Topical Article

SOCIETY FOR THE TEACHING OF PSYCHOLOGY

Teaching of Psychology
40(1) 15-19

DOI: 10.1177/0098628312465858
http://top.sagepub.com
SAGE

Cumulative Exams in the Introductory Psychology Course

Natalie K. Lawrence[1]

Abstract

Many teachers require their students to take cumulative exams, but there are surprisingly few studies that examine the benefits of such exams. The purpose of this study was to determine whether introductory psychology students who take cumulative exams throughout the semester would have better long-term retention than students who take a cumulative final exam after a series of unit (i.e., noncumulative) exams. As expected, the students who took cumulative exams throughout the semester did better on the cumulative portion of the final exam. This main effect evolved into an interaction on a follow-up test administered 2 months after the course ended. The long-term retention of good students was unaffected by the type of exams they took. However, low-scoring students remembered more of the course material when they took cumulative exams throughout the semester.

Keywords

cumulative exams, introductory psychology, formative assessment, summative assessment

Research suggests that students do not remember much from the introductory psychology course (e.g., Rickard, Rogers, Ellis, & Beidleman, 1988) and much of what they *do* remember is not highly relevant to the course material (VanderStoep, Fagerlin, & Feenstra, 2000). Thus, it is important to identify learning strategies that increase students' long-term retention of the material taught in the introductory psychology course. This is important for psychology majors who will encounter the material again in future courses and nonmajors who may never take another psychology course.

Many teachers advocate the use of cumulative exams to boost long-term retention. In a recent survey of introductory psychology teachers, 57% of the 206 respondents believed that students who take cumulative exams will have better long-term retention of the course material than students who do not take cumulative exams (Lawrence, 2012). Interestingly, 57% of the sample gave at least one cumulative exam in their most recent introductory psychology course. Of those teachers, 6% gave two cumulative exams and 5% gave more than two cumulative exams.[1] Research in cognitive psychology suggests that these teachers may be correct about the benefits of cumulative exams.

In a laboratory study, Szpunar, McDermott, and Roediger (2007) exposed participants to a series of word lists and then administered a free recall test after each list. They manipulated whether or not participants were aware of a final cumulative test that would be administered 30 minutes after the initial testing. As expected, participants who were aware of the final test outperformed those who were unaware. Szpunar et al. found evidence that the expectation of the final test encouraged the continued processing of the studied words, which made the words more accessible at the time of the final. The authors also proposed that *not* having an expectation of a final test serves as a cue to forget.

Szpunar et al.'s (2007) research suggests that the expectation of a cumulative final exam may enhance students' long-term retention of course material. However, the Szpunar et al. study was conducted in a laboratory setting very different than a typical classroom setting. Students usually take a final exam several weeks—not minutes—after learning begins. There is an obvious need for applied research, but there are surprisingly few studies that examine the benefits of cumulative exams in the classroom.[2]

The purpose of the present research was to investigate the effects of cumulative exams in an introductory psychology course. Though Szpunar et al. (2007) found that the expectation of a cumulative "final exam" encouraged participants to continue processing the material, this may not be true for students taking a final exam in a college course. In most cases, cumulative final exams cover several weeks of material. Many students are likely to "cram" for this type of exam rather than continually process the material over an entire semester. Decades of research on the spacing effect indicate that cramming is not the best strategy to enhance long-term retention (for reviews, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Dempster, 1989). A better strategy might be to have several cumulative exams throughout the

[1] Department of Psychology, James Madison University, Harrisonburg, VA, USA

Corresponding Author:
Natalie K. Lawrence, Department of Psychology, James Madison University, 91 E. Grace Street—MSC 7704, Harrisonburg, VA 22807, USA
Email: lawrennk@jmu.edu

semester—not just one cumulative exam at the end. Having multiple cumulative exams may motivate students to keep in the material in mind, which should lead to better long-term retention. The present research was designed to examine this possibility.

Method

Participants

Participants were introductory psychology students ($N = 105$; 59 women and 46 men) from a midsized, public university in the Southeast. The sample consisted of 53% first years, 14% sophomores, 26% juniors, and 7% seniors. The majority of the sample was Caucasian (87%) with a mean age of 19.14 ($SD = 1.22$) years. Students received extra credit for their participation.

Materials and Procedure

Students were enrolled in one of two sections of an introductory psychology course. I taught both sections back to back in a twice-weekly format. The only major difference between the two sections was the type of exams. Students in the first section—referred to as the *noncumulative section*—took three noncumulative exams plus a cumulative final exam. Students in the second section—referred to as the *cumulative section*—took four cumulative exams. Ideally, I would have had a control group (i.e., a section that had no cumulative exams), but I was unable to teach a third section of the course.

The first exam, which included 50 multiple-choice questions, was identical for both sections. For the noncumulative section, the second and third exams included 50 multiple-choice questions based on the most recently covered material. For the cumulative section, the second and third exams included 40 questions based on the most recent material and 10 questions based on earlier material (i.e., 20% of these exams was cumulative). The final exam was the same for both sections and included 40 questions based on the most recent material and 50 questions on previously tested material. Thus, the final exam was 55% cumulative.

Two months after the course ended, I administered a follow-up test that measured students' long-term retention of the course material. This test contained 50 multiple-choice questions that were all novel (i.e., they did not appear on any previous exams). As an incentive to complete the follow-up test, I gave students 3 points of extra credit on their final exam (worth 225 points) if they promised to take an "informal assessment of how much they remembered from the course two months after it ends." I also told students that they would be entered into a raffle to win one of two portable media players if they completed the test. Every student promised to take the test. Two months after the course ended, I e-mailed students with instructions for taking the test. I told students that they should not rely on outside sources (e.g., textbooks or other people) to answer the questions.

The primary dependent measures in this study were scores on the cumulative portion of the final exam and scores on the follow-up test. These can be considered as two separate measures of long-term retention. I predicted that students in the cumulative section would outperform other students on the final exam and follow-up test. This prediction is consistent with the work of Szpunar et al. (2007) and other cognitive psychologists.

In addition, I included measures of students' perceptions of the exams they took throughout the semester. Items included, "I thought the exams were easy" and "I am glad the exams were cumulative (not cumulative)." I also included measures of students' study habits throughout the semester. Students indicated the extent to which they agreed with statements such as "I crammed for the exams" and "After each exam, I disregarded any previously tested information and focused my attention on new information." The response scale for these items ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). Finally, I asked students to report how many hours they spent studying for the final exam.

Results

Preliminary Analyses

Preliminary analyses revealed that there were three outliers in the data set. These outliers were all students who scored 2½ to 3 *SD*s below the mean on the follow-up test. I removed the outliers, which reduced the data set to 102 students.

To ensure that there were no important differences between the two sections, I compared their scores on the first exam and their demographic information (age, sex, race, and semesters completed). Students in the two sections performed almost identically on the first exam, $t(100) = -.11$, $p > .05$, $d = -.02$. The two sections were similar on the demographic measures as well (all ps $> .05$). Therefore, it seems unlikely that the results I obtained are due to preexisting differences between the two sections.

Students' Perceptions and Study Habits

I ran a series of *t*-tests with the Bonferroni correction ($\alpha = .006$) on students' perceptions and study habits. I found a significant difference in responses to the statement, "I am glad the exams were cumulative (not cumulative)," $t(94) = -6.12$, $p < .001$, $d = -1.25$. Not surprisingly, students in the noncumulative section were happier with the nature of their exams ($M = 5.47$, $SD = 1.86$) than students in the cumulative section ($M = 3.16$, $SD = 1.83$). There were no other significant differences in students' perceptions of the exams, the way they studied for them, or the number of hours they spent studying for the final exam (all ps $> .006$).

Performance on the Final Exam and Follow-Up Test

Previous research suggests that low- and high-performing students respond differently to certain teaching methods and interventions (e.g., Forsyth, Lawrence, Burnette, & Baumeister, 2007; Saville, Pope, Truelove, & Williams, in press). For this reason, I performed a median split on the scores for the first exam and separated students into two groups: low scorers and high scorers. Grade point average (GPA) would have been a preferable criterion to split the sample, but the sample included mostly first years who did not have a college GPA at the time of data collection (because they had not yet completed their first semester of college). I found strong positive correlations for the scores on all four tests (all ps < .001), so the median split seemed to be a legitimate way of distinguishing between students.

I ran a 2 (type of student: low scorers vs. high scorers) × 2 (section: cumulative vs. noncumulative) factorial analysis of variance (ANOVA) on the cumulative portion of the final exam.[3] Scores were measured as percentage correct. Not surprisingly, there was a significant main effect for type of student, $F(1, 98) = 47.02$, $p < .001$, $\eta_p^2 = .32$. High scorers outperformed ($M = 83.80$, $SD = 6.54$) the low scorers ($M = 73.35$, $SD = 9.04$). There was also a significant main effect for section, $F(1, 98) = 4.55$, $p = .035$, $\eta_p^2 = .04$. Students in the cumulative section did better ($M = 79.92$, $SD = 9.46$) than students in the noncumulative section ($M = 77.02$, $SD = 9.34$). It is important to note that this difference is statistically and *academically* significant. For students, there is a big difference between a grade of 80% (B−) and a grade of 77% (C+). There was not a significant interaction between type of student and section, $F(1, 98) = .71$, $p > .05$, $\eta_p^2 = .01$. This suggests that both low- and high scorers benefited from taking cumulative exams throughout the semester.

I ran an additional 2 (type of student) × 2 (section) factorial ANOVA on the follow-up test scores. Unfortunately, the response rate on the follow up-test was only about 64%. Fewer of the low-scoring students completed the test, which resulted in unequal cell sizes (16 low scorers and 20 high scorers in the cumulative section, and 11 low scorers and 18 high scorers in the noncumulative section). I used an unweighted means solution to solve this problem. As expected, I found a significant main effect for type of student, $F(1, 61) = 7.85$, $p = .007$, $\eta_p^2 = .11$. High scorers did better ($M = 76.84$, $SD = 9.06$) than the low scorers ($M = 70.00$, $SD = 13.32$). There was no main effect for section, $F(1, 61) = .84$, $p > .05$, $\eta_p^2 = .01$, but there was a significant interaction between type of student and section, $F(1, 61) = 4.42$, $p = .04$, $\eta_p^2 = .07$ (see Figure 1). High scorers in the cumulative section performed just as well as high scorers as in the noncumulative section. However, low scorers in the cumulative section outperformed low scorers in the noncumulative section.

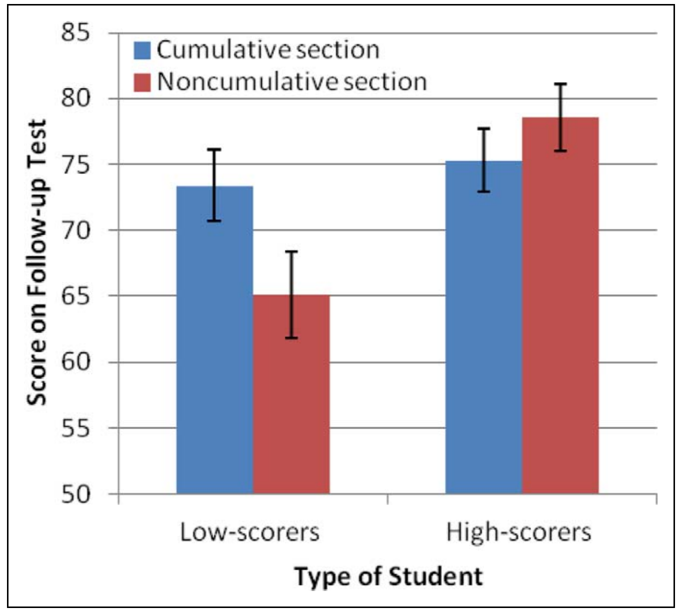

Figure 1. Mean percentage correct on follow-up test as a function of type of student (low scorers vs. high scorers) and section (cumulative vs. noncumulative).

Performance on Chapter Quizzes and Overall Course Grade

Because the results of students' perceptions and study habits were not very informative, I decided to explore whether the experimental manipulation affected other aspects of students' performance in the course. Students in both sections took an online chapter (i.e., noncumulative) quiz after we covered each new chapter. I performed a factorial ANOVA on the average quiz scores (measured as percentage correct) and found a significant main effect for type of student, $F(1, 98) = 20.62$, $p < .001$, $\eta_p^2 = .17$. High scorers did better on the quizzes ($M = 88.83$, $SD = 7.12$) than the low scorers ($M = 80.24$, $SD = 12.57$). I also found a significant main effect for section, $F(1, 98) = 7.86$, $p = .006$, $\eta_p^2 = .07$. Students in the cumulative section did better on the quizzes ($M = 87.03$, $SD = 8.73$) than students in the noncumulative section ($M = 81.87$, $SD = 12.59$). There was also a significant interaction between section and type of student, $F(1, 98) = 4.33$, $p = .04$, $\eta_p^2 = .04$. High scorers' quiz grades were unaffected by section, but low scorers in the cumulative section did better on their quizzes ($M = 84.79$, $SD = 9.39$) than low scorers in noncumulative section ($M = 75.33$, $SD = 13.84$). This difference (i.e., the difference between a B average and a C average) is academically significant as well as statistically significant.

A similar pattern emerged when I analyzed the total course grades (measured as percentage of points earned). There was a significant main effect for type of student, $F(1, 98) = 42.27$, $p < .001$, $\eta_p^2 = .30$. High scorers did better in the course ($M = 89.15$, $SD = 5.42$) than the low scorers ($M = 80.31$, $SD = 8.38$). There was a marginally significant effect of section, $F(1, 98) = 3.36$, $p = .07$, $\eta_p^2 = .03$. Students in the cumulative section tended to do better in the course ($M = 85.75$, $SD = 7.29$) than students in the noncumulative section ($M = 83.54$, $SD = 9.20$). There was also a marginally significant interaction between section and type of student, $F(1, 98) = 2.95$, $p = .09$, $\eta_p^2 = .03$. High scorers' course grades were unaffected by section. Low scorers in the cumulative section ended up with a B average ($M = 82.66$, $SD = 6.67$), whereas low scorers in noncumulative section ended up with a C+ average ($M = 77.78$, $SD = 9.39$).

Discussion

Many introductory psychology teachers require their students to take a cumulative final exam at the end of the semester. A few of these teachers require their students to take cumulative exams throughout the semester, not just at the end (Lawrence, 2012). The results of this study suggest that students may benefit more from the latter approach (i.e., requiring them to take multiple cumulative exams).

In this study, students who took cumulative exams throughout the semester performed better on the cumulative portion of the final exam than students who did not. Szpunar et al. (2007) found that the expectation of a cumulative final exam keeps the material in a more accessible state, making it easier to retrieve during a final exam. It is possible that having multiple cumulative exams increased the accessibility of the course material. Unfortunately, I did not include any direct measures of accessibility. It is also possible that the effect of taking multiple exams is mediated by study habits. I did not find any significant differences in students' self-reported study habits for the exams, but my dependent measures may not have been sensitive enough to detect real differences. I asked students to report how many hours they studied for the exams and found no differences between the two sections. However, it may be the case that students in the cumulative section studied *smarter* (e.g., they spaced their studying out over time). This explanation is consistent with the finding that students in the cumulative section performed better on the chapter quizzes than students in the noncumulative section.

It is interesting that an interaction—but no main effect—was found for the follow-up test, administered 2 months after the course ended. High-scoring students' long-term retention was unaffected by the experimental manipulation. In contrast, low-scoring students remembered more of the course material if they took multiple cumulative exams. This interaction was also evident in students' grades on the chapter quizzes and their final grade in the course. This pattern of results is consistent with previous research that shows that struggling students benefit more from certain pedagogical techniques than stronger students (e.g., Saville et al., in press). Most likely, having multiple cumulative exams motivates low-scoring students to engage in behaviors that promote better performance and long-term retention. High-scoring students probably already have the motivation to engage in these types of behaviors.

The data from the present study suggest that having multiple cumulative exams in an introductory psychology course enhances the long-term retention of the course material, at least for students who are not top performers. Future research should aim to replicate this effect and gain greater insight into its cause. In the present study, most of the cumulative exams were only 20% cumulative. It would be interesting to find out if a stronger experimental manipulation (e.g., making the exams 50% cumulative) would lead to larger effect sizes. It would also be interesting to explore whether having multiple cumulative exams has benefits in higher level psychology courses. Researchers also need to address a more basic question: Do students who take a single cumulative final exam have better long-term retention than students who do not? The results of Szpunar et al. (2007) suggest that this is the case, but researchers need to replicate it in the classroom. It is clear that students do not like cumulative exams, so teachers should make evidence-based decisions about their use.

Acknowledgments

I would like to thank the following undergraduate research assistants for their help organizing the data: Caitlin Davis, Christina Hargrove, Chelsea Johnson, Amanda Rothwell, and Katy Snyder. I would also like to thank Daniel Jurich, Bryan Saville, and Tracy Zinn for their advice concerning design and analysis.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The General Education Program at James Madison University provided grant support for this research.

Notes

1. In this unpublished study, I compared teachers who use cumulative exams and those who do not on a number of different variables (e.g., how long they have been teaching, the type of institution where they teach). I found no significant differences between these two groups of teachers.
2. Kouyoumdjian (2004) compared introductory psychology students' perceptions of unannounced quizzes and cumulative exams but did not examine the impact of cumulative exams on students' learning. Landrum (2007) measured his introductory psychology students' performance on a cumulative final exam after taking weekly quizzes, but he was not concerned with the effects of taking a cumulative exam.
3. Because there are problems associated with dichotomizing continuous variables, I also analyzed students' scores on their first exam as a continuous variable. Regression analyses yield the same results as the factorial ANOVAs. Students' scores on their first exam ($\beta = .68, p < .001$) and course ($\beta = -.16, p = .028$) were significant predictors of final exam performance, but the interaction was not significant ($\beta = -.001, p > .05$). When predicting scores on the follow-up test, scores on the first exam ($\beta = .44, p < .001$) and the interaction ($\beta = .24, p = .04$) were significant, but course was not ($\beta = -.13, p > .05$). I chose to report the results of the factorial ANOVAs rather than the results of the regression analyses because the interaction effects are easier to understand and visualize in ANOVA.

References

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380. doi:10.1037/0033-2909.132.3.354

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review*, *1*, 309–330.

Forsyth, D. R., Lawrence, N. K., Burnette, J. L., & Baumeister, R. F. (2007). Attempting to improve the academic performance of struggling college students by bolstering their self-esteem: An intervention that backfired. *Journal of Social and Clinical Psychology*, *26*, 447–459.

Kouyoumdjian, H. (2004). Influence of unannounced quizzes and cumulative exam on attendance and study behavior. *Teaching of Psychology*, *31*, 110–111.

Landrum, R. E. (2007). Introductory psychology student performance: Weekly quizzes followed by a cumulative final exam. *Teaching of Psychology*, *34*, 177–180. doi:10.1080/00986280701498566

Lawrence, N. K. (2012). [Survey of teachers' attitudes toward and use of cumulative exams]. Unpublished raw data.

Rickard, H. C., Rogers, R., Ellis, N. R., & Beidleman, W. B. (1988). Some retention, but not enough. *Teaching of Psychology*, *15*, 151–152.

Saville, B. K., Pope, D., Truelove, J., & Williams, J. (in press). The relation between GPA and exam performance during interteaching and lecture. *Behavior Analyst Today*.

Szpunar, K. K., McDermott, K. B., & Roediger III, H. L. (2007). Expectation of a final cumulative test enhances long-term retention. *Memory & Cognition*, *35*, 1007–1013.

VanderStoep, S. W., Fagerlin, A., & Feenstra, J. S. (2000). What do students remember from introductory psychology? *Teaching of Psychology*, *27*, 89–92.