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.
Effects of quizzing methodology on student outcomes: Reading compliance, retention, and perceptions

Carey Bernini Dowling

Department of Psychology, University of Mississippi, University, MS 38677, USA

(Received 29 October 2015; Accepted 28 November 2016)

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.

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 quizzing 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
grading time, and daily SRs quizzes plus written quizzes in only 25% of class sessions, which increase grading load but do not excessively.

The active learning strategies we have implemented include strategies that has included the impact on students' reading behaviors (Carlencron, 1994; Connnor-Greene, 2000; Mortling et al., 2008; Ryan, 2006). I am unaware of any studies that report the impact of quizzing on objective measures of how thoroughly students read the assigned material on a daily basis. Given the theoretical importance of reading assigned readings on time and the evidence that it suggests class performance (Clark, Middleton, Nguyen, & Zwick, 2014; 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 reported minimal student reading behaviors because student reports of reading behaviors 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” method 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 materials 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), there would be a majority of self-enhancement biases and a majority would lie on their self-reported reading compliance, but that students who failed the objective measure of whether, and how thoroughly, students read would be more self-enhancement biased and more likely to say they had read the entire syllabus.

METHOD

Participants

All students enrolled in my Spring 2014 Abnormal Psychology sections (taught at 8 AM and 9 AM MWF) were recruited for this study (56 students per section at the start of the semester). Although 101 students originally consented to participate in the course as a whole, only 98 students chose to participate in the study and only 97 students who consented in the 8 AM section, 49 completed the pre-post packet and 51 completed the post-post packet. Of the 46 students who consented in the 9 AM section, 45 completed the pre-post packet and 42 completed the post-post packet. Participants mean age was 21.5 years (SD = 3.35), they were primarily juniors (46.8%) or seniors in college (40.4%), Caucasian (55.2%), female (62.4%), and living off campus (65.7% of the objective measure). Previous research on predictors of students' success in college classes has found that students' performance goals (Elliot & Church, 1997; Elliot & Murayama, 2003), intrinsic motivation (Clark, Middlebro, Nguyen, & Zwick, 2014), conscientiousness, intelligence, and SAT performance (Conard, 2006; Kapp & van der Flur, 2012), all predict students' success in college classes. However, a number of studies examining the effectiveness of teaching methodologies that also examine whether groups of students in experimental conditions vary significantly on these or other variables of potential importance such as intrinsic motivation for the course or success in the course in general. Therefore, the second aim of the present study is to determine if the students in each quasi-experimental condition differed significantly on any of these other potential predictors of students' success in college classes. The fact that we decided to forgo comparisons between classes, but it was assumed that there would not be significant differences between students who chose to sign up for an 8 AM versus 9 AM class section means that the study does not compare students who chose to sign up at the same time. To assess this, I occasionally used clicker quiz questions that resembled exam questions for practice, but only after we had reviewed the relevant information. Thus, students who chose to sign up at the 8 AM class session did not receive any new information, but were prepared to complete the final exam along with a course evaluation (which was not part of this study), thereby ensuring that I was unaware if they were competing the final exam, the post-assessment packet, or the course evaluation. Students were told to complete 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.

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 competing the final exam, the post-assessment packet, or the course evaluation. Students were told to complete 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.

Quizzes

Students' quiz grades accounted for 15% of their final grade in the course. Throughout the semester for both sections, exam questions were created from all four questions embedded in the class plan for the day. Clicker questions were primarily designed to test their reading compliance and open class discussion. Each exam was created from the same pool of questions embedded in the class plan for the day. Clicker questions were primarily designed to test their reading compliance and open class discussion. Clicker questions were created from the same pool of questions embedded in the class plan for the day. Clicker questions were primarily designed to test their reading compliance and open class discussion. Clicker questions were primarily designed to test their reading compliance and open class discussion. Clicker questions were primarily designed to test their reading compliance and open class discussion. Students completing the pre- and post-packets were entered into a raffle for one of two $10 Amazon gift cards.
Assessments of Reading Compliance and Behavior

Assessments of Reading Compliance and Behavior during the semester included a daily clicker reading check, a retention assessment, and an exam. Daily clicker reading checks were conducted to determine if the intended dependent variable, daily clicker reading, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

Analytical explorations of daily clicker reading check. Following the analysis indicating a high level of dishonesty on the syllabus reading level checks, exploratory analyses were conducted to determine if the intended dependent variable, daily clicker reading checks, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

RESULTS

Assessments of Reading Compliance and Behavior

Assessments of Reading Compliance and Behavior during the semester included a daily clicker reading check, a retention assessment, and an exam. Daily clicker reading checks were conducted to determine if the intended dependent variable, daily clicker reading, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

Analytical explorations of daily clicker reading check. Following the analysis indicating a high level of dishonesty on the syllabus reading level checks, exploratory analyses were conducted to determine if the intended dependent variable, daily clicker reading checks, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

Analytical explorations of daily clicker reading check. Following the analysis indicating a high level of dishonesty on the syllabus reading level checks, exploratory analyses were conducted to determine if the intended dependent variable, daily clicker reading checks, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

Analytical explorations of daily clicker reading check. Following the analysis indicating a high level of dishonesty on the syllabus reading level checks, exploratory analyses were conducted to determine if the intended dependent variable, daily clicker reading checks, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

Analytical explorations of daily clicker reading check. Following the analysis indicating a high level of dishonesty on the syllabus reading level checks, exploratory analyses were conducted to determine if the intended dependent variable, daily clicker reading checks, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).

Analytical explorations of daily clicker reading check. Following the analysis indicating a high level of dishonesty on the syllabus reading level checks, exploratory analyses were conducted to determine if the intended dependent variable, daily clicker reading checks, could be trusted as valid. This was done because the daily clicker reading check was an identical clicker reading check question but was given at the start of all classes with a quiz. The correlation between students’ self-reported syllabus reading level responses and the remaining results are based only on those who responded. The majority of students who passed the objective syllabus reading check reported on both their homework (90.5%) and in-class (100%) that they read all of the syllabus. For the first self-reported syllabus reading level check, there was not a significant difference in students’ self-reported reading levels by group (passed/failed objective syllabus reading check). X² (2, N = 87) = 2.51, p = .29. However, for the second self-reported syllabus reading level check, there was a significant one-tailed, medium-sized effect of group on students’ self-reported reading levels. X² (2, N = 88) = 8.11, p = .05, Cramer’s ρ = .30, such that 100% of the students who passed the objective syllabus reading check reported “they read all” of the syllabus whereas only 70.1% of those who failed did (the remaining students chose four other options).
Students’ reports of when they typically read the assigned reading material. Normally, students reported that they read the material assigned for each class prior to that class. (which is questionable), a majority of students reported that they did not read prior to class even if they do not 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 due to 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 quizzes 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 made a difference in 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 permit comparison between daily quizzes to perform retention of the material. Based on the results of this study and the differences found on Exam 2 it appears that the manipulation made an impact on their course evaluations based on type of methodology. Additionally, it appears that students do not perceive a major impact of quizzing methodology on strengths/limitations, and future directions.

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 success. For instance, there were no significant differences on course evaluations, or the differences found on Exam 2 are the result of some other factor.

The differences in Exam 2 and no other assessments could be a result of changes in student behaviors across the semester. For instance, students may have noticed a difference in the grades they received on the quizzes at the end of the semester, and the students in the 9 AM section did receive higher grades in the quizzes than the students in the 8 AM section. However, the data collected by the authors of this study and the differences found on Exam 2 are the result of some other factor.

The differences in Exam 2 and no other assessments could be a result of changes in student behaviors across the semester. For instance, students may have noticed a difference in the grades they received on the quizzes at the end of the semester, and the students in the 9 AM section did receive higher grades in the quizzes than the students in the 8 AM section. However, the data collected by the authors of this study and the differences found on Exam 2 are the result of some other factor.
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 Britteney 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