

Examination Retest Procedures: Effects on Performance, Test Anxiety and Attitudes



Fichten: Professor of Psychology, Dawson College (B.S., McGill; M.A., Concordia). Adler: Professor of Psychology, Dawson College,



Montreal, Quebec; (B.A., McGill; M.A., Concordia). McGill; M.A., Concordia).

By CATHERINE FICHTEN
LILLIAN ADLER

USE of grading and examinations to enhance learning has recently received attention (McKeachie, pp. 1118-1164).¹² The goal of this study was to determine whether providing students with the option to take a second course examination soon after the first (retest condition) would result in (a) better performance, (b) superior retention, (c) less test anxiety, and (d) more positive attitudes toward the course than when only one writing (test-once condition) is allowed.

Investigations of the use of testing as an integral part of instruction have proceeded along three different paths. One has focused on the effects of administering practice tests and providing different types of feedback to the student on his performance both on practice and course examinations. The results of studies in this area have shown conclusively that giving practice tests improves performance on examinations and other test-like events,¹⁷ and that feedback, especially discussion type, is more effective than no feedback.^{18,19} However, the efficiency of feedback from actual examination situations when the student is aware of the opportunity to improve his performance in the very near future has not been thoroughly investigated. As motivation to achieve competence and to improve grades may be related to the well known goal-gradient phenomenon,¹⁶ feedback may be more effective when students are given an opportunity to make use of it in the very near future rather than at a much later date.

A second line of investigation has centered on evaluating the effects of various grading techniques on performance. Although results have been contradictory, a recent study by Kositsky & Franken⁷ in which students had to redo an assignment until they attained a predetermined level of proficiency has shown that this technique has improved performance. Whereas in the Kositsky and Franken study students were threatened with compulsory rewriting, in the present experiment

performance on tests rather than on assignments was studied and rewriting was optional, thus allowing students to act in accordance with their individual goals, needs, and aspiration levels.

A third approach has explored the connection between test anxiety, learning, and performance. The association of these variables, although the topic of numerous studies, is very complex and still poorly understood (Levitt, pp. 120-132)⁸ (Mechanic, 180).¹⁵ A frequent finding, however, is that the relationship between test anxiety and the other two variables is often curvilinear in heterogeneous populations, and that optimal anxiety level is inversely related to task complexity (Levitt, pp. 116-119).⁸ McKeachie, Pollie & Speisman¹⁸ found that anxiety during examinations builds up to the point that it interferes with memory and problem solving, and that reduction of stress results in improved test performance. In addition to other benefits, it was expected that the use of the optional retest procedure would reduce test anxiety to a preferable level.

A further consideration is examinee awareness of day to day variation in performance on test-like events (Cronbach, p. 128).⁴ Students are capable of monitoring their own performance levels,⁹ i.e. of assessing whether they have done well or poorly on an examination. As learning and achievement are motivated by the desire to attain^{2,6} and demonstrate,¹² competence, it should be assumed that students will retake an examination and possibly relearn course material when faced with the realization that they could have done better. Although the performance monitoring aspect of testing has been generally ignored in the literature, it was felt that students should be allowed a certain amount of autonomy in setting their own aspiration levels. It was hypothesized that if a student is allowed to retake an examination, he will do so should he feel that performance or prior learning was not optimal.

The present research focused on the consequences of enabling students to improve performance by allowing a retest on the same content. It was predicted that those in the optional retest condition would (a) obtain higher examination grades, whether they opt to be retested or not, (b) perform at least as well on a compulsory test, (c) obtain higher scores on a retention test, and (d) be less anxious than students in the traditional test-once condition. It was further predicted (e) that in the optional retest condition, those choosing to rewrite would improve their scores and (f) that the more student centered evaluation procedure coupled with better comprehension, would produce greater satisfaction with the course in general¹ when this is measured by a course evaluation.

THE STUDY

The subjects were 200 first and second year college students enrolled in a one semester Introductory Child, or Social Psychology courses. Two different instructors each taught four sections of approximately 25 students each. In all courses lectures, films, laboratory, group discussion, and a textbook were used. Some subjects, however, were omitted from certain analyses due to lack of data or oversubscription in some cells of the experimental design.

MATERIALS

Two measures of examination anxiety were used. One of these was score on items 29-39 of the College Form of the Test Anxiety Questionnaire (TAQ) developed by Mandler & Sarason.¹¹ This was computed in accordance with the scoring system recommended by Mandler & Cowen.¹⁰ The maximum possible score, 111, and the minimum score, 11, indicated extremely high and almost no anxiety respectively. Since it was felt that the TAQ score may only be a very general indicator, a second measure of examination anxiety (E_a) was formulated in order to assess anxiety experienced at the time of testing. This consisted of score on the four point Likert type item, "How anxious are you now?", responses ranging from "very anxious" to "not anxious at all" with values of one and four respectively.

Practice, compulsory (Test 1), optional retest (Test 2), and retention examinations were prepared by stratified random sampling without replacement from a pool composed of mutually exclusive items. In four of the eight sections all tests consisted of multiple choice questions, while in the others a number of short answer items were also included. Both types of examination items were designed to test memory as well as problem solving skills. Scoring of the latter was done by "blind" graders. In order to ascertain the equivalence of the compulsory and optional tests, four test-retest pairs were administered to students in two non-participating sections. Each pair was completed during 100 minute examination sessions, with Test 1 and Test 2 presented in counterbalanced order. As no significant differences in difficulty were found this way, it was assumed that all other examinations constructed by this method were also equivalent. The same sampling procedure was followed for both mid- and end-of-term testing times.

A course evaluation containing 13 five point Likert type items which were applicable to all three courses was also constructed. These items were designed to assess attendance and student reaction toward the course and subject matter.

At the beginning of the term, all students in each of the participating sections were randomly divided into two groups, A and B. At this time, students were informed of the basic experimental design, although not of the hypotheses. Two examination sessions were held, one in the middle and the other at the end of the term. Both of these had equal weighting on the final course grade, and there was no overlap in content. Students in both groups were encouraged to write a practice test one

to two days before each writing of Test 1. The practice tests, equivalent to the course examinations in format, difficulty, and loading on various content areas, were administered in order to partially equate "test-wiseness" in the two conditions. All students wrote Test 1 both at the mid- and end-of-term examination sessions. In order to counterbalance test-once and optional retest conditions, Group A had the option to be retested at the end-of-term writing, while Group B had this chance at the mid-term session. For those students in the optional retest condition who chose to rewrite, the higher score contributed toward the final course grade. Small group feedback meetings followed both the practice test and Test 1. The optional retest (Test 2) was administered during regular class meeting time, four to five days after Test 1.

Five minutes prior to the start of Test 1 and Test 2, the 11 item TAQ and the E_a question were administered to all students present. At the end of the 50 minute examinations, both questions and answers were collected. Approximately three weeks after Test 1, all students completed the course evaluation. A retention test was administered to both groups five weeks after Test 1 of the mid-term examination session. All students were told at this time that should their final course grade be borderline between two adjacent grade categories, a higher score on this test would be used to raise it. Although the two groups experienced the test-once and optional retest conditions in counterbalanced order, neither the retention examination nor the course evaluation was administered after the end-of-term writing session.

RESULTS

Due to the small number of students in each condition in each class, examination data from the various courses was pooled. In order to equate examination means and standard deviations, all raw data were transformed to standard (z) scores. This was done separately for each examination session in each class for each analysis.

In order to compare practice test scores of students in the test-once and optional retest conditions, practice-test raw scores were first transformed to z scores, and a 2 x 2 Latin Square analysis of variance was performed on the z scores. No significant differences between the scores of students in the two conditions were found. A similar comparison on Test 1 results also yielded no significant differences.

Of those students who had the option to be retested, the scores of 68.6% were higher on Test 2 than on Test 1; the scores of 5.8% remained the same, while the scores of 25.6% were lower on Test 2 than on Test 1. A normal curve test (Edwards, p. 39)⁵ indicated that the number of students obtaining higher scores on Test 2 was significantly different from that expected to occur by chance alone if students write two equivalent tests ($z = +3.45$, $p < .0006$). The Test 2 scores of students who chose to rewrite were significantly higher ($F = 36.190$, $df 1/84$, $p < .005$) than their scores on Test 1 (see Table 1). The z scores used for this comparison were calculated on the basis of the combined Test 1 and Test 2 data of students who rewrote. In addition, stu-

Table 1
MEANS AND STANDARD DEVIATIONS OF TEST SCORES IN
OPTIONAL RETEST CONDITION

Group	Examination			
	Test 1		Test 2	
	M	SD	M	SD
Group A	-.245	1.015	+.245	.893
Group B	-.303	.981	+.303	.808

Note. -Test grades were transformed to standard scores. N = 43 per Group.

dents in the optional retest condition obtained higher examination grades ($F = 16.628$, $df 1/182$, $p < .005$) than those in the test-once condition (see Table 2); this appears to be a consequence of the higher Test 2 scores obtained by those who rewrote. For this comparison, the Test 1 raw scores of students in both conditions were converted to z scores. In the case of students who received a higher score on Test 2 than on Test 1, the z score used was obtained by transforming the Test 2 raw score using the mean and standard deviation of the corresponding Test 1 raw data.‡

Table 2
MEANS AND STANDARD DEVIATIONS OF EXAMINATION GRADES

Group	Condition			
	Test-once		Optional retest	
	M	SD	M	SD
Group A	+.101	.972	+.366	.894
Group B	-.057	1.122	+.295	.912

Note. -Test scores were transformed to standard scores.

No support was found for hypothesis (c), as using a t test comparison, no significant difference was found between the two conditions on retention test z scores.

Repeated measures 2 x 2 analysis of variance comparisons revealed no significant differences between the Test 1 and Test 2 TAQ scores, nor between the Test 1 and Test 2 E_a question responses of students who chose to be retested.

No significant differences were found, using a 2 x 2 Latin Square analysis of variance comparison, between the TAQ scores of students in the test-once and optional retest conditions. However, in a similar comparison of E_a question responses, it was found that specific test anxiety was significantly higher ($F = 4.261$, $df 1/136$, $p < .05$) in the test-once than in the optional retest con-

Table 3
MEANS AND STANDARD DEVIATIONS OF EXAMINATION
ANXIETY QUESTION RESPONSES

Group	Condition			
	Test-once		Optional retest	
	M	SD	M	SD
Group A	2.406	.880	2.594	.975
Group B	2.290	.842	2.420	.812

Note. + The lower the score, the higher the anxiety. Group A had the option to be retested at the end-of-term session while Group B had this chance at the mid-term session.

dition (see Table 3). In the same analysis, a significant Condition x Order interaction ($F = 7.922$, $df 1/136$, $p < .005$) was also found, indicating that the difference between the two conditions was greater in the examination session at the end of the term than in the middle of the term.

Although a t test comparison revealed no significant difference between the course evaluation ratings of students in the test-once and optional retest conditions, it is interesting to note that there was a significant positive correlation (Pearson product-moment $r = +.488$, $df 128$, $p < .001$) between the students' happiness at participating and their perception of the relevance of the present research.

DISCUSSION

One of the major hypotheses of this study, that students in the optional retest condition would achieve higher examination grades than those in the traditional test-once condition, was confirmed. It was also shown that the superiority of this group was predominantly due to higher scores on the retest, since the Test 2 scores of those students who chose to be retested were significantly higher than their Test 1 scores, and since Test 1 results in the optional retest condition were not significantly different from those in the test-once condition. That no significant differences were found between the two conditions on either Test 1 or practice test scores indicates that students in the optional retest condition did not merely postpone studying, at least no more than the other students.

Superior scores on the retest were due not to any long term enhancement of learning, as there was no difference in retention between the two conditions, but may have been due to ameliorated achievement conditions resulting from effective use of performance monitoring and possibly of massed learning. Students experiencing the optional retest treatment effectively evaluated their own performance on Test 1, and had recourse to action if they considered that it was not at an optimum level.

The data on examination results cannot be explained by "test-wisness," as at the end-of-term Test 1 students previously in the optional retest condition had no advantage over those previously in the test-once condition. Feedback effects cannot explain the findings either, since no differences were found in retention between students in the two conditions. It is, however, possible that higher grades on Test 2 than on Test 1 are a consequence of two massed learning "cramming" sessions.²⁰ Should this be true, then two sessions are superior to one only immediately after the second one, and not five weeks later. More information concerning the study habits of students would be needed to substantiate this hypothesis.

No significant differences between the two conditions on the TAQ were expected or obtained, as the items comprising this test were designed to assess general test anxiety, rather than that felt just prior to a specific examination. These results simply confirm that Groups A and B did not differ initially in anxiety, and that the

significant difference found between the two conditions when anxiety was measured by the E_a item was, in fact, due to the experimental manipulation.

Allowing students the option to be retested resulted in lower specific test anxiety, especially at the end of the term. It is encouraging to note that as no significant difference was found between the Test 1 examination scores of students in the two conditions, anxiety in the optional retest condition had not been sufficiently diminished to impair performance. This is also implied by the absence of a significant difference between Test 1 and Test 2 anxiety scores of those students in the optional retest condition who chose to be retested, even though their examination performance was significantly higher on Test 2.

The absence of significant differences between the two conditions on course evaluation ratings may have been due to a methodological deficiency in the study. Had this measure been administered just prior to Test 1 instead of three weeks after the changeover in experimental conditions, and had items specifically designed to assess the presence and consequences of monitored performance been included, the predictions related to attitudes may have been substantiated.

Experiencing the optional retest treatment improved test performance and diminished examination anxiety, at least in this examination situation. Continued exposure may also help to provide the student with skills not ordinarily taught in the classroom, namely, the ability to evaluate his own performance effectively and to set himself realistic goals and aspiration levels.

FOOTNOTES

* This research was carried out during 1971-1972 and was supported by a grant from Dawson College. The authors thank Dolores Gold for her interest in the present study, and for encouragement and many useful suggestions. We are also indebted to Lewis B. Birch for reviewing an earlier version of this manuscript. Appreciation is also expressed to Zita Martos and Claudia Bourdon for their valuable assistance, and to Susan Raymond for her helpful comments.

† Reprint requests should be addressed to: C. S. Fichten, Dawson College, Department of Psychology, 350 Selby Street, Montreal 215, Quebec.

‡ If Test 2 raw scores had been included in the conversion to z scores, the z scores of students who wrote only Test 1 would have been artificially lowered, as Test 2 scores were higher than Test 1 scores.

REFERENCES

- ¹ Ash, M. J. Nondirective teaching in psychology. *Psychological Monographs*, 1951, 65 no. 4 (Whole No. 321).
- ² Butler, R. A. Discrimination learning by the rhesus monkey to visual-exploration motivation. *Journal of Comparative and Physiological Psychology*, 1953, 46, 95-98.
- ³ Caffrey, B. Lack of bias in student evaluations of teachers. In *Proceedings, 77th Annual Convention of the American Psychological Association*, 1969, 641-642.
- ⁴ Cronbach, L. J. *Essentials of Psychological Testing*. (2nd ed.) N.Y.: Harper & Row, 1960.
- ⁵ Edwards, A. E. *Experimental Design in Psychological Research*. (3d ed.) N. Y.: Holt, Rinehart & Winston, 1968.
- ⁶ Fiske, D. W. & Maddi, S. R. *Functions of Varied Experience*. Ill.: Dorsey, 1961.
- ⁷ Kositsky, N. & Franken, R. Effects of grading and rewrite procedures on an academic performance. *Psychological Reports*, 1970, 17, 244-246.
- ⁸ Levitt, E. L. *The Psychology of Anxiety*. Indianapolis: Bobbs-Merrill, 1967.
- ⁹ Luria, A. R. Frontal lobe syndromes. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of Clinical Neurology*. Vol. 2. New York: Wiley, 1969.
- ¹⁰ Mandler, G. & Cowen, J. Test anxiety questionnaires. *Journal of Consulting Psychology*, 1958, 22, 228-229.
- ¹¹ Mandler, G. & Sarason, S. B. A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 1952, 47, 166-173. American Documentation Institute, #5485.
- ¹² McKeachie, W. J. Research on teaching at the college level. In N. L. Gage (Ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.
- ¹³ McKeachie, W. J. *Teaching Tips: A Guidebook for the Beginning College Teacher*. (6th ed.) Massachusetts: Heath, 1969.
- ¹⁴ McKeachie, W. J., Pollie, D. & Speisman, I. Relieving anxiety in classroom examinations. *Journal of Abnormal and Social Psychology*, 1955, 50, 93-98.
- ¹⁵ Mechanic, D. *Students Under Stress*. N. Y.: Free Press of Glencoe, 1962.
- ¹⁶ Miller, N. E. Experimental Studies of Conflict. In J. McV. Hunt (Ed.), *Personality and the Behavior Disorders*. Vol. 1. N. Y.: Ronald Press, 1944.
- ¹⁷ Rothkopf, E. Learning from written materials: An exploration of the control inspection behavior by test-like events. *American Education Research Journal*, 1967, 3, 241-249.
- ¹⁸ Samuels, J. The effects of post-test relevant pre-tests and discussion type feedback on learning and retention. *Psychonomic Science*, 1969, 16, 57-68.
- ¹⁹ Sassenrath, J. M. & Garverick, C. M. Effects of differential feedback from examinations on retention and transfer. *Journal of Educational Psychology*, 1965, 56, 259-263.
- ²⁰ Woodworth, R. S. & Schlosberg, H. *Experimental Psychology*. N. Y.: Holt, 1954.