

Developments in Business Simulation & Experiential Exercises, Volume 15, 1988

GRADING AS A TEACHING AND FEEDBACK MECHANISM: INVOLVING STUDENTS IN THE GRADING PROCESS

Roger D. Roderick
California State University, Long Beach

Jim A. Wilterding
Boise State University

and

Donald L. Bates
California State University, Long Beach

ABSTRACT

The experiment reported here represents an effort to design a method which combines grading and feedback in such a way that they become contributing parts to the learning process. After taking an essay exam, students participated in the development of guidelines which the professor would use in grading the exam. A procedure was developed whereby (a) appropriate key answers were assured, (b) agreement and acceptance of "correct" answers were obtained, and (c) feedback and questioning focused on content. Participation increased in general course sessions as well as exam process; attitudes improved; grade levels rose; and overall interaction between faculty and students progressed.

INTRODUCTION

The conception of the experiment described in this paper emanated from a concern that the examination process be a learning experience. Certainly, the preparation for the examination should be such; but beyond that, our concern is that the grading itself and the post-examination feedback sessions be substantively productive events. It is our contention that students should be active participants rather than passive listeners, (or, worse yet, non-listeners) in those events. Specifically, we want to induce students to actively think about the "best" answer to the questions and to work enthusiastically and objectively to develop such responses even after they have taken the exam. There are two critical ingredients prerequisite to successful accomplishment of this goal: one is cognitive--understanding; the other is attitudinal-- acceptance. The experiment was designed to address each.

Both the motivation behind and the mechanical of the experiment were based on the belief that examinations, the grading of the examinations, and the feedback sessions that follow typically tap only the surface of their potential. The easiest and most comfortable explanation of this shortfall--i.e., easiest from the instructor's perspective--is to allude to the general lack of interest and intellect on the part of the students. There is doubtless some germ of truth to that allegation. However, it is argued here that there is almost always a substantial link between the limited productivity of the events and the way in which instructors conduct them.

In most instances, for example, only minima) effort appears to be devoted to the pedagogical aspects of

examination feedback sessions. There appear to be three principal barriers which inhibit productivity: (1) the amount of time which the instructor allocates to the explanation is too short (in some cases, even zero); (2) attempts to determine whether students are receiving that which is being transmitted during feedback are minimal or are entirely omitted (in some cases, the omission may reflect similar omissions in the regular class sessions); and (3) faculty defensiveness to student inquiries. (Student defensiveness is discussed below, but students clearly have no monopoly on defensive behavior).

The experiment reported here represents an effort to design a method which combines grading and feedback so that they become contributing parts to the learning process. As such, they would elicit student thought, concentration, participation, and acceptance even after the examination is over. This would be in contrast to a situation wherein students (a) pay no attention at all, (b) pay attention only to the grade they received, or (c) react defensively to any answers at variance with the ones they gave at exam time. Too often, those with high grades exit the process once they see that they have scored well, and, alternatively, those receiving low grades are either totally "turned off" or spend the time looking for loopholes, technicalities, and arguments, without regard for whether those "outs" have any substantive merit.

The fundamental purpose of the experiment, then, was to enhance learning. The premise is that this can be accomplished through a process which provides full and immediate feedback, elicits maximum participation from students, and in so doing generates strong commitment.

Description of the Experiment

This is the second in a series of experiments that the authors have used to address the above concerns. The first consisted of the following steps:

1. Instructor grades exams.
2. Grades and comments are written on separate sheets (not on exam papers).
3. Exam papers (without grade and comment sheets) are returned to students.
4. Students grade their own papers in class, using books, notes, etc.
5. Instructor provides feedback on the "right" answers.

Developments in Business Simulation & Experiential Exercises, Volume 15, 1988

6. Students grade their exams again, revising according to what the instructor has said.
7. Instructors' grade and comment sheets are returned to the students. (The original grade remains, as the purpose of this part of the exercise is learning, rather than grading).

Follow-up on that experiment indicated that (1) over time, the self-assigned grades and the instructor-assigned grades converged; (2) convergence occurred across all performance levels; and (3) the average grade for the class rose.

The current mechanism has expanded on that. It has a much greater degree of active participation. Students are involved in establishing the accepted answers to be used as a key in the final grading process. They also interact with each other as they do so, gaining an understanding of how their participation affects and is affected by their contemporaries. Further, the dimension of commitment is added. This creates a condition in which the students will be more attentive to the content of the feedback sessions and will receive the feedback in a more positive frame of mind. They will react better to the grades they are awarded and will approach the next segment of the course with a better attitude. These improved states, of course, create an environment more conducive to learning than would be the case were they absent.

The Setting. The class was an undergraduate course in Employee and Labor Relations. It met twice weekly from 9:00 to 10:30 a.m. There were 34 students, 16 of whom were Industrial Relations majors and 11 of whom were Construction Management majors. The course was required, then, for 27 of the 34 students. Of the 34, only three had previously been in any of the participating instructors' courses.

The experiment took place at the time of the second exam in the term. This was one of 2 interim examinations, each weighted at 15% of the course grade. The exam consisted of four essay questions (with the weights distributed as follows; 40%, 30%, 20% and 10%). Given the timing of the experiment - about 10 weeks into a 15 week semester--the students had already been assigned a substantial amount of group work. Also, students within the two predominant majors (IR and CM) tend to be fairly close. They were, then, rather well acquainted and accustomed to working together by the time the experience occurred.

The Process. The nature of this course is of a sufficiently subjective nature that the preponderance of the questions have no single "right" answers. Examination questions were of the essay and "short answer" types. This creates a situation in which feedback and understanding are most important if the examinations are to be learning experiences. It is also that in which they are most difficult. The difficulty factor is largely a function of the variety of approaches and answers which may be judged to be "right." A further complicating factor is that the grade for any single response is frequently assigned, in part, by the instructor's comparing

its quality to that of responses provided by other students. Thus, ex-post individual explanations of grading are inhibited by the then absence of those benchmarks against which the paper in question was originally assessed.

In the earlier experiment (referred to above), the instructor unilaterally graded the papers, and then used a multiple loop feedback process to develop understanding of the "right" answers (I.e., the answers which the instructor eventually provided). In the current iteration, the students are brought even further into participation in the grading of the exams. Throughout, the general guidelines for effective participation were followed:

1. The problem was clarified.
 - a. Students knew what was to be accomplished.
 - b. The problem clearly was one of student interest.
 - c. The instructor did not interfere by specifying solutions.
 - d. Procedures were clear.
2. Ideas from the students were accepted with judgments being made.
 - a. They were written on the charts so that all could see them.
 - b. They were discussed to assure that everyone understood them.
3. Alternative responses were evaluated by the class.
 - a. The instructor attempted to get everyone involved.
 - b. All responses were discussed and compared for quality.
 - c. The instructor pushed the class to develop the best possible answers.
4. A final set of responses was agreed upon.
 - a. Quality and acceptance were considered.
 - b. The students understood that the responses selected:
 - (1) Would be used to grade their papers.
 - (2) Were not subject to further appeal once accepted.

As for the specifics of the process, the first step was initiated on the day of the exam. At that time, the instructor told the students to bring their books and notes with them to the next class session, and pointed out that there should be no absenteeism at that session. When they came to the session they were told that they would be developing the key with which their exam papers would be graded. The full 90-minute class period was devoted to this experiment.

Going through the questions individually, students were asked to orally provide the major points germane to each. The procedure here resembled a nominal group technique exercise. For each question, the students first generated a series of key points or idea. These were recorded on a large flip chair at the front of the room. This continued until everyone was satisfied that all their ideas were out and had been recorded. Next, those who had contributed the points were allowed to clarify or elaborate upon them. After that, debate and negotiation occurred, where after final

Developments in Business Simulation & Experiential Exercises, Volume 15, 1988

statements were drafted. These were written in full on the flip charts and students had a last chance to object before formal acceptance.

Once having arrived at those lists, the class then turned to developing details for explanation and elaboration of each major point. The discussion, negotiation, and refinement in this round were essentially the same as when zeroing in on the major points. Finally, agreement was reached and each detail item under each major point was written up. These were, for the most part, a series of "one-liners."

To elicit true participation and to assure confidence in the sincerity of the professor, students were told that discussion of each point could continue until satisfactory resolution--even if such resolution might require extension to the next class meeting. It was explained, however, that once the key became finalized, no subsequent appeals would be heard.

One question that arose had to do with the extent to which the major points would be "fleshed out". How much detail would be developed? To focus on this, a very detailed "grade A" answer was established for the 40% question. Students were then allowed to vote on whether to repeat that for the remainder of the questions or to erect only skeleton guidelines for each. The vote (29-5) was to leave it as a skeleton guideline. They had already worked out one "detailed A" answer, so they knew it could be done. They had discussed the potential benefits and they knew the costs associated with going for the fuller, more prescribed set of guidelines. They knew what the results would be like and they knew that the professor was totally willing to follow that course, should they have chosen to do so. They chose not to do it.

The next step was for the professor to grade the set of exams using the written guidelines as the basis for the grading. This was done, and the results were distributed to the class at the following class meeting. Students were told that the professor would consider any question they might have about judgements on whether the response was close enough to the guidelines, but that no request that the guidelines be altered would be heard.

Discussion

The first sign that this experiment was "working" was that the class "disciplined" those who (a) did not take the task seriously; or (b) tried to slant the key in the wrong direction to fit any off-target answers that they themselves might have given. The better students kept the process on track and moving toward the "right" answers, and within the (preferred) time limits. The professor had very little to do in terms of providing answers. If the class had not done so, of course, the professor would have had to become more directive. (It was, after all, "participation," not "anarchy.")

Virtually everyone with grades of "D" or "F" came in to talk with the professor. That was encouraging (and, time-consuming). Even more encouraging was the fact that all were there to talk about content, what the professor expected, and how to study for the next exam. There was

one exception to that. One student disagreed with the key to one of the questions. She complained that the instructor had shut the door on discussion and firmed up the key for that question too quickly. In fact, the evidence was unmistakably clear that she (a) had had ample time to register a concern in the classroom discussion but had chosen not to do so; and (b) had been told that questions regarding guidelines were not going to be heard after agreement had taken place in the classroom. She was the only student who came in with a problem and left still voicing a basic lack of acceptance of or appreciation for the process.

Of all those who came in, the most vocal and aggressive were those who had not spoken up in the class--both throughout the term and in the development of the exam key. The results here were especially positive. Every person in this group admitted (indeed, some even volunteered self criticism) that they should have been more aggressive in the activity, and most assured the professor that they intended to participate more fully in the remainder of the course in general and particularly in any subsequent guideline establishment that might occur.

In the class period following the distribution of the grades, a questionnaire was circulated. Responses were kept anonymous. Thirty of the 32 present responded that they felt the grades they received were fair. Even those who had some disagreement with the grade assigned did not feel that it was unfair, since they had seen the different perspectives on the questions earlier as they participated in the construction of the key. They felt as though they had had some control in the process, and therefore did not view it as a totally arbitrary one.

Attitudes in the Class improved. Learning seemed to become more of a cooperative than an adversarial process. On the exam at which all this started, 60% of the class received grades of 60 or lower. On the next exam, the class average was 76, with only four scores of 60 or below. In the former exam, the question causing most problems was the one weighted at 40%. The class average there was 54. A question in that same topic area was included in the latter exam. In that case, the average was 92. The instructor also noted that the quality of the students' group case analyses improved over the remainder of the term. This is believed to in part reflect the attitudinal improvement and in part to reflect the fact that the students had undergone a group problem-solving experience (i.e., the development of the key).

A final note is that not only did students in this class want to continue the process, but students in other classes began lobbying for the process in their classes--a phenomenon that was met with varying degrees of enthusiasm by other professors.

This process would appear to be most readily transferable to other courses where the material under examination does not consist simply of sets of facts. At any time students have any degree of latitude as to which or how many "facts" to include in their answers, however, the process would

Developments in Business Simulation & Experiential Exercises, Volume 15, 1988

also be appropriate (even for some varieties of multiple choice questions).

REFERENCES

Crino, Michael D., and Michael C. White. "Feedback Effects in Intrinsic/Extrinsic Reward Paradigms." Journal of Management. Fall, 1982, pp. 95-108.

Gordon, Ronald D. "The Self-Disclosure of Interpersonal Feedback: The 'Dyadic Effect' in a Group Context." Small Group Behavior. August, 1985, pp. 411-13.

Heller, Scott. "Collaboration in the Classroom is Crucial if Teaching is to Improve." Chronicle of Higher Education. March 11, 1987, pp. 17-18.

McMullen-Patrick, Miriam, and Maryellen Gleason. "Examinations: Accentuating the Positive." College Teaching. Fall, 1986, pp. 68-91.

Nevo, Baruch, and Janet Sfez. "Examinees' Feedback Questionnaires." Assessment and Evaluation in Higher Education. Fall, 1985 pp. 236-49.

Ogilvie, John R., and Beth Haslett. "Communication Peer Feedback in a Task Group." Human Communication Research, Fall, 1985, pp. 79-98.

Perry, Raymond P., et al. "Perceived Control in a College Classroom: Limitations in Instructor Expressiveness Due to Noncontingent Feedback and Lecture Content." Journal of Educational Psychology. April, 1986, pp. 96-107.

Roderick, Roger D., and Jim A. Wilterding. "Grading as a Feedback Mechanism." Proceedings of the Midwest Division, American Institute for Decision Sciences, 1980.

Siders, James A. "Instructor, Self and Peer Review: A Formative Evaluation Triad." College Student Journal. Summer, 1983, pp. 141-44.