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 lags 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 in a first statistics course four months after course completion at a single institution, with some improvement 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 (680 students) and 16-months (95 students) after the completion of an algebra-based statistics 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. The potential advantages of a simulation-based curriculum are that students are exposed to principles of statistical inference from day one of the course, and are able to build 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 increased 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") were combined together for a total of 22 scores 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, at the conclusion of their course, and again four months after completing the course. A subsample of students also took the assessment again sixteen months after the conclusion of their statistics course.

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).

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 assessment. The student survey was administered pre/post/4-month retention over two different years, 2015 and 2016. In the year 2015, 261 (68.6% female) students completed the pre-test, post-test, and 4-month retention assessments. Of those 261 students, 88 also completed the 16-month retention assessment. In the year 2016-17, 350 students (72.8% female) completed the pre-test, post-test, and 4-month retention assessments. For both cohorts, the majority of students 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, but not using the ISI textbook. "Not SBI" groups courses that followed a more traditional, non-simulation based curriculum. Due to low intraclass correlation coefficients, 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 611 students from years 2015 and 2016. Overall, students in SBI courses have higher post-test and retention percentages correct (*p*-value for comparing SBI to not SBI for retention = 0.0318). This pattern generally holds up across the subscales, except for Descriptive Statistics. However, students in SBI curricula were also more likely to have taken a

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

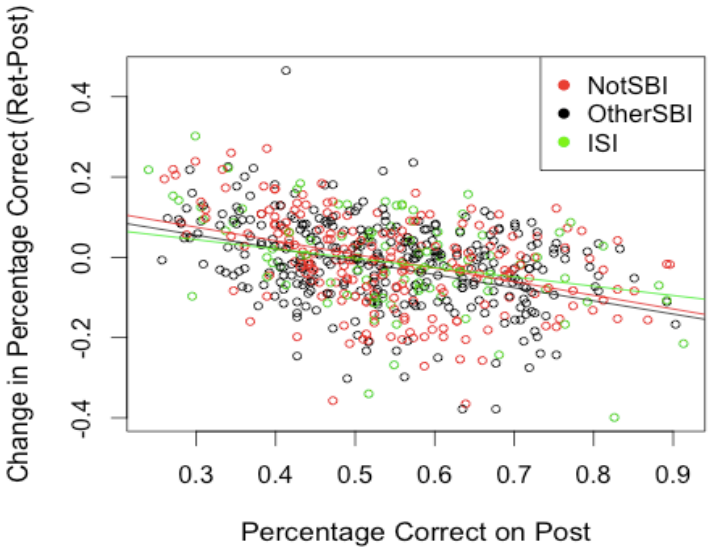
		Pre-test	Post-test	4-Month	Change	*p*-value[1]
Overall	ISI	0.458	0.570	0.551	-0.018	0.018
	OtherSBI	0.497	0.550	0.535	-0.015	0.139
	NonSBI	0.473	0.525	0.519	-0.005	0.434
Data Collection and Scope of Conclusions	ISI	0.489	0.605	0.609	-0.002	0.888
	OtherSBI	0.557	0.609	0.610	0.003	0.894
	NonSBI	0.515	0.530	0.569	0.037	0.009
Descriptive Statistics	ISI	0.476	0.570	0.545	-0.022	0.194
	OtherSBI	0.548	0.573	0.568	-0.005	0.803
	NonSBI	0.523	0.580	0.563	-0.016	0.206
Confidence Intervals	ISI	0.323	0.482	0.430	-0.053	0.003
	OtherSBI	0.380	0.462	0.426	-0.042	0.058
	NonSBI	0.344	0.476	0.425	-0.045	0.003
Significance Tests	ISI	0.586	0.703	0.685	-0.015	0.327
	OtherSBI	0.599	0.652	0.645	-0.009	0.659
	NonSBI	0.580	0.611	0.621	0.011	0.422
Simulation	ISI	0.325	0.377	0.396	0.019	0.340
	OtherSBI	0.301	0.348	0.317	-0.028	0.215
	NonSBI	0.314	0.326	0.307	-0.020	0.229

[1]*p-value for the paired t-test of percentage of correct concept questions from four-month retention to post assessments.*

follow-up statistics course by the time of the retention test (SBI: 35.8%; non-SBI: 26.2%). We also note that those in simulation-based courses tended to have larger decreases in their scores between the end of the course and four-months later. This could be due to certain questions on the assessment where students in non-SBI courses tend to improve from the post to 4-month retention tests, and the SBI students do not. For example, reduction in score from four-month retention to post-tests is significantly worse for students in an SBI course compared to a non-SBI course on a question requiring recognizing a correct interpretation of p-value (p-value = 0.0088). Conversely, students in SBI courses have a significantly higher four-month retention score compared to those in nonSBI courses especially in questions involving inference, like drawing a conclusion from a small p-value ($p < 0.0001$), and choosing an appropriate simulation process ($p < 0.0001$).

Although insignificant ($p = 0.328$), Figure 1 shows weak evidence of an interaction between type of curricula and post-test, 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. In general, students with larger post-test scores tend to revert closer to the average on the retention test. 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 simulation-based groups have generally higher post and four-month retention scores, but students' scores decrease more from post-test to 16-month retention assessments in the SBI groups. Questions where students in simulation-based courses had worse changes from post to 16-month retention assessments were primarily in significance test and descriptive statistics categories. When considering conceptual understanding gains from the post assessment to the 16-month retention assessment, students had the highest retention in questions regarding data collection and simulation, regardless of the textbook used in the course.

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 = 12)	0.434	0.568	0.553	0.386	-0.181	0.083	-0.433	-0.022	-0.372	0.065
OtherSBI (n = 26)	0.555	0.606	0.589	0.417	-0.189	0.070	-0.438	0.031	-0.452	0.128
NotSBI (n = 42)	0.481	0.522	0.550	0.499	-0.023	0.163	-0.133	-0.076	-0.064	0.050
Overall	0.498	0.554	0.560	0.447	-0.106	0.112	-0.285	-0.037	-0.25	0.093

[1]*Change in percentage correct from 16-month assessment to post assessment.*
[2]*Similar set up for five concept subscales.*

Instructors' level of familiarity with the American Statistical Association's Guidelines for Assessment and Instruction in a first Statistics course (www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf) is associated with conceptual 16-month retention (F test p-value = 0.002). The students in courses taught by an instructor who was "completely familiar" with the guidelines have significantly higher average 16-month conceptual retention compared to those in courses where the instructor was not familiar with the guidelines (t test p-value = 0.00014).

DISCUSSION

In this paper, we have presented an analysis of retention among students in introductory statistics courses. Four months after completion of an introductory statistics course, students were not found to have significantly different retention of statistical concepts across textbook categories. Although students in simulation-based curriculums were found to have marginally higher four-month retention rates in the each of the five conceptual subscales, these differences were not found to be significant. However, there is a pattern of higher post-test and retention scores for students with simulation-based curriculum, and 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. In particular, they show better retention for a handful of assessment questions asking about inference and simulation.

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 88 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, average conceptual understanding after 16 months was not found to be significantly higher or lower for those in different curriculums. Significant differences were also not found when considering achievement in the five concept subscales. However, attitudes about statistics 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 about statistics. This shows that students are able to accurately gauge their own understanding of statistical content.

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 familiarly 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 provide more face-time and interaction with experienced instructors tend to have higher retention rates 16 months after completion of the course and may enhance textbook effects.

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 Ed, 24*(3), 114-126.

delMas, R., Garfield J., Ooms, A., & Chance, B. (2007), Assessing students' conceptual understanding after a first course in statistics. *Statistics Ed 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).