ASSESSING RETENTION OF STATISTICAL CONCEPTS AFTER COMPLETING A POST-SECONDARY INTRODUCTORY STATISTICS COURSE

Stephanie Mendoza and Soma Roy
Department of Statistics, Cal Poly – San Luis Obispo, CA 93407
soroy@calpoly.edu

Educators fear that retention of course concepts in science and mathematics often decreases in the time after course completion. Few studies have explored patterns of retention in statistics courses. Tintle et al. (2012) found poor retention of statistical concepts among students at a single institution in a first statistics course four months after course completion, but with more persistence among students taking a modified version of the course centering on simulation-based inference. Here we report findings from a multi-institution study of retention both four-months (648 students) and 16-months (170 students) after completion of an algebra-based course, discuss retention patterns on a variety of key statistical concepts, and highlight similarities and differences in retention based on a variety of curricular choices (e.g., use of simulation-based methods).

BACKGROUND

"Simulation-based inference" (SBI) has been gaining in popularity in undergraduate statistics courses as an improved pathway to understanding key concepts of statistical inference and the statistical process as a whole. The potential advantages of a simulation-based curriculum are exposing students to principles of statistical inference from day one of the course, and building on this knowledge throughout the course. Preliminary assessment results have showed improved student conceptual understanding using SBI-curricula (e.g., *Introduction to Statistical Investigations, Unlocking the Power of Data*) compared to non-SBI courses at a single institution (e.g., Tintle et al., 2011), and in a cross-institutional study after adjusting for student and instructor characteristics (Chance, Wong, & Tintle, 2016). But few studies have focused on student retention (see Tintle et al., 2012 for one example). The question posed here is whether there is also evidence of improved retention of key concepts after courses using SBI compared to non-SBI courses.

METHODS

Assessment

The primary instrument consisted of attitude and conceptual components, along with additional demographics and background questions. The attitude component consisted of 36 questions from the Survey of Attitudes Towards Statistics (SATS-36; Schau, 2003), measuring students' perspectives on statistics for six subscales: affect, cognitive competence, perception of difficulty, expected effort, interest level, and value of statistics. The concept component of the survey consisted of 35 questions based heavily on the Comprehensive Assessment of Outcomes in a First Statistics Course (CAOS; delMas, Garfield, Ooms, & Chance, 2007). Multiple questions on the same concept (e.g., "check all that apply," series of valid/invalid statements) were combined together for a total of 22 points that each assessed separate statistical topics. These questions were also divided into five subscales: DC: data collection and scope of conclusions, DS: descriptive statistics, CI: confidence interval, TS: tests of significance, Sim: simulation and sampling distributions. The remaining 24 questions considered basic demographic information (e.g., gender, age, field of study), and mathematical and statistical background (e.g. SAT/ACT score, GPA, previous statistics courses). Each participant completed the assessment at the beginning of his/her introductory statistics course and at the conclusion of the course. Subsamples of students also completed the assessment again four and sixteen months later.

A separate survey was administered to instructors, consisting of questions about themselves, their institution, and their teaching experience. Information was also collected on logistics of the courses the instructors taught, including the type of institution, prerequisite for the class, textbook, class size, days and times the class met, amount of class time spent lecturing, and the incentive used to encourage students to take the survey (e.g., homework points or exam scores).

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

Participants

Instructors of introductory statistics courses across the nation were invited to have their class complete the assessment. Student participation was voluntary, though pre and post test completion was often awarded with course credit, and a ten dollar gift card incentivized students to take the retention assessments. The student survey was administered pre/post/4-month retention over two different years, 2015 and 2016. In the year 2015-16, 298 (73.4% female) students completed the pre-test, post-test, and 4-month retention assessments. Of those 298 students, 170 also completed the 16-month retention assessment. In the year 2016-17, 350 students (68.6% female) completed the pre-test, post-test, and 4-month retention assessments (16-month results will be reported later for this cohort). For both cohorts, the majority are from four-year colleges or universities, with a small amount from comprehensive universities and two-year colleges.

Statistical analysis

The main response variable used in this analysis was the percentage of concept questions correct in each assessment. Student scores on the concept subscales were also examined to explore differences in retention of different statistical topics. These scores are compared over three different curriculum types. "ISI" includes classes that followed a simulation-based curriculum using the *Introduction to Statistical Investigations* textbook. "Other SBI" included courses that followed simulation-based curriculum, using a non-ISI textbook. "Not SBI" courses followed a more traditional, non-simulation based curriculum. Due to low intra-class correlation coefficients and low response rates per section, in this paper we are considering observations within sections to be independent. Paired *t*-tests were used to explore the average difference in conceptual understanding from the post assessment to the four-month and 16-month retention assessments.

RESULTS

Four-month retention

Table 1 depicts the overall percentage correct for each time point for each type of curriculum and for the five concept subscales. Performance was found to be similar in both cohorts, so this table portrays all four-month post-test students from years 2015 and 2016 whose instructor identified a recognizable textbook. Overall, students in SBI courses have lower pre-test but higher post-test and retention percentages correct (*p*-value for comparing ISI, OtherSBI, not SBI retention = 0.0544). This pattern generally holds up across the subscales, except for Descriptive Statistics.

Table 1. Percentage of Concept Questions Correct (Change = 4-month – post)

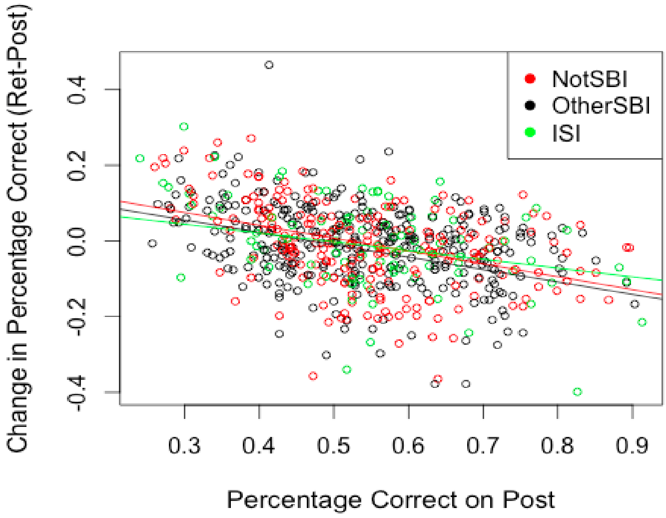
		Pre-test	Post-test	4-Month	*p*-value[1]	Change	*p*-value[2]
Overall	ISI (n = 197)	0.456	0.567	0.550	0.054	-0.017	0.028
	OtherSBI (n = 127)	0.496	0.549	0.533		-0.016	0.115
	NonSBI (n = 284)	0.473	0.525	0.519		-0.005	0.440
Data Collection and Scope of Conclusions	ISI	0.480	0.599	0.601	0.086	0.002	0.921
	OtherSBI	0.554	0.607	0.608		0.001	0.971
	NonSBI	0.518	0.532	0.568		0.036	0.010
Descriptive Statistics	ISI	0.476	0.564	0.547	0.652	-0.017	0.265
	OtherSBI	0.548	0.572	0.568		-0.004	0.836
	NonSBI	0.522	0.580	0.563		-0.017	0.206
Confidence Intervals	ISI	0.331	0.478	0.440	0.711	-0.038	0.027
	OtherSBI	0.379	0.468	0.423		-0.045	0.045
	NonSBI	0.345	0.471	0.425		-0.046	0.005
Significance Tests	ISI	0.580	0.695	0.681	0.006	-0.014	0.323
	OtherSBI	0.600	0.648	0.641		-0.007	0.697
	NonSBI	0.580	0.609	0.620		0.011	0.399
Simulation and Sampling Distributions	ISI	0.323	0.386	0.377	0.004	-0.009	0.622
	OtherSBI	0.298	0.345	0.315		-0.028	0.189
	NonSBI	0.313	0.325	0.305		-0.019	0.206

[1]*p-value for comparing curricula on 4-month retention concept scores*

[2]*p-value for paired t-test of percentage of correct concept questions (four-month retention – post-test)*

However, students in SBI curricula were also more likely to have taken a follow-up statistics course by the time of the retention test (ISI: 24%, other SBI: 30%, non-SBI: 20%). These higher averages at four-months are despite tending to have larger decreases in their scores from end of course post-test. The ISI students actually showed the largest drop in performance on a question related to designing a simulation (repeating the study vs. simulating outcomes under the null hypothesis) to finish on par with the non-SBI students. The ISI students did show improvement on questions related to simulation (including recognizing a proper simulation design) and sampling distributions, and all three groups shows the most loss in performance on the confidence interval subscale. Although insignificant ($p = 0.328$), Figure 1 shows weak evidence of an interaction between type of curricula and post-test on retention—students who score higher on the post-test tend to have larger negative changes (retention – post), but the ISI group shows less of a post-test effect. We continue to analyze how much of this is due to individual questions.

Figure 1. Interaction between post-test score and change in percentage correct (4 mos – post)



Sixteen-month retention

Table 2 shows similar trends for the subset of the 2015-2016 4-month retention students who chose to participate in the 16-month retention study. The students in ISI courses have lower pre-test but generally higher 16-month retention scores, although students' scores continue to decrease more from post-test to 16-month retention assessments in the SBI groups. Questions where students in simulation-based courses had large decreases from post to 16-month retention assessments were primarily two simulation questions and one confidence interval question. Students had the highest 16-month retention in significance tests (ISI and not SBI), descriptive statistics (other SBI).

Table 2. Percentages correct by textbook and subscale for 16-month retention students

Textbook	Pre	Post	Ret (4)	Ret (16)	Change[1]	DC[2]	DS	CI	ST	Sim
ISI ($n = 61$)	0.450	0.562	0.549	0.531	-0.031	-0.079	-0.041	-0.027	0.005	-0.045
OtherSBI ($n = 26$)	0.532	0.596	0.558	0.543	-0.053	-0.161	0.008	-0.092	-0.023	-0.013
NotSBI ($n = 61$)	0.463	0.510	0.539	0.499	-0.012	0.002	-0.066	-0.061	0.034	-0.004
Overall	0.475	0.550	0.546	0.517	-0.033	-0.061	-0.047	-0.050	-0.004	-0.025
p-value[3]	0.528	0.005	0.521	0.100	0.164	0.915	0.421	0.477	0.067	0.083

[1]*Change in percentage correct from 16-month assessment to post assessment.*
[2]*Similar set up for five concept subscales.*
[3]*p-value for two-sample t-test of percentage of concept questions correct between SBI and NotSBI groups.*

We are still exploring student and instructor level variables. For example, students who reported (at 16 months) they took this course "because it sounded interesting" have significantly higher 16-month retention scores than those who reported they took the course for a requirement ($p = 0.004$). In general, there is evidence of interaction with more positive attitudes (pre-course and 16-month) corresponding to higher 16-month retention scores (*effort* being the exception). There is also some evidence that students whose course met later in the day have higher average 16-month retention scores compared to those who met earlier in the day ($p = 0.097$).

DISCUSSION

In this paper, we have presented an analysis of retention among students in introductory statistics courses across 33 instructors. Four months after completion of an introductory statistics course, students in SBI courses were found to have higher retention of statistical concepts compared to non-SBI students, especially on questions related to significance tests and simulation. despite having lower pre-tests scores on average. However, there is a pattern of larger decreases from post-test to 4-month for the ISI students, primarily due to a question distinguishing repeating a study with modeling the null.

The results pertaining to 16-month retention rates is perhaps more telling of what students might retain of statistical concepts in the long-run. The 170 students used in this portion of analysis were similar in demographics, attitudes, and achievement on the pre, post, and four-month retention assessments compared to the full sample. Similar to the four-month retention results, there is a pattern of students with simulation-based courses having slightly higher overall performance, especially on questions related to significance and simulation, but with larger decreases post to 16-months. From this, we conclude that students enrolled in these simulation-based classes are learning and retaining at least an equivalent amount of statistical concepts compared to those in more traditional, non-SBI courses.

Although 16-month retention was not found to be associated with the curriculum used in the course, we continue to explore factors such as lecture time, experience of the instructor, and the instructor's GAISE familiarity to help us understand aspects of a course that produce higher conceptual gains for students. For example, there is preliminary evidence that students enrolled in classes that meets later in the day tend to have higher retention rates 16 months after completion of the course.

Also of note is that attitudes about statistics prior to the course and 16-months after completing their introductory course were significantly positively correlated with their scores on the concept test. This relationship is especially true for students' *cognitive competence* and *value* rating of statistics. The former indicates that students are able to accurately gauge their own understanding of statistical content.

ACKNOLWEDGEMENTS

This research was supported by NSF grants NSF/TUES/DUE-Phase II # 1323210.

REFERENCES

Chance, B., Wong, J., & Tintle, N. (2016). Student performance in curricula centered on simulation-based inference: A preliminary report. *Journal of Statistics Education, 24*(3), 114–126.

delMas, R., Garfield J., Ooms, A., & Chance, B. (2007), Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28–58.

Schau, C. (2003). *Survey of Attitudes Toward Statistics* (http://evaluationandstatistics.com/)

Tintle, N., Clark, J., Fischer, K., Chance, B., Cobb, G., Roy, S., Swanson, T., & VanderStoep, J. (to appear). Assessing the association between pre-course metrics of student preparation and student performance in introductory statistics: results from early data on simulation-based inference vs. non-simulation based inference. *Journal of Statistics Education*.

Tintle, N., Topliff, K., VanderStoep, J., Holmes, V., and Swanson, T. (2012). Retention of Statistical Concepts in a Preliminary Randomization-Based Introductory Statistics Curriculum. *Statistics Education Research Journal*, *11*(1).

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).