Abstract

The relationship between teaching and research is extremely important to colleges and universities especially with changing expectations of faculty performance. This piece summarizes the results of *empirical* studies that have tried to measure this association; most did not find a significant relationship. However, measures of teaching effectiveness are generally inadequate ignoring those aspects of teaching that are most likely to benefit from research activity.

John King Gamble, Jr. *

The Relationship Between Teaching and Research: Clear Answers to the Wrong Questions?[1]

I. Introduction

Few issues in higher education are more basic and important than the relationship between the two principal duties of faculty, teaching and research. The higher education climate of the 1980s and 1990s, characterized in most areas by a surplus of well-qualified faculty, will increase the importance of understanding this teaching/research interface. The last decade has seen a precipitous shift with research and scholarship given a greatly increased weight in decisions to hire, promote, grant tenure to and terminate faculty. The time is rapidly approaching when most colleges and universities in the U.S. and Canada will require a substantial amount of research and scholarship from faculty for the granting of tenure or promotion. This fundamental change in the "rules of the game" will have far reaching implications across higher education.

Thus there are two significant reasons for this inquiry. The first is very basic and practical — if we are to require research of many more faculty, most of whom traditionally focused their energies on teaching, we need to know how these different roles relate to each other. The second reason is loftier. If we are really to understand higher education and the contribution it can make, we need rigorous, well-thought-out models of the teaching/research interface. The chance of understanding this issue is enhanced because a considerable body of data already has been collected. As Table 1 shows, many empirical studies of the relationship between teaching and research have been completed. Measures of teaching are readily available because most colleges and universities routinely use a course evaluation questionnaire (CEQ). Standardized measures of research are not quite so prevalent, but they can be developed.

---

* Professor, Division of Business and the Social Sciences, The Behrend College, The Pennsylvania State University.

The Editor wishes to thank Professor Gamble for agreeing to have his fully-refereed article which would normally have appeared in another issue, included in this special issue. His subject was particularly appropriate for this theme.

TABLE I

EMPIRICAL STUDIES RELATING TEACHING AND RESEARCH

INVESTIGATOR	YEAR	SAMPLE	TYPE OF INSTITUTION	HOW RESEARCH MEASURED	HOW TEACHING MEASURED	RELATIONSHIP
McGRATH ‹2›	1962	75 "good teachers"	lib. arts colleges	pub. ? Y or N	admin. judge	"good teachers" publish
VOEKS ‹3›	1962	300	large univ.	res. society + wgt. pubs.	C.E.Q.	no relation
BRESLER ‹4›	1968	300	med. univ.	research support	C.E.Q. Q-tiles	moderate positive
McDANIEL ‹5›	1970	76	large univ.	hours + wgt. pub.	C.E.Q.	small pos & neg.
STALLINGS ‹6›	1970	250	2 large univ.	wgt. pub.	C.E.Q.	v. small positive
HARRY ‹7›	1972	230	large univ.	hrs. per week	C.E.Q.	none
ALEAMONI ‹8›	1973	500 "good teachers"	large univ.	wgt. pub.	2 CEQ	none
HICKS ‹9›	1974	450	large univ.	pub. ? Y or N	C.E.Q.	small positive
LINSKY ‹10›	1975	1500	16 col. + univ.	wgt. pub. + cit. ind.	critiques via CEQ	none
DENT ‹11›	1976	90	large univ.	citation	C.E.Q.	none
HOYT ‹12›	1976	180	large univ.	admin. + time + pubs.	C.E.Q.	none
MICHALAK ‹13›	1981	86	lib. arts college	sch'shp. scale	CEQ + admin.	moderate positive
TEAGUE ‹14›	1981	20 "good teachers"	large univ.	mean no. books, etc.	teaching awards	positive (loose)

I shall begin by examining the empirical research dealing with this subject over the last 20 years. Next, the intricacies of measuring teaching and research will be explored more critically. Suggestions about developing better studies are offered. Finally, some recommendations, based on the realities of the 1980s, are made.

II. Existing Empirical Research

Table 1 summarizes *empirical* studies completed in the U.S. during the last 20 years. A large number of worthwhile works have been omitted from the table because (1) they were anecdotal, not empirical; (2) they did not measure *both* teaching and research; and/or (3) they did not

explicitly measure the relationship. For example, an excellent study by Rodin and Rodin established the existence of two dimensions of teaching but did not attempt to measure research and, hence, is excluded from the table.[15] A recent article by Kasten did a thorough job of measuring teaching and research in order to relate them to rewards; but no correlation between teaching and research was calculated.[16] The table contains the most important, basic information: name of the first author, publication date, sample of faculty used, type of institution from which the sample was drawn, how teaching and research were measured, and overall results.

Time and space constraints do not permit a detailed critique of each study, but a few general observations can be offered. These are well-conceived, professionally-executed studies. While one might argue with certain details of each, none is fatally flawed. There is wide variation in sample sizes; a few of the studies are basically different from the others in that they selected "unusually good" teachers and examined the research records of them.[17] As such, they do not really compare, in the strictest sense, the teaching of research attributes. The most significant shortcoming of these studies is the over-reliance on student course evaluation questionnaires (CEQ). Eleven of 13 studies rely primarily or exclusively on these instruments; often they are accepted uncritically as *the* valid index of teaching quality. My point (on which I shall expand later) is not that CEQs are worthless but that they probably do not present the entire picture.

There was more variety in the techniques used to measure research performance. The most common method (and probably the best) was a system where the publications were counted and a weighting used, e.g., books = 15; books edited = 9; articles = 3; book reviews = 2; bulletins and technical reports = 3.[18] Another often-used indicator is citation indices.[19] It seems that a very persuasive case can be made that research output (performance) is best measured by a combination of a weighted publication count and a citation index.

Two other methods are less satisfactory. Two investigators[20] used a simple, dichotomous measure of research, e.g., has Prof. X published anything during the last five years? This is probably too insensitive to gradations in research performance. Even less satisfactory are those studies that measure research by asking faculty members how many hours per week they spend on scholarship and research.[21] The issue is not whether faculty will report the hours accurately (usually they will). But the goal is to measure and understand research *performance*, without getting mired in the concept of effort.

It is important to ask whether these studies are representative of higher education in North America. The sample tends to be widely drawn among faculty usually not excluding any groups. Most were conducted in large universities, but small liberal arts colleges are included in 3 of 13 studies. The most important point to make about the institutions selected is that they tend to be of high quality; some are elite. It is an important and difficult question whether these findings are generalizable across the range of 3,200 colleges and universities in the U.S. and Canada.

The most important and consistent conclusion emerging from these studies is that little or no relation exists between teaching and research. Nine of 13 studies found no relation or at most a very small positive relationship. None found what could be called a large positive association. This can lead to only two possible conclusions. First, the studies are essentially accurate; the two are unrelated. Second, teaching and research are related, but the way we go about measuring them disguises or underestimates the relationship. In order to address this issue, I shall examine the measurement of teaching and research.

III. Measuring the Basic Elements

With very few exceptions, writers acknowledge that teaching is more difficult than research to measure accurately and completely. Most discussion and disagreement about measuring research reduce to matters of fine-tuning and emphasis. For example, scholars may disagree about the appropriate weighting of a book versus an article or a book review. But most admit that some weighting system is needed to make entries "more nearly equivalent."[22] Many advocate employing a citation index as a measure of the impact the research is having on the scholarly community. Michalak and Friedrich put the matter succinctly when they wrote that "the ultimate test of a scholar's research is the degree to which it stimulates other scholars to think and write about it."[23]

Thus it seems that a reasonable index of research performance can be constructed relatively easily. It would need some flexibility to accommodate different disciplines because many have no citation index. Such an index of research performance would use publications with a weighting system as the principal measure supplemented by citation indices as appropriate and possible. In order to fit the context of a particular college or university, it might be desirable to integrate certain pre-publication activities into the system, e.g., conference presentations, draft manuscripts. Higher education seems to be moving towards a standard where publications are the clearest, perhaps the only clear, evidence of research.[24]

Measuring teaching effectiveness is much more complicated. There is a fairly common, yet no less disquieting, tendency for authors to acknowledge the limitation of CEQs yet to rely exclusively on them. Why has there been such reliance on CEQs? Several reasons stand out. First, they are readily available at most colleges and universities. Second, our consumerism mentality encourages their use. Bresler wrote that "a majority of the earlier reports suggest that the student, as a consumer, is in the best position to evaluate teacher effectiveness."[25] Finally, CEQs have been revised and improved so that we are reasonably certain they do not contain some of the most fundamental errors we might fear. For example, the evidence is quite conclusive that a faculty member cannot "buy" higher CEQ ratings by giving higher grades.[26]

What then are the shortcomings of concentrating only on the CEQs? Voeks' pioneering study identified three methods for evaluating teaching, achievement tests, colleague/administrator judgments, and CEQs.[27] She found the first two difficult or wanting and relied only on CEQs. This is unfortunate because she perceived the difficulty: "students can, for example, grossly misjudge the accuracy of material or even the usefulness of a concept or technique."[28] The same point was made forcefully by Rodin and Rodin who found an objective criterion of teacher effectiveness ("what students have learned") and a subjective element ("student evaluations").[29]

Herein lies the crux of the problem in relating teaching and research. We can achieve agreement about how to measure research, not easily, but manageably. However, the techniques commonly used for measuring teaching may be missing a sizable portion of the picture. CEQs can measure certain things that are vital to good teaching, e.g., how a faculty member relates with students, the clarity of her/his presentation (unless content is compromised to make the presentation neater and simpler), and how interest in the subject is stimulated. But CEQs cannot assure that the material taught is up-to-date, thorough and accurate. The consumer analogy can be carried only so far, because the students (the consumers in this case) often have very incomplete grounds for judgment and comparison. Especially when students take courses outside their major interest, they have no *substantive* grounds for evaluating the instructor. The content of the course may be deadwrong without the students getting an inkling!

IV. Recommendations

If our understanding of the relationship between teaching and research is inadequate, then obviously one solution is to undertake new and better studies. Funds should be made available for this purpose. The principal difficulty would be finding methods to measure other aspects of teaching, i.e., change in student achievement, not just student perception. This would be a formidable, but not an impossible task. The motivation for such a study should be high. If the results showed that research relates moderately positively with the "objective" portion of teaching (which seems probable), pressures would build for even more emphasis on research. If such a study could be undertaken, its validity and feasibility might vary from discipline to discipline. For example, it is one thing to measure learning in a calculus class, but quite another in political science. A point that can hardly be overemphasized is that all these conclusions must be adapted and molded to accommodate vastly different institutional contexts and disciplines. All these problems aside, the profession could surely do with several, more ambitious studies combining a more sophisticated measure of teaching with existing indices of research performance.

There may be a tendency to interpret the results of these various studies, i.e., the near-zero association between teaching and research, concluding that an increased emphasis on research is misdirected. It takes us away from our teaching and contributes nothing. This argument seeks a return to the "good old days" typified aptly by a 1958 study showing that only 39% of faculty surveyed indicated that it took *any* publishing to be promoted to *full professor.*[33]

My opinion is that there are five reasons that, in combination, make a convincing case for not returning to the era when research was de-emphasized completely.

(A) *Oversupply of good Ph.D.'s in most areas.*

From one vantage it may seem opportunistic, but the oversupply of well-educated Ph.D.'s in most fields means that, when hiring new faculty, institutions can easily attract people who are good teachers *and* good researchers. It is true that this surplus of faculty occurs only in certain areas while in others, e.g., accounting, engineering, computer science, there is an acute shortage of faculty. But if research is required of the sociologist, it also will be required of the engineer. In the context of shared governance and collective bargaining, increased research expectations have meant a major readjustment, a readjustment that will not be tolerated unless it is applied uniformly.

(B) *Research may enhance those aspects of teaching not measured by CEQ's.*

There is strong *prima facie* reason to believe that research performance does contribute to that portion (50%?) of teaching that is not measured by CEQs. Schmitt wrote "only the scholar can introduce continuing variety into his teaching. Only the scholar can escape the monotonous rut by projecting and teaching new courses and discarding what has become wearisome and insufferably dull."[34] Showalter expressed the opinion that those who do not publish lack the "courage" to test themselves before a wider, more critical audience.[35] In a similar vein, McGrath wrote that research "sharpens the critical faculties, prevents professional stagnation, and imposes an intellectual discipline lacking among those who restrict their activities to teaching."[36]

(C) *Research may help the teacher to reach fuller potential.*

Empirical studies might take a different tack. Virtually all such studies have asked how teaching and research relate to each other. In operational terms, this boils down to asking the macro question: Do high levels of teaching and research performance tend to exist in the same people? Perhaps a more important question is how teaching and research activities interrelate *within* the

individual and affect performance. We should at least examine the proposition that teaching and research are mutually reinforcing within the individual enabling faculty to reach their fullest potential. Voeks acknowledged the importance of the question.[37]

(D) *Research may reduce the chance of faculty burnout.*

Recently there has been increased concern about burnout in all professions including college teaching. It is already a serious problem and will be exacerbated as the number of tenured faculty with decades of service remaining increases. There is some evidence that CEQ scores increase for a while, but tend, if anything, to decline over the long run.[38] This may be evidence of burnout. Schmitt, in one of the most thoughtful pieces on the topic, wrote: "Teaching wears you out. You get tired of it. Research, in my experience, is less attritional."[39] Of course, Schmitt's experience may be atypical. But it is reasonable to assume that if a person has a wider variety of activities, burnout and stagnation are less likely to occur. Research provides variety from the routine of teaching. Hopefully, research will suggest new approaches to teaching and new courses that the faculty member may wish to teach. If faculty members are active in research, greater opportunities exist for funded research that can "buy" a reduction in teaching loads. External funds are very difficult to obtain, especially in some disciplines, but the potential should exist for teaching load reductions bought with internal ("hard") money if a strong, persistent case is made. It follows from this argument that the greatest need to do research in order to combat burnout exists at those institutions with heaviest teaching loads, precisely those schools that traditionally have cared less about research!

(E) *Research expands sensitivity to the discipline.*

A similar, but not identical, point relates to the need for sensitivity to one's discipline that can be achieved better through research. Good teaching requires a grasp of many subtleties and nuances of a field, a sensitivity for the strengths, weaknesses, limits and future direction of knowledge in that field. Especially in some disciplines it is probable that these sensitivities can be most easily and thoroughly gained by research activity. The best waiters have done at least some cooking, or at the very least, spent some time in the kitchen!

V. Conclusions

The interrelationship between teaching and research remains incompletely understood. Many empirical studies exist. These shed some light on the topic, but none does an adequate job of handling "objective" *and* "subjective" aspects of teaching. Further, none has delved systematically into the effect of the co-existence of teaching and research in the same people. One conclusion from these empirical studies is that research activity does not cause bad (or worse) teaching.

Given the realities, i.e., continued and often increased emphasis on research, colleges and universities must do a better job of explaining the rationale for research. Some of this explanation may be less than satisfying, e.g., the oversupply of Ph.D's. In other cases, we have to overcome basic misunderstandings about research. Many faculty (and, as often, their spouses) argue against research emphases on the grounds that the research product will probably not make a significant contribution. We must argue that the process of research is as important as the product, that the benefit is usually greater for the researcher than for the academic community as a whole. Of course, an ambiguity will exist because institutions link rewards to the tangible publication output rather than to the usually-more-important research process.

Empirical studies must be supplemented by careful thought about the relationship between teaching and research. Few of the empirical studies devote adequate time to the *logical* bases of

their investigations. Of course, in gross terms, teaching and research can relate positively, negatively, or not at all--but *why* might we expect to find a certain kind of relationship? It will remain important to look for reasonable explanations of possible relationships. Interestingly, the least amount of logical conjecture has dealt with the zero relationship, the one most of the empirical research work has uncovered! Many are satisfied with the almost tautological, research and teaching are separate capacities that happen to coexist in the same person! I have already discussed many possible explanations for a positive relationship.[40] Most of those who posit a negative relationship do so based on two arguments: (1) competing time demands[41] or (2) different personality traits.[42]

These issues will not be resolved quickly — neither will they go away. And they are too important to ignore. We must seek better empirical studies, but not lose sight of the root and reasons for our inquiries. In Berlin, more than a century ago, Humboldt described "an organic link between the creation of knowledge and its transference."[43] That is still the question today, the exact nature of the link along with the level on which the interconnection occurs. We will find different answers for certain disciplines than for others and for different kinds of institutions, but the answers must be sought.

VI. References

1. I should like to thank the editor of *The Journal of Educational Thought* for many helpful comments about an earlier draft of this work.
2. E.J. McGrath, "Characteristics of Outstanding College Teachers," *Journal of Higher Education*, Vol. 33, No. 3 (March 1962), pp. 148-52.
3. V.W. Voeks, "Publications and Teaching Effectiveness," *Journal of Higher Education*, Vol. 33 (April 1962), pp. 212-18.
4. J.B. Bresler, "Teaching Effectiveness and Government Awards," *Science*, Vol. 160, No. 3824 (April 1968), pp. 164-67.
5. E.D. McDaniel and J. F. Feldhusen, "Relationships Between Faculty Ratings and Indexes of Service and Scholarship," *Proceedings*, 78th Annual Convention, American Psychological Association (1970), pp. 619-620.
6. W.M. Stallings and S. Singhal, "Some Observations on the Relationship Between Research Productivity and Student Evaluation of Courses and Teaching." *American Sociologist*, Vol. 5 (May 1970), pp. 141-43.
7. J. Harry and N. S. Goldner, "The 'Null' Relationship Between Teaching and Research," *Sociology of Education*, Vol. 45 (Winter 1972), pp. 47-60.
8. L.M. Aleamoni and M. Yimer, "An Investigation of the Relationship Between Colleague Rating, Student Rating, Research Productivity, and Academic Rank in Rating Instructional Effectiveness," *Journal of Educational Psychology*, Vol. 64, No. 3 (June 1973), pp. 274-77.
9. R.A. Hicks, "The Relationship Between Publishing and Teaching Effectiveness," *California Journal of Educational Research*, Vol. 25, No. 3 (March 1974), pp. 140-146.
10. A.S. Linsky and M. A. Straus, "Student Evaluations, Research Productivity, and Eminence of College Faculty," *Journal of Higher Education*, Vol. 46, No. 1 (January/February 1975), pp. 89-102.
11. P.L. Dent and D.J. Lewis, "The Relationship Between Teaching Effectiveness and Measures of Research Quality," *Educational Research Quarterly*, Vol. 1, No. 3 (Fall 1976), pp. 3-16.

12. D.P. Hoyt and R.K. Spangler, "Faculty Research Involvement and Instructional Outcomes," *Research in Higher Education*, Vol. 4, No. 2 (1976), pp. 113-22.
13. S.J. Michalak, Jr., and R. J. Friedrich, "Research Productivity and Teaching Effectiveness at a Small Liberal Arts College," *Journal of Higher Education*, Vol. 52 (November/December 1981), pp. 578-597.
14. G.V. Teague, "Compatibility of Teaching and Research at a Large Land-Grant University," *Improving College and University Teaching*, Vol. 29, No. 1 (Winter 1981), pp. 33-35.
15. M. Rodin and B. Rodin, "Student Evaluations of Teachers," *Science*, Vol. 177, No. 4055 (September 1972), pp. 1164-1166.
16. Katherine L. Kasten, "Tenure and Merit Pay as Rewards for Research, Teaching and Service at a Research University," *The Journal of Higher Education*, Vol 55, No. 4 (July/August, 1984), pp. 500-514.
17. Aleamoni and Yimer, *op. cit.*; Dent and Lewis, *op. cit.*; and McGrath, *op. cit.*
18. Aleamoni and Yimer, *op. cit.*
19. Linsky and Straus, *op. cit.* and Dent and Lewis, *op. cit.*
20. McGrath, *op. cit.* and Hicks, *op. cit.*
21. McDaniel and Feldhusen, *op. cit.* and Harry and Goldner, *op. cit.*
22. Voeks, *op. cit.*, p. 215.
23. Michalak and Friedrich, *op. cit.*
24. I acknowledge the important exceptions in disciplines where other avenues exist, e.g., art and music.
25. Bresler, *op. cit.*
26. Rodin and Rodin, *op. cit.*
27. Voeks, *op. cit.*
28. *Ibid.*, p. 213.
29. Rodin and Rodin, *op. cit.*, p. 1164.
30. Goldner, *op. cit.*
31. *Ibid.*, p. 55.
32. Rodin and Rodin, *op. cit.*
33. H.A. Schmitt, "Teaching and Research: Companions or Adversaries?" *Journal of Higher Education*, Vol. 36, No. 8 (November 1965), pp. 419-27.
34. *Ibid.*, p. 426.
35. D.E. Showalter, "Publication and Stagnation in the Liberal Arts College," *Educational Record*, Vol. 59 (Spring 1978), pp. 166-72.
36. McGrath, *op. cit.*, p. 151.
37. Voeks, *op. cit.*
38. Linsky, *op. cit.*
39. Schmitt, *op. cit.*, p. 426.
40. Linsky, *op. cit.*

41. Hoyt and Spangler, *op. cit.*
42. Linsky, *op. cit.*
43. S. Kwiatkowski, "The Organisational Problems of Combining Teaching and Research: Humboldt, Weber, and the Polish Experience," *European Journal of Education*, Vol. 15, No. 4 (1980), p. 355.

Erratum: Book Review *Different People: Studies in Ethnicity and Education, JET* (August 1984), pp. 126-128. The length of the book is 99 pages and not 49.

Announcement: Available free with self-addressed, stamped envelope: (1) "How to Teach Values in the Schools" (A Wisconsin Education Association reprint), and (2) Description and reviews of the Teaching Students to be Critical Series. (Books for grades 4-12 on Emotion, Ethics, Lying, Time, Humor.) Send to: The Language Press, P.O. Box 342, Whitewater, WI 53190 U.S.A.