Research, Teaching, and Faculty Fate

John R. Hayes

The relation between the research activities of a faculty member and the quality of his teaching has been subject to considerable discussion, both in print (*1*) and over academic lunch tables. Empirical studies, however, have been scarce. Voeks (*2*), in a study conducted at the University of Washington, found that the ratings of teaching that faculty members received from their students bore no significant relation to such measures of their research activity as publication rate and membership in the university's Research Society. In a study done at Tufts University, however, Bresler (*3*) found that faculty members who held research grants received better teaching ratings from students than did faculty members who did not hold research grants. These two studies are the best available empirical studies on this topic, and they suggest very different relations between research activity and teaching. This study was undertaken to provide additional data bearing directly on the results of Voeks and Bresler, in an attempt to resolve their apparent differences, and to provide data which will place the teaching-research relation in the context of decisions about teaching assignments and promotion.

Design of the Study

Data were collected from 17 of the academic departments at Carnegie-Mellon University. Eleven of the departments (biology, chemistry, chemical engineering, civil engineering, computer science, electrical engineering, management science, mathematics, mechanical engineering, metallurgy, and statistics) are included in the College of Engineering and Science; five departments (economics, English, history, modern languages, and psychology) in the College of Humanities and Social Sciences; and one department (business and resource management) in Margaret Morrison Carnegie College, the women's college at C-MU. No data are included from fine arts departments. It was felt that several of the measures of performance in research and teaching, such as publication rate, did not have the same meaning when applied to departments in the fine arts as when applied to departments in the sciences and humanities.

Measures of performance in research and in teaching were obtained for 355 individuals in the 17 departments during a four-semester target period from fall 1967 to spring 1969. The following measures were sought (but not necessarily obtained) for each individual: (i) academic rank; (ii) teaching assignments (the courses taught and the hours per week devoted to each); (iii) publication list for the last 5 years (obtained from faculty biographies published by the university); (iv) grant status (a statement from the Proposal Information Office that the individual did or did not have an outside grant during the target period); and (v) student evaluation (the average of all of the teaching ratings received by the individual during the target period). Student evaluations at C-MU are conducted by a student organization and vary somewhat from semester to semester in form and coverage. The question that seemed best to reflect the ability of the teacher to motivate his students was chosen from each evaluation. In all cases, teachers were rated on a scale from 1, for best, to 5. In fall 1967 and spring 1968, the question "Was the class stimulating?" was used with alternatives ranging from "stimulating" to "deadly dull." In fall 1968, the quality rated was "teacher's presentation of the course material," with alternatives ranging from "stimulating and exciting" to "puts you to sleep." No student evaluations are available for spring 1969.

The chairmen of all cooperating departments were asked to make the following judgments about each department member: (i) research ability (the man's ability to do research or scholarship, irrespective of his productivity); (ii) research time (the amount of effort that the man devotes to research, irrespective of his ability or productivity); and (iii) teaching ability (the ability of the man as a teacher of undergraduates). In all cases, the judgments were to be made on a 5-point scale, from 1, for greatest, to 5; and they were to be made relative to the department rather than to the profession. Thus, the worst researcher in a strong department was to be rated 5, even though he might earn a 2 within his profession.

The department heads' judgments are of special importance here. This is not because I believe that these judgments are particularly accurate measures of research and teaching (which they may or may not be) but because I believe that they are instrumental in decisions about promotion and teaching assignment. Thus, a man's true abilities in research and teaching may be less influential in a promotion decision than what the department head believes about his abilities.

The data on teaching assignments were used to generate a level index for each teacher. All courses were classified on a 4-point scale: (i) elementary courses required of students in other departments (the so-called "service" courses); (ii) elementary courses primarily for students within the department; (iii) upper-level undergraduate courses; and (iv) graduate courses. The level index for a man was the average level of all the courses he taught during the 2-year target period.

The publications data was processed in two ways before it was used in the study. First, points were assigned for each publication according to the scale used by Voeks (*2*) (that is, 1 point for a small article, 5 points for being sole author of a book), to yield a point score for each faculty member.

Second, since the point score was based on publications appearing during the last 5 years, the score was adjusted for faculty members for whom fewer than five productive years had elapsed since they had received the Ph.D. The first year after receiving the Ph.D. was

The author is acting dean of the College of Humanities and Social Sciences and associate professor of psychology at Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

Table 1. Relationship between ratings of research ability and proportion of faculty members holding research grants. (Research ability is measured on a 5-point scale, with 1 being high; 5, low.)

Research ability	Science		Humanities	
	N	*P*	*N*	*P*
1	27	.593	21	.429
2	40	.500	31	.387
3	34	.432	18	.059
4	15	.267	10	.200
5	4	.000	10	.200

assumed nonproductive because of publication delays. Thus, a man who got his Ph.D. 3 years ago was assumed to have two productive years behind him. Faculty members with fewer than five productive years were given the point score that they would achieve if they continued to produce at the same yearly rate for 5 years.

Different numbers of faculty members are involved in the various aspects of the study. Comparisons involving student evaluation may include as few as 177 measures, whereas the correlation between research ability and research time is based on a sample of 334.

Results and Discussion

First, the relations among the measures of teaching quality and among the measures of research activity are described. Then these two classes of measurements are related to each other and to measures of promotion and teaching assignment.

Department heads' judgments of research ability (RA) and research time (RT) are highly correlated (product-moment correlation = 0.823; $N = 334$). RA and RT were studied separately in relation to the other variables. However, it was found that the relation of RA to any of the other variables closely resembles the relation of RT to that variable. For example, Fig. 1 shows the relation of both of these variables to publication index. The similarity of the relationships is obvious. All relations that are statistically significant for RA are also significant at the same level for RT, and vice versa. Hereafter, all discussion of RT is omitted as redundant. Figure 1 shows that faculty members judged by their department heads to be good researchers are also very productive publishers. This relation is significant at the 0.01 level by analysis of variance ($F = 10.767$; d.f. = 4, 279). Table 1 shows that, in both the science college and the humanities college, faculty members judged high in research ability are more likely to hold grants than those judged low.

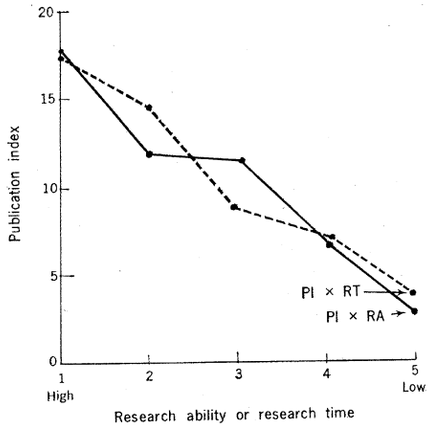

Fig. 1. Publication index related to research ability and research time.

Our two measures of teaching ability also are positively related. As Table 2 shows, faculty members whom department heads judge to be good teachers tend to rate well with students also. The product-moment correlation of these measures is + 0.62 ($N = 179$). Despite the general agreement between students and department heads about which faculty members are good teachers and which are not, our data reveal some consistent differences in the judgments of these two groups. Figure 2 shows that, in the department head's opinion, good teaching tends to be associated with high research ability. The relation is significant at the 0.01 level by analysis of variance ($F = 7.14$; d.f. = 4, 313). Students don't agree, however. The relation between RA and student evaluation of teaching is very weak and does not approach statistical significance ($F = 0.587$; d.f. = 4, 178).

These two results of student and department head opinions were obtained for populations of instructors that, while they overlapped, were not identical. In particular, many instructors who had been rated by their department heads for teaching quality received no student evaluations. The same results, however, were obtained when the study was restricted to individuals for whom both student evaluations and a department head's rating of teaching quality were available. The observed difference,

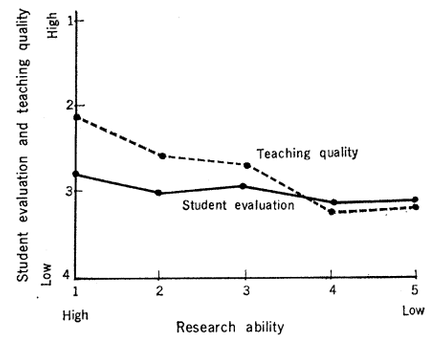

Fig. 2. Student evaluation and teaching quality related to research ability.

therefore, did not result from students and department heads' looking at different populations of teachers.

No significant relation was found between publication index and either teaching quality or student evaluations of teaching. The present results, then, confirm Voeks' (*2*) earlier finding of no relation between publication index and either teaching quality or student evaluations. This result contrasts with Bresler's earlier finding (*3*). Table 3 shows student evaluations. No significant relation was found between grant status and the C-MU data, together with the most comparable data from the Bresler study (*4*). The only explanation I can offer for the obvious differences in the data is a real difference between the two institutions.

Of the six relations between three measures of research activity (depart-

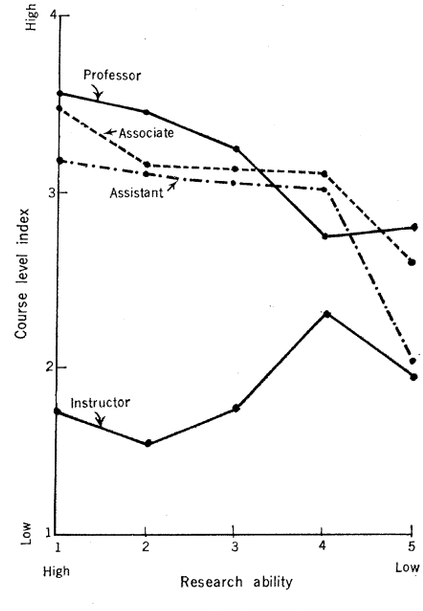

Fig. 3. Course level index related to rank and research ability.

ment head judgments, publication index, and grant status) and two measures of teaching ability (department head judgments and student evaluation) that have been examined, only one relation, that between department head judgment of research ability and department head judgment of teaching ability, was statistically significant. Publication rate failed to correlate with either student evaluations of teaching or with department heads' judgment of teaching quality. Likewise, status in regard to grants failed to correlate with either student evaluations of teaching or department heads' judgment of teaching quality.

These results can be interpreted in either of two ways. On one hand, research activity and the ability to teach may be considered, in fact, positively related. As an auxiliary hypothesis, to account for the fact that only one of the six relations was statistically significant, one might propose that department heads' judgments are "better" measures of research activity than are publication rate or grant status (the other measures used). On the other hand, one might believe that there is no correlation between research activity and the ability to teach. The statistically significant relation between the two judgments by the department heads may be accounted for as a "halo" effect (*5*), the familiar illusion in which judges rate an individual who is "good" in one trait as "good" in many.

Institutional Setting of Teaching and Research

The effect of the research-teaching relation on the student depends upon the way the university treats researchers and teachers. If good researchers are assigned only graduate classes, it really does not matter to undergraduates whether researchers are good teachers or not. If the university gives promotions only for research productivity, the student may face a faculty that was not selected for teaching ability and perhaps not particularly motivated to teach well.

Table 4 shows how the total number of hours that the faculty spent teaching each week (averaged over the four-semester target period) was distributed by the rank of the teacher and by course level. Some of the rank categories shown in the left-hand column of Table 4 require discussion. In the category of instructors and teaching assistants (TA's), almost all of the instructors teach in the humanities college and are working toward a graduate degree at some other university. The TA's, on the other hand, are mostly in the science college and are working toward graduate degrees at C-MU. The irregulars are neither professors, associates, assistants, instructors, nor TA's. They have titles such as visiting professor, adjunct professor, lecturer, and senior research engineer. Table 4 makes it clear that the higher an individual's rank, the less time he spends teaching undergraduates. Instructors and teaching assistants, low men in the academic hierarchy, do about one-third of all undergraduate teaching and teach more than half of the service courses. When the more prestigious faculty members do teach undergraduates, they tend to teach upperclassmen who are majors in the professor's area.

Table 2. Student evaluations related to department heads' judgments of teaching quality. (Teaching quality is measured on a 5-point scale, with 1 being high; 5, low.)

Teaching quality	*N*	Mean student evaluation
1	45	2.32
2	43	2.67
3	38	3.26
4	30	3.29
5	21	3.82

Figure 3 shows that, within the regular faculty ranks (professor, associate, assistant), individuals with high research ability are assigned to high-level classes. Instructors, on the other hand, teach only low-level courses, no matter what their research ability. Our statistical analysis of this phenomenon is restricted to the regular faculty because we have no research ability ratings for teaching assistants. In addition, instructors (associated primarily with the humanities) and irregulars (associated primarily with the sciences) teach in colleges which differ in the proportion of graduate teaching done.

The influence of rank and of research ability on course level index was tested by two-way analysis of variance with unequal N (*6*). The influence of research ability was significant ($P < .01$; $F = 13.73$; d.f. $= 4, 206$), but the influence of rank was not ($F = 0.75$; d.f. $= 2, 206$). There is a significant interaction between rank and research ability ($F = 2.29$; d.f. $= 8, 206$). Examination of Fig. 4 suggests that the interaction represents a tendency to assign assistant professors who are judged poor in research ability to low-level teaching. Taking into account the results shown in Table 4 and those shown in Fig. 3, it is clear that those of high rank and those of high research ability do considerably less undergraduate teaching than those of low rank and low research ability. There appears to be no relation between classroom assignment and department heads' judgments of teaching quality.

Table 3. Measures of teaching ability related to grant status.

Status	C-MU data		Bresler's data
	Mean teaching quality	Mean student evaluation (5-point scale)	Mean student evaluation (4-point scale)
Grant	2.50 ($N = 93$)	3.01 ($N = 48$)	1.92 ($N = 19$)
No grant	2.67 ($N = 144$)	3.04 ($N = 100$)	2.48 ($N = 42$)

Table 4. Total teaching hours per week classified by the rank of the teacher and the level of the course.

Rank	Time (hours) at course level:				Total
	1	2	3	4	
Professors	93.50	222.00	489.49	952.80	1757.79
Associates	99.00	117.64	394.17	470.25	1081.06
Assistants	313.30	224.32	1086.69	991.50	2615.81
Instructors and TA's	821.00	441.00	684.50	54.50	2001.00
Irregulars	192.00	82.04	425.65	798.55	1498.24
Total	1518.80	1087.00	3080.50	3267.60	8953.90

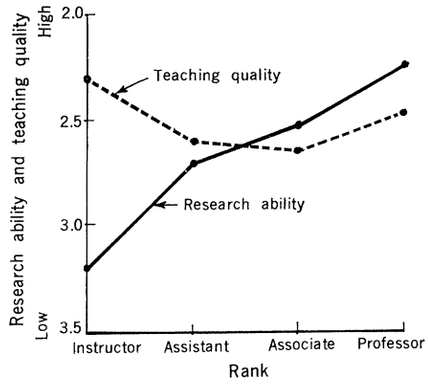

Fig. 4. Rank related to research ability and teaching quality.

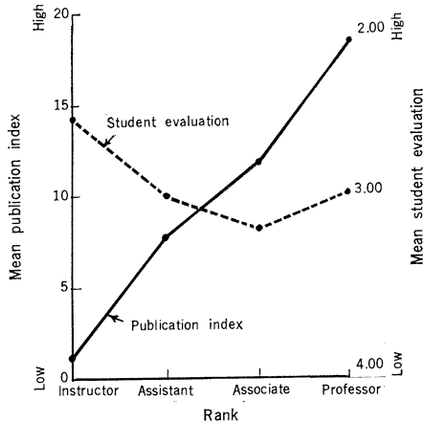

Fig. 5. Rank related to publication index and student evaluation.

If the university selectively promoted individuals on the basis of some trait (such as research activity or teaching ability), then we would expect that individuals of high rank would score better on the trait than those of low rank. We have used this "filtering principle" as an index of the extent to which research and teaching ability influence promotion decisions at C-MU.

Figure 4 shows that department heads rate the higher-ranking faculty members as better researchers than those in lower ranks. This does not seem to be true of department head judgments of teaching quality. Analysis of variance on the three highest ranks showed that the relation of rank and research ability was significant at the 0.05 level, but that no significant relation existed between rank and teaching quality. We repeated the analysis above, substituting publication index and student evaluation for the department heads' judgments of research ability and teaching quality. The results, shown in Fig. 5, are parallel to those in Fig. 4. Analysis of variance on the three highest ranks indicates an association between rank and publication index that is significant at the 0.01 level ($P = 3.56$; d.f. $= 2, 236$). No significant relation between rank and student evaluation was found.

The evidence seems clear that research activity influences promotion decisions at C-MU. There is no clear evidence that the quality of teaching influences promotion.

Conclusions

I have tried to answer three questions here: (i) Are research activity and teaching ability related to each other? (ii) In what way do research activity and teaching ability influence classroom assignment? (iii) In what way do research activity and teaching ability influence promotion?

The answer to the first question is not wholly clear. If one takes department heads' judgments at face value, there is evidence of a strong positive relation between research ability and teaching quality. If, on the other hand, one interprets the correlation in the department heads' judgments as a "halo" effect, then there is no evidence in the six relations measured that research activity and teaching ability are related.

Answers to the second and third questions appear quite clear: individuals with high research ability and high rank tend to be assigned to high-level classes. Teaching quality is unrelated to classroom assignment. Promotion is strongly related to measures of research activity but appears to be unrelated to teaching ability.

The assertion sometimes is made that a policy of hiring and promoting primarily for research ability is beneficial to undergraduate education because good researchers are good teachers. Before we accept such an assertion we should ascertain: (i) whether good researchers are, in fact, good teachers—a proposition not clearly supported by the data of this study; (ii) whether the classroom assignment policy brings good researchers into contact with undergraduates with the same probability as bad researchers (this clearly is not the case at C-MU, where good researchers do considerably less undergraduate teaching than bad researchers)—if research and teaching *are* positively correlated, such a policy differentially assigns *bad* teachers to undergraduates; and (iii) whether a policy that promotes individuals on the basis of research and not quality teaching discourages interest in teaching.

References and Notes

1. See, for exam[ple] E. J. McGrath [*J. Higher Educ.* **33**, 148 (1962)]; J. R. Killian (*Atl. Mon.*, December 1965, p. 53); and E. R. Hutchison (*Amer. Scholar*, summer 1966, p. 430).
2. V. Voeks, *J. Higher Educ.* **33**, 212 (1962).
3. J. B. Bresler, *Science* **160**, 164 (1968).
4. In Table 3, we show Bresler's data for the faculties of science, engineering, and social science. We have not included his data for the faculties of arts and humanities because these data were not reported in a form comparable to that of the present study. The reader should note that student evaluations are measured on a 4-point scale in Bresler's study rather than on a 5-point scale, as in the present study. The actual differences between the two studies is then even larger than is at first apparent.
5. For a discussion of the halo effect see J. P. Guilford, *Psychometric Methods* (McGraw-Hill, New York, 1954), p. 279.
6. We used the procedure described in A. L. Edwards, *Experimental Design in Psychological Research* (Holt, Rinehart & Winston, New York, 1968), p. 264.

7. I thank H. Robbins for tabulating the data and for typing and editing the final copy for publication, and G. Forehand and J. Korn for criticisms of the manuscript.