July 2017

Effects of Quizzing Methodology on Student Outcomes: Reading Compliance, Retention, and Perceptions

Carey Bernini Dowling
University of Mississippi Main Campus, cbdowlin@olemiss.edu

Recommended Citation
Available at: https://doi.org/10.20429/ijsotl.2017.110203
Effects of Quizzing Methodology on Student Outcomes: Reading Compliance, Retention, and Perceptions

Abstract
This study set out to replicate and extend research on students' reading compliance and examine the impact of daily quizzing methodology on students' reading compliance and retention. 98 students in two sections of Abnormal Psychology participated (mean age = 21.5, SD = 3.35; 72.4% Caucasian). Using a multiple baseline quasi-experimental design the daily quizzing methodology was changed at different points in the semester from Clicker questions to Clicker questions plus random written quizzes. The classes did not differ significantly on predictors of success and only differed significantly on one demographic variable. 77.6% of students failed Sappington et al.'s (2002) objective measure of reading compliance and the majority lied about their reading compliance. There was mixed evidence for the impact of quizzing methodology on learning outcomes. Daily quizzing appears to be effective, but adding written quizzes may not improve learning outcomes enough to justify increased grading time.

Keywords
Quizzing methods, reading compliance

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.

Cover Page Footnote
I would like to express special thanks to the research assistants involved with this project: Billy Rush, Teresa Davis, Mia Kloth, and Brittney Stone. I would also like to express sincere appreciation to Conor Dowling and the reviewers for their feedback on previous versions of this manuscript.
INTRODUCTION
An undergraduate college education in psychology has multiple desired learning goals (APA, 2012). In order for students to meet these goals, it is necessary for them to actively participate in their education. As educators who desire to help students succeed in college we must understand what predicts their success and what we can do to help them succeed.

One of the first ways students can actively participate in their education is to prepare for their classes by completing reading assignments. Research suggests this preparation is important because it is associated with overall class performance (Sappington, Kinsey, & Munsayac, 2002) and students report lack of preparation for class is a barrier to their class participation (Karp & Yoels, 1976). However, recent research suggests that a majority of college students do not complete reading assignments prior to coming to class (Burchfield & Sappington, 2000; Clump, Bauer, & Bradley, 2004; Connor-Greene, 2000; Sappington et al., 2002). Sappington et al. (2002) found only 22% of students passed their objective measure of reading compliance. Unfortunately, this trend of lack of preparation for class might be increasing (Burchfield & Sappington, 2000). Yet, it is possible that students’ reading compliance varies by the testing schedule of the course, with students reporting they are more prepared for classes with daily quizzes than classes with exams only (Connor-Greene, 2000).

If students’ reading compliance is declining and consistently at levels below 30%, it is important to determine effective strategies for increasing and maintaining student reading compliance across the semester. Multiple strategies have been implemented to increase student reading compliance and course performance, such as completion of out-of-class assignments that require reading (Carkenord, 1994; Ryan, 2006), daily written quizzes (Connor-Greene, 2000), and randomized reading quizzes (Ruscio, 2001).

Although reading is not required to complete in-class quizzes, quizzes may be an effective means of improving reading compliance (Connor-Green, 2000; Ruscio, 2001) while also improving course performance. Quizzing has been found to positively impact exam grades when done in a manner to simulate basic research on the testing effect (see Nguyen & McDaniel, 2015). Research on the testing effect suggests that testing itself and testing with feedback are powerful means to improve the learning of material (Butler, Karpicke, & Roediger, 2008; Roediger, Agarwal, McDaniel, & McDermott, 2011; Roediger & Karpicke, 2006). Immediate feedback after testing allows the learner to correct erroneous knowledge as well as correct metacognitive errors regarding low confidence in correct answers (Butler et al., 2008). Therefore, it is not surprising that previous research has found utilizing student response systems (SRS) during class to quiz and provide immediate feedback to students improves students’ course and examination performance (Brady, Seli, & Rosenthal, 2013; Hall, Collier, Thomas, & Hilgers, 2005; Morling, McAuliffe, Cohen, & Di Lorenzo, 2008) and increases course engagement and motivation (Hall et al., 2005).

Although SRS and written quizzing have shown positive benefits for students, these methods are not without concerns. First, there are multiple time demands on professors that may make grading of written quizzes impractical, especially in large sections. Additionally, multiple time demands are a large source of stress for faculty (Gmelch, Lovrich, & Wilke, 1984), so it is especially important to examine if assessments that require grading confer enough of a benefit to justify the grading time. Second, while utilizing SRS during class reduces (or eliminates) grading time, it is easier for students to guess the correct answer even if they have not read the material, thus potentially reinforcing students who did not read and perpetuating their perception that they can succeed without coming to class prepared. A third concern with utilizing both forms of quizzing has to do with potential negative ramifications on student evaluations. Individuals responsible for evaluating teaching effectiveness rate student evaluation scores and written comments among the top three most important measures to use for evaluating teaching effectiveness (Shao, Anderson, & Newsome, 2007). Thus, it is pragmatic for professors to be concerned about poor student evaluations.

Given the multiple time demands for professors as well as concerns over poor student evaluations, it is beneficial for professors to determine the best methods to simultaneously achieve multiple goals (encouraging students’ reading compliance, engagement with the material, and learning of the material; avoiding an unduly difficult grading load; and avoiding unfavorable student evaluations). Therefore, I set out to determine whether a combination of the use of SRS with pop written quizzes would achieve all of these goals. This study utilized daily SRS quizzes, which require minimal...
are likely to be invalid (see Sappington et al., 2002). Thus, the first student reading behaviors because student reports of reading than students who passed the objective measure. Following Sappington et al. (2002) it is important to determine whether active learning strategies, such as quizzing, also impact how frequently students read on time and how thoroughly they read assigned readings. Previous research has rarely reported student reading behaviors because student reports of reading are likely to be invalid (see Sappington et al., 2002). Thus, the first aim of the study was to explore the validity of students’ self-reports regarding how thoroughly they read assigned readings to determine if they could be a valid dependent variable. Sappington et al. (2002) utilized an objective measure of student reading compliance utilizing a dichotomous “yes/no” option for students to report whether they read the entire syllabus. Thus, students who had skimmed the entire syllabus or read most of the syllabus were forced to decide whether they felt what they did counted as “reading the entire syllabus” and potentially increased the chances of students engaging in self-enhancement bias. Thus, I set out to determine if when students were given multiple options regarding how thoroughly they read assigned readings such as “read all, read at least some, skimmed all, did not look at any,” whether they might show less self-enhancement bias and more valid responses. I hypothesized that, similar to Sappington et al. (2002), students will engage in self-enhancement bias and a majority would lie on their self-reported reading compliance, but that students who failed the objective measure of reading the entire syllabus would report lower levels of reading compliance for the objective measure. Previous research on predictors of students’ success in college classes has found that students’ performance goals (Elliot & Church, 1997; Elliot & Murayama, 2008), intrinsic motivation (Clark, Middleton, Nguyen, & Zwick, 2014), conscientiousness, self-reported high school GPA was a 3.69 (SD = 0.48) reported current GPA was 3.20 (SD = 3.35), they were primarily juniors (46.8%) or seniors in college (40.4%), Caucasian (24.3%), and a majority (74.5%) identified as the Clicker only method in order to than the Clicker only method. In order to adjust the depth of coverage needed on a topic. For instance, following a question such as “Which of the following is definitely in a manifold envelope. I then left the room and the envelope was sealed by the research assistant and was kept sealed until after final grades had been submitted. For all classes the manifold envelope was designed to allow students and keep me blind to percentages participating in each course until the end of data collection. Students who were absent were given individual alternative exams. For all exams students were asked to return the envelope sealed with the packet completed if they consented and blank if they did not. I am unaware of any studies examining the effectiveness of teaching strategies, such as quizzing, also impact how frequently students read on time and how thoroughly they read assigned readings. Following Sappington et al. (2002) the following line appeared near the bottom of the syllabus: “Students who have read this far in the syllabus will receive I point added to their final average if they e-mail me at […] with the subject line: ‘FY 311 Section 3 Syllabus bonus’ by the time Homework #1 is due.” Students were asked to report their compliance with reading the syllabus first at the end of homework #1 and again using their clickers at the class of starting reviewing the syllabus/homework #1. Students were asked for the e-mail address they use the most on their homework assignment to verify they had access to an e-mail address. Students experienced the course as if the study was not being conducted. However, the quizzing method was modified according to a multiple-baseline quasi-experimental design (the 9 AM section was randomly chosen prior to the semester to receive the manipulation first). At the start of the semester both sections participated in daily multiple-choice clicker quizzes (referred to as the Clicker only method). Following Exam 1, the 9 AM section continued daily clicker quizzes but now had a 25% chance of having a written quiz to start the class, which would replace their clicker quiz points for the day (referred to as the Clicker plus written method). Halfway through new material coverage for Exam 3, the 9 AM section was switched to the Clicker plus written method. The class lectures and clicker questions were identical with the exception of variations in student responses to my questions and student questions prompting varying responses from me. All students were given the post-assessment packet with their ID number at the start of the final examination and asked to complete the packet of questionnaires (if participating) at the end of the final exam along with a course evaluation (which was not part of this study), thereby ensuring that I was unaware if they were completing the final, the post-assessment packet, or the course evaluation. All students were told to place the post-assessment packet in a manila envelope regardless of whether they completed it or not and I sealed the envelope at the end of the final exam period. Students completing the pre- and post-packets were entered into a raffle for one of two $10 Amazon gift cards. Quiz questions (both types) were intentionally basic, recognition questions designed to be difficult to answer without reading the assigned reading but not so difficult they required students to do study outside of class. After I read over the questions, I showed a list of words I occasionally used clicker questions that resembled exam questions for practice, but only after we had reviewed the relevant material in class. The questions were designed to primarily evaluate students’ understanding of the material and ability to apply knowledge gained (e.g., by correctly answering a question about a strategy of which they had taken a portion of exam questions also assessed recall and recognition of important facts. Demographic questionnaire. The demographic questionnaire created for this study requested information on major demographic characteristics such as participants’ age and year in college. Predictors of Student Success. Four questionnaires were given to students to evaluate whether the sections differed significantly on variables found in previous studies to significantly predict student success in AMPS classes. Revised (AGQ-R; Elliot & Murayama, 2008) was included in the pre-assessment packet to measure mastery-mastery, mastery-avoidance, performance-avoidance, and performance-avoidance.” I occasionally used clicker questions that resembled exam questions for practice, but only after we had reviewed the relevant material in class. The questions were designed to primarily evaluate students’ understanding of the material and ability to apply knowledge gained (e.g., by correctly answering a question about a strategy of which they had taken a portion of exam questions also assessed recall and recognition of important facts. Measures
RESULTS
Assessments of Reading Compliance and Behavior

The vast majority of students did not read the assigned syllabus. For example, at least 30.1% of students did not report reading all of the syllabus when both sections were quizzed using the Clicker only method. Exam 2, the first half of Exam 3, and the second segment of the Final Exam assessed material covered when only the 9 AM section was being quizzed using the Clicker plus written method. The second half of Exam 3, Exam 4, and the third segment of the Final Exam assessed material covered when both sections were quizzed using the Clicker plus written method. A series of independent means t-tests were performed to determine if students in each section differed significantly on any of the measured potential predictors of students' success in college classes. There were no significant differences found for self-reported current or high school GPA, competency expectations, intrinsic motivation for the course, AGG-R achievement goals, or TIPI personality factors.

Assessments of Retention of material
As shown in Table 2, the sections did not differ significantly in their scores on the behavior questions which they had themselves chosen as their preferred learning method (i.e., pre-manipulation and post-manipulation in both sections). There were mixed results for the hypothesis that the 9 AM section would score higher when they had the Clicker plus written quizzing method than the 8 AM section. As hypothesized, the 9 AM section scored significantly higher on Exam 2. However, contrary to hypotheses, the 9 AM section did not score significantly higher on the first portion of Exam 3 and did not perform significantly higher on any of the segments of the Final Exam, including segments assessing material covered when only the 9 AM section had the Clicker plus written quizzing method.

DISCUSSION
Assessments of Reading Compliance and Behavior

Among the students who did not report reading all of the syllabus, a majority of students did not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and responded to reading compliance and extends its results by finding that even when students are given a more nuanced opportunity to be honest in their report of whether they read all of the syllabus, a majority still lie. These results were further supported by the finding that the vast majority of students do not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and concluded that the majority of syllabus reading compliance and extends its results by finding that even when students are given a more nuanced opportunity to be honest in their report of whether they read all of the syllabus, a majority still lie. These results were further supported by the finding that the vast majority of students do not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and concluded that the majority of students do not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and concluded that the majority of students do not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and concluded that the majority of students do not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and concluded that the majority of students do not read assigned readings at all, or at least not thoroughly (Burchfield & Sappington, 2000; Clump et al., 2004; Connor-Greene, 2000; Sappington et al., 2002) and concluded that
Students’ reports of when they typically read the assigned reading materials. Although this correlation is statistically significant at $p = .008$, practically speaking it does not provide much confidence in students’ self-reports. Over 25% of students who failed the objective syllabus reading check during the semester did not report prior to class even if they did not read the entire assigned reading prior to class.

Predictors of student success

The sections did not differ significantly on any of the predictors of student success, thus increasing confidence in the likelihood that any differences found on assessments of retention of material by section were because of the manipulation rather than pre-existing differences on these predictors.

Assessments of retention of material

Overall, there was mixed support regarding whether the quizzing methodology made a significant impact on students’ retention of the material. Based on Exams 1 and 2, it appeared that the manipulation in the 9 AM section did cause students to perform better on the assessment of material covered when they were quizzed using the Clicker plus written quiz methodology. However, none of the other predicted differences were found between the sections’ performance on the remaining assessments of their retention of the material tested.

Given the mixed support for the hypotheses regarding the assessments, it is unclear at this point whether the change in the quizzing methodology impacted student performance on the daily quizzes to perform retention of the material. One possible reason for the mixed results is that the quizzing methodologies did not differ in a manner necessary to cause the students to learn and remember the material (i.e., percent of reading completed, thoroughness of their reading time spent preparing for classes, class participation, and understanding of course material was not different). It is possible that both quizzing methodologies did not address sufficient student behaviors across the semester. For example, students in the 9 AM section did indeed begin coming to class and open up relevant class discussions, both written and clicker quizzing sections. Therefore, neither quizzing methodology primarily addressed the important question of whether students enrolled in the sections involved differed significantly on other variables that may impact student performance other than the quasi-experimental manipulation. Indeed, student evaluations of the course as a whole did not differ by section (even though the 9 AM section had the Clicker plus written quizzing methodology for 75.6% of the semester and the 8 AM section had only the Clicker plus written methodology for 39% of the semester). This was likely the result of the students’ quizzing methodology, which was similar across the semester. Future research will have to contend with these important questions regarding which quizzing method would lead to the most learning was the Clicker plus written method, the majority of students chose other options. Taken together, these results suggest that professors who want to incorporate daily quizzing in their course do not need to be excessively concerned about the impact of their course evaluations based on type of quizzing methodology. Additionally, it appears that students do not perceive a major impact of quizzing methodology on strengths and limitations, and future directives.

One major strength of the current study is that it was able to answer the important question of whether students enrolled in the sections involved differed significantly on other variables that may impact student performance other than the quizzing methodology. Although only one of the hypothesized differences in exam grades was found, we can be relatively sure that the difference was not due to pre-existing differences between the sections. Future research on the impact of the difference in the quizzing methodologies used for potentially important pre-existing differences to ensure they are due to the manipulation only.

In reality, five of the written quiz classes were randomly chosen and the other seven classes were randomly removed. Therefore, it is possible that the additional grading time required by written quizzes is warranted for obtaining outcomes not currently obtained by the use of daily clicker reading comprehension quizzes.
indicated they received the e-mail or were concerned the e-mail had been intended for them, nor did they indicate they remembered it when I announced the changed policy in their section.

ACKNOWLEDGEMENTS

I would like to express special thanks to the research assistants involved with this project: Billy Rush, Teresa Davis, Mia Kloth, and Britteny Stone. I would also like to express sincere appreciation to Connor Dowling and the reviewers for their feedback on previous versions of this manuscript.

REFERENCES


