. *International Journal of Mathematical Education in Science and Technology*, 32:2, 267-290. DOI: 10.1080/00207390010022176.

Starnes, S. D., Yates, D., and Moore, S. D. (2010). *The Practice of Statistics*, 4th Ed. New York, NY: W. H. Freeman and Company.

Tan, C-K. (2012). Effects of the application of graphing calculator on students' probability achievement. *Computers & Education*, 58, 1117-1126. Doi: 10.1016/j.compedu.2011.11.023.

