OPERANT LEARNING PRINCIPLES APPLIED TO TEACHING INTRODUCTORY STATISTICS¹

WILLIAM A. MYERS²

WISCONSIN REGIONAL PRIMATE RESEARCH CENTER

A grade of A was given in an introductory statistics course for meeting a set of contingencies that included no work outside of class (except by request), near-perfect performance on exams following each unit of work in a programmed text, correction of all exam errors, self pacing of work, and the chance to finish the course early. A grade of incomplete was given otherwise. Correlations among performance measures failed to show any meaningful relationships between time taken to finish the course, errors made on exams, and errors made in the programmed text. Responses to a five-part questionnaire were overwhelmingly favorable to the course, but did not vary as a function of grade point average, time taken to finish the course, or number of errors made on exams. The uniformly high level of performance, the students' lack of interest in social contact with the instructor during class, and the absence of drop-outs are all attributed to the contingencies employed, chief among which, according to the instructor's judgment and student rankings, were self-pacing, frequent non-punitive exams and a guaranteed grade of A for near-perfect work at every stage.

Persons interested in the application of operant methods to instruction are familiar with such implied criticisms as: does the student feel deprived of personal contact with the instructor? Does the student feel he is being treated as an object or is not recognized as an individual? Was the material so easy that a sense of excitement and adventure was totally missing? Does the student think the grading system was fair? Would the student voluntarily take another course taught by similar methods? Were the bright students bored an appreciable part of the time? In addition to these attitudinal questions, there is the question of the quality of objectively measured performance in a course taught by operant methods.

Answers to the above questions have so far been favorable for an operant approach to instruction (Keller, 1968; McMichael and Corey, 1969), but additional data are needed, particularly from students.

There is another whole set of attitudinal questions which, since they do not impugn operant techniques, are not asked by critics; namely, are there features of the operant approach to instruction which students actively like?

The main purpose of the introductory statistics course described here was to arrange the conditions of learning so that a group of students with varied background and ability would master the course content. Performance and attitudinal data were collected to determine whether that purpose was realized and to evaluate the set of learning conditions employed.

The methods used were similar enough to those described by Keller (1967) to be considered as a partial replication of his methods. Certain departures from his methods were deliberately introduced in the hope that they would reduce the number of students failing to complete the course, the major flaw in Keller's approach.

The course was given at the University of Wisconsin in an eight-week summer session in 1968. Of the 38 students enrolled, there were seven sophomores, 15 juniors, 10 seniors, and six adult specials. Most were in the college of liberal arts, with a scattering from business, agriculture, and education.

¹I am grateful to Barry Bragg (now Dr. Bragg) who, as teaching assistant in the course, worked harder than I had a right to expect.

²Now at the Minneapolis, Minnesota, Veterans' Hospital. Reprints may be obtained from the author, 2648 Humboldt Avenue South, Minneapolis, Minnesota 55408.

METHOD

Contingencies

At the first class meeting, the instructor explained the rules of the course and gave a brief rationale for them. Students were told that they could work only in class; that the time allowed for work was continuous 2.5 hr period each weekday; that they could arrive and leave whenever they chose, but the text book (Elzey, 1968) could not be taken out of class; that each unit of the book and the exercises following it had to be completed in writing before taking the exam for that unit; that each exam must be passed with fewer than three errors before beginning the next unit; that all errors on exams had to be corrected whether the exam was passed or not; that a grade of A would be given for passing all 24 unit exams and a grade of Incomplete for anything less; that individual tutoring could be had anytime by raising a hand or while an exam was being graded, but it had to be initiated by the student himself; that the instructor and teaching assistant were to be considered equals insofar as grading exams was concerned, and that the course was over when the last unit exam in the book was passed.

Unit exams varied in length and type of question depending on the nature of the material covered in the corresponding unit of the text. An effort was made, however, to include a number of questions requiring calculation from each unit. Students moved freely between the classroom and another room containing desk calculators.

Daily Study Sessions

A session began when the instructor unlocked the storeroom where the texts and exams were kept between sessions. Students found their own books and began work. When they were ready for an exam, they took it from a folder after showing the instructor or teaching assistant that the unit and the book exercises had been done. When they finished an exam, they had the option of bringing it either to the instructor or the teaching assistant for grading.

When they left for the day, books were left in the classroom, except in cases where the instructor gave the student the option of taking the book home for the weekend. This op-

tion was offered when a student needed more time to complete the work.

When asked on several occasions if they wanted lectures to supplement the text, only a handful said yes. Accordingly, no lectures were given.

RESULTS

Performance

Figure 1 indicates that there were large individual differences in the time taken to complete the course, in the number of errors made on exams, and in the number of errors made in working through the text. The Pearson Product-Moment correlation between number of days to finish the course and errors made on the exams was 0.414 ($P < 0.05$); between number of days to finish and errors made in the book, $r = 0.43$ ($P < 0.05$); and between errors in the book and errors on exams, $r = 0.04$ (n.s.).

Fig. 1. Performance among students in time taken to finish the course, number of errors made on unit exams, and number of errors made in the programmed text.

Work done in class was generally intensive. Although it was not prohibited, there was virtually no eating, sleeping or magazine reading. When asked why they did not take advantage of the freedom to leave class for a break, most students said they were acutely conscious of a time pressure, even though, objectively, there was more than enough time.

Questionnaire

There were five parts to the questionnaire. In the first part were 15 items comparing the

course as taught against what they imagined it would have been under the lecture method. Each item was rated on a seven-point scale whose midpoint was "no difference between the course as it was given and a lecture course" and whose end points were "a lot more by the course as it was given" and "a lot less by the course as it was given." Since they did not actually take the course by the lecture method (except for four students who had previously failed or dropped the course when it was given by the lecture method), their judgments were necessarily between the course as it was taught and their impression of the lecture method from past experience.

Table 1

Features of the course receiving the first five ranks as a function of students' cumulative grade point average (on a four-point grading scale).

GPA		
3.0 and Above	2.5 and Above	Below 2.5
8	8	8
3	3	11
11	12	12
12	11	3
13	13	1

Note—The numbers designate the following items in the list of features to be ranked.

1. Being able to move around or take a break when you felt like it.
3. Taking an exam on each set.
8. Being able to work at your own pace.
11. Not being penalized for mistakes by lowering the grade.
12. Knowing you would get an A if you did everything asked of you.
13. Being asked to correct mistakes on exams.

Out of 18 items in Part I, 16 elicited relatively extreme judgments, all favoring the course as taught. Some of the judgments most favorable to the course were (1) "the amount of material fully understood" ($x = 2.03$: the verbal equivalent lies between "a lot more" and "a moderate amount more" by the course as given); (2) confidence in solving statistical problems ($x = 2.11$); (3) amount of time bored in class ($x = 6.62$: the verbal equivalent lies between "a lot" and "a moderate" amount less by the course as given); (4) the extent to which natural differences of ability among the students were taken into account ($x = 2.00$); (5) the extent to which the course failed to

recognize you as an individual ($x = 6.22$). All of these findings agree with Keller's (1967).

In Part II, the students ranked a list of 13 features of the course in the order of "their importance to you". The fact that students have very different past histories of academic performance is reflected by the high variability of the rankings. The item receiving the highest mean rank ($x = 2.65$) and the lowest variability was "Being able to work at your own pace". The item receiving the next highest rank ($x = 4.89$) was "Taking an exam on each set", followed by "Not being penalized for mistakes on exams by lowering the grade ($x = 5.54$)".

Table 1 shows which features of the course received the first five ranks as a function of cumulative grade point average. All groups agree that self-pacing is most important. Although agreement is close on the remaining four items shown in the table, there are important differences. While all groups consider grading contingencies important, those who have been well rewarded in the past value means (Item 3) above ends (Items 11 and 12), whereas those who have been poorly rewarded in the past value ends (Items 11 and 12) above means (Item 3).

Section III asked for absolute judgments on a five-point scale concerning the difficulty of exams, amount of work required for exams, adequacy of book, and amount of vocal instruction. The mid-point was "about right" and the extremes were "much more than necessary" and "much less than necessary". The means of the judgments were all within half a scale point of the midpoint and the variability was very small (range of the Average Deviation was 0.05 to 0.63).

Section IV asked for personal evaluations of the teaching assistant and instructor on strictness of grading, extent to which they were helpful and encouraging, and their skill in explaining material. With one exception, the means of the judgments were close to but on the favorable side of the mid-point of a seven-point scale. The exception occurred on the item concerning the instructor's grading of exams. The mean judgment was 3.54, which deviates slightly from the midpoint towards the "too strict" end of the scale.

The results of Section V, which asked the students to predict their future choices between the type of course as given and the lec-

Table 2

The number of students willing to take future courses taught by operant versus lecture method.

Question	Yes	No
Would you take another course like this one in statistics?	29	6
Would you take another course by the lecture method in statistics?	10	26
Would you take another course in psychology by the lecture method?	34	1
Would you take another course in psychology taught like this one?	27	6
Would you take another course like this one in subjects besides statistics and psychology?	29	6

Note—The responses do not add to the same total for each question due to several missing answers.

ture method, are shown in Table 2. Their responses favored the operant method, but did not show a willingness to give up the lecture method entirely.

In answer to the question, "What did you like most about the course?", 19 of 37 mentioned self-pacing. Other frequent responses were "receiving positive reinforcement for each unit" and "having immediate knowledge of results". Only one of the students mentioned something other than a programmed contingency.

By contrast, many answers to the question, "What did you like least about the course?", were irrelevant to the teaching method used (*e.g.*, "having to take the course in the psy-

chology building, which is extremely ugly and sterile").

The final section, entitled simply "Comments", contained many remarks favorable to the course, the instructor, and the teaching assistant. The most frequent theme of these comments and also of conversations before class was the need for more courses taught by these methods. The enthusiasm of students for the method used, although it is not as easily interpreted as numerical data to specific questions, was astonishing in its volume and repetitiveness. One student asked three days in a row if he was really going to get an A for meeting the instructor's demands.

In order to answer questions concerning the effects of individual performance differences on attitudes towards the course, the students were divided into groups based on their cumulative grade point averages, the number of class days required to complete the course, and the number of errors made on the unit exams.

The data in Table 3 do not show any discernible effect of grade point average or number of class days to complete the course on questionnaire ratings. The number of errors made on exams is a different story. On most of the items in Part I, the groups (C2 and C3) that made the most errors on the exams had mean ratings more favorable to the course than the group (C1) that made the fewest errors on exams. To complete the symmetry of this relationship between number of errors and attitude towards the course, the group (C3) with

Table 3

Comparisons of mean ratings on Part I of the questionnaire as a function of grade-point average, time taken to complete course, and number of errors on exams. Entries are the number of items from a total of 18 on which the mean rating of one group favored the course more than the mean rating of another.

Grade-Point Average		Time to Finish Course		Errors on Exams		
$A_1 > A_2$	$A_1 > A_3$	$B_1 > B_2$	$B_1 > B_3$	$C_1 > C_2$	$C_1 > C_3$	$C_2 > C_3$
9	11	9	9	3	4	7

A_1 —grade-point average of 3.0 or greater

A_2 —grade-point average between 2.5 and 3.0

A_3 —grade-point average less than 2.5

B_1 —up to 35 days

B_2 —between 36 and 40 days

B_3 —between 41 and 45 days

C_1 —between 4 and 15

C_2 —between 16 and 25

C_3 —between 26 and 53

the most errors gave ratings more favorable to the course than the group (C2) that made an intermediate number of errors on 11 of the 18 items of Part I.

Item 6 of Part I deserves special mention because of the high frequency with which boredom is ascribed to operant methods of instruction. The item asks "... what per cent of the time were you bored in class?" A rating of seven means "a lot less by the course as given". The group with the highest grade point average had a mean rating of 6.92 with an average deviation of 0.15 ($N = 12$). The middle and lowest groups had mean ratings of 6.38 ($A.D. = 0.95$) and 6.58 ($A.D. = 0.62$) respectively.

Of 38 students taking the course, 37 received As, one received a B³, and no one received an Incomplete.

DISCUSSION

In Keller's course (1967) 20% of the students took Incompletes and another 19% failed or dropped the course. In McMichael and Corey's course (1969), 12% of the contingently-managed group dropped the course. The fact that there were no dropouts in the course described here is probably due in part to its being given in summer school when many students attempt to meet curriculum requirements.

Another reason may have been the contingencies used in grading. The over-all requirement for an A may be thought of as a second-order multiple schedule in which the student had to complete a fixed number of work units (24), each requiring a variable amount of work [FR24(VRx)]. The consequence of completing each unit may be a combination of events including (1) the satisfaction of passing the unit exams (2) the opportunity to proceed, (3) the confidence produced by near-perfect performance and (4) advancing faster than other students. Meeting the overall requirement produced three academic credits of A and, to judge from the questionnaires, various pleasant affective states. Failure to complete the course not only produced an undesirable grade, but it negated all the work done before quitting. Many students said they would have

been satisfied with a grade lower than A, so the contingencies should thwart the option of mediocre performance, providing the instructor has chosen to require A-quality work and has made this clear in his work contract with the students. The perseverance usually seen in ratio schedules (Ferster and Skinner, 1957, pp. 39-41) coupled with an unwelcome event contingent on quitting may have been sufficient to prevent dropping-out in this course.

The frequently heard charge that programmed texts are boring, particularly to students may or may not be true, but a distinction should be made between the boredom attributable to the text itself and that attributable to a set of contingencies that includes a programmed text. As a record of feelings about the set of contingencies used in this course, the questionnaire data indicate that no one was bored, least of all the students with the highest grade-point averages. A related criticism is the charge that programmed learning is too easy, but all three grade-point groups judged that they had done a moderate to a lot more work than they would have done in a lecture course, and none of the groups thought the grading of exams was lenient, so it is safe to conclude that none of the groups thought the course too easy. The questionnaire items dealing with (1) the amount of personal contact with the instructor, (2) the extent to which natural differences of ability was taken into account, and (3) the extent to which the student was treated as an individual all received highly favorable ratings regardless of grade point average, time taken to finish the course, or number of errors made on exams. The agreement among students, given these differences, was the most surprising finding in the questionnaire data. It seems highly likely that the favorable attitudes reported were a function of the success that all but one student experienced. Providing that the success is earned in a way that meets acceptable academic criteria, this would seem to be an uncontroversial outcome. The experience of success may also explain the paradoxical result recorded in Table 3. Those who made the most errors liked the course most; perhaps because they were successful in spite of their errors.

The behavior of the students in this course also bears on the much discussed issue of student-faculty contact. Although they were friendly and talkative (especially about the

³Of 38 students, 37 received an A and one, who went through the course in his own way without consulting the instructor, received a B.

teaching method being used) outside of class hours, they chose to limit their contacts with the instructor during class hours to matters concerning the content of the course, in spite of the fact that they could have had conversation with the instructor at any time simply by raising their hand. The great majority refused the offer of supplementary lectures and most students cut short attempts by the instructor to talk about anything but the subject matter. As one student put it, "With the chance of finishing the course early, why should we waste time talking to you?" Keller (1965) also reported that attendance at lectures given as supplements to his course in introductory psychology were poorly attended. This raises the possibility that the amount and kind of contact that students seek from faculty is not some sort of behavioral invariant, but rather a function of the learning conditions operative in a given course. At the very least, this finding falsifies the generalization that social contact between student and instructor during class is a necessary condition for effective learning.

The correlations between performance measures were much smaller than would be expected on the view that these measures are largely determined by the students' personalities. Knowing how many errors a student made on the programmed text, for example, was not useful in predicting his rate of progress or his exam performance. The fact that planned contingencies can produce low correlations among performance measures shows that performance is not an invariant beyond the instructor's ability to change, but rather an outcome that can be made sensitive to classroom conditions. This implies that the instructor bears a good part of the responsibility for arranging the conditions that produce high-level performance.

The view that the student bears some responsibility for his own performance is equivocally supported by the present data. Table 1 shows that students can identify the major conditions of learning that led them to high levels of performance. Of 13 features of the course given to them for ranking, the students ranked them in nearly the same order of importance as the instructor. This finding is equivocal, however, because it is possible that students do not know how to put these conditions of learning into an effective combination. It is also an open question whether they would

have identified the important contingencies at all without at least one exposure to their effectiveness.

There were, of course, sizable individual differences in performance, but these were not considered in assigning grades. Their value to the students is evident, however, in that the brighter and harder working students finished the course earlier. As part of the instructor's explanation to the students of the contingencies in the course, it was announced that performance differences were relevant to grades only if a student failed to do quality work even under optimal conditions. It could then be justifiably said that a student was not capable of earning an A. It was further explained, however, that (1) the conditions for this course were probably not optimal, and (2) given university entrance requirements, everyone present could do A-level work. In consequence of this rationale, it was decided that a grade of B or C would mean only that a student had not done the work, not that he could not do it. Under the conditions of the course, therefore, the practice of employing grades to designate ability levels could not be justified and was not used for that purpose. Instead, either a student did all the work required and received an A, or he failed to do some of the work and got an Incomplete.

The number and role of staff in this course with 38 students was very different from the staffing of similar efforts aimed at large groups. Both Keller (1967) and Malott (1969) employed a hierarchy of personnel while, in the present case, there were only two staff members and the hierarchical distinction between them was deliberately minimized. It would appear that staffing can take many forms, providing basic contingencies are met.

The most frequent argument in favor of operant methods in education is that with ever-increasing student populations they are a fiscal necessity. The fact that an operant approach worked well in a small group should help to highlight its educational value as well.

REFERENCES

- Elzey, F. F. *A programmed introduction to statistics*. Belmont, Calif.: Wadsworth, 1966.
- Keller, F. S. "Goodbye Teacher . . ." *Journal of Applied Behavior Analysis*, 1968, 1, 79-89.
- Keller, F. S. Neglected rewards in the educational

- process. Proceedings 23rd American Conference of Academic Deans, Los Angeles, Jan., 1967. Pp. 9-22.
- Keller, F. S. In a talk given to the Center for Research in Human Learning at the University of Minnesota, 1965.
- Ferster, C. B. and Skinner, B. F. *Schedules of reinforcement*. New York: Appleton-Century-Crofts, 1957.
- Malott, R. W. and Svinicki, J. G. *Contingency management in an introductory psychology course for 1000 students*. Western Michigan University. Unpublished.
- Skinner, B. F. *The technology of teaching*. New York: Appleton-Century-Crofts, 1968.
- McMichael, J. S. and Corey, J. R. Contingency management in an introductory psychology course produces better learning. *Journal of Applied Behavior Analysis*, 1969, 2, 79-83.

Received 27 August 1969.

(Revised 7 June 1970.)