The Bio-de-grade-able Class

By EILEEN T. BROWN

Here are some things they said:

"I think the idea stinks."

"I am surprised, delighted, glad, and curious at the same time."

"Maybe you'll lose your job."

"I think it's great."

What possibly could have elicited such divergent reactions from a class of college-bound high school sophomores? These comments were in response to the announcement of an introspective experiment in which *they* were to be the guinea pigs. For one school semester, conventional grading was eliminated from the biology classroom. At the outset of the experiment the students were told that regardless of what they did they would receive an A for that term. Under these conditions they were to subject their own attitudes, behavior, and performance to rigorous analysis. It is the purpose of this paper to discuss the motivation, procedure, and results of this experiment, as well as to consider the circumstances that hampered it and the questions it raised.

Background

Grades, whether numbers or letters or any other symbol hung about the students' necks and recorded for posterity, have long been part of the American educational system. Only recently, with the advent of student activism, have the conventional methods of evaluation come under intense fire. Many colleges have found it worthwhile to institute a pass–fail system in a limited number of courses taken by each student. In some experimental primary and secondary schools advocating open, more flexible education, grades as we know them have been eliminated or minimized. The Murray Road Annex to the Newton (Mass.) High School is an example. In October 1968, in a summary of developments of the first year of this experimental high school, the teachers (who also acted as administrators) reported as follows:

> From the beginning, the faculty attempted to focus student attention on the processes they considered essential to self-education: *self-assessment of one's personal needs, interests and goals,* consideration of alternative courses of action, choice of a course of action, and *self-evaluation of one's progress. It appeared clear that giving grades to students would be inconsistent with this emphasis.* The faculty felt that awarding grades to the students would encourage them to direct their efforts toward pleasing the teacher rather than toward setting and meeting their own learning goals. The faculty feared that the progress of students toward self-directed learning would be severely inhibited if they felt that the only thing that really counted for the record was the teacher's appraisal of their work. Consequently the faculty proposed that no grades be awarded at Murray Road School, but that instead the student's record consist of (a) an evaluation of the student's progress written by the teacher, together with (b) an evaluation of his progress written by the student himself. The high school principal accepted the proposal, and rescinded his original directive that grades be given in all major courses. [Emphasis added.]

Best-selling books such as *Teaching as a Sub-*

The author graduated from Barnard College in 1960 and received a master's degree in teaching from Harvard University the following year. She has taught general science and biology in the schools of Newton, Mass., and at present is a part-time teacher of 10th-grade biology at Newton High School, Newtonville, Mass. 02160. Mrs. Brown recently was on a year's leave of absence in Jerusalem, where her husband taught at Hebrew University.

Table 1. Correlation of performance and attitude toward the experiment. Figures are number of students.

Evaluation of performance	Initial attitude toward the experiment	
	Favorable	*Unfavorable*
Did work all the way through	2	0
Did work at beginning but not at end	9	1
Did work at end but not at beginning	1	0
Did almost no work at all	5	2

versive Activity (Postman and Weingartner, 1969) and *Education and Ecstasy* (Leonard, 1968) have suggested that a more open-minded approach to evaluation is in order. "Open educators" have said they believe that objective measures of performance may be an actual hindrance to the learning process; that the qualities of education most easily measured are not necessarily the most important; that errors are a necessary part of learning; that the preferred source of verification for the learner should be from the material itself, not from some outside authority; and that it is next to impossible to design a truly objective and reliable measure of learning.

On the other hand, most teachers and administrators, as well as students, believe that some of the functions served by grades are necessary and cannot be filled in any other way. This group would maintain that grades are indeed essential in indicating to a student his own progress or lack of it; that forces in the outside world (colleges, employers) require this measure in evaluating potential candidates; and that without grades a large part of what motivates students to learn is gone. It can be argued that these functions of grades apply only to schools as they now exist and that they will eventually be obsolete. But most schools do not change that rapidly. In the here and now, testing and grades are very much a part of the educational picture.

In view of the controversy about grades, I felt my students should have an opportunity to examine more intensively, and in a meaningful way, some of the variables involved.

Procedure

After an initial discussion of the functions served by grades as they saw them, the students were informed that for the second marking period (from the beginning of November through the end of January) they would all be guaranteed a grade of A. This new situation was to be treated as an experiment from the point of view of scientific method. In order to have adequate data at the conclusion of the experiment, each student would keep a diary—daily, if possible—in which he would record his behavior, attitudes, and performance over the course of the experiment. I, too, would keep a diary, recording in it my changing attitudes and reactions.

Some of the curriculum for that semester was directly pertinent to the grading experiment itself. For example, the class was assigned readings in *Pygmalion in the Classroom* (Rosenthal and Jacobson, 1969), in which the question of a teacher's expectation and its effect on students' performance is raised. The research carried out by these authors was analyzed from the point of view of scientific method, and the more general phenomenon of the self-fulfilling prophecy was considered both in terms of human relations in the world at large and in terms of its applications to the students' own education. For the specific questions that guided the discussion see list 1.

Similarly, the class went on a field trip to Murray Road, the experimental high school mentioned earlier in this paper; there, evaluation was handled by means of student–teacher conferences and accumulation of credits, rather than by more standard methods. My students were encouraged to observe classes, sit in on school meetings, and ask questions of Murray Road students. Thorough discussions preceded and followed the visit.

During the remainder of the term the standard curriculum (which happened to deal at this point with bacteriology and then with population biology) was followed. Of course, in our attempts to discover how we must now deal with such things as tests, homework, laboratories, and student reports under this new regimen, numerous discussions dealing with the original aims of the grading experiment ensued. These ran the gamut from whether or not students elected to tear up a particular quiz to the nature of the learning process itself and of the role of such factors as competition, guilt, motivation, fun, and tension.

At the conclusion of the semester the students

Table 2. General opinions: effects of grading or lack of grading; value of the experiment. Figures are number of students.

Student opinion	*Held by*
Did better with grades	14
Did better without grades	1
Grades (or lack of grades) make no difference	4
No opinion	4
Experiment was a worthwhile experience, however painful	12
Didn't learn much from the experiment (because it merely reinforced former ideas or because the sources of error were too formidable)	4
No opinion	4

were to integrate their experiences and reactions in a paper thoroughly analyzing the experiment. I distributed a list of 16 questions, which served as a guide for this paper; see list 2. Aspects of most of these questions were discussed in class before the writing of the analyses.

Results

A considerable number of the results of the grading experiment might have been predicted. (These are summarized in tables 1 and 2.) For example, it might have been expected that while most students generally tended to do less work as the term progressed (as indicated in table 1), the students who consistently performed well were those who had received high Bs or As the first semester. Generally speaking, the girls were more conscientious than the boys. Those who stopped working in biology usually blamed (i) excessive work in and pressure from other classes that *were* being graded and (ii) time given to extracurricular activities. A majority of the class performed less satisfactorily on tests and quizzes, although they were given the option of not handing them in. Many students felt unsettled by not seeing a number or letter grade on their assignments and tests, in spite of the fact that I found myself giving more time and effort to the numerous comments I put on their papers to compensate for the lack of a grade.

It is interesting to note that the student in category C in table 1, whose work pattern differed completely from that of the majority of the class, reported that as the experiment was drawing to a close he realized he was losing the opportunity to learn for its own sake rather than for a grade, and he did not want the opportunity to completely slip through his fingers.

Somewhat less predictable was the fact that students began to transfer their experiences in biology to other subjects, where they began to question the value of particular assignments and sometimes decided, on this evaluation, not to do a given assignment. Similarly, many in the class experienced "culture shock" when returned to a conventional grading system in third-term biology: they found it really difficult to readjust to old routines.

The students' response to the grading experiment in general, and their own performance in the presence or absence of grades, is summarized in table 2.

List 1. Questions guiding the discussion of Pygmalion in the Classroom.

1. What is a self-fulfilling prophecy? Give one example from the book. Give one example from your own experience.
2. How is the concept of the self-fulfilling prophecy relevant to your education, if at all?
3. Examine table 1–1 on page 7 [of the book]. Under which category does the personal example you gave in question 1 fit?
4. What is an anti-self-fulfilling prophecy? Have you ever run across one in your educational experience so far?
5. What general conclusions do the authors draw about the survival value of "expectancy" and "predictability" for the human species?
6. What possible explanation is there for the results obtained in the animal experiment described on page 174 of the reprint?
7. What is meant by the "ethical grounds" referred to in the second paragraph of page 175?
8. Describe briefly the procedures used in the experiment that was performed.
9. Analyze the experiment performed according to scientific method. State the original hypothesis (procedure already described), the control, the major significant variable, sources of error, and conclusion.
10. What is meant by the "halo effect"?
11. What possibilities for further research are suggested by this experiment?
12. What possible "theories" explain the results obtained?
13. What are the implications, if any, of this research to:
    a. our second-term experiment on grading?
    b. the grading system currently in use at Newton High School?
    c. the tracking system currently in use at Newton High School?
    d. racial prejudice in the schools?
    e. anything else you would like to comment on?

List 2. Questions to be answered in the students' analysis of the grading experiment.

1. Analyze the "experiment" from the point of view of scientific method; *i.e.*, discuss purpose, procedure, observations, controls, variables, sources of error, and conclusions.
2. Include diary.
3. Discuss the significance of latency to this experiment; any experiment.
4. What, if anything, did you learn about yourself from this experiment?
5. What, if anything, did you learn about the educational system in which you are functioning?
6. Would you consider the experiment a success or a failure? Explain.
7. Would you recommend it be tried again with future classes? Explain.
8. Describe your own feelings about being back on a conventional system of grading in biology.
9. What, if any, insights did the reprint *Pygmalion in the Classroom* give you about your education in general; about this experiment in particular?
10. What, if any, insights did the Murray Road trip give you about your education in general; about this experiment in particular?
11. How was your attitude towards tests, labs, homework, and class discussions affected by the experiment?
12. At the beginning of the experiment, this class listed the following as the functions served by grades:
    a. lets a person know how he's doing
    b. measurement for college, jobs, etc.
    c. motivation
    How have your thoughts about these functions changed as a result of the experiment?
13. Discuss any problems that arose in analyzing this experiment from your role as a participant–observer.
14. What do you predict will be the implications of our second-term experiment on your third-term behavior (performance, attitude, etc.) in biology; in other subjects?
15. How has this experiment affected your relationship with your parents; your friends; your other teachers; me?
16. Anything you would like to add.

Paradoxically, perhaps one of the more important findings was that no definite and all-encompassing conclusions could be drawn at all, because of the large number of sources of error involved. Some of the major errors—discovered by the students themselves—are listed below.

1. The time limit on the experiment (only one term) introduced the concept of latency in research procedures. Several students maintained that the true results of the experiment required more than three months to make themselves felt and that after an initial slacking-off period they might have worked effectively in the absence of grades. However, because the experiment lasted only three months there was not sufficient opportunity to "find out."
2. Students were still being graded regularly in other classes. Only one pressure valve (biology) was released, while the others continued at full steam. This made it difficult to assess a truly "gradeless" learning experience.
3. Some students felt that in the long run they were indeed being graded in biology, due to such external measures as departmental finals and College Board achievement exams.
4. Much of what was observed during the experiment could have been due to the so-called Hawthorne effect; that is, the newness and uniqueness of the situation was a more significant causal factor than the variables of the experiment itself.
5. As a scientific experiment, it was almost impossible to draw any but the most tentative of conclusions, because of the lack of adequate controls. In such a complex situation it is almost impossible to delineate and control all the extraneous variables impinging on the subjects of the "experiment."
6. Students felt that since the time between Thanksgiving and Christmas is often a period of laziness anyway, it was putting the experiment at an unfair disadvantage.
7. My own attitude as a teacher may have affected the experiment in a way I had not originally intended. I discovered that as the term progressed I began to expect students not to do homework in biology, and thus I may have introduced a self-fulfilling prophecy. In fact, it was pointed out to me that I actually somewhat decreased the number of assignments I gave towards the end of the term.

In spite of all these serious (perhaps devastating) limitations, it is still possible to take another look at the experiment, in terms of its major positive and negative effects on the biology class.

Positive Points

1. It brought under investigation the whole function of grades and evaluation.
2. It created a more relaxed, freer atmosphere in our classroom.
3. It made students question what their own responsibility is for their own education.
4. It caused them to question the relationship between output and effort (as well as time and intensity) in their education.
5. It caused them to question the value of particular assignments in biology as well as in other subjects.
6. It entailed a thorough analysis of scientific method, the uses to which it can be put, and its limitations. What, for instance, is important to look for in an experiment when human behavior is involved?
7. It made the students question their own nature. They saw whether predictions they made about themselves were really true and valid.
8. It encouraged a kind of sincerity and honesty with themselves and their teacher, which unfortunately doesn't always exist in schools.
9. It made them question the definition and nature of learning, although no clear answers were obtained.
10. They learned more about the nature of motivation in the learning process.

Negative Points

1. Several students reported that their enthusiasm for biology as a subject decreased as a result of the absence of grades. However, in a significant minority of cases the opposite was true.
2. Some students found out things about themselves that they were not happy to know, and this resulted in ego-deflation and loss of self-confidence.
3. Many in the class experienced strong feelings of guilt.
4. The quantity of strictly biologic subject matter that was dealt with in the time period of the grading experiment was obviously less than it might have been under "normal" circumstances.
5. The situation in our biology class created feelings of envy on the part of students not in my class.

In neither of the above lists has an attempt been made to arrange the items according to their importance. This is an individual and highly subjective decision, which each reader can make for himself.

Conclusion

The grading experiment obviously raised more questions than it answered. I, for one, believe these questions needed to be raised. Many of them have already been discussed in the body of the paper and will not be repeated here. I will merely add that when students begin to question the validity of any experimentation involving human behavior, start to ask fundamental questions such as whether it is necessary to be educated for freedom, and wonder what actually is involved when learning occurs—then, I believe, worthwhile activity is taking place in the classroom.

Not all students need a grading experiment such as the one we performed in order to raise such issues. Not all students are capable, at this point in their educational journey, of dealing with them, even in a superficial manner. That most of the stu-

(Concluded on p. 74)

The acute effects of inhalation of cigarette smoke on pulmonary function have been demonstrated (Chiang and Wang, 1970). Although no difference was measured in the lung volumes and flows of smokers and nonsmokers, breath wash-out studies indicated significant differences of residual volume, lung clearance index, nitrogen wash-out time, size of lung compartments, and alveolar dilution factors. In fact, these authors recommended that no pulmonary-function test be conducted until at least an hour after the patient's last cigarette.

These three "now" effects of smoking—oxygen-carrying capacity of the blood, depression of the oxygen–hemoglobin disassociation curve, and acute effects on pulmonary function—can be effectively used to show the cause-and-effect relationship of smoking on the physiology of the smoker.

REFERENCES

AMERICAN COLLEGE OF CARDIOLOGY [convention report]. 1970. Benefits of exercising, hazards of smoking. *Modern Medicine* 38 (8): 47-49.

CHIANG, S. T., and B. C. WANG. 1970. Acute effects of cigarette smoking on pulmonary function. *American Review of Respiratory Diseases* 101 (6): 860-968.

NATIONAL AIR POLLUTION CONTROL ADMINISTRATION. 1970. *Air quality criteria for carbon monoxide.* U.S. Government Printing Office, Washington, D.C.

SCIENTISTS' INSTITUTE FOR PUBLIC INFORMATION. 1970. *Air pollution.* Privately printed, New York.

SPAIN, D. M., and V. A. BRADESS. 1970. Sudden death from coronary heart disease: survival time, frequency of thrombi, and cigarette smoking. *Chest* 58 (2): 107-110.

NEW STATE RESPONSIBILITIES

The Woodrow Wilson Center for Scholars, at the Smithsonian Institution, has published *Managing the Environment: Nine States Look for New Answers.* Prepared by Elizabeth Haskell and Smithsonian staffers, the report tells how nine representative states—Illinois, Minnesota, Washington, Wisconsin, New York, Vermont, Maine, Maryland, and Michigan—have begun to modernize their governmental structures to deal with ecologic problems. New responsibilities are being undertaken by land-use and waste-management agencies, the state courts, and consolidated environmental departments. *Conservation News* calls the report "a valuable document for any group . . . interested in environmental problems and the practical details of modifying 'the system' to solve them." A who-does-what feature is a list of persons interviewed in each state. The report is available from the Woodrow Wilson Center for Scholars, Smithsonian Institution, Washington, D.C. 20560.

Bio-de-grade-able . . . . . from p. 70

dents in this particular class believed that the experiment was a worthwhile experience and that it should be repeated with future classes is some indication of its success.

REFERENCES

HOLT, J. 1968. *John Holt on testing.* School Services Group, Sterling Institute, Boston.

LEONARD, G. B. 1968. *Education and ecstasy.* Delacorte Press, New York.

POSTMAN, N., and C. WEINGARTNER. 1969. *Teaching as a subversive activity.* Delacorte Press, New York.

ROSENTHAL, R., and L. JACOBSON. 1968. *Pygmalion in the classroom.* Holt, Rinehart and Winston, New York.

LIFE ON THE SEABED

Ecologic theory says a stable environment should lead to the presence of many kinds of life. The ocean floor is a highly stable environment. It is constantly dark, constantly cold, and constantly low in available food. Such an environment, stable or not, has been thought to be relatively inhospitable to life; the ocean floor has even been called a biologic desert.

But Howard L. Sanders, a Woods Hole (Mass.) Oceanographic Institution scientist, working with National Science Foundation support, has been writing even the deep-ocean floor into ecologic theory. He has been finding that even though the density of life may be low, the diversity of species on the deep-ocean floor "is about the same as that in the physically stable, shallow, tropical marine environment," where life abounds.

Earlier efforts to sample life forms from the seabed 450 to 4,500 m beneath the surface met with relatively little success; the limited number of samples brought up in dredges led inevitably to the conclusion that there were few kinds of life to find. But through the use of improved collecting equipment of their own design, Sanders and his colleagues have been able to find tens of thousands of organisms where their predecessors found few or none at all. In 19 samples, for instance, they retrieved 3,257 specimens of a single bivalve species, only one specimen of which had ever been seen before. Of another almost unknown species, they retrieved 255 specimens in 10 samples. Sanders has found that deep-sea-floor species, of which there appear to be thousands, vary far more with depth than they do with geography, and that the sea-floor temperatures are often far more critical to their survival than are pressures.