IMPLEMENTING A SIMULATION-BASED INFERENCE CURRICULUM IN INDONESIA: A PRELIMINARY REPORT

Kie V. I. Saputra[1] and Olivia Couch[2]
[1]Universitas Pelita Harapan, Tangerang, Indonesia
[2]Dordt College, Sioux Center, IA, USA
kie.saputra@uph.edu

Simulation-based inference methods are growing in popularity around the world, but to date much of the innovation and assessment has been focused in a few, English speaking countries (e.g., U.S., Australia, Canada). We recently began a multi-step initiative to implement and assess a simulation-based inference curriculum at an Indonesian university. In this paper we will discuss our implementation and provide assessment data comparing student performance on standardized conceptual assessments of student understanding using a traditional introductory statistics curriculum as compared to a simulation-based inference approach.

INTRODUCTION

Introductory Statistics (Statistics 101) is taught as a part of almost every major in Indonesia whether it is in engineering, natural science or social science. It can have many names such as Elementary Statistics, Introductory to Statistics or Biostatistics, but much of the core statistical content is the same. Importantly, most students are still afraid of such subject and statistical literacy is not getting better, so there has to be an effort to change either the way of teaching or the content of statistics course. As far as the author knows, there have not been much local nor national efforts in Indonesia addressing this issue.

The typical Indonesian statistics curriculum still follows the algebra-based introductory statistics curriculum, which consists of four parts, i.e. descriptive statistics, data collection and design, probability and sampling distributions, and inferential statistics. Cobb (2007) and others (e.g., Tintle et al. 2011, Tintle et al. 2012, Chance et al. 2017) have discussed potential problems as a result of this sequencing and proposed simulation-based inference (SBI; the use of computationally intensive methods like simulation, bootstrapping and permutation tests to enhance student understanding of the core logic of inference). They argue, and provide some preliminary data illustrating improved student learning outcomes about the core logic of inference when using SBI.

While much of the existing literature to date has focused on the U.S. and other English speaking countries, we believe that much of the challenges in teaching introductory statistics hold true in Indonesia, though limited quantitative data exists. The need of new curriculum in Indonesia is also motivated by the fact that some statistical concepts were already taught in high school. In Indonesia, high school education has already seen much of the material in the first and second parts of the traditional statistics course. Indonesian high school education currently uses the 2013 curriculum (see Permendikbud No 69/2013) which exposes students to learn descriptive statistics. They will learn as much as center data, variability, histogram and other concept of presenting data. This means that it is reasonable to restructure the statistics 101 course in the first year university.

In this research study, we want to address the following issues. (1) To introduce the new simulation based inference (SBI) in elementary statistics to statistical instructors all over Indonesia, (2) To assess our statistical education, which still uses the traditional curriculum, (3) To compare the assessment result between the traditional curriculum and the simulation based curriculum, (4) To explore the possibilities on expanding the new curriculum across Indonesia. This article mainly focuses on issue 2 and issue 3. Note that for the course material, we will be closely following the book by Tintle et al. (2014), where students first see a simulation based approach to analyze the data, followed by the equivalent theory-based test. The assessment that we use is conceptual inventory that has been used previously in similar studies (Chance et al. 2017). In the last section of this paper, we will also discuss problems on the implementation of the new curriculum and this is somewhat related to issue 4 above.

METHODS

Data was collected over a three semester period, across three instructors, and five class sections. Table 1 illustrates the distribution of students, instructors, semesters and curricula. The simulation-based curricula used Tintle et al. (2014), a course that focuses on SBI, whereas the other sections used traditional statistics such as Walpole et al. (2012) or similar textbooks. ~~XXXXX~~

Table 1. Descriptive statistics on the sample

Instructor	Curriculum	Semester	Sample size	Response rate (total class size)
A	Simulation (Sim 1)	Fall 2016	66	90% (66/73)
A	Simulation (Sim 2)	Spring 2017	22	76% (22/29)
B	Traditional (Trad 1)	Spring 2016	7	31% (7/23)
B	Traditional (Trad 2)	Fall 2016	22	69% (22/32)
C	Traditional (Trad 3)	Spring 2016	13	46% (13/28)

Students completed a 30 question conceptual inventory (Chance et al. 2016). The score is indicated by percentage, in which a 100% score means that student answered all questions correctly. Students were not given credits for completing the test. In Spring 2016, the pre and the post test were paper-administered and conducted in the first week and in the last week inside a class. In Fall 2016 and Spring 2017, students completed the same survey in the first week and in the last week of class. This time, the test was administered electronically. As noted in the table above, only one instructor used the simulation based curriculum and two different instructors used the traditional curriculum.

RESULTS

Table 2 illustrates the overall change in average conceptual score by section and by curriculum. Both of the simulation-based inference courses showed significant improvement, compared to only one of the three traditional courses. However, the gain for one of the traditional courses was substantial (11.4%).

Table 2. Overall mean and SD Pre and post by section

Section	Pretest Mean (SD)	Posttest Mean (SD)	Change (Post-Pre) Mean (SD)
Simulation-combined (n=88)	42.3% (9.9%)	45.8% (9.0%)	3.5% (9.3%)**
Sim 1 (n=66)	44.6% (8.4%)	47.3% (9.8%)	2.7% (9.1%)*
Sim 2 (n=22)	35.5% (7.1%)	41.3% (9.0%)	5.8% (9.8%)*
Traditional-combined (n=42)	40.3% (14.0%)	46.1% (8.2%)	5.8% (12.1%)***
Trad 1 (n=7)	40.0% (4.0%)	39.2% (9.3%)	-0.8% (8.7%)
Trad 2 (n=22)	41.6% (14.5%)	53.1% (11.4%)	11.4% (11.9%)***
Trad 3 (n=13)	38.0% (8.6%)	37.8% (8.2%)	-0.2% (10.0%)

Statistical significance determined by paired t-tests. ***$p<0.001$; **$p<0.01$; *$p<0.05$.

Figure 1 illustrates the overall change (post-test minus pre-test) for each of the five sections illustrating Trad-2 performed much better than other sections, while the two simulation-based inference courses showed modest improvement and the other two traditional curricula sections showed little to no improvement for many students.

Table 3 stratifies the sample into three groups based on pre-test scores. Simulation-based students outperformed traditional students in the lowest group, while the reverse was true for more typical students. Students doing the best on the pre-test showed evidence no improvement in either curriculum. Table 4 gives the estimated effect of the simulation based curriculum relative to the traditional curriculum on change in scores, stratified by pre-test score group. Positive numbers represent better performance (larger pre to post-test changes) for the simulation-based curriculum.

Negative numbers represent better performance by students in the traditional curriculum. The model equation was:

Change in Score = Curriculum (1=Simulation; 0=Traditional) + Random Effect (Section)

where we ran the model separately on each of the three pre-test groups, and included a random effect for the five different sections.

Figure 1. Boxplots of change (Post-test minus Pretest) by section

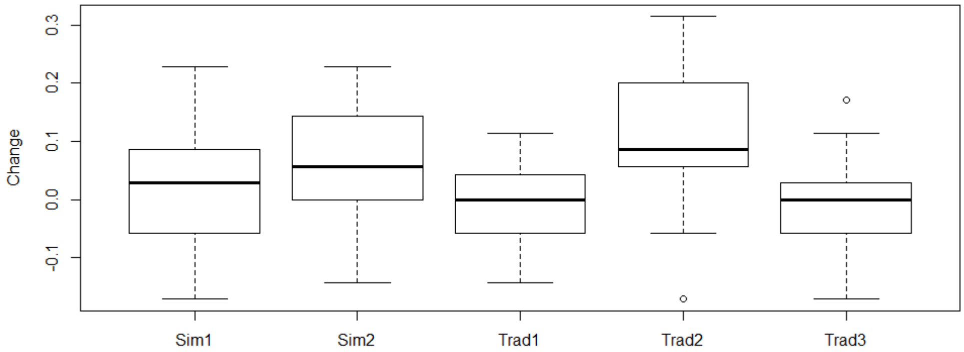

Table 3. Overall mean and SD by Pre-and post by pre-test group and curriculum

Pre-test strata	Pre-test Mean (SD)	Post-test Mean (SD)	Change (Post-Pre) Mean (SD)[1]
Lower (Below 40%)			
Simulation	34.1% (5.2%)	41.7% (8.9%)	7.6% (9.4%)***
Traditional	34.7% (4.5%)	39.8% (9.1%)	5.0% (8.6%)**
Middle (40-50%)			
Simulation	46.1% (2.4%)	47.1% (8.9%)	1.0% (7.7%)
Traditional	45.2% (2.5%)	56.9% (15.8%)	11.7% (15.6%)*
Highest (Over 50%)			
Simulation	54.6% (3.9%)	53.2% (9.4%)	-1.4% (8.3%)
Traditional	54.3% (1.8%)	52.4% (14.4%)	-1.9% (14.6%)

Statistical significance determined by paired t-tests. ***p<0.001; **p<0.01; *p<0.05.

Table 4. Estimated effect of the simulation based curriculum vs. traditional curriculum on change in score (Pre-post)[1]

Content area	Lowest pre-test group Effect (SE)	Middle pre-test group Effect (SE)	Highest pre-test group Effect (SE)
Overall	0.03 (0.04)	-0.10 (0.11)	0.04 (0.14)
Confidence intervals	0.03 (0.06)	-0.21 (0.17)	0.09 (0.35)
Data Collection	0.12 (0.07)	0.08 (0.05)	-0.07 (0.15)
Descriptive Stats	0.05 (0.05)	-0.01 (0.15)	-0.01 (0.11)
Scope of Inference	0.11 (0.12)	-0.31 (0.23)	-0.06 (0.60)
Significance	0.05 (0.06)	-0.09 (0.13)	0.14 (0.47)
Simulation	-0.17 (0.10)	-0.20 (0.24)	0.10 (0.10)

Notably, the simulation-based curriculum did not show significant evidence of an impact overall or in any of the three subgroups compared to the traditional curriculum, though the majority of effects estimates were positive for the lowest group and negative for the middle group.

Finally, to investigate where the simulation-based inference curriculum showed the most evidence of a positive effect we ran similar models (see Model Equation above) for each of the 30 questions on concept inventory for the lowest pre-test group. Four of the thirty questions showed significant effects of the simulation based curriculum ($p<0.05$). Three of questions showed significant improvement for the simulation-based curriculum (misconceptions about confidence intervals; inference from a randomized experiment; and inference from dotplots). The question that showed significantly poorer performance in the simulation-based curriculum students was on predicting variability in a sample proportion.

DISCUSSION AND CONCLUSION

Early papers (e.g., Tintle et al. 2011, 2012, 2014; Chance et al. 2017, among others) showed promising results from the implementation of the simulation based curriculum. We sought to investigate whether this potentially could be the case in Indonesia. In this preliminary implementation at a single institution we did not find strong statistical evidence of widespread student improvement when using the simulation based inference curriculum. However, notably, four separate questions on the survey did show significant differences between curriculums, with three showing significantly better performance by students taking the SBI curriculum and in areas of focus of SBI curricula.

The lack of statistically significant findings could be due to a number of factors. Foremost among these reasons is that there were only a small number of instructors and sections, meaning that instructor effects and curricular effects were highly conflated. Furthermore, we have identified numerous ways we would like to improve the implementation of the simulation-based inference curricula in future semesters including (a) additional faculty training for instructors teaching simulation-based methods, (b) more relevant data contexts for Indonesian students, (c) laboratory/active learning opportunities for all students (the SBI students in this sample did not have those opportunities, while students in the traditional class did), and (d) better ways to incentivize students to participate in and take seriously the assessments. This last point is worth particular comment because better response rates were obtained in the SBI course, which could lead to better performing students participating in the post-test for the traditional curriculum as has been noted in prior studies of this type (Tintle et al. 2012). After addressing these areas of improvement we plan to redo the assessments to continue to monitor and evaluate the impact of the SBI curriculum in Indonesia.

ACKNOWLEGEMENTS

We acknowledge PEER-USAID grant AID-OAA-A-11-00012 and NSF-DUE-1323210.

REFERENCES

Chance B, Wong J, Tintle N (2017) Student Performance in Curricula Centered on Simulation-based Inference: A preliminary report. *Journal of Statistics Education*, *24*, 114-126.

Cobb, G. (2007). The introductory statistics course: a Ptolemaic curriculum? *Technology Innovations in Statistics Education*, *1*(1).

Ministry of Education and Culture. (2013). *Kompetensi Dasar dan Struktur Kurikulum SMA*. Permendikbud No. 69/2013.

Tintle, N. L, Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., & VanderStoep, J. (2014). *Introduction to statistical investigations* (Preliminary edition). Wiley.

Tintle, N., VanderStoep, J., Holmes, V. L., Quisenberry, B., & Swanson, T. (2011). Development and assessment of a preliminary randomization-based introductory statistics curriculum. *Journal of Statistics Education*, *19*(1).

Tintle, N. L., Topliff, K., VanderStoep, J., Holmes, V. L., & Swanson, T. (2012). Retention of statistical concepts in a preliminary randomization-based introductory statistics curriculum. *Statistics Education Research Journal*, *11*(1), 21.

Tintle, N. L., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., & VanderStoep, J. (2013). Challenging the State of the Art in Post-Introductory Statistics: Preparation, Concepts, and Pedagogy. In *Proceedings of the 59th ISI World Statistics Congress*, (Session IPS032), 295–300.

Tintle, N. L., Rogers, A., Chance, B., Cobb, G., Rossman, A., Roy, S., ... & VanderStoep, J. (2014). Quantitative evidence for the use of simulation and randomization in the introductory statistics course. In *Proceedings of the International Conference on Teaching Statistics*. Flagstaff, Arizona.

Walpole, R.E., Myers, R.H., Myers, S.L.,Ye, K. (2012), *Probability and statistics for engineers and scientists* (9th Edition). Prentice Hall.