PROGRESSIVE MASTERY TESTING: DOES IT INCREASE LEARNING AND RETENTION? YES AND NO

Dalene Stangl, David Banks, Leanna House, and Jerry Reiter
Duke University, United States
dalene@stat.duke.edu

Student mastery of material from an introductory statistics course was compared across courses that differed by whether a progressive mastery testing method was used. Students taught using progressive mastery methods showed increases in exam scores. The increases were small but of the same order of magnitude as increases associated with earning a course grade of A versus B+ and as increases associated with 100 points increases on the SAT math test. The increase in exam scores was about 75% of the increase associated with having taken a high school advanced placement course in addition to the college introductory course. However these increases were short lived as these students showed double rates of decline when tested in follow-up semesters. All benefit of the method vanished within two semesters.

INTRODUCTION

In the last twenty years considerable literature has appeared examining the impact of mastery learning. The operationalization of "mastery learning" differs from study to study, but commonalities include specification of clear learning objectives and methods for evaluation, individualized pacing, frequent evaluation and feedback with avenues for remediation, and evaluation of whether final learning objectives have been achieved. While many researchers argue that mastery learning improves teaching effectiveness (Block, Efthim and Burns, 1989; Slavin, 1987), to our knowledge no test of its effectiveness in introductory statistics courses has not been demonstrated to date. However, in hopes of counteracting the pervasive norm of cramming, a faculty member at Duke (DS) thought it worth a try. Because mastery learning programs require considerable resources to implement, it is important to know whether these costs are returned in terms of student learning and retention.

METHODS

Duke students were recruited from the pool of undergraduates who had or were currently taking Statistics 101, an introductory algebra-based course. The pool consists primarily of students majoring in the social sciences, although as many as twenty different majors are represented in any semester. Most students take the course as a requirement for their major or to fill a quantitative studies curriculum requirement. The format for the course includes 150-minutes of lecture and 50-minutes of data-analysis computing lab. Class sizes were between 120 and 150 students, with lab sections of 25-30 students. The text book used in all sections is Freedman *et al.* (1998). The course covers all topics in the text—which includes the standard topics in an introductory course—plus introductions to multiple regression and Bayesian inference.

Professors who did not use the progressive mastery method gave a combination of homework, weekly quizzes, two midterm exams, a data project, and a final exam. Under the progressive mastery method, students were given daily warm-up exercises that covered material in the preceding lecture, three practice exams given during the first eight weeks of the semester, and then a sequence of four progressively more challenging exams corresponding to mastery levels of D, C, B, and A. Students did not progress to the next level until they had achieved 75% correct on the preceding level.

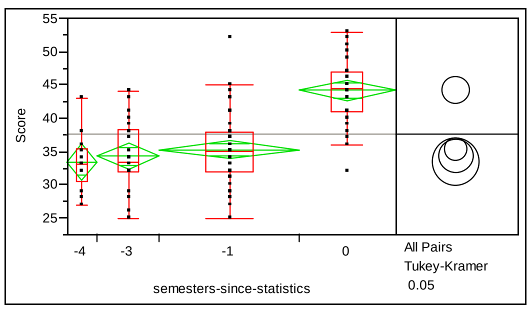
Students who had taken Statistics 101 in any one of eight sections and received a final grade of B- or higher were invited to participate. Because the exam was administered during the reading period before final exams in Spring 2004 and Fall 2005, for students currently enrolled, only those students who were expected to achieve a B- or higher were invited. As an incentive to participate, students were offered $40 to take the voluntary exam. Confidentiality of all test scores was assured and maintained. Students were instructed not to study for the exam. The goal was to measure retention rather than what could be relearned.

The questions in the exam covered study design, simple graphical interpretations, basic probability problems, confidence intervals, hypothesis testing, and other topics. Exact question wordings are available from the authors. During the second test (Fall 2005) students were given the option of using a crib sheet for the exam. Forty eight percent of the students did so, so this variable is included as a covariate in analyses.

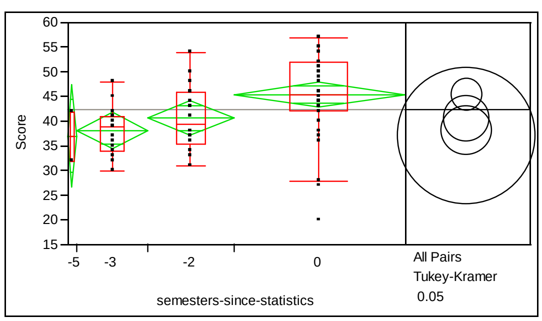
RESULTS

Data was examined using analysis of variance and multiple regression. Figure 1 gives the results for the Spring 2004 and the Fall 2005 exams separately. Students taught via the progressive mastery method are those with semesters-since-statistics coded as 0 in Spring 2004, because they were currently enrolled in the course, and they are coded as -2 in Fall 2004, since they took the course 2 semesters previously.

Figure 1 shows Spring 2004 means (SD) of 33.38 (4.23), 34.31 (5.45), 35.29 (4.95), and 44.18 (4.95) for each cohort respectively. It shows Fall 2004 means (SD) of 37 (7.07), 38.13 (4.96), 40.56 (6.64), and 45.33 (8.49) for each cohort respectively. Both rounds, Spring 2004 and Fall 2004, show statistically significant overall differences (R^2=43%, F=34.02, p<.0001 and R^2=16%, F=4.28, p=.008). For the first round, Spring 2004, the mean score for students currently in introductory statistics (and using the progressive mastery method) differed from all other cohorts at α=.05. For the second round, Fall 2005, while overall differences were statistically significant, the mean of current students differed only from those who had taken statistics three semesters previously at α=.05. The mean for students who had used the progressive mastery method (semesters-since-statistics=-2) did not differ from the current students at α=.05.



Spring 2004: Progressive Mastery Method = 0



Fall 2004: Progressive Mastery Method = -2

Figure 1: Spring 2004 and Fall 2005 exam score by semesters-since-statistics

To control for differences across sections in student abilities, exposure to advanced placement courses, recruitment, and use of crib sheets, a multivariate regression was run. Each of these variables along with number of semesters since taking statistics, an indicator for whether the student was in a section that used the progressive mastery method, and an interaction between the progressive mastery method and semesters since statistics, was included in the model. The results are shown in Figure 2.

The model yielded an overall R^2 of 56%. All variables were statistically significant at α=.05. Each unit increase in class grade (coded from 8=B- to 13=A+) was associated with an increase of 1.37 points on the exam. Having taken an advanced placement statistics course in high school was associated with a 3.35 point increase, using a crib sheet a 4.66 point increase, 100 points on the SAT math section a 2.07 point increase, each unit increase in current GPA a 3.05 point increase, and exposure to the progressive mastery method a 2.45 point increase.

For students not exposed to the progressive mastery method, each additional semester since taking statistics was associated with a drop of 1.68 points, while for students exposed to the progressive mastery method each additional semester since taking statistics was associated with a drop of 3.41 points. Hence all benefit vanished after two semesters.

Term	Estimate	Std Error	*t* Ratio	Prob>|*t*|
Intercept	-0.83	5.63	-0.15	0.8835
class.grade	1.37	0.35	3.96	0.0001
AP.0/1	3.35	0.84	4.00	<.0001
crib.sheet.0/1	4.665	1.05	4.46	<.0001
SAT.math	0.025	0.01	3.02	0.0029
current.GPA	3.05	1.26	2.42	0.0167
semesters-since-statistics	1.68	0.30	5.57	<.0001
mastery.method.0/1	2.46	1.06	2.30	0.0224
sem-since-stat*mastery	1.73	0.82	2.09	0.0382

Figure 2: Regression Parameter Estimates

CONCLUSION

This paper explored the impact of progressive mastery testing on the learning and retention of introductory statistics at Duke University. The costs of such teaching were great for the professor and students. For the professor there was extensive exam construction, feedback to students, and grading. For the students there was the weight of perceptions that the course was unfair and required more work than other sections resulting in considerable stress and a backlash in course evaluations. While the method showed short-term benefits of higher exam scores, the benefits were lost within two semesters of taking the course.

REFERENCES

Angelo, T. A. and Cross, K. P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco, CA: Jossey-Bass.

Block, J. H., Efthim, H. E. and Burns, R. B. (1989). *Building Effective Mastery Learning Schools*. New York: Longman.

Carroll, J. B. (1989). The Carroll model: A 25 year retrospective and prospective view. *Educational Researcher*, 18(1), 26-31.

Cox, W. F. and Dunn, T. G. (1979). Mastery learning: A psychological trap? *Educational Psychologist*, 14, 24-29.

Levine, D. (1985). *Improving Student Achievement Through Mastery Learning Programs*. San Francisco, CA: Jossey-Bass.

Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research*, 57(2), 175-214.