2023 HAWAII UNIVERSITY INTERNATIONAL CONFERENCES
SCIENCE, TECHNOLOGY & ENGINEERING, ARTS, MATHEMATICS & EDUCATION JUNE 7 - 9, 2023
PRINCE WAIKIKI RESORT, HONOLULU, HAWAII

PRE-SERVICE ELEMENTARY TEACHERS' CONCEPTUAL UNDERSTANDING OF STATISTICS

CHAPHALKAR, RACHEL M.
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN - WHITEWATER
WHITEWATER, WISCONSIN

Dr. Rachel M. Chaphalkar
Department of Mathematics
University of Wisconsin – Whitewater
Whitewater, Wisconsin

Pre-service Elementary Teachers' Conceptual Understanding of Statistics

Synopsis:

This study compared two different methods of teaching a unit on statistics taught over a four-week portion of a content course for future elementary teachers. One of the units focused on the statistical process and the other was consistent with a standard textbook. This study used the Levels of Conceptual Understanding of Statistics Test in addition to regular coursework to measure students' conceptual understanding.

Pre-service Elementary Teachers' Conceptual Understanding of Statistics

ABSTRACT

Statistics is one of the many content areas of mathematics taught at the university level for future elementary teachers. The GAISE report and Statistical Education of Teachers report give guidelines to focus on the statistical process. However, textbooks for these courses and for elementary school students often focus on disconnected statistics topics. This study compared two different methods of teaching a unit on statistics taught over a four-week portion of a content course for future elementary teachers. Student's conceptual understanding of statistics was measured by the Levels of Conceptual Understanding of Statistics Test in addition to regular coursework.

INTRODUCTION

In the United States, future elementary teachers often complete content courses in mathematics related to content taught in elementary school before taking mathematics methods courses. The credit total of this content requirement typically ranges from 3 to 12 credits, however, the recommendation made by The Mathematical Education of Teachers II (Conference Board of Mathematical Sciences, 2012) is 12 credits. These courses often include units about number and operations, algebra, geometry, measurement, statistics, and probability, and are often taught by mathematicians, mathematics educators, and statisticians. The goal is to strengthen students' conceptual understanding of the topics and to use some of the methods students will later learn about in their methods courses. Often these students have had negative experiences in mathematics courses in the past and so a secondary goal of the course is to help to give them a more positive outlook on mathematics.

Textbooks appropriate for these courses cover all areas of mathematics that the elementary teachers will teach, with a large emphasis on number and operations. Jones and Jacobbe (2014) examined the textbook exercises in five of the most popular textbooks in the United States for these courses and found that most questions focused on data analysis. Few textbook questions asked about formulating questions, collecting data, or interpreting results. In addition, many of these exercises fall into the lower (A and B) levels in the GAISE standards for K-12 students, and certainly, not all exercises are assigned for students to complete.

As an instructor of these courses, I became frustrated in trying to give my students an understanding of statistics that will help them in their future classrooms in a very short period of time. This was a concern because most of my students will not take an additional statistics courses but will be teaching statistics in their future classrooms. I also became familiar with the Statistical Education for Teachers (SET) report (Franklin, Bargagliotti, Case, Kader, Schaeffer, & Spangler, 2015), and wanted my course to align with these recommendations. Together, these experiences prompted my revision of the statistics unit the course to focus more heavily on using the statistical process, as well as statistics related to the future (teaching) careers of my students.

The SET report encouraged a focus on the statistical process as well as active learning and use of technology for the content course(s) for future elementary teachers (Franklin et al., 2015). The report provided an example of an appropriate scenario of a statistical study that would be of interest to future elementary teachers and showed a variety of analysis options using descriptive statistics and graphs. Additional inspiration came from teacher researchers who described activities completed with middle school students in statistics including Smith and Kenlan (2016) whose sixth grade students evaluated how digital games help them learn math.

Roscoe (2016) had future teachers work with eighth grade students investigating bivariate data about cars. Lim, Rubel, Shookhoff, Sullivan, & Williams (2016) had middle school students research and use simulation to learn about winning the lottery. These studies reported that students found these projects interesting and appreciated that they were able to explore topics of interest.

In order to explore the impact of the revised statistics unit, both in general performance in the course and in the conceptual understanding of statistics, the following research questions were posed:

(1) Did pre-service elementary teachers in the revised unit perform better on classroom assessments?

(2) Did pre-service elementary teachers in the revised unit have a greater conceptual understanding of statistics?

METHODOLOGY

This study had a non-equivalent control group design since students were not able to be randomly assigned to course sections.

Participants and Courses

This study took place at a mid-sized four-year comprehensive university in the United States. The participants were pre-service elementary and early childhood teachers who were enrolled in their second and final 3-credit required mathematics content course. This course served as the only content course for these students covering statistics, probability, geometry, and measurement. All students enrolled in the four classes taught by the author over two semesters were invited to participate in the study on the first day of class. 125 students agreed to participate. The revised content unit within this course was taught during the first four weeks of the semester. All data collected were part of the coursework assigned to all students in the class regardless of their willingness to participate in this study. Both sections of the course in Spring 2017 were designated the control group and both sections in Fall 2017 were assigned to the treatment group with the revised statistics unit.

The control sections were taught as a hybrid course with one out of three 50-minute meetings per week as an online day with a mixture of online lecture videos with notetaking, reading, and practice problems. The treatment sections were taught completely face to face, with two 75-minute class meetings per week, and all lecture notes taken during class. The control sections activities included: asking a statistical question discussion, graphing M&M (candy)-related variables, mean as a fair share activity adapted from activity manual for *A Problem Solving Approach to Mathematics for Elementary School Teachers, Tenth ed.* (Billstein, Libeskind, & Lott, 2010), and an activity using boxplots to make comparisons. The treatment sections included the same statistics content and used the same textbook, however, nearly all of the activities were revised to focus on the four parts of the statistical process and data related to teaching. The one activity that did not change was the mean as a fair share activity. One of the new activities was adapted from the SET report's statistical process activity comparing students' standardized test scores based on those who ate breakfast and those who did not (2015). The treatment sections also used one 75-minute class period to use a computer lab to learn how to use Google Sheets in order to make graphs and calculate weighted averages.

Data Collection

Both quantitative and qualitative data were collected to answer the research questions. Data collected included pre and post Levels of Conceptual Understanding of Statistics (LOCUS) assessments, the first in-class exam, the cumulative final exam, and course grades. All students in the course received all of these assignments, regardless of their participation in the study. Students took the beginning/intermediate short form of the LOCUS tests outside of class time both at the beginning of the course and around the first exam, which covered statistics and probability. The LOCUS assessment was chosen to measure conceptual understanding due to its validity in measuring statistical understanding (Jacobbe, Case, Whitaker, & Foti, 2014), endorsement by the American Statistical Association and National Council of Teachers of Mathematics, and cost and availability (free, and available online at www.locus.statisticseducation.org.) To encourage students to take the assessment, they were given participation credit for class each time they took the LOCUS assessment, however, they did not receive their scores on this assessment.

Data Analysis

The identifying characteristics of participants were removed from data and participants were assigned pseudo-IDs. Since the participants were not randomly assigned to the treatment and control groups, data were analyzed using descriptive statistics. Participants who spent less than ten minutes taking the 45-minute LOCUS assessment were excluded from the analysis because it seemed unlikely that they read all of the questions in that amount of time. The remaining test times and scores were plotted in a scatterplot with no clear association between these variables. Participants who did not take both LOCUS assessments were also excluded from the analysis since it was not possible to pair their pre-tests and post-tests.

RESULTS

The first research question looked at the differences results of the control and treatment groups on two classroom assessments: the first test and the final exam. Both of these exams included topics not covered in the statistics unit. The final exams were nearly identical, however the first exams were not.

Test 1 Results

Test 1 had slightly different formats for the two groups based on the assigned class schedule: the control group had a 50 minute exam, while the treatment group had a 75 minute exam. The treatment group had more multiple-choice questions on this exam, which may have made it more difficult. In addition, this exam also tested the probability unit in addition to the statistics unit. Table 1 gives the means and standard deviations of students' scores on the first exam.

Table 1. *Overall Scores (out of 100) on Test 1*

	Test 1 Mean	Standard Deviation	Number of students
Control	76.2	13.8	65
Treatment	71.0	11.6	59

Some of the questions from the first exam that were similar between the two semesters were collected for additional analysis. Table 2 looks at the results of a question identifying whether a variable was numerical or categorical. On the test for the control group, this question also asked for an explanation, while on the test for the treatment group, this was a multiple-choice question without an explanation. 91.5% of the students in the treatment group got this question correct.

Table 2. *Test 1 Question on Identification of Variable Type*

	Control	Treatment
Percentage Correct	67.7	91.5
Percentage Incorrect	32.3	8.5
Number of Students	65	59

Both tests contained a question that asked students to create a graph from categorical data provided. The two coders looked for an appropriate graph type (bar graph, pictograph, or pie chart), correct labeling or the axes and an appropriate title, and also that there were no other errors in the graph, such as bars in a bar graph that were touching. Nearly all students drew appropriate graphs, and over 75% in both groups gave an appropriate title and labeled all axes correctly.

Table 3. *Test 1 Question on Graph Creation Student Work*

	Appropriate graph type	Appropriate labels of axes and title	Correctness of graph apart from title and axes labels	Number of students
Control	98.5	76.9	92.3	65
Treatment	100	78.0	89.8	59

In addition, on both tests, students were asked to use a table to find the relative frequency of one of the variables. The results are given in Table 4. In the treatment group, 89.8% of students found the relative frequency, while the other 10.2% of the students had major errors.

Table 4. *Test 1 Question on Calculation of Relative Frequency Student Work*

	Control	Treatment
Percentage correct	70.8	89.8
Percentage with minor error(s)	10.8	0
Percentage major error(s)	18.5	10.2
Number of students	65	59

Final Exam Results

The only other exam in the class that contained material from the statistics unit was the final exam. The final exam only had about 20% on statistics content and was nearly identical for both groups of participants. The overall final exam scores are given in Table 5.

Table 5. *Overall Scores (out of 100) on Final Exam*

	Final Exam Mean	Standard Deviation	Number of Students
Control	71.6	13.1	65
Treatment	72.6	10.5	60

The final exam questions that were related to statistics were scored using a rubric. There were eight questions, some of which were part of multi-part questions. Table 6 gives the mean number correct on this subset of the final exam questions. The treatment group had a mean of 4.53 questions correct out of the 8 statistics-related questions on the final exam, with a standard deviation of 1.49. This was lower than the mean number correct of the control group by approximately 0.5 questions.

Table 6. *Final Exam Questions Related to Statistics out of 8*

Final Exam	Mean Number Correct	Standard Deviation	Number of Students
Control	5.00	1.64	65
Treatment	4.53	1.49	60

LOCUS Assessment Results

To answer the second research question, which focused on the conceptual understanding of students, the LOCUS assessment scores of the control and treatment groups were compared.

The LOCUS assessment provided overall scores, as well as scores in the four parts of the statistical process (formulate questions, collect data, analyze data, and interpret results.) Each part was scored out of 100%. Table 7 gives the pre-test percentage scores of the control group and treatment group. The treatment group scored lower than the control group on each area except *Collect Data*, where they were stronger on average by just over 5%.

Table 7. *Pre-test LOCUS Scores*

	Overall Mean (S.D.)	Formulate Questions Mean (S.D.)	Collect Data Mean (S.D.)	Analyze Data Mean (S.D.)	Interpret Results Mean (S.D.)	Number of Students
Control	61.2 (12.7)	73.1 (19.7)	54.2 (20.3)	49.0 (19.3)	80.4 (17.0)	52
Treatment	56.7 (16.4)	73.9 (20.9)	59.5 (22.2)	41.6 (20.0)	66.8 (26.9)	44

A summary of the mean differences from the pre and post LOCUS assessments of the control and treatment groups in found in Table 8. The treatment group had a 7.77% increase in the overall score from the pre-test to the post-test, while the control group had a decrease of 1.33% from the pre-test to the post-test. The only increase made by the control group was in the data collection portion, which was an increase of 10.77%. This was also the only part of the pre-test that the control group had scored lower than the treatment group (Table 7.) The largest increase for the treatment group was made in the analyze data portion, with a 15.32% increase, while their smallest gain was made in the collect data portion (0.45%).

Table 8. *Pretest and Posttest Score Differences on LOCUS Assessment*

Score Category	Control (*n*=52)		Treatment (*n*=44)	
	Mean Difference	Standard Deviation	Mean Difference	Standard Deviation
Overall	-1.33	13.26	7.77	15.21
Formulate Questions	-3.85	24.94	4.55	25.45
Collect Data	10.77	27.85	0.45	29.09
Analyze Data	-5.63	23.88	15.32	18.03
Interpret Results	-3.85	22.76	5.45	28.40

DISCUSSION AND CONCLUSIONS

Overall, there was a small increase in the students' conceptual understanding as measured by the LOCUS assessment for the treatment group, as well as small differences between the treatment and the control groups in the in-class assessments. The questions from the in-class assessments were often measuring disjoint topics or pieces within the statistical process and often included calculations. On the other end of this spectrum were the questions on the LOCUS assessment, which were very conceptual and contained no calculations.

The lack of improvement in the treatment group as compared to the control group on the classroom assessments indicated that the time spent working on the statistical process did not harm their learning from the textbook content of the course. This is not the most encouraging result; however, the class time was spent in ways that did not always fit the textbook (and the "old") questions on the test. The lower scores on the LOCUS pre-test might indicate that the treatment group was a bit weaker than the control group at the start of this study. In addition, as the instructor of the course, I think that the treatment group's Test 1 was harder overall because there were more multiple-choice questions instead of short answer questions, which may have contributed to lower test scores on Test 1.

The LOCUS assessment scores did show a small increase in the conceptual understanding of the students in the treatment group, though the increase was smaller than desired. After excluding students who either spent less than 10 minutes on one of the LOCUS assessments or did not take both, only about 75% of the students in the class took both of the pre- and post- LOCUS assessments. In addition, students never received their scores on the assessment because I did not want low scores to discourage them or to give them a negative attitude toward statistics or the class. I also avoided telling the students answers directly to these assessment questions to keep students from trying to memorize the test. Finally, it is likely that some students may not have given their best effort because this assessment was not graded and completed outside of class.

Implications for Teaching

While only small gains in conceptual understanding were seen in the revised course, my instructor buy-in was greater because I felt that students were getting more of the content knowledge for teaching statistics that they needed. After this first revision of the unit, there were additional changes that could be made such as having students complete a small statistics study. While test results may not show it, it did seem like the students in the treatment class saw the usefulness of learning statistics as future teachers.

Research has also shown that in-service elementary teachers lack understanding of statistical content. Jacobbe and Horton (2010) found that even excellent elementary teachers struggled with numerical and categorical comparisons, and Begg and Edwards (1999) found that preservice teachers were better than in-service teachers at answering questions about mean, median, mode. Professional development activities for in-service teachers in statistics should also focus on the relevance of statistics in the classroom and on conceptual understanding.

Limitations and Future Work

This was a small quasi-experimental mixed methods study, which limits the population for which results can be generalized to. In addition, there was a lack of diversity in the participants, in part due to the University location and student body, and also due to the course being designed for a particular group of students. Finally, the statistics unit of this course should undergo additional revisions, with additional time put into making statistics relevant for future teachers.

ACKNOWLEDGEMENTS

Special thanks three students who are now teachers who assisted with this study: Amanda Alms helped the author to design this study, and Alexander Crogan and Jake Krueger assisted with data analysis.

REFERENCES

Billstein,R., Libeskind, S., & Lott, J. (2010). *A Problem Solving Approach to Mathematics for Elementary School Teachers* (Tenth edition). Boston: Addison – Wesley.

Begg, A., & Edwards, R. (1999). Teachers' Ideas about Teaching Statistics. Paper presented at the Joint Conference of the AARE and NZARE, Melbourne.

Conference Board of the Mathematical Sciences (2012). *The Mathematical Education of Teachers II*, Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.

Franklin, C. A., Bargagliotti, A. E., Case, C. A., Kader, G. D., Schaeffer, R. L., & Spangler, D. A. (2015). The Statistical Education of Teachers. Commissioned by the American Statistical Association. www.amstat.org/asa/files/pdfs/EDU-SET.pdf

GAISE College Report ASA Revision Committee. (2016). *Guidelines for Assessment and Instruction in Statistics Education College Report 2016*. Retrieved from The American Statistical Association. www.amstat.org/education/gaise

Jacobbe, T., Case, C., Whitaker, D., & Foti, S. (2014). Establishing the validity of the LOCUS assessments through an evidenced-centered design approach. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics*, Flagstaff, Arizona, USA. Retrieved from: www.iase-web.org

Jacobbe, T. & Horton, R. M. (2010). Elementary School Teachers' Comprehension of Data Displays. *Statistics Education Research Journal, 9*(1), 27 – 45. Retrieved from: www.stat.auckland.ac.nz/serj

Jones, D. L. & Jacobbe, T. (2014). An analysis of the statistical content in textbooks for Preservice Elementary Teachers. *Journal of Statistics Education, 22*(3), p. 1-14. Retrieved from: www.amstat.org/publications/jse/v22n3/jones.pdf

Lim, V., Rubel, L., Shookhoff, L., Sullivan, M., & Williams, S. (2016). The Lottery is a Mathematics Powerball. *Mathematics Teaching in the Middle School, 21*(9), 526-532.

National Governors Association Center for Best Practices and Council of Chief State School Officers (2010), *Common Core State Standards for Mathematics.* Retrieved from http://www.corestandards.org/Math

Roscoe, M. (2016). A Vehicle for Bivariate Data Analysis. *Mathematics Teaching in the Middle School, 21*(6), 348-356.

Smith, C.P. & Kenlan, K. (2016). Invention Activities Support Statistical Reasoning. M*athematics Teaching in the Middle School, 21*(6), 358-363.

This is Statistics. (2016). "Why you need to study statistics." [video file] Retrieved from: www.youtube.com/watch?time_continue=180&v=wV0Ks7aS7YI