AN EXPLORATORY STUDY OF TEACHER EFFECTIVENESS IN BEHAVIORAL SCIENCE

Harriet Seligsohn
University of South Florida

The subject of teaching effectiveness is well represented in print. To date, it appears that the primary investigative method used is that of having current students evaluate the behavioral qualities of the classroom performance of teachers. Other factors which have been explored are student attitudes toward the educational process; observers' ratings; student progress; and former student evaluations. Most of these approaches yield teacher characteristics in the form of personality traits such as warm, friendly, permissive, etc. It appears that many studies of teaching effectiveness lack adequate operational definitions.

In some of the literature on this topic, objective approaches to determining teacher effectiveness have been urged. However, these avenues have rarely, if ever, been followed.

This study was conducted to (a) determine relative effectiveness of teachers of behavioral science on the basis of student progress, (b) relate this progress index to student-teacher agreement on instruction practices, (c) identify critical teaching practices, and characteristics of the teachers and students of the high and low gain groups, and (d) investigate subject matter retention in relation to the progress index.

METHOD

The progress index was based on initial and common final examinations over the course material and was determined in two ways. The first method consisted of computing z-scores for the section means on both the initial and final tests and subtracting the former from the latter. A difference of plus or minus one z-score defined the high or low progress groups. The sections were ranked according to the difference in z scores. The second rank order was developed by computing (1) the regression equation to predict the final examination raw score from the initial test score; (2) the mean predicted and actual scores by section; and (3) the difference between the predicted and actual means. Those sections at the extremes by both ranking methods were used for further analysis.

The extent of agreement between the students and teachers for each section was derived from questionnaires completed by the teachers and students. This analysis was based on eight questions common to both instruments and three from the student questionnaire. A minimum of fifty per cent was defined as agreement. That is, on the common items, fifty per cent of the students in a section had to choose the same response as the teacher and on the three student questions, fifty per cent had to choose the response favorable to the teacher. Other items in the questionnaire dealt with various teacher or student characteristics and were used to describe high and low gain groups. Chi-square, Fisher's Exact Probability Method or the Median Test was used in the analysis.

Finally, eight items from the final examination were included on the final examination of the follow-up course and the raw score difference was used as the measure of retention.

RESULTS

Student-Teacher Agreement

The ranking methods mentioned above yielded nine high gain and seven low gain sections out of thirty-six classes. The hypothesis that the number of agreements and disagreements were the same in both the high and low gain groups was tested by the Median Test using Fisher's Exact Probability Method. The probability of having this or a more extreme distribution of medians on either agreements or disagreements in the high and low gain groups was .02. Thus it appears there is a relationship between learning gain as defined by the tests used in this investigation and student-teacher agreement on instruction methods.

The eleven items used to determine agreement were analyzed individually in two ways. First, a 2X2 chi-square analysis was used to find which items contributed to the significant difference in student-teacher agreement for high gain and low gain groups. The number of agree and disagree responses was added across all sections by high gain and low gain groups in the following format.

Group	Agree	Disagree
High gain		
Low gain		

For example, if there were 5 possible responses for an item, the number of students who chose the same response as their teacher were tallied in the agree column and the number who chose the other four responses were tallied in the disagree column for the appropriate gain group. Note that the actual response did not have to be the same across sections. The results of this analysis are in the middle column of Table 1.

In the second case, a 2XC (C equals number of choices available) chi-square was used to ascertain if the responses of the high and low gain groups differed. Here, the agreement variable was ignored.

	Response Number				
Group	1	2	3	4	5
High gain					
Low gain					

The right-hand columns of Table 1 and Table 2 contain these results.

Teacher Practices and Characteristics

The number of instructors represented by the high and low gain sections is small but there are trends evident in our measure of teaching practices between the high and low gain group. These are shown in Table 3.

Group Characteristics

Additional characteristics which were investigated for the two groups are shown in Table 4.

There is a tendency for the teachers of the high gain sections to be younger, to be female and to have less education than the teachers of the low gain sections. The high gain sections are more likely to meet three rather than two times a week (where the total number of meeting hours are the same).

As for retention measured by the eight items common to the final examinations of this course and its follow-up, there was no significant difference between the two groups ($z = -.25$).

Table 1

Significant results of the chi-square analysis between the upper and lower groups on the eleven items used to determine instructor-class agreement. Eight items involved matching questions on both questionnaires. The last three were determined from the student questionnaire alone.

Subject	p	Student-teacher agreement of groups	p	Distribution of responses of high and low gain groups on student questionnaire
Satisfaction with accomplishment in course	.001	More upper group agreed with instructor and were satisfied; more of lower group disagreed and dissatisfied.	.02	More of upper group were satisfied--lower group divided 50-50.
Preferred type of tests	.001	More of upper group agreed and preferred objective tests; more of lower group disagreed.	.05	Both groups preferred objective tests but more than expected of lower group preferred combination tests.
Responsibilities defined as generalized or specific expectations			.001	Majority of both groups felt responsibilities were defined as generalized expectations but fewer than expected of lower and more than expected of upper group felt responsibilities were defined as specific expectations.
Necessity of class attendance	.01	Although the majority of both groups disagreed, proportionately more of lower than upper group disagreed with instructor.	.001	The most popular response for both groups was that class attendance was helpful but fewer than expected of upper group and more than expected of lower group felt it was unnecessary.

Desired strictness of definition of responsibility	.01	Both groups were satisfied with definition of responsibilities but proportionately more of lower group than of upper group preferred a stricter definition than was given them.
Comprehension level of lectures and discussions	.05	Both groups felt lectures and discussions were at their level of comprehension, but proportionately more of lower group felt it was either above or below.
Availability of faculty for clarification of readings	.001	Proportionately more of lower group felt faculty members were not available although both groups agreed they generally were available.

Table 2

Differences in Responses of High and Low Progress Groups:
Questions Not Used to Determine Agreement

Subject	p	Distribution of responses on student questionnaires
Level of work	.01	Both groups did an average level of work, but proportionately fewer of lower group did their best work and proportionately more did not work to capacity.
Independent reading	.05	Both did what was required, but more of the lower group did less than required.
Preference of help from faculty or students	.001	Upper group preferred to get help from faculty and lower group from fellow students.
Accomplishment compared to independent study	.001	Upper group felt they would not have accomplished as much if they had taken the course by independent study. Proportionately more of lower group felt they would have accomplished as much.
Number of formal conferences with instructor	.001	Proportionately fewer of lower group had 1-2 conferences while most of both groups had none.
Interest in course content	.02	Upper group had high interest; lower group had moderate interest.
Choice of section	.05	Proportionately more of upper group got their first or second choice of sections.

Table 3

Differences in Teaching Practices Between the High and Low Gain Groups

Observation	Probability of distribution of responses at least as extreme as those observed.[a]
Upper group teachers gave objective tests, lower group teachers more inclined toward combination objective-essay tests.	.01
Upper group teachers more likely to discuss tests in class the next class period.	.05
Upper group teachers were more inclined to present the students' responsibilities as specific expectations while lower group teachers were more inclined to represent them as generalized expectations.	.05
Lower group teachers had more than 3 years college teaching experience while upper group teachers had less.	.05
Upper group teachers more inclined to have lectures by other faculty members.	.12
Upper group teachers less inclined to use a panel discussion by students than lower group teachers.	.15
Upper group teachers less likely to use a common class project than lower group teachers.	.15
Lower group teachers less inclined to read quotations and ask for student criticisms of them.	.07
Upper group teachers more inclined to ask students to write a report of their reaction to their class experiences.	.03
Lower group teachers more likely to divide class into subgroups for discussion.	.05

[a]Fisher's Exact Probability Method--all comparisons were reduced to a 2 x 2 table.

Table 4

Characteristics of High and Low Gain Groups

Characteristic	High gain sections	Low gain sections	Probability of distribution at least as extreme as that observed[a]
Teacher's Age	N = 7 $\bar{X}$ = 31 Mdn = 29	N = 5 $\bar{X}$ = 40 Mdn = 39	.12[b]
Teacher's Sex	N = 7 5 female 2 male	N = 5 4 male 1 female	.12
Teacher's Education	N = 7 5 Master's 2 Doctor's	N = 5 3 Doctor's 2 Master's	.28
Class Meeting Times	N = 9 7 - 3 meetings per week 2 - 2 meetings per week	N = 7 6 - 2 meetings per week 1 - 3 meetings per week	.02[c]

[a]Fisher's Exact Probability Method--all comparisons were reduced to a 2 x 2 table.

[b]Median Test

[c]Includes 3 sections for each of two instructors.

DISCUSSION

There are a number of factors which make these results difficult to interpret. The use of objective tests as the basis for the progress index might well be responsible for the high gain sections being those which prefer these types of tests since generally preference for and performance on tests are closely related. However, it was found that the Florida Twelfth Grade Test scores of both the high and low gain groups were the same. It is possible that the additional practice of taking objective tests made the difference. In conjunction with this finding, the tendency for the teachers of

the high gain group to be less experienced suggests that perhaps they may be more inclined to follow the text closely. Another possibility is that the objective tests are not measuring the particular intellectual processes presumably cultivated by the more experienced teachers. This does not preclude the fact that different results might have been obtained had the tests been different, but it seems clear that further investigation based on more comprehensive measures of learning is needed.

The finding that students in the three one-hour class meetings showed greater gain than those in two meetings of equivalent total hours is consistent with the outcomes of some experiments in learning--generally that frequent rest periods and short practice sessions result in greater learning of certain types of subject matter. Future studies in this area should make provision for the distribution of class meetings as well as sex, age, and possibly types and extent of non-teaching activities of instructors. There is also a need for a more comprehensive measure of student-teacher agreement.

SUMMARY

By means of an objective measure of student progress, the relative effectiveness of teachers in a behavioral science course has been examined. It was found that there was a greater extent of student-teacher agreement on various aspects of instruction in the high gain classes. The students of the high gain group may be described as follows: (1) satisfied with their accomplishment in the course, (2) prefer objective tests, (3) feel class attendance is necessary to a greater extent than the low gain group, (4) prefer to get help needed from faculty rather than students, and (5) feel they would not have accomplished as much if they had taken the course by independent study. The seven teachers of the high gain group may be characterized as young females with less than three years college teaching experience who give objective tests, discuss tests in class the period after they are given and present the students' responsibilities as specific rather than generalized expectations.

After a period of one trimester, there was no difference in retention between the two groups.

It should be noted that these data are derived from an exploratory study and the generalizations made must be interpreted as merely a basis for further investigation.