
Using Your Tests to Teach: Formative Summative Assessment

Steven R. Wininger
Western Kentucky University

Two studies supported the efficacy of a structured method for providing students with feedback on exams and means of using the teaching potential of exams, collectively referred to as formative summative assessment (FSA). In the first study, students responded positively to the method. In the second study, students enrolled in 2 sections of educational psychology taught by the same instructor received either the FSA method or their exams to look over in class. Students retook the exam 1 week later. The FSA method resulted in an improvement of comprehension of 10%, significantly more than the control group, whose scores improved by only 2%. I also present other benefits of this method and suggestions for variations of the method as well as future research ideas.

Formal assessment and evaluation of students in higher education has a long history in the United States dating back to 1793 (Milton, Pollio, & Eison, 1986). The original purpose of formal testing was to differentiate among students with regard to level of achievement. Formal testing of students is traditionally conceptualized as summative assessment (SA). SA is an assessment that teachers use on the completion of instruction for the assignment of grades (i.e., the purpose is to report achievement). However, testing can and does serve many more purposes in educational settings. Formative assessment (FA) provides feedback to both student and teacher about learning progress with the goal of improving learning and instruction. Wininger and Norman (2005) defined FA as the measurement of student progress before or during instruction for the expressed purpose of modifying instruction and improving student performance. Reviews of studies on FA reveal that classroom use of FA results in substantial learning gains (Black & Wiliam, 1998; Fuchs & Fuchs, 1986). Typical effect sizes for these gains range from .4 to .7 *SD*. Means of conducting FA include tests (practice tests or quizzes), tasks (homework or group activities), and observation (nonverbal communication or questions asked; Linn & Gronlund, 2000; McMillan, 2001).

This article presents a marriage of FA and SA, which I refer to as formative summative assessment (FSA). Quite simply, FSA consists of going over exams in class with students and garnering both quantitative and qualitative feedback from the students about their comprehension. Individual techniques that constitute the FSA method are not novel. However, what is unique is the composite use of these techniques to create a method (FSA) that I propose in this article.

I conducted two studies to evaluate the FSA method. The first study examined students' perceptions of the method. I ex-

pected that students would have positive perceptions of the FSA method. A second study examined the effectiveness of the method with regard to improvements in comprehension. I expected students receiving FSA to demonstrate greater improvements in comprehension compared to a control group.

Study 1

Method

Participants. Thirty-eight students enrolled in a university educational psychology class served as participants. Students enrolled in this course were mostly sophomores and juniors. There were no freshmen. The majority of the students were women ($n = 26$).

Materials. The materials consisted of visual aids (e.g., Microsoft® PowerPoint® slides) of exam items, removing exam item numbers. This strategy allows the instructor to change the order of the exam items to prevent copying of answer sequences and alleviates the problem incurred from handing back multiple exam versions to go over in class. Create slides for your scoring rubric or example responses when you use open-ended items. I limit the number of questions on each slide to keep students from reading ahead and thus losing their attention.

Procedure. I posted the students' grades prior to going over the exam in class. Thus, they could have already known their scores but not which items they missed. The final exam in this class is cumulative, which makes exam feedback much more influential on student learning (Crooks, 1988).

Prior to going over the exam items, I reminded students of my syllabus policy on "arguing for points," which is "students who believe that they deserve more points for any item will need to provide their rationale in writing," a method endorsed by McKeachie (1999). I emphasized to students that going over exam items is a learning opportunity and that they should ask questions for clarification and understanding; I also encouraged them to provide feedback about items that they believed were misleading or unclear.

Individual students read multiple or alternate choice items, and then the class answered aloud in unison. Heavier prompting to facilitate understanding followed items that students answered hesitantly or with disparity. Such prompting included asking a student to explain why the correct an-

swer was correct. Another example would be asking students why incorrect distracters were not correct, which modeled critical thinking for students who missed the question. For short-essay items, I usually asked for volunteers to provide an answer. I then provided exemplars of full-credit and partial-credit answers (Carkenord, 1998). Appropriate “wait-time” (Rowe, 1986) between items facilitated student questions. Last, I presented the item analyses for the exam and informed students about how I used this information to make decisions about dropping certain items or curving exam scores. I tried to leave extra time at the end of class for students to ask follow-up questions and to look over their Scantrons to check for scoring errors. The procedure took about 30 min for a 50-item exam. After receiving feedback on their second exam, students anonymously completed a five-item survey (see Table 1).

Results

Table 1 shows descriptive statistics for all items. The feedback revealed that students’ attitudes toward the exam-feedback method were positive. All item responses except for Item 5 were approximately normally distributed. Item 5 indicated that students did not believe that simply handing back exams would be as effective as the FSA method.

Study 2

Method

Participants. Seventy-one students enrolled in two sections of educational psychology served as participants. There was an 8:00 a.m. ($n = 34$) and a 9:30 a.m. section ($n = 37$), which met on Tuesdays and Thursdays for 75 min per class. The same professor taught both sections and the students were predominantly sophomores and juniors. There were no freshmen in the classes. The majority of the students were women ($n = 57$).

Table 1. Descriptive Statistics for Student Evaluation Questions

Questions	<i>M</i>	<i>SD</i>
1. Going over the exam helps me to clarify and understand items that I missed or didn't understand.	4.66	.48
2. I would like this type of exam feedback in other classes.	4.29	.57
3. Going over the exam is a worthwhile use of class time.	4.37	.63
4. I feel that there is an opportunity to ask questions when going over the exam.	4.34	.75
5. I feel that handing back individuals their exams would be just as effective as the method currently used.	2.68	.97

Note. The scale anchors were 1 (*strongly disagree*) and 5 (*strongly agree*).

Procedure. Students took a 50-item exam consisting of true–false, multiple-choice, labeling, and matching items as the measure of achievement. The exam content spanned four chapters. This exam was the students’ first in the class. I used the FSA method for the 8:00 a.m. section. Students in the 9:30 a.m. section (i.e., control group) received a copy of the exam and individual Scantrons with the correct answers indicated for all items missed. I instructed students to ask questions for clarification or understanding. Each section received the same amount of time (about 30 min) to go over the exams.

One week from the day students reviewed the exam, they took the same exam again but for extra credit; scores on the exam determined the amount of extra credit awarded. They could earn up to 5 extra credit points toward their original scores, calculated by taking a student’s percentage score on the retake and multiplying it by the 5 possible extra credit points (e.g., 80% on the retake = 4 points). Neither of the two sections knew about this retake for extra credit prior to the class period in which they took it. However, the 8:00 a.m. section received the FSA method to combat the potential critique that the 9:30 a.m. section may have heard about the retake from the 8:00 a.m. section, prompting students in the 9:30 a.m. section to briefly study for the retake, thus allowing the FSA section to potentially score higher. In addition, because of a cumulative final exam, to be fair to the 9:30 a.m. section, they did receive the FSA method after the retake.

Results

The average scores for each class on the initial administration of the exam were 81.41% ($SD = 10.22$) for the FSA class and 81.62% ($SD = 10.16$) for the control class. Average scores on the retake were 90.82% ($SD = 6.55$) for the FSA class and 83.72% ($SD = 9.50$) for the control class. There were no significant differences between the two classes on the initial exam administration, $t(69) = -.09, p = .93$. ANCOVA revealed that scores from the initial administration were a significant covariate for retake scores, $F(1, 68) = 44.28, p < .01$. After factoring out the variance of the initial scores, students subjected to the FSA performed significantly better than the control group, $F(1, 68) = 22.10, p < .01$. In other words, students who received the FSA method scored significantly higher on the retake. Eta-squared values indicated that scores from the initial administration accounted for 39% of the variance in the retake scores and exam feedback method accounted for an additional 25% of the variance.

Discussion

Students’ responses in the first study indicated that they believe the FSA method helped them to clarify and understand missed exam items. Students also believed that the method was a good use of class time and that they would like this type of feedback in other classes. For me, the results of using this method have been an improved process, in terms of attention, and a less argumentative demeanor from students. I also have observed that my students ask more clarification questions during exam reviews, potentially resulting in an increase in learning.

The results of the second study supported the effectiveness of the FSA method with regard to student comprehension. This finding is interesting because recent survey data of faculty at my university revealed that only 48 of 215 faculty (22%) believed going over exams with students was a valuable aid to student learning. Yet, students exposed to the FSA method demonstrated an improvement of almost 10% in their test scores, whereas scores for students in the control group improved by only 2%. More important, further analysis revealed that the exam feedback method accounted for 25% of the variance in exam retake scores above and beyond variance accounted for by initial scores.

One potential nuisance variable in this study was student motivation. Specifically, a critic might infer that students who enroll in an 8:00 a.m. section are more motivated. To combat this potential critique, I examined the differences between the two sections on a behavioral measure of motivation (extra credit activities engaged in during the semester). There was no significant difference between the two sections, $t(69) = 1.54, p = .13$, which casts doubt on the motivational hypothesis as an explanation for the group differences on the retake.

The underlying cause of the achievement increase is unknown—it may be a result of enhanced attention when reviewing the exam or improvements in test-taking skills; however, identification of the underlying cause is a focus for future research. The FSA method outlined in this article is not a “set” method. An instructor could employ a less structured version of FSA, particularly if students receive exams to keep. A consideration of other factors commonly discussed in classroom assessment literature might lead to improvements in the FSA method. One possible modification of the method is to limit discussion to items that were at least moderately difficult. As Wexley and Thornton (1972) pointed out, discussing easy items has instructional value for only a small number of students. If deleting easy items resulted in equivalent learning gains, then such modifications may make FSA more attractive to instructors who are concerned about cutting into instructional time.

Future research should examine potential differences in the effectiveness of the FSA method for low- versus high-achieving students and in different subject areas. Low-achieving students should benefit most. Future research also could examine the effects of using the FSA method on students’ self-assessment abilities. Future research should examine students’ and professors’ emotional responses to returning exams using the FSA method versus a more traditional method such as handing back exams.

The FSA method has the potential to serve additional purposes beyond enhancing student learning. First, the instructor receives feedback about which concepts require ad-

ditional coverage and where there are gaps in teaching methods or student learning. Second, FSA provides a medium for garnering feedback about exam items. I have received invaluable feedback from my students about test items that has subsequently resulted in improvement of my tests. The FSA method eliminates many of the concerns that professors cite as their reasons for not providing exam feedback (e.g., emotional distress, copies of tests “getting out,” complexity of multiple forms). In the past I experienced some anxiety about providing exam feedback. However, since I have been using this method to provide the exam feedback, I actually look forward to it.

References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice*, 5, 7–74.
- Carkenord, D. M. (1998). Assessing the essay feedback technique of providing an example of a full-credit answer. *Teaching of Psychology*, 25, 190–192.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438–481.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53, 199–208.
- Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice Hall.
- McKeachie, W. J. (1999). *McKeachie’s teaching tips: Strategies, research, and theory for college and university teachers* (10th ed.). Boston: Houghton-Mifflin.
- McMillan, J. H. (2001). *Classroom assessment: Principles and practice for effective instruction* (2nd ed.). Boston: Allyn & Bacon.
- Milton, O., Pollio, H. R., & Eison, J. A. (1986). *Making sense of college grades: Why the grading system does not work and what can be done about it*. San Francisco: Jossey-Bass.
- Rowe, M. B. (1986). Wait-time: Slowing down may be a way of speeding up. *Journal of Teacher Education*, 37, 43–50.
- Wexley, K. N., & Thornton, C. L. (1972). Effect of verbal feedback of test results upon learning. *The Journal of Educational Research*, 66, 119–121.
- Wininger, S. R., & Norman, A. D. (2005). Teacher-candidates’ exposure to formative assessment in educational psychology textbooks: A content analysis. *Educational Assessment*, 10, 19–37.

Notes

1. I thank Amy Wininger for her willingness to read earlier drafts of this article. In addition I thank the six anonymous reviewers, Adrienne Y. Lee, and Randolph Smith for their comments on earlier versions of this article.
2. Send correspondence and requests for additional details or clarification on any of the steps outlined in this article to Steven R. Wininger, Department of Psychology, 1 Big Red Way, Bowling Green, KY 42101; e-mail: steven.wininger@wku.edu.

Copyright of Teaching of Psychology is the property of Lawrence Erlbaum Associates. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.