RESEARCH

Review of the Academic Evidence on the Relationship Between Teaching and Research in Higher Education

Mohammad Qamar uz Zaman

department for
education and skills
creating opportunity, releasing potential, achieving excellence

Research Report RR506

Research Report
No 506

Review of the Academic Evidence on the Relationship Between Teaching and Research in Higher Education

Mohammad Qamar uz Zaman

ISBN 1 84478 162 3

CONTENTS

NON-TECHNICAL SUMMARY

Objective

- This document reviews the empirical academic evidence on the link between teaching and research in higher education. The findings of the most reliable studies are highlighted, and future research directions are suggested.

Three Contrasting Positions

- It is possible that teaching and research are complementary and hence mutually supportive roles. It is also possible that they tend do detract from one another. As a third possibility they may have a null relationship. Determining the direction and extent of the association through a review of quantitative studies is the primary focus of this report.

The Empirical Evidence

- Most studies attempting to measure the link use simple correlation analysis and originate in the U.S. Results of such studies fluctuate with correlation coefficients between research and teaching varying from -.4 to +.8. Three meta-analyses have summarized the statistical relationship found in such studies (about 60 investigations). They reveal that the results converge to a modest, positive correlation of about 0.10.

- Almost none of these correlation studies accounts for other factors which may affect both the teaching and research variables (discipline, institution type, stage of academic career, class size, department size, level of study, and sex). This major shortcoming along with numerous limitations of individual studies characterizes these results as doubtful.

- In the U.K., three studies show strong correlation between RAE and TQA scores. However these too fail to account for the effect of resource and reputational factors. Once the effect of these factors is controlled for, the partial correlation between RAE and TQA scores is found to be 0.10.

- Studies collecting academics' views find academics' overwhelmingly think the roles are mutually supportive. Research is thought to enhance teaching to a greater extent than teaching benefits research. The link is also considered by academics' to be much stronger at postgraduate than undergraduate levels. This though fails to provide conclusive empirical proof of a substantial link, as it may represent false perceptions.

- Two U.K. studies and one investigation conducted in Australia demonstrate strong, positive student perceptions of staff research. At the undergraduate level perceptions of courses, where staff research is integrated, as current and intellectually exciting is observed. This is tempered with criticism of course curriculums being distorted towards staff research and researchers being less available to students in terms of time. At the postgraduate level the association is strongly perceived to be positive. The scope of all three studies though is very small, and it is difficult to confidently generalize these findings.

- It is possible that initially the roles of research and teaching may enhance each other until a threshold level is reached where increasing effort spent on one operates to reduce the quality of the other due mainly to the limitation of time. A single study examining this proposition using U.S. faculty data finds that up to 8 hours per week of teaching are indeed facilitative of research. This detection of a non-linear relationship is significant and could partly explain why a strong association is not found between the two activities by the bulk of studies which measure linear relationships.

- The evidence gathered for this document suggests that research and quality teaching are not contradictory roles. However, we can not conclude from the information at hand that the link is strongly positive. The evidence indicates the relationship may be modestly positive, though it is likely to be stronger at postgraduate than undergraduate levels. The overall quality of the statistical analyses on which these conclusions are based is not high.

Further Work

- Given the severe limitations of the studies reviewed it is advised that some further work be undertaken before drawing firmer conclusions about the nature of the link. Three possible routes for further investigation are suggested.

- Firstly further examination of the possible link between research and time spent on teaching should be attempted. This would help corroborate or reject the findings of the solitary report evidencing a curvilinear relationship. If such a relationship is found levels of optimal distribution between time spent on both activities could be calculated.

- Secondly, and especially if a curvilinear relationship is not detected, a regression analysis at the level of the individual incorporating factors such as discipline, level of study, institution type, stage of academic career etc. which may have an effect on both roles is advised.

- Thirdly, supplementing the regression analysis at the level of the individual, a departmental level analysis measuring the link between RAE and QAA scores, while accounting for resource and reputational factors should be considered. This has the advantage of data being readily available.

INTRODUCTION

REVIEW OF THE ACADEMIC EVIDENCE ON THE RELATIONSHIP BETWEEN TEACHING AND RESEARCH IN HIGHER EDUCATION

INTRODUCTION

This report reviews the empirical academic evidence on the interaction between teaching and research in higher education. While there is considerable research literature on this issue much of this is critically flawed due to defects in methodology. Care has to be taken when drawing conclusions from most of these studies. The report aims to provide some visibility into the weaknesses of the literature, to identify the most robust studies and highlight their findings, and suggest avenues for further research. The context within which this advice is provided is U.K. higher education institutions.

This introductory section starts by examining the arguments supporting and counteracting the presence of a positive link between research and teaching. It then provides an overview of the empirical research conducted in the area. Finally, possibilities for further research to strengthen the evidence base regarding the relationship between teaching and research are discussed.

The rest of the paper deals with in depth reviews of studies identified as being the most important.

THE CONTRASTING PERSPECTIVES

The relationship between research and teaching has been examined at length in the context of higher education. There are three contrasting perspectives – Positive, Negative, and Null - on the relationship. The arguments supporting each position are described as follows:

Positive Relationship

There are several different arguments hypothesizing a positive relationship between research and teaching. This note identifies sixteen such contentions from the substantial existent literature on the topic.

Research to Teaching

- Research helps in expert and contemporary knowledge being passed onto the student. In certain institutions and disciplines, it is important for students to experience being at the cutting edge of their subject. This is a relationship where the excitement of engaging with the development of the knowledge base of the discipline itself contributes to student learning.

- Textbooks may not be current in many rapidly developing areas. Lectures by active researchers aware of the newest perspectives in their field may be the first point of contact for students with the latest developments. Additionally, results from one's research can be used to clarify, update, and amend the teaching of a topic.

- Every higher education student can potentially benefit from exposure to the methods and attitudes associated with well-developed forms of scholarly activity by developing the attitude of inquiry, the use of data to test theories and ideas, and the transferable skills of critical analysis and presentation of findings based

on evidence. Active researchers are more effective at instilling an actively critical approach rather than a passive acceptance of facts.

- Students appreciate teachers who present research that the teachers have actually conducted. This provides an authenticity to the presented material that differs from presentations by teachers who are only discussing the work of others in which they have no active involvement.

- Research leads to credibility enhancement. Students have the desire to learn from people respected in their fields.

- There is an important role for research in helping institutions to attract, reward and retain high caliber staff, who might otherwise not be available for undergraduate teaching.

- Successful research can increase lecturer confidence, leading to better classroom performance.

Teaching to Research

- Teaching can be particularly good for young researchers because it can reinforce their ability to expound and clarify their thinking.

- Directly, teaching provides a stimulus to individual academics. Discussions in class may produce ideas for further research. And some student projects may produce data, which could feed into published research or grant applications.

- The process of teaching the subject matter of a discipline forces academics to clarify the big picture into which their specific research specialization fits, hence providing a positive impulse for their research. Preparation of teaching materials can elucidate gaps in the academic's knowledge base.

Teaching to Research & Research to Teaching

- Research is also thought to be good for staff development, institutional image and reputation, and student recruitment. These factors could assist the setting up of an environment most conductive for learning.

- An active research interest is important for good university teaching because there are common abilities underlying both research and teaching. There is a correlation between the two because the attitudes, values and competencies that lead to excellence in research (dedication, hard work, imagination, originality and critical analysis) are also likely to lead to excellence in other spheres of academic activity.

- Exposing students to research makes them more likely to consider doing research themselves which could be important in areas where there is a shortage of researchers (as in medicine).

- The personal learning of researchers can make them more able to identify with their students' learning. Staff and students can learn together about research rather than the beginning researchers feeling inhibited by their inexperience.

- Sharing the results of one's research efforts with an appreciative audience provides reinforcement for having done the research and pursuing further research.

Negative Relationship

A review of literature also provides a list of arguments supporting a negative relationship between quality in research and teaching.

- There is limited time, energy, and commitment, for faculty to do both teaching and research. With academics usually prioritizing research over teaching and significantly increased pressures on staff time particularly during academic terms, students suffer.

- Teaching and research require contrary personality characteristics unlikely to be found in the same person. For example teaching success may depend on attributes such as gregariousness that might tend to be inversely correlated with attributes associated with research success such as intellectuality.

- According to the divergent rewards model different obligations and rewards are allocated to each activity. Teaching does not contribute significantly towards overall salary, and therefore suffers in comparison to research, which does bring monetary gain.

- Promotion for faculty on the basis of research alone sends a signal to young academics to reduce the time and effort spent on teaching to a minimum so that they can get on with churning out publications. It provides a clear incentive for faculty to neglect teaching in favor of research. Research-active staff being able to "buy themselves out of teaching" works along similar lines.

- Academics may attempt to distort the curriculum toward their own research at the expense of a broader program of study.

- Active research in an area might lead to strongly held views and consequently poor tolerance of alternative viewpoints raised in the classroom.

- Researchers might pitch their classes at too high a level.

- It is likely that research tends to be much more specialised than teaching and this would produce disparities between research content and teaching content.

Null Relationship

The zero relationship between teaching and research is analyzed through four hypotheses.

- Research and teaching are different enterprises, and research is seldom driven by curricular considerations.

- The notion that researchers and teachers are different types of people.

- If research and teaching were separately funded, the former could follow the interests of funding agencies and teaching could focus on the needs of the students.

- Both sides of the argument are correct and incompatibilities and complementarities tend to cancel one another out in which case we would observe a zero relationship between teaching research.

SUPPORTING EVIDENCE

A fundamental query arising from the existence of these three contrasting perspectives is the extent to which they are supported by empirical research. While this section provides a brief overview and general direction of the empirical evidence, the studies discussed here are examined in greater detail in the literature review section of the report.

Reviews of empirical research

Five substantial reviews of empirical research have been conducted on the relationship between teaching and research (Faia: 1976, Feldman: 1987, Allen: 1996, Hattie & Marsh: 1996, Braxton: 1996).

Faia's (1976) analysis is dated, covering 11 empirical studies of teaching-research linkages published been 1952 and 1975. He finds that of the set of studies, 4 show no relationship between teaching and research proficiency, whereas almost all the others show weak positive associations.

Three of the research reviews are meta-analyses (Feldman: 1987, Allen: 1996, Hattie & Marsh: 1996). A meta-analysis describes the statistical integration of separate studies. The method is superior to alternative reviews of research. A traditional narrative review has the disadvantage of being subjective and therefore prone to bias and error. Meanwhile, the "vote-counting" method, which chooses the view supported by most studies, ignores sample size, effect size, and research design. The meta-analytic approach overcomes these problems and is therefore seen as an attractive alternative.

The overall relationship between quality of teaching and research is found to be slightly positive in each of the three meta-analyses. Feldman (1987) detects an overall correlation of 0.12. Allen (1996) finds a relationship of 0.11. Hattie & Marsh (1996), which is the most recent and the most comprehensive of the three studies, find the smallest linkage of the three. Their correlation coefficient being 0.06. This rises to 0.11, in line with the findings of the other two investigations, when the overemphasis of their results on a few studies with a large number of reported correlations is removed. Thus, all three meta-analyses find some support for a positive (though modest) relationship between teaching and research.

This method of averaging the correlations obtained from different studies is appropriate for deriving an aggregated measure of association between teaching and research. However, the process does not permit the simultaneous appraisal of the three contrasting perspectives because it masks variation in findings among the individual studies. To overcome this, Braxton (1996) employs a vote-counting method which entails tallying the proportion of studies that support each of the three contrasting perspectives. Analyzing 30 studies which focus on the relationship between research productivity and student appraisals of teaching effectiveness, Braxton finds moderate support for both the positive relationship perspective (37% of the studies) and the null relationship perspective (60% of the studies).

Two conclusions can be drawn from these reviews with regard to support for each of the three positions. First, research and quality teaching are not contradictory goals.

There is no perceptible evidence in any of the reviews of the presence of a negative relationship between teaching and research. Second, the evidence does not point towards a strong positive link. While the meta-analyses in particular indicate that research on the relationship on average suggest a positive link between teaching and research, the magnitude of the link is found to be modest.

A different interpretation is possible which could put a different spin on the conclusions stated above. It is possible that both sides of the argument are correct to a degree and that the incompatibilities and complementarities may tend to largely cancel one another, due to which the reviews observe only a weak positive relationship between teaching skill and research productivity.

Correlation analyses form the bulk of the studies investigated in these reviews. There are various critical problems inherent in using the process to measure the linkage between research and teaching. The weaknesses of these individual studies are severe enough to cast significant doubt on the findings of any reviews based on them.

Correlation studies

Most empirical research in this area is limited to the evaluation of simple correlations and has not taken advantage of advances in statistical methodology. Of the studies reviewed, Linsky & Strauss (1975), Faia (1976), and Centra (1983) were chosen for having the largest sample sizes among all the studies in Hattie & Marsh's (1996) sample. Kremer (1990 & 1991) were chosen for being two of the most recent studies in the sample and for having a fairly unique methodology. Noser *et al* (1996) was not covered in Hattie & Marsh's sample and was included as an example of more recent research following the same methodology as older studies. In addition a number of shorter reviews dealing mainly with easily located empirical studies from the Hattie & Marsh sample are provided. These did not warrant longer reviews because the results and methodology were highly similar to the studies for which more substantial reviews are provided. These were included to illustrate that almost all studies identified in the three meta-analyses based their findings on simple correlations. There are several serious problems inherent in using the technique to gauge the relationship between research and teaching.

A methodology based entirely on zero-order correlations does not consider the effects of other variables, which could affect the relationship. For example Linsky and Strauss (1975) demonstrate that enrolment relates oppositely to overall teaching ratings. It is only by statistically removing the effect of this variable (and other significant variables that influence the linkage) that a true value of the basic relationship between teaching and research can be obtained. A study based entirely on zero-order correlations cannot manage this. It is possible to use partial correlations that control for the influence of specified variables. However, most of the studies reviewed did not apply this process (Stallings & Singhal: 1970 is an exception).

Two further points to be noted about correlation studies of this nature. Firstly, even if the correlation coefficient between the factors analysed is zero it does not mean that the two variables are independent. And this leads directly to the second point, which is that correlation in fact is a measure of linear association or linear dependence only. Deviations from linearity will curb the correlation coefficient even if there is a very close relationship between the variables. For example it would not pick up any non-linear effect of time spent on teaching on research productivity as found by Mitchell & Rebne (1995).

Finally, though it is a measure of linear association between two variables correlation does not imply any cause and effect relationship. Causal relations cannot be proved based on correlation coefficients. The associations being identified through correlations may in fact be spurious, i.e. correlations may be due mostly to the influence of other variables.

While several of the reviewed studies identified variables, which could influence teaching, research, and their relation (e.g. class size, type of institution, discipline, stage of academic's career), and some even calculated the correlations between several of these variables and measures of either teaching or research, almost none reverted to partial correlations to control for their influence.

A problem observed among some of the correlation studies reviewed is the calculation of the link between research measures and teaching measures that cover different time frames. For example Faia (1976) relates publications in the preceding two years, to a lifetime receipt of a teaching award. It is obvious that some academics may have teaching awards simply by virtue of having been in the profession for a longer period of time. It is not certain how having received a teaching award 10 or 15 years ago would relate to a publication in the past two years. Similarly Noser *et al* (1996) relate a two-term measure of teaching performance with a lifetime measure of research output. While this is not a problem singularly associated with correlation studies it has cropped up in several of the correlation papers reviewed. Linsky & Strauss (1975) is similarly problematic.

Given these problems it is questionable how valuable/trustworthy the simple associations obtained in these analyses are toward shedding light on the link between teaching and research.

RAE and TQA/QAA scores

As seen in the correlation studies review section, there have been extensive studies at the level of the individual. The compilation of Research Assessment Exercise (RAE) and Teaching Quality Assessment (TQA) scores allows the opportunity to measure the relationship between the two indices at institutional and departmental levels. Given its visibility to policy, surprisingly the primary research covering this issue is limited.

A study by Ellis (2001) finds a compelling match between English departments which scored highly on the RAE and those who do well on the TQA. Examining Scottish universities Drennan (1999) calculates that over 70% of the variation in mean TQA scores can be explained by RAE scores. She also discovers that the match between the two indices is strong for science subjects and weak for social sciences. Neither of these studies however accounts for the effects of other factors which could influence the relationship. Like the zero-order correlation studies at the level of the individual described before, this limitation characterizes these results as doubtful. HEFCE's (2000) report is similarly constrained.

A more comprehensive statistical analysis by Drennan & Beck (2001) who investigate the relationship for all U.K. universities, finds when reputational and resource factors are accounted for the level of variation in TQA scores explained by RAE ratings falls to around 10%. This partial correlation analysis is the most reliable investigation among the four studies reviewed. Even then there are several problems associated in attempting to quantify the link between teaching and research quality using TQA and RAE scores as attempted by the studies reviewed.

One of the biggest problems is averaging scores across all departments of an institution, as Drennan (1999) and Drennan & Beck (2001) do. This masks the true relationship between the scores. An institution could have high TQA scores for some departments but low RAE scores and vice versa. Also an institution could have had most of its departments assessed in the TQA, but could have submitted only a few departments in the RAE review.

Another significant problem is the subjectivity inherent in the peer review process of the RAE and in assessors' evaluations in the TQA. The latter could involve a case of research reputation and resources giving "halo effects" to perceived teaching quality.

The two indices of assessment do provide an opportunity to provide a measure of the strength of the link, though there are limitations to matching the two indices accurately and problems with the nature of the assessment processes themselves. None of the four studies reviewed is a perfect piece of research, even given the limitations. It is suggested that a much less flawed measure of the relationship can be calculated from the indices. This option is discussed in the section covering further avenues of research.

Faculty & administration perceptions

The aim of the papers covered under this heading is to look more closely at the ways in which there is interaction between research and teaching by asking university faculty members and administrators their opinions. It may be questioned to what extent faculty members' evaluation of the interaction between research and teaching are a good measure of the actual symbiosis between these activities. There is obviously the issue of self-interest in administrators emphasising the value of research for teaching, if research awards and grants bring in resources. Jenkins *et al* (2003) suggest however that it could also be argued that administrators have in certain ways a more informed and sophisticated view as to what is teaching quality than maybe obtained by student questionnaire studies. The argument could be extended for faculty members.

The quantitative data in Leslie *et al* (1998) indicates that though there are some differences in strength of support dependent on institutional characteristics, chief academic officers overwhelmingly support the idea that research enhances teaching effectiveness. These results are similar to two linked studies (not covered in this report), one by Neumann (1993) in Australia, and the other by Rowland (1996) at Sheffield University which demonstrate that department heads and administrators see strong positive correlations between staff involvement in research and the intellectual currency of their courses.

Smeby (1998) finds in a survey of Norwegian academics that more than 95 percent of faculty at PhD level and more than 90 percent at major subject level thought their teaching was affected by their research 'a great deal' or 'some extent', while there were few who thought this at a undergraduate level (about 50 percent). Faculty also thought that teaching gave positive impulse to research to a lesser extent than they thought research enhanced teaching.

The strength of these studies lies mainly in shedding light on the complexity of the interaction and hence advancing the understanding of the linkage. It is difficult though to get a clear answer to the question of the true linkage between research and teaching from these studies. This can only be obtained if the measure of teaching performance is obtained completely independent of the measure of research

performance. In failing to meet this critical requirement all this type of studies can provide are perceptions of faculty and administrators.

Elton (2001) provides an interesting hypothesis as to why academics may believe strongly in the presence of a positive link between teaching and research. He suggests that “academic teachers think of students in terms of their own student experience and rarely if ever verify how typical it is from the point of view of their own students. Since only a very small proportion of students ever become academics, it is of course the very opposite of typical; yet it is the experience which for a long time has been dominant in the minds of academics”.

Student perceptions

Student perceptions on the effects of lecturer research on learning are evidently relevant to the debate, but only three studies have investigated the issue by asking students directly.

Neumann (1994) carried out an exploration at a research oriented Australian institution. She found that there were tangible benefits to students of staff research, mainly through students perceiving that their courses were up to date and that staff demonstrated interest in what they were studying. However, many students were also critical of subjects in which a teacher’s individual research and research interests were seen to dominate, particularly at the expense of the aims of the course.

Jenkins *et al* (1998) arrive at similar conclusions after an investigation at Oxford Brookes University. The authors found that students felt that they benefited from staff research, in that the teaching was more up-to-date and more scholarly. However, they saw disadvantages in that research oriented teachers tended to be less available to them, were often preoccupied with their research at the expense of their teaching, excluded their students from stake holding in their research and at times distorted the curriculum towards their research. Nevertheless, the authors conclude, that from the student perspective there is a largely positive teaching-research link, while the main adverse impacts can at least in part be resolved through effective management.

The Lindsay *et al* (2002) study reports a quantitative analysis of the data discussed by Jenkins *et al* which focused on undergraduate students, and then reports findings from a new investigation which replicates the original study using postgraduate students. When discourse contributions are related with the Research Assessment Exercise Score of a department, it is observed that as RAE rating goes up, the percentage of positive comments about the effects of research on teaching tends to increase. For postgraduates, the percentage of negative comments also tends to decrease contrary to an increase observed for undergraduate students.

This last described study is the only one of the three which makes a serious attempt at producing quantitative results. The methodology of tallying up discourse contributions does have several flaws however (discussed in review of paper). Also since the findings are from a limited number of disciplines, in one institution only, and based on a limited sample of students, the results cannot be generalised with any degree of confidence.

Examining the time spent on teaching & research, and other studies

Under this heading there are four analyses of interest. Oliveras *et al* (2003), Gottlieb & Keith (1998) and Euwals and Ward (2000) all focus to some extent on U.K. academics. The first two studies examine the relationship between time spent on research and time spent on teaching. Both run linear standard multiple regression models, but end with different results. One suggests a tension between times devoted to the two activities, while the other suggests the possibility of complementarity.

The most important of these studies is Mitchell & Rebne (1995). It supports the view that a combination of the complementarity role and the economy of time variables could give rise to a curvilinear relationship, between teaching and research. The study tests the proposition that moderate amounts of faculty times spent on consulting and teaching are facilitative of research productivity by fitting continuous piecewise-linear regression models to 1980 data of U.S. faculty. The analysis supports the view that it is inappropriate to regard academic job content in zero sum terms when research productivity is the outcome of interest. The results indicate that up to four hours per week of consulting and up to eight hours per week of teaching are indeed facilitative of research productivity.

This is a most interesting finding. The problem with most of the studies reviewed is they see teaching and research in zero sum terms. It is entirely possible that initially research enhances teaching (and vice versa) until a threshold level is reached where research efforts operate to reduce the quality of teaching. This non-linear effect would not be picked up in correlation studies, which form the bulk of the literature on the topic and the basis for the three meta-analyses described.

Surprisingly only this one study has attempted to pick up this non-linearity. The attempt of further research should be to attempt to replicate the general methodology of the study (albeit with some changes, as noted in the discussion of weaknesses in the review of the paper). By examining how time spent on alternative activities affects research, support for, or refutation of Mitchell and Rebne's (1995) results could be provided. This is also important given the outdated nature of the data employed in that particular study (a 1980 sample).

FURTHER AVENUES OF RESEARCH AND PITFALLS TO AVOID

As suggested in the section above one attempt of further research should be to examine the possible non-linear relationship between research and time spent on teaching. While time spent on teaching is not equivalent to good performance in teaching, the calculation of this association would help confirm or deny whether research initially benefits from teaching before time constraints cause a conflicting relationship to emerge between the two roles. If this non-linear role is confirmed it would be obvious why a multitude of correlation studies had failed to pick up the true link between research and teaching.

Also, if the relationship is curvilinear, another attempt of further research should be to identify what levels of research and teaching are optimal to enhance the complementarity between them. Optimal redistribution would be mutually beneficial for the roles, and would lead to a higher association than found in the meta-analyses and correlation studies.

A point to be noted if this avenue is explored. Mitchell & Rebne measure the relationship between research productivity and time spent on teaching, but do not

investigate the link between teaching effectiveness and time spent on research. It is suggested that if a replica study is carried out, this latter relationship be calculated also to gain a fuller understanding of the nature of the link.

None of the correlation studies reviewed deals with data from the U.K. Only four of all types of studies reviewed in this report use U.K. data. Two of these Oliveras *et al* (2003) and Gottlieb & Keith (1998) use only time spent on research or teaching as dependent variables in their analysis.

Gottlieb & Keith (1998) use the First Carnegie Institute Survey (CIS) of academics (1991-93) covering 13,984 faculty members from 18 studies. No new edition of the survey has been attempted. It should be possible to replicate Mitchell & Rebne's (1995) study to examine the presence of a non-linear relationship between teaching and research using the CIS. The survey includes data on time spent on teaching and research, research productivity, and individual and institutional characteristics. If an analysis is required for the United Kingdom alone, it should be possible to use only U.K. specific data from the CIS. Three points to note though. First, though more recent than the 1980 data set employed by Mitchell & Rebne, the CIS is already more than 10 years old. Second, from Gottlieb & Keith's analysis it looks likely there is no measurement of time spent on service/administrative duties (though this is not confirmed). And third, the measure of research used by Gottlieb & Keith is a straight publication count. It is possible however, that an indicator of quality could be derived from the CIS which Gottlieb & Keith failed to do.

Euwals & Ward (2000) do analyze the hypothesis that *productive* researchers are also the best teachers employing data from five British universities. However their measure of teaching effectiveness is suspect. Apart from this analysis, no other study attempts to quantify the link for the U.K.

If an examination of a possible non-linear relationship between teaching and research refutes the findings of Mitchell & Rebne (1995), a regression analysis improving on Euwals & Wards (2000) study is also advised. As is apparent a primary consideration when quantifying the relationship between teaching and research requires the use of legitimate measures to gauge the performance of both activities. Possible measures to use, and some commonly made mistakes to avoid are discussed in the section on measurement problems.

In addition to a replication of the Euwals & Wards (2000) study as described, another avenue could be measuring the link between RAE and QAA scores. This has the considerable advantage of measuring the relationship at a departmental (instead of individual) level with **data that is readily available**.

Drennan & Beck (2001) is the finest of the four RAE/TQA studies reviewed. The attempt of the suggested study should be a replication of the regression run in that paper with several modifications.

Firstly, the RAE and QAA scores should **not** be averaged across all departments of an institution. Regressions should be at the level of the department. Where departments have been assessed for teaching quality but not entered for review by the RAE, it is suggested they be given either a 0 or the lowest rating on the RAE scale. If possible the percentage of research active staff in departments should be accounted for.

Drennan & Beck (2001) ran into a multicollinearity problem and as a solution when regressing RAE on TQA they remove the collinear variable (student entry standards).

As explained in the review of the paper that invariably led to specification error causing further problems. They further calculated high order partial correlation coefficients as an alternative. While the process calculates the incremental contribution to the explained variance of TQA scores by RAE rating, a regression analysis would also be worthwhile through estimating the value of TQA score given RAE rating. The multicollinearity problem could be overcome by a simple transformation of variables. RAE results from 2001 and 1996 could be used to correspond with the latest round of QAA scores and the round prior respectively, and a regression be run in the first difference form. Although the levels of RAE and student entry scores may be correlated, there is no a priori reason to believe that their differences will also be highly correlated. The first difference regression model should therefore reduce the severity of multicollinearity. Before this step is taken the degree of multicollinearity should be examined however. It is possible that while the multicollinearity shows up at the institutional level it is less problematic at the subject level (though this is not likely).

This still leaves the problems inherent in the processes of both assessment indices. The subjectivity inherent in the assessment method cannot be removed. A close relationship between the two might scales might be indicative of nothing more than "halo effect" in the teaching assessment, though reputation and resource factors included in the regression should account for most of this. In addition it could also be argued that since the previous system of TQA run by HEFCE has been replaced by the QAA arrangement which is independent in nature, this "halo effect" could have reduced, which further suggests an updated study.

There seems to be no case for attempting to replicate or update either the correlation analyses reviewed or inquiries from practicing academics which in reality investigate in the main their personal views. The avenue of student perceptions though is interesting. A replication of Lindsay *et al* (2002) at a larger scale over several institutions and more departments could be valuable. However such a project is bound to be time, effort and resource consuming. It is likely a body of work covering several institutions may eventually be built through independent exploration by different researchers, given the approval such work has gained in several academic papers considering the link.

Measurement problems

Trying to quantify the relationship between teaching and research is complicated by two related problems:

- research and effective teaching are subject to varying definitions, and
- regardless of which definition is used, the two factors are not measured easily.

The first stated issue of definition is beyond the scope of the present report. However, the issue of measurement problem arising from the review of empirical research will be discussed here.

At the level of the individual, several methods of evaluating research productivity and teaching effectiveness, i.e. what counts as good within these two domains, have been used in different studies. Measures of research used in the literature reviewed include:

i) number of publications
ii) grants awarded
iii) number of citations

iv) peer or chair rating of research
v) time spent on research
vi) faculty membership in a university research society
vii) awards for research
viii) combination of grants and publications, and
ix) the research creativity of the scholar as rated by other faculty

Output as such has an ambivalent meaning. A clear distinction should be made between productivity and quality. The primary interest of researchers should be to produce publications of high scientific *quality*. The most commonly used measure though has been the number of publications. Brew and Boud (1995) suggest ratings of research productivity by citation counts come closest to a measure of research quality. One problem with citation counts though is the time period elapsing between a work actually being published and other works referring to it. Citation analysis is bound to be retrospective: it measures *past* performance. It has been argued that this should not be an obstacle, in most cases researchers performing excellent in the past will be excellent in the future as well, and *vice versa*. However it could also be argued that an academic's research profile changes over his/her career. In this case citation scores should only ideally be used for research quality measurement over a long time period (or life time) in conjunction with a measure of teaching effectiveness covering a similar time period. Additionally some disciplines may not have citation indices which allow easy collation of citation scores.

Apart from citation analysis, to evaluate the *quality* of research, several parameters can be used: Invitations for reviews or chapters in books, etc. invitations for international key-note lectures, editorships of international journals, and capacity to obtain external funding. Perhaps the most convenient measure of quality (especially where measures being related cover a short period of time) would be a weighted composite which allows for the quality of journals in which articles are published, and some grading of meets where papers are presented.

One limitation of most research measures is the need to assume that levels of activity were fairly consistent for the years preceding the survey. Given the often lengthy period between the time research is initiated and the time it reaches print, the effect being sought is more properly that of a lagged variable reflecting research activity a few months or years back. This consideration of the lagged period of research has not been picked up in most studies, and should ideally be corrected for when attempting to quantify the link between research and activity.

The bigger problem is getting a good measure of teaching. Studies have used:

i) student evaluations
ii) peer evaluation
iii) time spent on teaching activities
iv) a nomination or receipt of an award for teaching, and
v) measurement of teaching related activities

Studies most commonly use student and staff ratings of teachers. These are based on their perception of teaching effectiveness. As Brew and Boud (1995) argue this gives an incomplete and partial indication of what constitutes good teaching. Further weaknesses are apparent in both measures.

Staff (peer) evaluation of teaching is argued to represent a "halo" effect created by the knowledge of an academic's research record. These are considered inherently less trustworthy than student evaluations. However peer reviews can be valuable

through highlighting areas of teaching which students are ill-equipped to notice. It should be noted that compared to students, faculty possess only the most indirect and sparse information about the teaching activities of their colleagues.

A contrary argument could also be made that it is conceivable that students' awareness of the scholarly eminence of a faculty member might lead them to entertain unexpectedly high expectations regarding teaching performance. And therefore their evaluations could very well be "contaminated" too. Cahn (1987) suggests, students know if instructors are likeable, not if they are knowledgeable; they know if lectures are enjoyable, not if they are reliable. However, since students are the direct consumers of the teaching product and have the most first-hand information concerning their instructors' teaching behavior, they may be in the best position to evaluate their professors' performance.

Faia (1976) makes an interesting case for the use of teaching awards as a measure of teaching skill. His argument is that teaching awards often involve collaborative decisions made by student, faculty, and administrators. And so therefore it is safe to assume that faculty members who have received a teaching award have satisfied two or three constituencies. However this measure can only be used when the time frame is fairly long and needs to be related with a research measure covering the same time frame.

A common questionable practise in studies covering several institutions is the use of student evaluation scores across institutions without control. The format of the instrument may differ from institution to institution, which could affect respondents' scoring on the evaluation sheet. Ideally some corrective method (e.g. converting all evaluation scores to a common unit of measurement carrying out a z-score standardisation for each institution) should be used to standardise scores from each institution.

Another problem is scales, which use number of publications, are discipline-specific; different disciplines have different publication rates. Within a discipline this causes no difficulties, however there are problems of translating findings across disciplines. A simple solution would be to standardize publication counts from each discipline.

Additional factors

A final point not to be forgotten is to collect data on discipline, institution type, stage of academic career, class size, department size, level of study, and sex and to use these explanatory variables in the regression analysis of the effect of teaching on research and vice versa. While these are mostly self evident and have been discussed at greater individual length in reviews of studies where they have been used, the point on the level of study is especially important to reiterate.

Studies measuring student perception have shown a stronger positive association between research and teaching at the postgraduate level than at the undergraduate level. Similarly academics' overwhelmingly support the position that the strength of the link is far greater at higher levels. Most studies either fail to mention what level of study their sample covers, or else focus only on undergraduate courses. Simply measuring the link at the undergraduate level may not be reflective of the true association, given a situation where more and more students opt to undertake postgraduate study.

One limitation of this exercise evolves from its prescribed function of gathering evidence on the link between research output and teaching quality. It is likely that

teaching and research both benefit from the large element of common facilities in the form of laboratories, libraries and computer facilities which may not be as extensive in the absence of one of the two roles. Some estimation of cross subsidization would need to be calculated to see the benefit of shared facilities.

References:

Brew, A and Boud, D (1995) Teaching and research: establishing the vital link with learning, *Higher Education*, **29**, pp 261-73

Cahn, S. (1987, October 14). Faculty members should be evaluated by their peers, not by their students. *Chronicle of Higher Education*, p. B2

Elton, L (2001) Research and teaching: conditions for a positive link, *Teaching in Higher Education,* **6**, pp 43-56

Jenkins, A, Breen, R, Lindsay, R and Brew, A (2003) *Reshaping teaching in higher education: linking teaching with research*, Kogan Page Ltd., London

Neumann, R (1993) Research and scholarship: perceptions of senior academic administrators, *Higher Education*, **25**, pp 97-110

Rowland, S (1996) Relationships between teaching and research, *Teaching in Higher Education*, **1** (1), pp 7-20

The reference list includes only those publications mentioned which are not covered by separate literature reviews in this report.

REVIEWS OF EMPIRICAL RESEARCH

RESEARCH PRODUCTIVITY AND SCHOLARLY ACCOMPLISHMENT OF COLLEGE TEACHERS AS RELATED TO THEIR INSTRUCTIONAL EFFECTIVENESS: A REVIEW AND EXPLORATION
*Research in Higher Education, **24**, 1987*

Ken Feldman

THE RELATIONSHIP BETWEEN RESEARCH AND TEACHING: A META-ANALYSIS
*Review of Educational Research, **66** (4), 1996*

John Hattie
Herbert W Marsh

RESEARCH PRODUCTIVITY AND POSITIVE TEACHING EVALUATIONS: EXAMINING THE RELATIONSHIP USING META-ANALYSIS
*Journal of the Association for Communication Administration, **2**, 1996*

Mike Allen

Meta-analysis is a set of statistical procedures designed to accumulate experimental and correlational results across independent studies that address a related set of research questions. As research results accumulate, it becomes increasingly difficult to find the knowledge in this flood of information. Meta-analysis is the statistical analysis of a collection of individual studies for the purpose of integrating their findings. There have been a number of studies attempting to quantify the linkage between research and teaching. Feldman, Hattie & Marsh, and Allen have conducted meta-analyses to more fully assess the findings on this relationship.

The process has the advantage of providing a rigorous alternative to the casual, narrative discussions of research studies, which typify attempts to make sense of rapidly expanding research literature. In a meta-analysis, research studies are collected, coded, and interpreted using statistical methods. The result is an integrated review of findings that is more objective and exact than a narrative review.

Feldman's analysis is the earliest such attempt to collate information on the linkage between teaching and research. It includes 29 studies, all of which are also included in the meta-analyses of Hattie & Marsh, and Allen. Allen's study is based on 46 quantitative studies, while Hattie & Marsh expand the number of studies used in Feldman to a total of 58. Three studies are unique to Allen's sample, all the rest are included in Hattie & Marsh. We expect the results to be similar across all three meta-analysis given the significant amount of overlap among the quantitative studies used. That there are three meta-analyses increases the reliability of results if they confirm each other. All research requires replication before accepting the results as definitive or authoritative. If any error exists in the conducting of the analysis, the replication would reveal the error and permit an assessment of it.

To locate studies Allen conducted a literature search using both a manual and CD ROM search of *ERIC, Psychlit, and CommIndex*. A manual search was conducted on the *Educational Research Index*. The key words used were "faculty evaluation" and

"faculty promotion" as well as "productivity". All manuscripts' reference sections found by this search were examined for possible additional sources of information.

Relating a similar mode, Hattie & Marsh report that their process of locating studies required much systematic searching of CD-ROM data bases, tracking through references in other articles, and personal contact with many of the leading workers in the area.

Feldman does not describe the literature search method used, making it impossible to know to what extent the author searched the relevant literature and under what conditions. This is in fact one of the reasons that Allen gives for attempting to replicate the initial study.

For inclusion in this meta-analysis a manuscript had to relate to universities or similar higher education institutions and include:

a) a measure of teaching effectiveness
b) a measure of research productivity, and
c) statistical information permitting the calculation of an effect size

A number of studies (as noted before) were identified according to this criteria in each of the three papers. Given the number of studies it is likely that there are differences in the way each study is structured. If the method used was a simple review of the findings, it would be difficult to determine if the differences between the study outcomes are due to chance, to inadequate study methods, or to systematic differences in the characteristics of the studies. The process of meta-analysis used by these papers helps investigate the relationship between study features and study outcomes. Accordingly, all three authors code their studies' features according to relevant moderator variables to review the relationship and to assess the effects of various moderators. Moderators used include year of study, measure of research, measure of teaching, institution type, and discipline.

The statistical analysis requires two steps:

- First, the statistical information within the primary investigations becomes converted to a common metric for comparison. The metric chosen in each study was the correlation coefficient.
- Second the separate estimates from each report are averaged. The averaging process in Allen's study uses a weighted average by sample size of academics. Hattie & Marsh in contrast use all correlations (498) found between any measures of research and teaching in the 58 studies they include in their paper. Feldman uses an average effect size from each study.

A final step employed by Hattie & Marsh, and Allen, but not Feldman involves testing the average estimate for homogeneity.

A complete listing of the studies and the average correlation found for each appear in table 1 from Allen's paper and table 2 from Hattie & Marsh's manuscript. Feldman's listing is not included since all the studies he employs are covered by the other two meta-analyses as well (all 29 studies are also covered in the table included in the review of Braxton: 1996). The coefficients reported are of corrected correlations averaged across multiple measures. It can be observed that the correlations differ in many cases between the two listings. These differences could have arisen if methods used to correct for artefacts of measurement such as attenuated measurements or

dichotomization of variables was not consistent across both studies. Generally though there is a high degree of agreement between the reported correlations.

The overall relationship between quality of teaching and research was found to be slightly positive in each of the three studies. Feldman detects an overall correlation of 0.12. Allen finds a relationship of 0.11. Hattie & Marsh find the smallest linkage of the three. Their correlation coefficient being 0.06.

Hattie & Marsh base their analysis on all 498 correlations discovered in the 58 studies they used. The problem with the approach is the overemphasis on one or a few studies which report large numbers of correlations. That a study reports many different correlations seems a strange reason to give it greater weight. One study alone (Centra: 1983) accounts for 118 of the correlations used by Hattie & Marsh. If instead of all correlations reported, the overall average correlation from each study is used Hattie & Marsh's correlation coefficient rises to 0.11. The same result (0.11) is obtained averaging by sample size instead. The latter two methods are used by Feldman and Allen. All the correlation coefficients then fall into line (0.11 to 0.12).

The process used by Allen also has its weakness. He weights the correlations according to sample size. The 46 studies he includes cover 64,925 academics. 53,034 of these come from one study alone (Faia: 1976). It is obvious using this method that the correlation of this one study alone will determine the overall correlation found to a great extent. Not weighting for sample size increases the correlation to 0.13. Removing the Faia study from the analysis and using the weighted average from before, however, decreases the detected association to 0.10. Either way, in practise the variation is not large.

It can then be inferred from this that the overall size of the correlation between quality of teaching and research, as measured by the studies included in the three meta-analyses is close to 0.10-0.13.

Allen finds on the basis of a chi-square test that the relationship is not based on a homogeneous set of correlations. This indicates that the average effect should be interpreted with caution since at least one moderating variable probably exists. Hattie & Marsh also discover some degree of heterogeneity but conclude that the resounding indication is that there is very little variance, anywhere, between research and teaching. They also improve their data by removing 10% of correlations making the largest contribution to the heterogeneity finding. Feldman conducts no test of homogeneity for the average correlation produced. The effects of the various moderators is discussed below.

A feature of interest in the analysis is whether the year of data collection would moderate the analysis. The argument runs that the changing emphasis on research within the academic community may change the underlying relationship between research and teaching. Allen finds that the analysis shows a small negative correlation (-.04) between the size of association and the year of data collection. A secondary analysis examining whether the correlations changed over the various decades also show that the general trend is negative. Hattie & Marsh discover a similar negative relationship between the year of publication and the teaching and research relationship (-0.16) which indicates that more recent studies tended to find the lowest relationships.

The method of teaching evaluation is considered because of the potential differences in assessing instructional effectiveness. Both Allen and Hattie & Marsh discover that the average correlation for peer evaluations were the largest among all methods. The

significant variability found between the various types of teaching productivity is also indicated in both studies to be primarily due to the greater heterogeneity within peer ratings of teaching. It is hypothesised that some type of halo effect may exist for faculty rating other faculty which makes the measure unreliable, or possibly indicate that one professional judgement of another professional stems from an understanding of the content and technique beyond that of a naive student observer.

Research productivity can and has been measured in a variety of ways. Hattie & Marsh find the indicators of quality of research lead to slightly higher correlations than for more quantity-based indicators. They temper this by noting that this increased correlation was primarily due to two much larger correlations. In these two cases quality was measured by research awards and ratings of research merit. Allen in contrast finds the highest correlation for the measure “number of grants”, for which Hattie & Marsh discover the lowest correlation.

The relationship between teaching and research is also observed, by Hattie & Marsh, to be greater for the social sciences than for the humanities, followed by the natural sciences. This agrees with Feldman who finds that there are higher relationships in the social sciences than in the humanities, and zero average relationships in sciences. Allen does not investigate differences between disciplines.

Hattie & Marsh and Feldman also locate several correlations pertaining to the effects of time on teaching and research. Their findings are identical. They find time on research is positively related to articles published, but surprisingly the amount of time or effort devoted to teaching and closely related activities does not seem much related to teaching effectiveness. Time on research is not related to quality of teaching, but time on teaching is negatively related to publication outputs. They observe that it appears there is a tension between the times devoted to the two activities, but this tension may not be translated into differential outcomes. Feldman speculates that time on research probably comes from non-teaching times and that there is, at best, not a one-to-one trade off between time on teaching and time on research.

Feldman closes by saying: “the likelihood that research productivity actually benefits teaching is extremely small or that the two, for all practical purposes, are essentially unrelated”.

Hattie and Marsh conclude: “the common belief that research and teaching are inextricably entwined is an enduring myth. At best, research and teaching are very loosely coupled.” They note, however, that productivity in research does not detract from being an effective teacher.

Allen, in contrast sees the results in a positive light, stating “the correlation indicates that as either teaching effectiveness or research productivity increases the other variable does so as well.” He concludes by stating:

The practical implications of the finding deserve some consideration. While research is not a perfect indication of high quality teaching, clearly productive research is not inconsistent with quality of teaching. More than likely there is at some point a level of diminishing returns where research efforts operate to reduce the quality of teaching but that point is not developed in this data. The data do clearly support the idea that research productivity and quality teaching are not contradictory goals, the degree to which they are compatible or complementary goals could still be argued.

This is a very interesting interpretation. And it lays the seeds for seeing the relationship between research and teaching as curvilinear, similar to the findings of Mitchell and Rebne (1995). The problem with most studies is they see teaching and research in zero sum terms. It is entirely possible that initially research enhances teaching (and vice versa) until a threshold level is reached where as Allen states, research efforts operate to reduce the quality of teaching. This non-linear effect would not be picked up in correlation studies, which form the bulk of the basis for these meta-analyses. If in fact the relationship is curvilinear, the attempt of further research should be to identify what levels of research and teaching are optimal to enhance the complementarity between them. Optimal redistribution would be mutually beneficial for the roles, and would lead to a higher association than found in these meta-analyses.

Some further points to note about these studies:

Only a correlation exists, no ability to evaluate any causality between the features is possible given the restricted set of conditions of the data. While arguments exist about the nature of the connection, the exact causal connection remains unclear.

This method of averaging the correlations obtained from different studies is appropriate for deriving an aggregated measure of association between teaching and research. However, the process does not permit the simultaneous appraisal of the three contrasting perspectives – positive, negative, or null relationship - because it masks variation in findings among the individual studies.

The average correlation in Allen's study comes from a sample of heterogeneous effects, therefore any interpretation must be cautious. Feldman conducts no test of homogeneity so the condition is not noted. The inability to generate a homogeneous solution using the moderators provides some uncertainty about the ability to generalise the average observed. The direction of the correlation though would not change.

The bulk of publications have often been found concentrated among a few academics. E.g. Ramsden (1992) finds that 14% of academics in his sample of 18 Australian institutions produced half of all output. If the range of publications is restricted this would suppress the correlation between research and teaching.

Departmental characteristics may influence teaching, research, and their relation. The low correlations between teaching and research may represent an amalgamation of positive relations in some departments and negative relations in others. Most studies in the meta-analyses to not distinguish between departments.

Almost all the quantitative studies included in the three papers are limited to the evaluation of simple correlations and have not taken advantage of advances in statistical methodology. Problems inherent in correlation studies are discussed in the main section and the reviews of Linsky & Strauss (1976) and Centra (1983), two studies used in all three meta-analyses.

In light of the non-linear relationship detected between time spent on teaching and research productivity by Mitchell & Rebne (1995), the reliability of the zero sum correlations located by Feldman and Hattie & Marsh concerning time spent on different activities and their effects is particularly doubtful.

References:

Centra, J A (1983) Research productivity and teaching effectiveness, *Research in Higher Education,* **18**(2), pp 379-89

Linsky, A S and Strauss, M A (1975) Student evaluations, research productivity, and eminence of college faculty, *Journal of Higher Education,,* **46** (1), pp 89- 102

Mitchell, J E and Rebne, D S (1995) Nonlinear effects of teaching and consulting on academic research productivity, *Socio-Economic Planning Sciences*, **29** (1), 47-57

Ramsden, P (1992) *Learning to teach in higher education*, New York, Routledge

CONTRASTING PERSPECTIVES ON THE RELATIONSHIP BETWEEN TEACHING AND RESEARCH

*New Directions for Institutional Research, **90**, Summer 1996*

John M. Braxton

Braxton assesses empirical support for three contrasting perspectives on the relationship between teaching and research: Null, Conflict, and Complementarity. He examines thirty studies which empirically calculate the link between teaching and research. 29 of these come from Feldman's (1987) meta-analysis. The one additional study (Voeks: 1962) was reviewed by Feldman but not included in his analysis. The additional benefit from Braxton's investigation comes through the use of vote-counting.

Braxton argues that the method of averaging the correlations obtained from different studies – as applied by Feldman (1987) in his meta-analysis – while suitable for deriving a summed measure of association between teaching and research, does not allow the simultaneous appraisal of the three divergent perspectives. This happens because the process masks variation in findings among the individual studies. By applying a vote-counting method, Braxton is able to overcome this and simultaneously test contrasting positions.

The vote-counting method entails tallying the proportion of studies that support each of the three contrasting perspectives. In fuller detail: all studies linking teaching and research are examined. Three possible outcomes are defined. The relationship between teaching and research is either significantly positive (complementarity), significantly negative (conflict), or there is no specific relationship in either direction (null). The number of studies falling into each of these three categories is then simply tallied. If most of the studies fall into any one of these three categories, with fewer falling into the other two, the modal category is assumed to give the best estimate of the direction of the true relationship between the two activities.

Braxton uses the following criteria to categorise each study according to the perspective supported:

- Studies reporting a statistically significant average correlation between teaching and research of +.10 or higher support the complimentarity perspective.
- Studies indicating a statistically significant average correlation between teaching and research of -.10 or lower affirm the conflict perspective.
- Studies reporting statistically non-significant average correlations, or significant average correlations between -.09 and +.09 confirm the null perspective.

Braxton also defines the extent of support for each position:

- Strong support is noted when 66 percent or more of the total number of studies reviewed establish a given perspective.
- Modest support is noted if between 34 and 65 percent of the studies confirm a position.
- And weak support is noted if less than a third of the studies corroborate a given position.

The results of the analysis are exhibited in table 1. Moderate support is evident for both the complementarity (37%) and null perspective (60%). There is scant support

for the conflict perspective, given that only one of the thirty studies reviewed sustains this position.

Braxton concludes that research does not interfere with teaching effectiveness. But also that there does not appear to be a systematic relationship between teaching and research.

In common with the three meta-analyses reviewed elsewhere in this report, Braxton's study harbours several weaknesses due to the type of studies analysed in the reviews. For a fuller elaboration of the deficiencies listed here see the review for the meta-analyses.

- Most empirical studies included are limited to the evaluation of simple correlations. Problems inherent in correlation studies are highlighted in the main section of the report.
- No causality can be evaluated from this empirical analysis.
- It is possible the range of publications in the studies used is restricted, suppressing the correlation between teaching and research.
- Most of the studies in the analysis do not distinguish between departments. Some do not account for differences in institution type and career stage of academic. The average correlations reported for each study, on which the voting-count analysis is based, do not control for any of these factors.

Further specific to the process applied, the vote-counting method does not incorporate sample-size into the vote. It is well known that as sample size increases, the probability of finding a statistically significant relation between independent and dependent variables also increases. Also, to give the results of a study, which covers a handful of academics of a single discipline from a single university, the same importance as the findings from a national sample covering thousands of academics across many disciplines and types of institutions is, would not be good practice.

References:

Feldman, K (1987) Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: a review and exploration, *Research in Higher Education*, **26**, pp. 227-98

Allen (1996): Table 1. Effect Sizes Relating Research Productivity and Teaching Effectiveness

Author	*Date*	*Correlation*	*N*
Ahern	1969	.238	75
Aleamoni	1973	.000	360
Aleamoni	1973	.033	28
Bausell	1972	.061	105
Braunstein	1973	.040	349
Braxton	1983	.325	174
Bresler	1968	.227	106
Centra	1983	.099	2,968
Centra	1983	.071	1,623
Clark	1973	.255	45
Cornwell	1974	.000	70
Dent	1976	.022	90
Faia	1976	.110	53,034
Freedman	1979	.242	129
Frey	1978	.070	42
Goldsmid	1977	.172	90
Grant	1971	.000	685
Harry	1974	.190	77
Hayes	1984	.210	250
Hicks	1974	.192	459
Hoffman	1984	-.250	65
Hoyt	1974	.086	173
Hoyt	1976	.170	183
Lasher	1974	.000	873
Linsky	1975	.009	1,091
Marquardt	1975	.286	91
Maslow	1956	.640	86
McCullagh	1975	.045	52
McDaniel	1970	.043	76
Michalak	1981	.260	86
Plant	1970	.000	32
Ratz	1975	.000	15
Richardson	1992	.260	67
Riley	1950	.220	389
Root	1987	.199	27
Rossman	1976	.327	122
Rushton	1983	-.066	52
Siegfried	1973	.039	45
Stallings	1970	.260	128
Stallings	1970	.260	121
Stavridis	1972	.163	32
Teaque	1981	.000	16
Usher	1966	.230	26
Voeks	1962	.000	198
Wood	1976	.395	69
Wood	1978	.023	22

Hattie & Marsh (1996): Table 2. Summary of the 58 studies

Author	*Year*	*No. of academics*	*No. of correlations*	*Correlation*
Aiken	1975	360	2	-.030
Aleamoni & Yimer	1973	360	10	-.015
Baird	1980	233	3	.129
Bausell & Magoon	1972	105	3	.074
Braunstein & Benston	1973	349	5	.042
Braxton	1983	174	3	-.168
Bresler	1968	106	2	.224
Centra	1983	4,596	118	.012
Clark	1973	45	1	.310
Cornwell	1974	101	6	.085
Dent & Lewis	1976	90	15	.039
Faia	1976	265,682	2	.111
Freedman & Stumpf	1979	129	6	.169
Frey	1978	42	2	.077
Friedrich & Michalak	1983	74	25	.174
Genn	1980	796	1	-.424
Goldsmid & Gruber	1977	60	1	.000
Guthrie	1949	233	1	.000
Harry & Goldner	1972	211	8	.038
Hayes	1971	355	1	.000
Hicks	1974	459	1	.255
Hoffman	1984	65	1	-.255
Hoffman	1984	65	2	.208
Hoyt	1974	222	1	-.020
Hoyt & Spangler	1976	183	16	.071
Kremer	1990	89	20	-.007
Kremer	1991	90	12	.041
Lasher & Vogt	1974	120	1	.388
Lewis & Gregerio	1984	408	1	.213
Linsky & Strauss	1975	1,439	72	.047
Marquardt & McGann	1975	91	1	.255
Marsh & Overall	1979	183	22	.168
Maslow & Zimmerman	1956	86	2	.792
McCullagh & Roy	1975	52	2	.045
McDaniel & Feldhusen	1970	76	14	.057
McGrath	1962	50	1	.299
Michalak & Friedrich	1981	125	12	.368
Plant & Sawrey	1970	32	1	.000
Ramsden & Moses	1992	869	18	-.094
Ratz	1975	75	1	.000
Riech & Rosch	1988	20	5	.517
Riley & Ryan	1950	382	3	.196
Rosanna	1977	244	6	.378
Rossman	1976	122	3	.189
Rothman & Preshaw	1975	25	4	.341
Rushton & Murray	1983	52	2	-.072
Siegfried & White	1973	45	4	.030
Stallings & Singhal	1970	248	4	.170
Stavridis	1972	32	11	.171
Tanner and Manakyan	1992	182	4	-.030
Teague	1981	25	4	-.008
Usher	1967	26	1	.000

Voeks	1962	193	11	.120
Volkwein & Carbone	1991	27	5	.050
Wood & DeLorme	1976	69	1	.412
Wood	1978	23	4	-.008

Braxton (1996): Table 1. Overall Support for the Three Contrasting Perspectives on the Relationship Between Teaching and Research

	Perspective Supported	R
Null	Aleamoni & Yimer (1973)	.002
	Bausell & Magoon (1972)	.07
	Braunstein & Benston (1973)	.04
	Centra (1983, 1)	.07
	Dent & Lewis (1976)	.02
	Frey (1976)	.07
	Friedrich & Michalak (1983)	.18[a]
	Hoyt & Spangler (1976)	.17[a]
	Linsky & Strauss (1975)	.01
	McCullagh & Roy (1975)	.05
	McDaniel & Feldhusen (1970)	.04
	Rushton & Murray (1983)	-.07
	Siegfried & White (1973)	.03
	Stallings & Singhal (1970, 2)	.11[a]
	Stavridis (1972)	.24[a]
	Usher (1966)	.23[a]
	Voeks (1962)	no relationship
	Wood (1978)	-.07
Conflict	Hoffman (1984)	-.25
Complimentarity	Bresler (1968)	.23
	Centra (1983, 2)	.10
	Clark (1973)	.30
	Faia (1976)	.11
	Freedman & Stumpf (1979)	.23
	Harry & Goldner (1972)	.19
	Hicks (1974)	.25
	Marquardt & McGann (1975)	.25
	Marsh & Overall (1978)	.14
	Stallings & Sinhal (1970, 1)	.26
	Wood & DeLorme (1976)	.39

[a] Not statistically significant ($p < .05$)

CORRELATION STUDIES

RESEARCH PRODUCTIVITY AND TEACHING EFFECTIVENESS
*Research in Higher Education, **18**(2), 1983*

John A. Centra

Centra seeks to investigate the teaching-research relationship by using two samples, one of 2,973 academics and the other of 1,623 from a variety of U.S. institutions. The study is included by Hattie & Marsh in their 1996 meta-analysis. The sample size is the second largest of all studies included by Hattie & Marsh.

In his review of past research Centra cites several U.S. studies from the 1960's and 1970's including Linsky and Strauss (1975) which indicate that teaching and research are independent functions with performance in one unrelated to the other. Centra also quotes several studies, which find modest associations between research productivity and student ratings of teacher effectiveness. Almost all the studies cited by Centra are correlation studies. In the light of the existence of analyses, which indicate that many good teachers are also good researchers Centra attempts to explain the reasons behind such a hypothesis. His arguments are as follows:

1) Teaching effectiveness and research productivity are both likely to be affected by the general ability and energy levels that individual faculty members possess.
2) A spill over effect: Research could influence teaching when the excitement and involvement of research is communicated to students and they are able to see knowledge as a constantly growing thing.
3) Participation in research could help maintain the faculty member's interest in the subject matter
4) Teaching might spill over into research when stimulating discussions with students lead to productive avenues of research.

The arguments of spill over in both directions, and the general ability argument have been stated before by Linsky & Strauss (1975). The realisation of the link between teaching and research as a two-way street in the final point is often ignored in other studies. Gruner (1995) has explored this avenue in greater detail.

Centra's study seeks to investigate the teaching-research relationship further by considering faculty members at different career stages and in different academic fields. Centra's argument is that the failure of many past studies to reveal a relationship between teaching and research may in fact reflect inadequate design rather than the independence of the two roles. By including faculty members from a variety of academic fields in a single analysis, it is possible other studies may have minimised significant relationships. Centra also hypothesises that teaching and research would be more likely to be significantly related for faculty members in their middle or later years. This is because while younger faculty members may concentrate on research at the expense of teaching in an effort to improve their chance of winning tenure, older faculty members who publish may be more effective teachers because of the spill over effects discussed earlier.

Two samples of faculty members were studied in order to test these hypotheses. Faculty members covered by the sample had administered the Student Instructional Report (SIR) in one of their courses. This provided the basis for a measure of teaching performance. For both samples the measure of research productivity used was the self- reported number of publications during the most recent five-year period. The first sample, which was administered from 1976 to 1978, consisted of 2,973 faculty members from 61 four- year institutions. Many of these institutions did not put a heavy emphasis on research. The second sample consisted of 1,623 faculty

members at 10 four-year colleges and institutions who had administered the SIR in their courses in 1979 or 1980. An effort was made to include institutions with more emphasis on research productivity in the second sample. As a group the faculty members in the second sample averaged more publications during the five-year period than did the previous sample (2.5 vs. 1.7). Faculty members were classified into the social sciences, natural sciences, or humanities. The second sample also included teachers of professional areas (this included teachers mainly of engineering and business education). In addition teachers were grouped according to the number of years of their teaching experience.

Student ratings of teaching effectiveness were gathered from the Student Instructional Reports in both samples. Two global rating items on the SIR, the value of the course to the students and the overall effectiveness of the teacher were analysed in both samples. In addition for the first sample six additional factors of teaching performance were also analysed. These student ratings of instruction were correlated with the self-reported number of publications over the most recent five years for each faculty member. The results of the first sample are included in table 1, while table 2 exhibits the results of the second sample.

Looking at the results of both analyses teachers of social science courses were the only group for which there were consistent significant relationships between the number of published articles and student ratings of instructor effectiveness or course value, the two global student ratings. Even these are not particularly high, ranging from 0.15 to 0.29. For natural science teachers the correlations were either insignificant or negative. While for teachers of humanities and professional courses mostly the correlations were insignificant. Centra explains that 49% of the faculty members in the first sample and 44% in the second sample did not publish any articles in 5 years. He hypothesises that it is possible that this restricted the range in publication rates enough to depress the correlations found.

The hypothesis that teaching and research would more likely be correlated for faculty members in their middle or later years was generally not supported with the exception of social science teachers in sample 1. Centra concludes therefore the career stage of teachers does not appear to be an important factor in the teaching research relationship.

On the other hand subject field differences seem to be critical in the study. Centra argues that it is possible that the spill over effect of research and teaching, or the general ability factor could in part account for the correlations for social science teachers. This though does not explain why this happens in this discipline and not in others. Centra reasons it could be possible that time spent on research would lead one to expect the low or negative correlations found for natural science teachers in the two samples of the study. This of course would only be a valid argument if time spent on research for natural science faculty was significantly more than that spent by faculty members in other fields. Euwals & Ward (2000) do find for a sample of academics from 5 British universities that the number of published articles was on average substantially higher for Science faculty than Arts and Social Science colleagues.

Centra concludes that student ratings of teaching are unrelated or only modestly related to research productivity. In the light of these results he argues that while no one would question the need for teachers to keep up with current knowledge in their fields, whether they must actually carry on research in order to do this is questionable.

Two points to be noted about correlation studies of this nature. Firstly even if the correlation coefficient between the factors analysed is zero it does not mean that the two variables are independent, i.e. zero correlation does not necessarily imply independence. And this leads directly to the second point, which is that correlation in fact is a measure of linear association or linear dependence only. It has no meaning for describing non-linear relations. For example it would not pick up any non-linear effect of teaching on research productivity or vice versa as found by Mitchell *et al* (1995). Finally, though it is a measure of linear association between two variables correlation does not imply any cause and effect relationship.

References:

Euwals, R and Ward, M (2000) What matters most: teaching or research? Empirical evidence on the remuneration of British academics, *Centre for Economic Policy Research Discussion Paper no. 2628*, London

Gruner, C R (1995) The teaching/research symbiosis: a two-way street, paper presented at the Southern States Communication Conference, New Orleans, April

Hattie, J and Marsh, H W (1996) The relationship between research and teaching: a meta-analysis, *Review of Educational Research*, **66** (4), pp 507-42

Linsky, A S and Strauss, M A (1975) Student evaluations, research productivity, and eminence of college faculty, *Journal of Higher Education,,* **46** (1), pp 89- 102

Mitchell, J E and Rebne, D S (1995) Nonlinear effects of teaching and consulting on academic research productivity, *Socio-Economic Planning Sciences*, **29** (1), 47-57

TEACHING AND RESEARCH: RAPPORT OR MESALLIANCE
*Research in Higher Education, **4**, 1976*

Michael A. Faia

Faia attempts to answer the question: are teaching and research mutually supportive or do they tend to detract from one another? The sample size of 53,034 U.S. faculty members used in the study is the largest data set employed in past research of this kind (this has been wrongly identified as 265,682 by Hattie and Marsh in their 1996 meta-analysis).

Faia considers two positions taken by previous researchers. One, that teaching and research tend to be incompatible, the other that teaching and research are mutually reinforcing, or complementary. He argues that it is of course possible that both sides of the argument are correct and that the incompatibilities and complementarities may tend to cancel one another out in which case we would often observe no significant relationship between teaching skill and research productivity.

In his review of past literature Faia lists 11 empirical studies of teaching-research linkages published been 1952 and 1975. He finds that most of these are of limited scope with 8 studies having gathered data from a single institution, one study using two institutions, and one study using a sample of 16 colleges and universities. Out of the listed studies 4 show no relationship between teaching and research proficiency, whereas almost all the others show weak positive associations. He believes that this indicates that complementarity can exist between teaching and research but that conditions may arise that tend to reduce it.

Faia makes a meaningful observation on the matter of expectations which he considers to be of substantial importance. This arises from Linsky and Strauss's (1975) contention that studies of teaching skill using peer evaluation are inherently less trustworthy than those using student evaluations, because colleagues are likely to be influenced in judging teaching ability by their knowledge of a professor's research record. Faia argues to the contrary that it is conceivable that students' awareness of the scholarly eminence of a faculty member might lead them to entertain unexpectedly high expectations regarding teaching performance. And therefore their evaluations could very well be "contaminated" too.

In his discussion of explaining a positive relationship between teaching and research Faia focuses on Linsky and Strauss's reasoning that spill over effects may exist between the two activities in both directions. Faia garners this as involving role complementarity, i.e. the notion that roles may be mutually reinforcing. He goes on to state that a complementarity is said to exist to the degree that roles are similar (this is not necessarily true). He proceeds to examine several different dimensions of the two roles of research and teaching to determine the degree of similarity between them. As a caveat he notes that time is a limited resource and speculates that the greater the disparity between relative amounts of time spent on the two activities the less the complementarity. Some aspect of similarity/dissimilarity among roles he examines are as follows:

1) The area of specialisation. Faia argues that academics teach and conduct research in highly specialised realms, and that they tend to confine both activities to the same specialities hence complementarity exists. But that it is likely that research tends to be much more specialised than teaching and this would actually produce disparities between research content and teaching content.

2) The nature of the values around which the roles are organised. As a similarity both teaching and research have the goal of furthering knowledge. But whereas research emphasises the creation of new knowledge, teaching emphasises the diffusion of established knowledge. Recognising this as an oversimplification he gives the example of an academic who confines his teaching to a consideration only of established knowledge, in which case his own research would contribute almost nothing to role complementarity. And he also states the opposite extreme, i.e. an academic teaching nothing but his own research, with the same consequence for complementarity.
3) Looking at incentive structures around the two roles Faia states that if such reward disparities exist at institutions where research is emphasised they generate a special form of compensatory selection. He hypothesis that a lack of complementarity may be due to the existence of reward disparities. And that if a lack of complementarity is an undesirable circumstance then one of the most effective means of mitigating that circumstance would be to reduce reward disparities.

A final interesting point Faia makes in his hypothesis is that it could turn out that the degree of role complementarity between teaching and research is influenced primarily by the sort of philosophies academics hold about the incompatibility or complementarity of the goals of knowledge creation and knowledge diffusion. In other words, he says that we could be in the presence of a self-fulfilling prophecy. A review of the research in the area shows, however, that beliefs among academics that research and teaching are strongly linked are not borne out by student perceptions of teaching linked with research productivity. If it were the case that the degree of role complementarity depended on the sort of philosophy that academics held about the compatibility of research and teaching, we would expect their views to be more consistent with what the wider literature says on the existence of such a link. Unless of course if the link has not been measured robustly.

Faia's study is based on data collected in a national survey during the 1972-73 academic year by the American Council on Education. As mentioned before the total sample covers 53,034 academics which makes it the largest sample used in this sort of a study. The measure of teaching proficiency used is based on the question: "Have you ever received an award for outstanding teaching?" Faia makes an interesting case for the use of this measure of teaching skill. His argument is that teaching awards often involve collaborative decisions made by student, faculty, and administrators. And so therefore it is safe to assume that faculty members who have received a teaching award have satisfied two or three constituencies. Almost 17.5% of all faculty members in the sample had received such a teaching award.

The measure of research productivity used is the number of self-reported publications or acceptances for publications in the preceding two years. A shortcoming of the measure used is that it provides no way of assigning different weights to articles, books, edited volumes etc. Also Brew and Boud (1995) would argue that this measure confounds quantity and quality. And that ratings of research productivity by citation count would come closest to a true measure of research quality.

The findings of the study are shown in table 1. The results are grouped by institutions with a weak research emphasis (these include comprehensive universities and colleges not offering doctorate degrees, liberal art colleges, and two-year institutes), and institutions strong on research emphasis (research universities and doctoral granting universities).

In institutions classified as weak on research emphasis faculty members who publish extensively are twice as likely as non-publishers to have received teaching awards. At the more research oriented institutions a similar but weaker relationship exists. Faia explains this as indicating that while a lack of complementarity cannot be tied directly to the existence of reward disparities, it tends to occur at institutions where, presumably, there are large reward disparities (institutions where research productivity is more generously rewarded than good teaching). His overall conclusion is that teaching and research tend to be mutually supportive especially at schools where research is not emphasised.

Faia's statistical analysis is simplistic. It is obvious that some academics may have teaching awards simply by virtue of having been in the profession for a longer period of time. It is not certain how having received a teaching award 10 or 15 years ago would relate to a publication in the past two years. Also if teaching awards were more prevalent a few decades prior to the study it would skew the receipt of teaching awards toward older faculty members. It is possible that if younger appointees concentrate on research in an effort to improve their chances of tenure the study would be underestimating the relationship between the two roles. Of course the contrary argument could also be true where teaching awards could be more prevalent more recently in which case Faia would be overestimating the link. This measurement of teaching through an outstanding teaching award throughout a lifetime of academia is greatly problematic. The problem is confounded by trying to establish a link over all disciplines and without attempting to control for the stage of the career of the responding faculty members. While the sample size is impressive in its range it is questionable whether the simple associations obtained in the analysis are of any value at all toward shedding light on the linkage between teaching and research.

References:

Brew, A and Boud, D (1995) Teaching and research: establishing the vital link with learning, *Higher Education*, **29**, pp 261-73

Hattie, J and Marsh, H W (1996) The relationship between research and teaching: a meta-analysis, *Review of Educational Research*, **66** (4), pp 507-42

Linsky, A S and Strauss, M A (1975) Student evaluations, research productivity, and eminence of college faculty, *Journal of Higher Education,* **46** (1), pp 89- 102

CONSTRUCT VALIDITY OF MULTIPLE MEASURES IN TEACHING, RESEARCH, AND SERVICE AND RELIABILITY OF PEER RATINGS

*Journal of Educational Psychology, **82** (2), 1990*

John F. Kremer

Kremer determines the construct validity of measures in teaching, research and service using a multitrait-multimethod matrix, employing data from the Science faculty of a large U.S. university. Though it is not the study's aim to determine empirically the link between research and teaching, this is a byproduct of the analysis. Kremer also investigates the reliability of peer evaluations.

Construct validity seeks agreement between a theoretical concept and a specific measuring device or procedure. This can be thought of as a labelling issue. In this case when "research" (or teaching or service) is measured, the question is whether this actually what is being measured.

Kremer obtained data for all full-time, tenure track faculty in the School of Science of a large mid-western university. The sample covered 89 faculty members representing seven departments: biology, chemistry, computer information sciences, geology, mathematics, physics, and psychology. The number of publications, total grant money, number of awards, and number of service contributions for the past 3 years were obtained from the school's annual reports. For each of the first three measures, a listing was classified as being either research, teaching, or service. An overall index of student satisfaction for each faculty member was calculated, based on student evaluations for the 3 years corresponding to data collected from the annual report.

Peer raters were selected with the following attributes: knowledgeable about evaluation criteria; experience of making important personnel decisions; employed full-time at the school at least 3 years; published at least 1 refereed article within the past 3 years; had above average student ratings in teaching; participated on at least one school committee within the past 3 years; and had been the member of at least one promotion and tenure committee. Confidentiality was ensured. Two or three members of each department who met these criteria were selected. Kremer speculates that these characteristics would increase the reliability of peer ratings. This may be true, and perhaps similar rationale should be employed when using peer evaluations as measures of teaching or research effectiveness to improve reliability. However, the results on reliability obtained by Kremer after this selection process cannot be generalized to <u>all</u> peer evaluations.

Each peer rater evaluated faculty members in his or her department answering the following question:

The overall contribution of the faculty member over the last several years in the area of _____ (teaching, research, or service) has been...

1 2 3 4 5

Minimal Outstanding

Please use perception only. Do not look up any information (e.g. annual review) and do not consider any other point of view (e.g. administration).

Peer raters were also asked to do a rating of how confident they were in each of their judgements, on a scale from 1 to 3.

The tool employed for assessing construct validity by Kremer is a multitrait-multimethod matrix (MTMM). This is simply a matrix or table of correlations arranged to facilitate the interpretation of the assessment of construct validity. In order to argue that the measures have construct validity under the MTMM approach, both convergent and discriminant validity has to be demonstrated. Convergent validity is the principle that measures of theoretically similar constructs should be highly intercorrelated. Discriminant validity is the principle that measures of theoretically different constructs should not correlate highly with each other. In the one matrix it is possible to examine both convergent and discriminant validity simultaneously.

Table 1 presents the results. The figure shows three concepts (research, teaching, and service) each of which is measured with different methods. Evidence for convergent validity of a variable is demonstrated by significant correlations between that variable and other methods of measuring the same trait. Discriminant validity has two criteria:

1) the average convergent validity for a variable should be greater than the correlations between that variable and any other variable having neither trait nor method in common,
2) the convergent validity correlations for a variable should be greater than the correlations between that variable and other variables having the same method but measuring different traits

Kremer finds that the data provides strong evidence for the construct validity of all the research methods. There is also moderate support for the construct validity of three teaching measures: peer ratings, student evaluations, and teaching awards. But there is no evidence for the construct validity of the service method.

For each department in the study there were two or three peer raters. Within each area of research, teaching, and service all inter-rater correlations were combined to get an average inter-rater reliability. Kremer finds across all raters, the reliability for research was substantially higher than for teaching and service. Peers were also more confident in their research rating (60% of the ratings) than service (43%) or teaching (33%).

The study provides data on the correlation between several measures of research and teaching. Interestingly student evaluations are seen to be positively correlated with all the measures of research (peer rating, publication, grants, awards) used in the study. Teaching awards and peer ratings of teaching are also observed to be positively correlated with peer ratings of research. The highest of any of these associations though is a modest 0.22. Modest negative associations are apparent between teaching grants and all measures of research. The same negative relationship is evident for pedagogic publications.

These findings are similar to those observed in many other correlation studies (Centra: 1983, Linsky & Strauss: 1975). In common with such studies it is doubtful these results reveal the undistorted magnitude of the linkage between teaching and research.

Correlation measures a relation between two variables only to the extent to which it is linear. Deviations from linearity will curb the correlation coefficient even if there is a very close relationship between the variables. Causal relations cannot be proved based on correlation coefficients. The associations being identified through correlations may in fact be spurious, i.e. correlations may be due mostly to the

influence of other variables. It is possible through partial correlations to control for the influence of variables, which we believe influence the relationship.

Apart from the problems inherent with correlation analyses (examined in more detail in the reviews of several of the other correlation studies), the range of this study, one school in one institution, is not sufficient for the results to be generalised. Also it has been demonstrated by Centra (1983), who finds higher correlation between research and teaching for Social Science faculty than other departments, among others that discipline is seen as an important factor in the relationship. Results from a Science faculty cannot be generalised for all disciplines.

Of special interest is the correlation between peer ratings of teaching and student evaluations of teaching. The correlation found (0.57) is fairly high. This would seem to give some credibility to using either measure as a gauge of teaching performance. It should be noted though that the peer raters in this study were selected according to a number of attributes which are deemed to increase their reliability. Also only a small proportion (33%) of peers were confident in rating teaching.

References:

Centra, J A (1983) Research productivity and teaching effectiveness, *Research in Higher Education,* **18**(2), pp 379-89

Linsky, A S and Strauss, M A (1975) Student evaluations, research productivity, and eminence of college faculty, *Journal of Higher Education,* **46** (1), pp 89- 102

IDENTIFYING FACULTY TYPES USING PEER RATINGS OF TEACHING, RESEARCH, AND SERVICE

*Research in Higher Education, **32** (4), 1991*

John Kremer

Kremer investigates the relationship between teaching, research, and service by classing academics into types based on peer ratings in each of these three areas. His data set of 90 full-time, tenure track faculty from the School of Science of a major U.S. mid-western university, is the same used in Kremer (1990). Further description of the data can be found in the review of that paper.

Kremer's approach is unique. He postulates that faculty can potentially be divided into eight types (see table 1). He then focuses on identifying the dominant types of faculty and the percentage of faculty in each type. A strong positive relationship between teaching and research suggests four faculty kinds:

Those who are good researchers and good teachers
- All Stars
- Researchers and teachers

Those who are not good researchers and not good teachers
- Good Citizens
- Uninvolved

A strong negative relationship also suggests four kinds:

Those who are good teachers but not good researchers
- Teachers and Good Citizens
- Teachers

Those who are good researchers but not good teachers
- Researchers and Good Citizens
- Researchers

A non-significant relationship between teaching and research effectiveness would indicate that faculty can be classified into all eight of these categories.

In the study, faculty types were determined by cluster analysis with peer ratings of performance in teaching, research, and service as classification variables. Proximity was calculated using squared Euclidean dissimilarity coefficients. Two cluster analyses were performed. For the first analysis unweighted pair group average was used as the linkage rule. To validate the findings from this method, a second analysis where variables were grouped using the Ward's method was performed. The two cluster analyses yielded similar results. A five-group classification solution was indicated by a fourfold increase in the error term occurring between the fifth and sixth clusters.

The faculty types developed from the cluster analysis and their comparison with additional variables is presented in table 2. Examining the question of the linkage between teaching and research using information about types yields the following results: The statement that the best researchers are also the best teachers (or poor teachers are also poor researchers) described 31 percent of the faculty (All Stars and Uninvolved). On the other hand, the principle that the best researchers are not the best teachers fit 69 percent of the faculty (Teachers, Teachers and Good Citizens, and Researchers).

Kremer's methodology is fascinating. The approach has not been attempted in any other study in the area. His hypothesis of types and how these would indicate the existence of a positive, negative, or no relationship between teaching and research seem sound. His choice of using Euclidean dissimilarity coefficients to calculate proximity has the advantage of using level, as well as shape and scatter of the data, as the basis for clustering. Both methods he uses for cluster analysis, the unweighted pair group average and Ward's method, are regarded as very efficient linkage rules. He reports they indicate a 94 percent agreement, which further supports the validity of the classification Kremer finds. The main problem in the analysis appears to be the use of peer ratings to measure research, teaching, and service. This is especially dubious in the case of teaching, where only 33% of peer raters report confidence in their rating (Kremer:1990). Examining table 2 it is apparent that those classified as Researchers also had high student evaluations of classroom teaching. Considering this it is likely had Kremer used this measure (student evaluations) of teaching he would have found greater support for the hypothesis that good researchers are also good teachers. Another primary limitation of the study is that it is limited to one faculty (Science) in one university. It does not address cross-site, cross-discipline variability.

Also of interest is the final column in the table, the percent increase in salary for each type of faculty over the preceding three years. All Stars and Researchers are observed to have received the highest increases, averaging 33 percent across the three-year period. While this suggests that there is reward disparity between the activities of research and teaching, part of this distinction could be explained by the flawed use of peer ratings as a measure of teaching as described above. In any case the results are different to those observed by Euwals and Ward (2000) who find greater dependence of salary increases on teaching skill than research publications in their analysis of data from 5 British universities.

References:

Euwals, R and Ward, M (2000) What matters most: teaching or research? Empirical evidence on the remuneration of British academics, *Centre for Economic Policy Research Discussion Paper no. 2628*, London

Kremer, J F (1990) Construct validity of multiple measures in teaching, research, and service and reliability of peer ratings, *Journal of Educational Psychology*, **82** (2), pp 213-18

STUDENT EVALUATIONS, RESEARCH PRODUCTIVITY, AND EMINENCE OF COLLEGE FACULTY

*Journal of Higher Education, **46**, 1975*

Arnold S. Linsky and Murray A. Strauss

This study determines empirically the relationship of academics' involvement in research to their classroom performance for a sample of 16 U.S. colleges and universities. The sample size of 1439 academics is the third largest among all studies identified by Hattie and Marsh in their 1996 meta-analysis.

Linsky and Strauss start by examining the rationale for either a positive or negative correlation between teaching and research, or no correlation at all. The arguments they present supporting each of these three positions are as below:

Positive correlation

1) One of the two activities directly influences the other. In other words a spill over effect occurs. The excitement of engaging with the development of the knowledge base of the discipline may be communicated to students and can contribute to their learning. In the other direction a spill over effect might occur if stimulating teaching could produce good feedback from students which could lead to productive lines of research.
2) The abilities that lead to excellence in research are also likely to lead to excellence in other spheres of academic activity.

Negative correlation

1) The two roles may interfere with each other. There is limited time and energy for faculty to do both teaching and research.
2) It is also possible that teaching and research require contrary personality characteristics unlikely to be found in the same person. For example teaching success may depend on attributes such as gregariousness that might tend to be inversely correlated with attributes associated with research success such as intellectuality.

No correlation

1) The skills involved are in fact randomly distributed among college professors making any one combination as likely as any other.

In their review of past literature Linsky and Strauss uncover 9 empirical studies. They find that one is grossly defective in design. A methodological problem they find arising in four of the other studies is the use of peer evaluation for measures of either teaching performance or indications of research performance. These studies in general find positive correlations between the two roles. Linsky and Strauss doubt the robustness of these results noting that these relationships may be spurious since they are based on the judgements of colleagues and administrators. They consider that measures of teaching performance especially are likely to be contaminated by colleagues knowledge of the professor's research record. Linsky and Strauss consider the other four studies in their review adequate in meeting minimal technical requirements in regard to data and methodology. The findings of these studies differ with two showing small positive statistically significant relationships between research productivity and teaching competence (Bresler: 1968, Stallings & Singhal: 1970), one study showing no relationship between the two measure (Voeks: 1962), and one showing a small negative association (McDaniel & Feldhusen: 1970).

Linsky and Strauss's study is based on data from 16 colleges and universities. The schools in the sample are Boston College, Columbia College, Barnard, McGill University, University of Southern California, Yale, University of Minnesota, SUNY, University of New Mexico, UCLA, University of Texas (Austin), Ohio State University, University of Washington, University of Utah and University of New Hampshire. This sample includes various types of colleges and universities including large and small, public and private, prestigious and less distinguished institutions. However, the sample as a whole is over representative of larger and better known universities.

As a measure of teaching performance, student course evaluations are used. Since the ratings were of faculty from 16 different colleges, the authors convert all teaching scores to a common unit of measurement carrying out a z-score standardisation for each college. Two measures are used for involvement in research.

1) A publication score was computed based on a weighted summary score for articles and books written over an approximately 20-year period. This score was limited to 8 disciplines: one applied field, engineering; one physical science, physics; one biological science, biology; two humanities, English and philosophy; and three social sciences, anthropology, sociology, and psychology.
2) A citation score was based on the number of times a scholar's work was cited by others over a ten-year period. This is felt to reflect scholarly impact or research quality rather than output. The data for this were obtained from the Science Citation Index. Since English and philosophy are not covered by the index, therefore these are omitted from the citation scores.

The correlations of student evaluations of teaching to a number of research productivity variables are given in table 1. Overall measure of teaching fail to show much association with any of the components of research involvement. The rating for instructor's knowledge does seem to be correlated with the indices of research performance. Publications are also consistently correlated with student evaluations of course content. Linsky and Strauss think the two variables "course content" and "instructor's knowledge" can be grouped conceptually under "intellectual competence" in the teaching role. They consider therefore there is certain face value validity in the correlation of these items with research performance. None of the other indices of teaching performance are consistently correlated with any of the measures of research performance. Surprisingly citation scores are negatively (and significantly) correlated with four items of teaching performance. These associations though, are quite small. Although some of the correlations in the table indicate meaningful patterns of association, the absolute value of all of these correlations is generally low. Considering this Linsky and Strauss conclude that there is little correlation between teaching and research roles as measured in their analysis.

In a further section of the paper the authors consider several other factors which may be associated with teaching evaluations. These results are displayed in table 2. The correlations indicate that student ratings of faculty tend to be somewhat higher in more advanced courses than at lower levels. Enrolment or class size has a negative association with overall teaching rating. While academic rank is uncorrelated with overall teaching score, students evidently attribute higher degree of knowledge to higher ranking professors. Whether this represents a "halo" effect created by the knowledge of the professor's rank or an unbiased judgement on the part of the students cannot be ascertained from the analysis. Teaching performance appears to decline over the course of an academic's career as indicated by the negative correlation between teacher rating and PhD years. Linsky and Strauss remark however that the low correlation masks a higher but curvilinear relationship.

The second section on associations between teaching performance and other variables illustrates why correlation studies of this nature cannot provide adequately robust results on the linkage between the two roles. A methodology based entirely on zero-order correlations does not consider the effects of other variables which could affect the relationship. For example Linsky and Strauss demonstrate that enrolment relates oppositely to overall teaching ratings. It is only by statistically removing the effect of this variable (and other significant variables that influence the linkage) that a true value of the basic relationship between teaching and research can be obtained. A study based entirely on zero-order correlations cannot manage this. Linsky and Strauss suggest that some of their data indicates there may be an inverted U curve relationship between classroom performance and academic experience (years since PhD). A correlation analysis cannot decipher this curvilinear association. Apart from the problems inherent in the use of a correlation study, Linsky and Strauss fail to control the period of publications and citation scores beyond 20 and 10 years respectively. It is obvious that some academics will have high publication scores simply by virtue of having been in the profession for a longer period of time. In disregarding this problem, the measures of research performance used fail to accurately act as a gauge for research quality. In this case where it is very likely that the measures of research quality and quantity are corrupted, it is uncertain whether the associations obtained are of any substance, even as simple correlations.

References:

Bresler, J B (1968) Teaching effectiveness and government awards, *Science*, **160**, pp 164-68

Hattie, J and Marsh, H W (1996) The relationship between research and teaching: a meta-analysis, *Review of Educational Research*, **66** (4), pp 507-42

McDaniel, E D and Feldhusen, J F (1970) Relationships between faculty ratings and indexes of service and scholarship, Proceedings of the 78th American Psychological Association Annual Convention, Washington

Stallings, W M and Singhal, S (1970) Some Observations on the relationships between research productivity and student evaluations of courses and teaching, *American Sociologist*, **4**, pp 1-4

Voeks, V W (1962) Publications and teaching effectiveness, *Journal of Higher Education*, **33**, pp 212-18

RESEARCH PRODUCTIVITY AND PERCEIVED TEACHING EFFECTIVENESS: A SURVEY OF ECONOMICS FACULTY

*Research in Higher Education, **37** (3), 1996*

Thomas C. Noser
Herman Manakyan
John R. Tanner

This study examines the relationship between research involvement and teaching productivity for a sample of U.S. faculty, to shed light on the question of whether performance in one area enhances performance in the other. Additionally, the study examines the impact of a variety of individual and institutional characteristics which may influence research output as well as teaching performance. Because of the focus on a single field of study it is possible that the results obtained may be specific to the particular discipline of economics.

The sample of economics faculty members studied came from a mail questionnaire sent out in 1992 to 1000 economics academics randomly selected from the membership of the American Economic Association (AEA). The response rate was 34.4%. Noser *et al* while admitting the possible existence of non-response bias do not attempt any control, arguing they did not have knowledge of the characteristics of the entire economics faculty population. They could, however, have at least corrected for gender, from the AEA academic membership list, and institutional type.

Noser *et al* use student course evaluations to gauge teaching performance. The two measures used are respondents' average score over all courses taught in the last two terms:

- on the question that addresses overall teaching effectiveness
- and on the average score on all questions on evaluation forms.

The authors note that student evaluation scores have been found to be related to a number of factors including organisation of the presentation, relevancy of the material, and interpersonal skills of the instructor. Additionally, they have been found to be positively related to expected grade in the course and biased by the nationality of the instructor (Kassaye: 1984). While admitting that there are questions about the validity of student evaluations in judging teaching performance, Noser *et al* argue that they are relatively easy to administer, and produce quantifiable results.

Two measures of research are also used in the study, both constructed from the number of publications in different categories of journals, presentations at national and regional meets, and books. Two research scores that are standardised:

- the first by dividing the sum of all publications by the length of the respondents' career
- and the second by dividing a weighted value by the length of the respondents' career

are computed by the research output described. By including the second method the study includes a measure of research quality (distinct from research quantity which the first stated measure gauges).

The relationship of these research measures with teaching is problematic. The teaching measures gauge teaching performance over the preceding two terms. Meanwhile the research measures are a lifetime average. It is likely that research

activity changes over the stage of an academic's career. Not controlling for stage of career while using the measure complicates the problem. It is not evident what the association between the two measures would actually mean for the debate on the link between research and teaching.

In common with many other studies in the area, Noser *et al*'s analysis is based on the correlation between activities. Their results are presented in table 1. Only a marginal relationship is detected between research activity and teaching effectiveness. Surprisingly modest positive associations are detected for those faculty with primary teaching responsibilities at the undergraduate level, but not at a higher level. This is in conflict with the widely held view that the teaching-research nexus is stronger at postgraduate level. Studies of both the perception of academics (Smeby: 1998) and students (Jenkins *et al*: 2002) clearly indicate that the link is viewed to be stronger at higher levels.

In order to identify any institutional or individual characteristics that may affect research or teaching performance, Noser *et al* employ standard analysis of variation techniques. Their main findings from this process are listed below:

- faculty at universities of larger size have significantly higher research output
- the size of the economics program yields similar results
- faculty from major state institutions have higher research scores than faculty at regional universities
- faculty with 6-10 years of full-time experience have significantly higher research scores than more experienced faculty
- those with least teaching load have highest research output
- teaching effectiveness is seen to be significantly impacted only by one of these characteristics. For faculty with 11 to 15 years of experience, their teaching scores are significantly higher than their least experienced colleagues

Noser *et al* consider these findings to be consistent with the view that some factors are related to teaching effectiveness and others to research productivity, with little correlation between the two.

As noted earlier in the review there are several problems associated with the study. Especially in relation to the association between a two-term measure of teaching performance and a lifetime measure of research output.

The use of student evaluation scores across institutions without control is also questionable. The format of the instrument may differ from institution to institution, which could affect respondents' scoring on the evaluation sheet. Ideally some corrective method (e.g. z score) should have been used to standardise scores from each institution.

Common with other correlation studies there are serious problems inherent in using the process to measure the linkage between research and teaching. These have been discussed in some detail in the main section and the reviews of Linsky and Strauss (1975) and Centra (1983).

One final point to note from the paper's analysis of individual characteristics. An inverted U shaped relationship is evident between teaching effectiveness and years of experience. This backs up Linsky and Strauss's (1975) observation of a similar relationship.

References:

Centra, J A (1983) Research productivity and teaching effectiveness, *Research in Higher Education,* **18**(2), pp 379-89

Kassaye, W W (1984) Some sources of personal prejudice in student-instructor evaluations, *Journal of Financial Education*, **13**, pp 33-39

Linsky, A S and Strauss, M A (1975) Student evaluations, research productivity, and eminence of college faculty, *Journal of Higher Education,,* **46** (1), pp 89- 102

STUDENT AND DEPARTMENT CHAIRMAN VIEWS OF THE PERFORMANCE OF UNIVERSITY PROFESSORS

Journal of Applied Psychology, ***58*** *(2), 1973*

Daniel N Braunstein
George J Benston

Braunstein & Benston compared student course evaluations taught over two years by 347 professors to rankings made by department chairmen of their faculty at the University of Rochester. Among the ranking criteria, chairmen sorted faculty on the basis of current research and overall teaching performance. The relationship between the chairman evaluations of research and teaching performance was analysed through median rank-order correlations. The correlation was observed to be 0.54 one year, and 0.34 the next. The faculty evaluations made by departmental chairmen were related to evaluations of overall teaching effectiveness made by students for each faculty member. Correlations between chairmen ranking of research and student evaluations of teaching ranged from -.31 in the Management department to +.36 in the Engineering department. The other disciplines investigated had the following associations: Social Science -.05, Natural Science -.04, Humanities +.24. The average correlation found was +.04. This indicates a null relationship between teaching and research. It is possible a halo effect is responsible for the chairmen rankings of teaching performance and therefore the higher association observed in the median rank-order correlation. It is quite clear from these data that university departmental chairmen share only moderate agreement with students in evaluating teaching. It should be noted that that the results pertain to one institution only.

STATISTICAL TREATMENT OF DATA FROM STUDENT TEACHING EVALUATION QUESTIONNAIRES

Journal of Chemical Education, ***51*** *(3), 1974*

C D Cornwell

Cornwell examines student course and teaching evaluations in the discipline of Chemistry. One of the aims of the study is to examine the effect of several factors other than teaching skill of the lecturer in order to learn the extent to which such factors might systematically affect the evaluation results. One of the factors included in the study is research. Cornwell uses data for 70 different Chemistry lecturers from 20 institutions collected through an American Chemistry Society questionnaire. The results are likely to be discipline specific. The measure used to gauge research is to simply categorise academics into two groups: those who are actively engaged in research, and those who are not. Cornwell does not declare what percentage of faculty falls into either group. However, it is most likely that much of the sample were research active, Chemistry being a hard/pure area (see Moses: 1990). Given this, and the small sample size, it would be difficult to detect a relationship between teaching and research. As expected, Cornwell finds no association between the two.

VALIDITY OF THE COURSE-FACULTY INSTRUMENT (CFI): INTRINSIC AND EXTRINSIC VARIABLES

*Educational and Psychological Measurement, **39**, 1979*

Richard D Freedman
Stephen A Stumpf
Joseph C Aguanno

Freedman et al determine the construct validity of student ratings of instructors. Though it is not the study's aim to determine empirically the link between research and teaching, this is a byproduct of the analysis. The methodology is similar to that applied by Kremer (1990) and Kremer (1991). The data set used for the analysis covers 129 instructors at a large north-east U.S. university. An overall instructor-rating index based on student appraisals for each faculty member was calculated. This was based on a mean class rating across all classes taught by each instructor over one year. Research was defined through publication productivity. A weighted bibliographic count of the past three years of research publications was used, where the process applied was unit weighting across various categories of publication. The study provides data on the correlation between the measures of research and teaching. The association found – correlation coefficient of 0.23 – though larger than the overall 0.10-0.12 range suggested by the meta-analyses of the literature, is similar to that observed in several other studies (Kremer: 1990, Stallings: 1970). The range of this study, one school in one institution, is not sufficient for the results to be generalised.

SOME OBSERVATIONS ON THE RELATIONSHIPS BETWEEN RESEARCH PRODUCTIVITY AND STUDENT EVALUATIONS OF COURSES AND TEACHING

The American Sociologist, May 1970

William M Stallings
Sushila Singhal

Stallings & Singhal seek to investigate the teaching-research relationship by using two samples, one of 128 University of Illinois academics and the other of 121 Indiana University instructors. The investigators used a unit weighted bibliographic count as the criterion for research productivity. Teaching performance was measured through student evaluations. Both measures covered a period of one year. The research productivity index data was correlated with course level, academic rank, field of study (science or non-science) and course evaluation scores. With the research measure, course evaluation of instructor correlated .26, and total course evaluation score correlated .20 for the University of Illinois. The corresponding associations for Indiana University were observed to be .13 and .18 respectively. Stallings & Singhal also found that the research productivity index score increased with course level and academic rank, but was not related with course type. Further, the teaching evaluation scores were found to increase with academic rank. The authors computed first-order partial correlations between research productivity and teaching evaluation scores controlling for these factors (This is one of very few studies to do this). They found none of the zero-order correlations increased notably.

FACULTY EVALUATION: RELIABILITY OF PEER ASSESSMENTS OF RESEARCH, TEACHING, AND SERVICE

*Research in Higher Education, **26** (1), 1987*

Lawrence S Root

Root analyses the peer assessment of faculty performance for the determination of salary increases. While that central objective of the paper is distant from the aim of the current review, pursuing a side lane Root also undertakes an analysis to examine the relationship between peer assessments of performance in teaching and research. It has been noted that the robustness of findings based on peer evaluation is dubious. It is argued these relationships may be spurious since they are based on the judgements of colleagues and administrators and a halo effect is likely to be present. Based on a very limited sample (6 peer raters assessing each other in one U.S. institution) Root calculates the ratings both as interval and ordinal measures. He finds a weak positive association between the ratings of individual faculty members for teaching and research (Pearson's correlation coefficient: .185, Rank-order correlation coefficient: .130). Given the extremely small sample size it would be dangerous to infer anything from these results.

IS SCIENTIFIC ACHIEVEMENT A CORRELATE OF EFFECTIVE TEACHING PERFORMANCE

*Research in Higher Education, **3**, 1975*

A I Rothman
R Preshaw

The purpose of this study is to test the hypothesis that there would be a positive relationship between the scientific productivity of faculty members and their effectiveness as teachers in the context of a centre heavily committed to scientific activity. Two courses given by the Department of Physiology at a Canadian university were used to test the hypothesis. Both were at an advanced level and were given to third and fourth year students. 15 instructors gave lectures in the third-year class and 10 in the fourth-year class. An index of teaching effectiveness was derived from students' perceptions of teaching effectiveness. A citation index of scientific productivity used was derived from the Scientific Citation Index covering a six year period prior to the courses for which student evaluations of teaching effectiveness were collected. A simple count of number of publications was also collected. Correlations were calculated between the measures of teaching and research. The results indicate significant positive relationships between student ratings of teaching proficiency and the two indexes of research productivity (.55 correlation with citation score, and .60 with number of publications). Significant results were not observed for the third-year class. Rothman and Preshaw hypothesize the advanced level of the fourth-year class provided opportunities for the active researchers to utilize specialized knowledge in their teaching. Also it is observed that class size in the fourth-year was 39, while enrolment in the third year was 118. It could be the larger class size at the lower level prevented interaction between teachers and students. Thus, particularly inhibiting more intangible tacit aspects of knowledge and learning. The results of the study provide support to the notion that research and teaching are not separate and adversary phenomena each working to the disadvantage of the other, but under certain conditions are interrelated with research activity supportive of effective teaching. The restricted sample size and the single discipline examined limit the generalisability of the findings.

Centra (1983): Table 1. Sample 1: Correlations Between Number of Publications and Student Ratings of Instruction, by Years of Experience

Student Ratings of Instruction	2 Years or Less	3-6 Years	7-12 Years	13-20 Years	Over 20 years
Humanities	*N=157*	*N=280*	*N=327*	*N=288*	*N=204*
Instructor effectiveness	.09	.16**	.15**	.15**	.06
Course Value	.12	.07	.02	.16**	.02
Faculty/student interaction	.11	.10	.11*	.03	.03
Course organization and planning	.02	.16**	.07	.09	.02
Communication	.11	.10	-.01	.04	-.02
Textbooks and reading	.13	.13*	.02	.04	.00
Course difficulty and workload	-.09	-.13*	-.19**	-.23**	-.15*
Tests and exams	.13	.03	-.10	.09	.06
Natural Sciences	*N=93*	*N=126*	*N=136*	*N=111*	*N=43*
Instructor effectiveness	.14	-.15	-.06	.05	-.33**
Course Value	.18	-.09	-.06	-.03	-.19
Faculty/student interaction	.21*	-.20*	-.04	-.01	-.40**
Course organization and planning	.20*	-.17*	-.11	.00	-.49**
Communication	.09	-.13	-.16*	-.13	-.13
Textbooks and reading	-.10	-.22**	-.06	.08	-.24
Course difficulty and workload	-.04	.14	-.24**	.07	.15
Tests and exams	.24**	-.02	-.12	-.18*	-.06
Social Sciences	**N=188**	*N=349*	*N=340*	*N=172*	*N=154*
Instructor effectiveness	-.08	.09	.16**	.21**	.29**
Course Value	-.07	.07	.17**	.15*	.28**
Faculty/student interaction	-.09	.03	.17**	.22**	.16*
Course organization and planning	-.16*	.08	.16**	.01	.23**
Communication	-.10	.04	.11*	.09	.24**
Textbooks and reading	-.16*	.03	-.02	.04	-.06
Course difficulty and workload	-.01	-.12*	-.06	-.21**	-.23**
Tests and exams	-.02	.12*	.11*	.15*	.11

Statistically significant at the following levels: *p<.05 **p<.01

Centra (1983): Table 2: Sample 2: Correlations Between Number of Publications and Student Ratings of Instructor Effectiveness and Course Value, by Years of Experience

	Teaching Experience		
Student Ratings of Instruction	**6 Years or Less**	**7-12 Years**	**13 Years or More**
Humanities	*N=104*	*N=48*	*N=91*
Instructor effectiveness	.09	.09	-.09
Course Value	.14	.09	-.04
Natural Science	*N=107*	*N=94*	*N=122*
Instructor effectiveness	-.19*	-.03	.04
Course Value	-.28**	.03	.09
Social Science	*N=96*	*N=83*	*N=65*
Instructor effectiveness	.23**	.24**	.23*
Course Value	.25**	.19*	.22*
Professional Areas	*N=354*	*N=179*	*N=280*
Instructor effectiveness	.01	.17*	.13
Course Value	-.01	.07	.07

*Statistically significant at the following levels: *p<.05 **p<.01*

Faia (1976): Table 1. Receipt of Teaching Award by Recent Publication Rate and (Institutional) Research Emphasis, Full-Time Faculty Members, 1973 (In percentages)

	Publication Rate				
Receipt of teaching award	**None**	**1-2**	**3-4**	**5-10**	**More than 10**
Weak research emphasis					
No	83.3	77.7	75.9	80.1	68.9
Yes	16.7	22.3	24.1	19.9	31.1
Total	100	100	100	100	100
Weighted N	193,090	51,242	12,786	6,444	2,120
Strong research emphasis					
No	85.2	83.0	83.4	82.2	80.0
Yes	14.8	17.0	16.6	17.8	20.0
Total	100	100	100	100	100
Weighted N	69,292	58,628	40,684	29,636	8,739

Kremer (1990): Table 1. Multitrait-Multimethod Matrix for the Constructs of Research, Teaching, and Service

	Research				Teaching					Service				
Construct/method	PR	P	G	A	PR	P	G	A	SE	PR	P	G	A	ES
Research														
Peer rating (PR)	-													
Publication (P)	.72	-												
Grants (G)	.56	.48	-											
Awards (A)	.35	.24	.56	-										
Teaching														
Peer rating (PR)	.18	-.06	-.04	.01	-									
Publication (P)	-.25	-.21	-.12	-.04	.06	-								
Grants (G)	-.07	-.10	-.05	-.03	.13	.15	-							
Awards (A)	.12	-.01	-.09	.00	.44	.23	.48	-						
Student Evaluations (SE)	.22	,11	.08	.11	.57	.02	.06	.21	-					
Service														
Peer rating (PR)	.12	-.02	-.12	-.12	.31	-.08	.04	.23	.29	-				
Publication (P)	.01	.09	.05	.14	-.13	-.10	-.03	-.12	.02	-.18	-			
Grants (G)	.20	.06	-.07	-.01	.14	-.06	-.02	.14	.08	.16	-.05	-		
Awards (A)	-.23	-.18	-.09	-.05	.03	.08	-.02	.06	.19	.04	-.05	-.04	-	
External service (ES)	-.16	-.13	-.17	-.14	.23	.26	.18	.33	.20	.22	-.09	.11	.30	-

Kremer (1991): Table 1. Potential Faculty Types

	Area of High Performance		
Faculty Types	**Research**	**Teaching**	**Service**
All Stars	X	X	X
Researchers & Teachers	X	X	
Teachers & Good Citizens		X	X
Researchers & Good Citizens	X		X
Researchers	X		
Teachers		X	
Good Citizens			X
Uninvolved			

Kremer (1991): Table 2. Description of Five Faculty Types Developed from Cluster Analysis[x]

Variables for Cluster Analysis					Additional Variables						
	Peer Ratings[y]				Teaching		Research			Service Salary	
Faculty Types	N	Teaching	Research	Service	No. of awards	Stud. Evals.[y]	Grant Dollars[z]	No. of Pubs.	No. of Awards	No. of Awards	Percent Increase
All stars	11	3.94[c]	3.68[c]	3.94[c]	.60[b]	4.05[b]	7.60[a]	2.50[a]	.00[a]	3.60[b]	.33[b]
Teachers and Good Citizens	13	3.51[c]	1.37[a]	3.46[c]	.31[b]	3.86[b]	.09[a]	.46[a]	.00[a]	3.31[b]	.21[a]
Researchers	40	2.85[b]	3.81[c]	2.53[b]	.15[a]	3.76[b]	54.62[b]	4.92[b]	.19[b]	1.04[a]	.33[b]
Teachers	9	3.43[c]	2.24[b]	1.31[a]	.00[a]	3.81[b]	15.55[a]	1.45[a]	.00[a]	1.64[a]	.15[a]
Uninvolved	17	1.97[a]	1.54[a]	2.13[b]	.08[a]	.40[a]	.09[a]	1.00[a]	.00[a]	2.00[a]	.14[a]

For variable in each column, faculty types that have letters in common are not significantly different from each other. Faculty types that do not have letters in common are significantly different. [x]Data covers three years; [y]High = 5, Low = 1; [z]In thousands of dollars.

Linsky & Strauss (1975): Table 1: Correlation of Selected Indexes of Teaching and Research Performance

	Measures of Research Performance					
Indexes of Teaching Performance	**Total Publications**	**Articles**	**Ed. Books**	**Joint Books**	**Solo Books**	**Citations**
Overall teaching rating	.04 (1422)	.02 (1423)	.03 (1436)	-.01 (1439)	.05* (1439)	-.05 (760)
Course content	.12* (302)	.11* (302)	.18** (304)	.03 (303)	.10* (304)	.16* (159)
Instructor's personalization	-.05 (1046)	-.05 (1047)	.00 (1055)	-.03 (1059)	-.04 (1058)	-.07* (578)
Instructor's knowledge	.27** (333)	.22** (333)	.24** (332)	.09 (334)	.17** (334)	.07 (185)
Course coherence	.03 (512)	-.04 (512)	.13 (84)	.30** (84)	.20* (84)	.07 (53)
Instructor's interest in his subject	.07 (83)	-.04 (83)	.13 (84)	.30** (84)	.20* (84)	.07 (53)
Motivates students	.00 (943)	-.04 (943)	.06* (954)	.00 (955)	.05 (955)	-.11** (524)
Values of readings	.06* (1080)	.00 (1081)	.04 (1094)	.01 (1095)	.12** (1095)	-.02 (581)
Exam quality	.00 (974)	-.02 (975)	.00 (985)	.00 (989)	.01 (988)	-.07* (585)
Value of papers	.12* (264)	.07 (265)	-.01 (265)	.08 (266)	.16** (265)	-.06 (98)
Fairness	.02 (430)	.04 (430)	.05 (439)	.07 (440)	.06 (440)	-.07 (232)
Recommend course	.10* (308)	.17** (308)	.04 (306)	-.05 (308)	.07 (308)	.09 (170)

*Statistically significant at the following levels: * = p < .05 ** = p < .01*

Linsky & Strauss (1975) Table 2. Correlations between Teaching Ratings and Selected other Variables

Indexes of Teaching Performance	Course Level	Enrolment	Rank	Highest Degree	PhD Years	Carter Rating
Overall teaching rating	.09** (4646)	-.01** (4257)	.00 (3530)	-.06** (2758)	-.08** (1729)	-.04* (2172)
Course content	.13** (1126)	-.01 (882)	.09** (905)	.05 (332)	-.11** (424)	-.05 (594)
Instructor's personalization	.10** (3443)	-.22** (3015)	-.08** (2549)	-.09** (2351)	-.05* (1303)	-.05* (1740)
Instructor's knowledge	.17** (1180)	-.03 (1151)	.25** (1093)	-.07 (417)	-.04 (601)	-.02 (696)

*Statistically significant at the following levels: * = p < .05 ** = p < .01*

Noser et al (1996): Table 1. Research Output and Teaching Effectiveness Correlations

	Overall Teaching Effectiveness Score		
	Undergraduate	*Graduate*	*Combined*
TARS	0.1189**	-0.1151	0.0858*
WARS	0.0967*	-0.1261	0.0715
	Average on All Questions on Evaluation Form		
	Undergraduate	*Graduate*	*Combined*
TARS	0.0698	-0.0292	0.0450
WARS	0.0360	-0.0082	0.0332

**Significant at the 10% level*
*** Significant at the 5% level*
TARS = total annual research score
WARS= weighted-annual-research-score

RAE AND TQA/QAA

TEACHING QUALITY ASSESSMENT SCORES AND THEIR INFLUENCES – THE SCOTTISH EXPERIENCE

2nd International, Inter-disciplinary, Biennial conference on Evidence-Based Policies and Indicator Systems, Durham, July 1999

Lynn T Drennan

TEACHING QUALITY ASSESSMENT SCORES: MEASURING QUALITY OR CONFIRMING HIERARCHY?

The Sixth Quality in Higher Education Seminar, The End of Quality?, Birmingham, May 2001

Lynn Drennan and Matthias Beck

DOES RESEARCHING HELP OR HINDER YOUR TEACHING?

English Subject Centre Newsletter, Royal Holloway, Issue 1, May 2001

R. J. Ellis

THE INTERACTIONS BETWEEN RESEARCH, TEACHING AND OTHER ACADEMIC ACTIVITIES – ANNEX B: ACADEMIC ACTIVITY

HEFCE Fundamental Review of Research Policy and Funding, 2000

J M Consulting Ltd.
Institute of Education (University of London)
Commonwealth Higher Education Management Service
Higher Education Consultancy Group

To judge the quality of research taking place in universities, HEFCE conducts a UK-wide Research Assessment Exercise (RAE) every few years. RAEs have taken place in 1992, 1996 and 2001. Institutions may submit research outputs for review in up to 69 subject areas. Outputs can be publications, products or even artistic performances, which are assessed through peer review by panels of experts in the subject. Each submission is given a quality rating, judged against standards of national and international excellence, according to a seven-point scale ranging from 1 at the lowest through 2, 3b, 3a, 4 and 5, to 5* at the top.

Teaching Quality Assessments (TQA) have been carried out by the Funding Council assessing the quality of teaching in higher education institutions since 1993 until 1998. Only a limited number of subject areas are assessed in each round of TQAs. A new system of independent academic review operated by the Quality Assurance Agency, was introduced in 2001. The assessment involves a close examination of all aspects of a department's teaching practice.

Unlike the RAE which is a UK wide assessment, the exact style of the TQA varied between the four parts of the United Kingdom. However, they all encompassed common elements, including institutional or departmental self-assessment and visits by a team of academic peer reviewers. Most of the assessors were nominated by the higher-education institutions themselves, with the inclusion of some independent practitioners. Initially the scale used in England was on three points: unsatisfactory, satisfactory and excellent. This was later changed to a four point scale with the

inclusion of a highly satisfactory banding. In Scotland teaching was graded from the start in four categories, making direct comparison with England difficult.

As seen in the review sections covering correlation studies, there have been extensive studies at the level of the individual. The compilation of RAE and TQA scores allows the opportunity to measure the relationship between the two indices at institutional and departmental levels. The primary research covering this issue is reported in this section of reviews.

Drennan (1999) in the earliest of these studies investigates the relationship between TQA and RAE scores of the 13 Scottish universities. The data used for this study is that produced by the Scottish Higher Education Funding Council following the completion of cognate area assessments. The analysis focuses on a comparison of these TQA scores, averaged out over the five-year period from 1993 - 98, with the mean RAE scores from the 1996 exercise, representing the quality of research activity over a four-year period from January 1992 to December 1995.

The first RAE took place in 1992, when four of the five new Scottish universities had only just come into existence. There was limited involvement in the 1992 assessment by these institutions and it was not until the 1996 round that there was sufficient data on which conclusions could be drawn about the performance of each of 13 universities. Given the time lag involved between actual research taking place and its publication, it is perhaps better that the 1992 RAE results were not employed in the study.

In order to establish the relative contribution of research performance on teaching assessment outcomes, Drennan (1999) explores what other elements might be influencing the teaching assessments scores. To this aim she compiles data for student entry standards, student-staff ratios, and library spending from the annual Times Higher Education league tables.

Drennan estimates four two-variable linear regression models, each with average annual TQA score as the dependent variable. The regressions employ the following factors as independent variables: RAE score, entry standards of students, staff-student ratio, and library spending, with each factor being used as the solitary independent variable in each of the four regressions. The results are included as table 1.

A significant positive relationship is apparent when the mean TQA scores are regressed with raw average RAE scores, with over 70% of the variation being explained by the independent variable. When weighted RAE score (accounting for the size of academic departments) is used the variation explained is even larger. Taken at face value this would suggest a strong link between research and teaching quality.

Realising the possibility that TQA scores are influenced by a number of factors other than those related to research output, the author also ran single variable regressions using three other selected variables. Regressing average TQA scores with student entry standards reveals a highly significant positive association with over 80% of the variation in teaching assessment scores being explained on the basis of student entry standards.

Drennan notes that the process inherent in this relationship could be path dependent, with high scoring universities – defined in terms of their teaching and research score – attracting the best students with their scores and reputation. She also considers

another interesting prospect. It is likely that assessors when visiting institutions as part of the assessment inspection, would have contact with students through classroom observation of teaching and learning, and individual discussion with selected students. Drennan hypothesises that students who entered university with higher grades may be more articulate and more confident, thereby creating a better impression with the assessors leading to higher TQA scores for the institution.

As expected it is observed (through the negative relationship perceived in the regression) that universities with lower staff to student ratios have higher teaching quality assessment scores than their more encumbered counterparts. This can be explained by the possibility of the benefit from more individual attention being greater in smaller classes.

Finally, it is observed that over 70% of the variation in the TQA scores can be explained by library spending, with better resourced institutions achieving higher scores. This could reflect that assessors were influenced in their assessment of teaching quality by the standard of resources available within institutions. It could be argued though, that since these resources support student learning they should be having an impact on TQA scores. While that may be a valid argument, it would suggest that TQA does not provide an accurate measure of *quality of teaching* as far as the individual teacher is concerned. If the attempt is to quantify the relationship between quality of teaching and quality of research, then TQA scores (affected by resource availability) may not be the best measure.

While these two-variable regressions provide information on the relationship between TQA scores and individual factors, it is mystifying why a multi-variable regression was not run. It is obvious that each of the independent variables is theoretically associated with TQA scores. Each two-variable regression is in fact an under-fitted model. We expect all the independent variables to be correlated, especially standard entry scores and RAE scores. Consequently the regression coefficients derived are most likely biased as well as inconsistent. Also the measured variance of the coefficients is likely to be biased. In consequence, the measures of statistical significance used cannot be trusted. The mistake is elementary, and spoils the chance of observing the actual link between RAE scores and TQA scores.

In the second part of the analysis Drennan examines the relationship between RAE and TQA scores in selected cognate areas for the Scottish universities. Two-variable regressions are run between TQA and RAE scores, for 8 specific disciplines. Results are noted in table 2. The subject groups where the research assessment exercise score closely predicts teaching quality assessment outcomes are physics, chemistry, biology and business and management. The worst predictions are for sociology, accounting, politics and history. Drennan believes this indicates that universities with poorer research assessment exercise scores may have made some inroads in achieving good teaching quality outcomes in specific areas, but this has proved difficult to do so in the sciences. Once again by omitting other relevant explanatory variables the models are most likely under-fitted and suffering from the consequences of the specification error discussed before.

Drennan concludes that institutional research prestige and its resources are giving “halo effects” to perceived teaching quality. And that rather than being an objective and value free process of quality assessment, the TQA exercise brings a subjective focus into play. Assessments are less likely to be mission sensitive and more likely to be influenced by pre-existing research reputations, and levels of facility, which favours the older universities and disadvantage the new. Drennan also hypothesises that teaching quality outcomes are subject to similar assessment criteria as those for

research assessment and that the expectations of teaching quality assessors are strongly influenced by their own experiences as active researchers, as well as the research reputation of the institution or subject group being assessed.

It is possible Drennan's reasoning is correct, but this research does not prove as a causal relation the "halo effect". It is true that resource and reputation factors are found to affect TQA scores, but with significant flaws in the methodology of the analysis it is difficult to infer anything conclusively.

Drennan and Beck (2001) extend Drennan's (1999) study of Scottish universities, to examine the relationship between TQA scores and RAE scores for all UK universities. They use data produced by the funding councils and published in the Times Higher Education Supplement and league tables for the year 2000. The authors focus on a comparison of the mean of all TQAs, for all subjects across each institution, as published in January 2000 (representing data collected since 1993), with the mean RAE scores from the 1996 exercise (representing the quality of research activity from January 1992 to December 1995). The impact of reputational and resource variables is also explored. The variables used in the analysis are described below:

- *Teaching*: mean of all subject reviews across institutions by the funding bodies, quality assurance agency and Ofsted and published on the QAA web-site at January 2000 (TQA).

- *Research*: mean research assessment exercise score per staff member, based on the 1996 exercise on a seven point scale (RAE).

- *Entry standards*: mean average A-level scores on entry to first-year, first degree students (SES).

- *Student/staff ratio*: mean ratio for full-time equivalent total student numbers on non-franchised courses and total teaching only, plus teaching/research staff (SSR).

- *Library and computer spending*: spending averaged over three years on the academic services; central libraries and information services and central computers and computer networks per FTE student numbers. Includes expenditures on all libraries, learning resource centres and none – administrative computing (LCS).

The authors define universities into four types: ancient, pre-1945, modern and post-1992. They find that there is a pronounced clustering of RAE scores, based on the age based reputational hierarchy. The ancient universities achieve, on average, by far the highest RAE scores. These differ drastically from the scores for post 1992 institutions. Mean TQA scores for the four types of universities display smaller but still clearly discernible differences (see table 3).

Drennan and Beck calculate correlations using each possible pairing of the variables selected. The matrix displayed in table 4 shows the zero-order correlation coefficients derived. The strongest interaction can be detected between the research score and student entry standards. A very high correlation of 0.746 is apparent between TQA and RAE scores.

In order to identify which variables best predict a university's TQA score, TQA scores - as dependent variable - are regressed with RAE scores, student/staff ratio, and

library and computing spending. Drennan and Beck note that due to the likelihood of high multicollinearity between student entry standards and RAE scores the analysis excludes the former variable. The result of the regression is included in table 5. RAE scores are seen to be the strongest predictor of TQA scores, followed by the much weaker staff/student ratio.

When faced with strong multicollinearity, one of the simplest things to do is to drop one of the collinear variables. In this case Drennan and Beck choose to drop the student entry standard variable. But in dropping a variable from the model they may be committing a specification error. In this case the RAE coefficient will be a biased estimate of the true value, and further an overestimation. Hence it is possible the remedy chosen may be worse than the disease because, where multicollinearity may have prevented precise estimation of the parameters of the model, omitting the SES variable may seriously mislead us to the true values of the parameters. Knowing this, it would be ill-advised to trust the results of this regression.

Drennan and Beck, however, further extend the investigation by applying a partial correlation analysis to control for all variables included in the data set. Partial correlation allows us to determine what the correlation between any two of the variables would be (hypothetically) if the other variables were held constant. As shown in table 6 higher order partial correlation coefficients confirm the strength of the relationship between TQA and student entry scores. To a lesser degree the negative relationship between TQA and staff-student ratios is also established. Most interestingly, however, the higher order partial correlation between TQA and RAE scores drops to less than 0.1. This indicates that RAE scores themselves correlate with other, in this case, resource related variables.

This partial correlation analysis is the most robust investigation among the four studies reviewed. Unlike all the other results in these which indicate strong associations among research and teaching, this analysis produces a small positive relationship between the two, when other factors affecting the relationship are accounted for. Even then there are problems with accepting these results due to the general difficulties/problems/flaws described at the end of this section.

Drennan and Beck (2001) conclude:

The analysis of data indicates that the determining factors in relation to high TQA scores are age of institution and reputation, in terms of attractiveness to students with high entry qualifications, and in terms of peer assessed research. These variables are, in turn, strongly related to other variables such as staff-student ratio and library and computer spending. Since reputations are established over a long period of time and resources appear to be most easily obtained by institutions with strong reputations, post-1992 universities appear to be at a considerable disadvantage.

Rather than representing an independent indicator of teaching quality, external quality of reviews, as evidenced by the UK TQA scores, has largely serve to confirm existing hierarchies, which may be determined by factors other than teaching.

Drennan (1999) had carried out a comparative analysis at the institutional level in the Scottish context. Drennan and Beck (2001) expanded this to all UK universities. In contrast **Ellis (2001)** has performed an analysis of those English departments which were entered in both the 1996 RAE and the 1994/95 TQA. The sample covers 72 Departments.

In the TQA, departments had been rated as excellent, satisfactory, or unsatisfactory. Ellis transforms these into scores:

- Excellent = 5
- Satisfactory = 2.5
- Unsatisfactory = 0

To correspond with these three categories, Ellis groups together firstly RAE 5* and 5 departments, then RAE 4, 3a and 3b departments, and finally RAE 2 and 1 departments.

The analysis undertaken is a simple match of the TQA and RAE categories the results of which are included in table 7. Departments rated 5 or 5 * in the research assessment exercise turn out to average dead on five, since all of them were rated excellent in the teaching quality assessment. The teaching quality assessment average consistently rises as research assessment exercise ratings rises. Ellis further reports exactly three-quarters of departments rated 4,5 or 5* in the RAE obtained an "excellent" rating on the TQA. This compared with only one in five for those departments rated 3, 2 or 1 in the RAE.

Ellis concludes from this that *plainly good research and good teaching go hand-in-hand. The implication must be: research does indeed help you teach better.*

While the matching is compelling a few points to note. Firstly the analysis allows no account for other factors which may be affecting both RAE and TQA scores. As Drennan and Beck (2001) demonstrate the association found between the two through partial correlation analysis, accounting for the effects of reputational and resource variables, is drastically smaller than that found in zero order correlation. Secondly, the results may be specific to the discipline, English, and should not be generalised. Third, the implication derived by Ellis that "research does indeed help you teach better", may not necessarily be true. Theoretically the relationship could be the other way around, with teaching proving beneficial for research (though Ellis's rationalisation may be more likely).

As part of **HEFCE's (2000)** fundamental review of research policy and funding, a short section investigates the statistical relationship between RAE and TQA scores. The analysis covers sociology and electrical and electronic engineering. The study reports that:

In electrical and electronic engineering there was a good correlation between the two scores, and regression coefficient suggested a 1% higher RAE score was associated with a 0.7% higher total TQA score. In sociology the correlation was much lower and a 1% higher RAE was associated with only 0.4% higher TQA score.

The study also reports that it was examined whether size of department made any difference to these results and found that in general it did not.

The type of model used, or additional explanatory variables employed, or a summary table of the regression results is not provided in the study. It can be inferred from the wording that a log-log two-variable regression model was employed for each section, though the use of the term correlation in the sentences reporting the regression coefficients is confusing. It could in fact be possible that the associations reported are correlation coefficients, and the usage of the term "regression" is flawed.

In any case there are certain points to be noted. First, the stronger relationship observed for the science discipline in comparison to the social science subject is similar to the findings of Drennan (1999). Second, exactly like that study the findings are compromised by omitting the effects of reputation, resource and other variables which are likely to affect both RAE and TQA scores. Third, like Ellis (2001) the results are subject specific.

There are several problems associated with trying to quantify the link between teaching and research quality using TQA and RAE scores as attempted by the studies reviewed in this section. Specific issues with individual studies have been dealt with briefly in their reviews. More general problems associated with all such kind of studies are discussed below. Some of these difficulties arise due to complications in accurately matching TQA and RAE scores. Others originate from the processes of assessment themselves.

In the TQA all department members contribute to the ratings, whereas the RAE rating is based only on selected “research active” staff. The funding council’s criteria for RAE do not allow clear penalization of departments returning low levels of research active staff. Departments can enter only a few staff plainly of international standing to obtain first-rate research outcomes. Matching these with the teaching performance of all faculty in a department as measured by the TQA cannot provide robust evidence as to the actual relationship between research and teaching quality.

An additional problem with the RAE is the subjectivity inherent in peer review. Ellis (2001) notes several case of what he terms “unedifying human behavior” in peer review. Ellis lists: risk aversion, misplaced confidence in members of the club, unattributable and unsustainable attacks on competitors, and subliminal collusion between review panel members. And further notes:

It's amazing how successful one is when one serves as a member of a grant-awarding body, even though one leaves the room when one's own applications are being discussed. It has happened to me. There's nothing Machiavellian in this - it is inevitable human behaviour.

Similar problems may be present in assessors’ evaluations in the TQA process. It could be difficult for assessors from ancient universities to really understand what a post-1992 institution is trying to do with a wide range of students, including many non-standard candidates. Similarly it could be difficult for assessors from both new and old universities not to be impressed by the research reputations and facilities of ancient universities in comparison with the post-1992 institutions. It is hard to view these scores as being objective and value free.

The division of what is classified as research and what is not in the RAE is also an issue. Pedagogic research, for example, was not classified as research according to the exercise until the 2001 round. Even then the policy was announced too late in the cycle to have considerable impact on what was submitted to subject panels for the 2001 RAE.

Analyzing average TQA and RAE scores across institutions as both Drennan (1999) and Drennan and Beck (2001) attempt is potentially flawed. Theoretically an institution could have high TQA scores for some departments but low RAE scores, similarly there could be departments with low TQA scores and high RAE ratings. Averaging the scores across all departments of an institution would mask the actual linkage between the scores (which are in fact collected at the department or subject level). Also an institution could have had most of its departments assessed in the

TQA, but could have submitted only a few departments in the RAE review. Averaging scores for the institution and the then matching them, as done by the two identified studies, could lead to misleading results.

(Drennan: 1999) Table 1. Regressing mean TQA results with RAE, resource and reputational factors

Dependent Variable	AVERAGE ANNUAL TQA SCORE
Independent Variable	RAW RAE
Coefficient (unstandardised)	.248
Significance probability (P)	.000
R^2	.731
Adjusted R^2	.706
Independent Variable	WEIGHTED RAE
Coefficient (unstandardised)	.178
Significance probability (P)	.000
R^2	.750
Adjusted R^2	.727
Independent Variable	STUDENT ENTRY STANDARD
Coefficient (unstandardised)	.060
Significance probability (P)	.000
R^2	.822
Adjusted R^2	.806
Independent Variable	STAFF STUDENT RATIOS
Coefficient (unstandardised)	-.114
Significance probability (P)	.028
R^2	.369
Adjusted R^2	.312
Independent Variable	LIBRARY SPEND PER FTE
Coefficient (unstandardised)	.005
Significance probability (P)	.000
R^2	.714
Adjusted R^2	.688

(Drennan: 1999) Table 2. Regressions for individual cognate areas

Dependent Variable	AVERAGE ANNUAL TQA SCORE		
Independent Variable	ACCOUNTING RAE SCORE		
Coefficient (unstandardised)	.750	R^2	.197
Significance probability (P)	.319	*Adjusted* R^2	.036
Independent Variable	BIOLOGY RAE SCORE		
Coefficient (unstandardised)	.280	R^2	.614
Significance probability (P)	.021	*Adjusted* R^2	.550
Independent Variable	BUSINESS AND MANAGEMENT RAE SCORE		
Coefficient (unstandardised)	.140	R^2	.420
Significance probability (P)	.031	*Adjusted* R^2	.356
Independent Variable	CHEMISTRY RAE SCORE		
Coefficient (unstandardised)	.275	R^2	.634
Significance probability (P)	.018	*Adjusted* R^2	.572
Independent Variable	HISTORY RAE SCORE		
Coefficient (unstandardised)	1.050	R^2	.350
Significance probability (P)	.162	*Adjusted* R^2	.220
Independent Variable	PHYSICS RAE SCORE		
Coefficient (unstandardised)	1.860	R^2	.779
Significance probability (P)	.020	*Adjusted* R^2	.724
Independent Variable	POLITICS RAE SCORE		
Coefficient (unstandardised)	1.200	R^2	.300
Significance probability (P)	.261	*Adjusted* R^2	.125
Independent Variable	SOCIOLOGY RAE SCORE		
Coefficient (unstandardised)	1.000	R^2	.273
Significance probability (P)	.478	*Adjusted* R^2	-.091

(Drennan and Beck: 2001) Table 3. Mean TQA scores (1993-2000) and mean RAE scores (1996) by age-type of university

Age-type	Mean TQA Scores	Mean RAE Scores
Ancient	21.4	4.9
Pre 1945	20.9	4.3
Modern	20.5	3.9
Post 1992	19.7	0.9

(Drennan and Beck: 2001) Table 4. Zero Order Correlation Matrix

	RAE	**SES**	**SSR**	**LCS**
TQA	.746	.763	-.653	.606
RAE		.904	-.666	.675
SES			-.672	.606
SSR				-.561

(Drennan and Beck: 2001) **Table 5. Multiple Regression of TQA with RAE, SSR and LCS**

Dependent Variable	TQA		
Independent Variables	RAE	SSR	LCS
Coefficient			
Unstandardised	.227	-.072	.000
Standardised	.486	-.253	.137
Sign. Probability	.005	.000	.130
R^2	.610		
Sign Probability F	.000		

(Drennan and Beck: 2001) Table 6. Higher Order Partial Correlation Analysis, TQA, SES, RAE, SSR, LCS

TQA with	**0 order**	**3rd order (all variables)**
RAE	.746	.091
SES	.763	.274
SSR	-.653	-.229
LCS	.606	.179

(Ellis: 2001) Table 7. Matching TQA average Rating with RAE Rating

RAE Rating	**Average TQA Rating**	**Sample Size**
5* & 5	5.00	8
4, 3a & 3b	3.45	37
2 & 1	2.87	27

FACULTY AND ADMIN PERCEPTIONS

CHIEF ACADMIC OFFICERS' PERCEPTIONS OF THE RELATIONSHIP BETWEEN FACULTY RESEARCH AND UNDERGRADUATE TEACHING

*Sociological Spectrum, **18**, 1998*

Paul L. Leslie
Lynn K. Harvey
George J. Leslie

The purpose of this paper is to examine how college and university chief academic officers perceive the relationship between teaching and research. Leslie *et al* contend that chief academic officers' perceptions of scholarship significantly affect the academic climate and that they are instrumental in establishing institutional definitions of scholarship and creating faculty reward patterns.

Data was gathered for the study through a mail questionnaire sent to 300 chief academic officers from U.S. colleges and universities. A stratified random sampling technique was used to establish a sample representative of the various types of institutions of higher education. 160 questionnaires were returned complete and usable for analysis. There is observed regional bias in the returned questionnaires. It is not certain if non-respondents differ significantly from respondents. This limits the ability of the results to be generalised to the entire population of U.S. 4 year institutions. In their analysis of the quantitative data Leslie *et al* find that sex of academic officer, region of the country, size of student body, type of institution (public or private) and degree of research emphasis of institution, have no effects on chief academic officers' understanding of the relationship of research to teaching. Three factors are significantly related to chief academic officers' perception of the effect of research on teaching. The results are displayed in table 1.

Of all participants 92.5% responded yes and 7.5% responded no to the question "Do you believe that faculty research activity enhances teaching effectiveness at your institution?". Officers of schools with fewer than a hundred faculty are less likely to perceive that research enhances teaching effectiveness than their counterparts from institutions with more faculty members. Officers from schools receiving least amount of research funds are almost ten times as likely as respondents from schools receiving more grant money to report that research does not enhance teaching. Officers from schools that do not offer graduate degrees are three times as likely to report that research activity does not improve teaching effectiveness than their counterparts from schools that do offer graduate degrees. In all cases though an overwhelming majority of chief academic officers reported that research by faculty positively affected teaching.

Further insight can be gained by looking at the responses to the request for comments at the end of the questionnaire. 52 respondents commented on the relationship between research and teaching effectiveness. 40% of these (n = 21) commented that research positively affects teaching. Comments reflective of this opinion included:

Doing some kind of research is the best, and often the only way to keep from becoming stale.

I believe students are best serves by being taught by instructors who are helping to create new knowledge.

While this is not meant to be a research institution, research activity is extremely important in maintaining us as a good teaching institution. Teaching without research becomes stale and irrelevant over time.

23% of those who wrote comments (n = 12) stated that there was no conflict between the two roles. Comments reflective of this opinion include:

Effective teaching requires faculty to be alive and vigorous intellectually, and research is a means to that end. However, many good researchers pay no attention to their teaching or how their research enhances their teaching.

The only people who insist on separating teaching from research are people who have done very little of either.

17% of those who commented (n = 9) were of the dissenting camp. One response suggested:

Frankly, for most faculty, doing research is not terribly important. While we would like to believe that there is a correlation between research and teaching effectiveness – there are no data to suggest that this is the case.

The remaining 19% of those who commented (n = 10) largely suggested that the effect of research on teaching depends on multiple factors. One response states:

I think that it enhances teaching effectiveness up to a point. It helps professors stay current and helps them use current examples in their teaching. But much of the effect can also be gained by reading current literature.

The quantitative data in the study indicates that though there are some differences in strength of support dependent on institutional characteristics, chief academic officers overwhelmingly support the idea that research enhances teaching effectiveness. These results are similar to two linked studies, one by Neumann (1993) in Australia, and the other by Rowland (1996) at Sheffield University which demonstrate that department heads and administrators see strong positive correlations between staff involvement in research and the intellectual currency of their courses. There is obviously the issue of self-interest in administrators emphasising the value of research for teaching. Jenkins *et al* (2003) suggest however that it could also be argued that administrators have in certain ways a more informed and sophisticated view as to what is teaching quality than maybe obtained by student questionnaire studies.

Leslie *et al* themselves speculate that the call for more accountability for teaching effectiveness gives chief academic officers little choice but to contend that their schools employ stellar teachers. Without systems in place which objectively measure teaching chief academic officers revert to research publications as a central criterion of evaluations. They term this process as “regressive determination”. However, even if officers’ are aware of some other from of teaching evaluation, it is conceivable that their judgement of teaching performance are likely to be contaminated by their knowledge of the professors’ research record. In such a circumstance they may say the linkage between research and teaching as much stronger than it might be. A clear answer to the question of the true linkage between research and teaching can only be obtained if the measure of teaching performance is obtained completely independent of the measure of research performance. In failing to meet this critical requirement all this study can provide us is what its title suggests i.e. the “perception of chief academic officers”.

References:

Jenkins, A, Breen, R, Lindsay, R and Brew, A (2003) *Reshaping teaching in higher education: linking teaching with research*, Kogan Page Ltd., London

Neumann, R (1993) Research and scholarship: perceptions of senior academic administrators, *Higher Education*, **25**, pp 97-110

Rowland, S (1996) Relationships between teaching and research, *Teaching in Higher Education*, **1** (1), pp 7-20

TEACHING, RESEARCH AND SCHOLARSHIPIN DIFFERENT DISCIPLINES
*Higher Education, **19**, 1990*

Ingrid Moses

The aim of this paper is to look more closely at the ways in which there is interaction between research and teaching by asking university faculty members their opinions. The study is based on quantitative survey data gathered from faculty at a traditional Australian university.

Moses argues in the paper that firstly without the notion of scholarship the debate about research and teaching functions causes more confusion and misconception than clarification of the issues. Secondly, that culture and conventions within disciplines influence the conceptions of both research and scholarship. The strength of this study is that it sheds light on the complexity of the interaction especially in the context of different disciplines. However, it is based upon interviews from a sample of disciplines. This makes it difficult to generalise on its basis.

Moses takes her definitions of the concept of research and scholarship from the 1988 White paper, by the Australian Minister for Employment, Education and Training. Research is defined as “a systematic and rigorous investigation aimed at the discovery of previously unknown phenomenon, the development of explanatory theory and its application to new situations of problem and the construction of original works of significant intellectual merit”. The White paper describes the concept of scholarship as “the analysis and interpretation of existing knowledge aimed at improving through teaching or by other means of communication the depth of human understanding”.

After defining these concepts, Moses presents categorisation of disciplines according to various theories. Against the background of what has been done in previous studies (Becher: 1984, Whitley: 1984), Moses categorises disciplines:

- by the development of a paradigm in to hard/soft, with the sciences and engineering hard (i.e. having high level of paradigm development) and the humanities soft.
- by the concern with application into pure/applied, with the physical sciences and the arts pure and education and engineering for example applied.

While these categories may help in the understanding of disciplinary cultures, it is to be noted that there are many variations in sub-groups. The study builds upon quantitative survey data. The survey data were generated by a 1987 questionnaire study among 400 faculty members in the departments of law, chemistry, English and engineering at a traditional Australian university. The response rate was 81 percent.

Faculty were asked about the extent to which they thought research enhanced their teaching. Most teachers (90 percent) among all respondents across the university agreed with the statement. The percentage was lowest among law faculty (80%). The faculty were also asked if they thought their research was enhanced by their teaching. The majority of the staff in all the departments again agreed with the statement. However, the faculty thought that teaching gave positive impulses to research to a lesser extent than they thought research influences teaching. There were also more significant differences between fields of learning. A larger percent of chemistry and English faculty thought that teaching positively impacted their research than colleagues in other fields of learning. For many staff in all departments teaching in areas in which they were not specialists was seen to be a welcome challenge.

Interestingly, while most faculty agreed that there were significant positive impulses from teaching to research and from research to teaching, a significant number agreed that the statement that one can be an excellent teacher without doing actively research. Results are summarised in Table 1.

It is interesting to compare these findings with Smeby (1998), who examines data from faculty from four Norwegian universities. Smeby found similar percentages of faculty agreeing with the statements of enhancement between the two activities for faculty teaching at postgraduate levels. The percentage was lower at undergraduate level. Smeby also found that there was a higher percentage of teachers in the humanities and social sciences than in the natural sciences, medicine or technology who thought that their teaching (research) was influenced by their research (teaching) to a great or some extent. The lower levels of role complementarity perceived by law (social science) faculty and the higher levels viewed by chemistry (natural science) faculty in Moses' study could possibly be explained by:

1. differences across countries, or
2. that law does not adequately represent all social science faculty and similarly chemistry is not reflective of entire natural science faculty.

Several studies have pointed out that the interaction between teaching and research is often of an indirect nature. Elton (1986) thinks the reason why it is difficult to find simple and clear connections between research and teaching is because interaction is mediated through scholarship. This concurs with Moses' assertion in the paper that scholarship overlaps with both teaching and research. She examines the notion of scholarship across the four different disciplines in her study through sixteen different variables. The results of these are given in Table 2.

Chemistry staff were observed clearly to be scholarly within their research context. That they frequently communicate with colleges and graduate students through personal interaction, seminars and conferences, and by participating in the validation of research problems and strategies by reviewing articles is seen by Moses as typical of hard/pure areas.

Compared with other departments on the scholarship variables no activities or attitudes are found that distinguishes *Engineering* staff. This does not mean that engineering staff did not perform scholarly activities, just that their involvement was less marked than staff in chemistry. Comparing staff in engineering (hard/applied area) from those in the physical sciences (hard/pure area) Moses observed that engineering staff more often engaged in systematic study to gain new knowledge or acquire new research techniques, and they agreed more with: "I keep up to date with developments in the subject areas I am teaching". They were less inclined though to have informal discussions with colleagues in the department about common research interests or to participate in seminars. Considering the characteristics of hard/pure areas and hard/applied areas Moses considers that it should be expected that engineering staff communicate less often, but score high on scholarship variable related to problem solving compared to other staff.

English faculty less often held informal discussions with colleagues about research interests, delivered conference papers or reviewed articles for a professional journal. As a pure/soft area scholarship in English is seen by Moses as mainly the reinterpretation of existing work. Moses argues that there is little urgency to communicate frequently through journals, which is why we would expect arts faculty to score low on scholarship variables related to research. For a pure/soft area being up to date is not an issue. It is not necessary that students are up to date with the

subject content. There are often schools of thought and necessary skills are taught by evaluating within these context.

The *Law* department is a professional department teaching a degree that provides the basis for professional accreditation for practising lawyers. Law respondents had a scholarly approach to their work with regard to teaching. They claimed that they kept up to date with new developments in the subjects they were teaching and when revising a course they did research to make the content up to date. Moses also does not consider it surprising that participation in seminars played such a small role for law staff. She considers the variables relating to research shoe law staff as working in isolation and that this is quite typical for soft fields.

There are several issues of that should be noted concerning the paper. Firstly it illuminates the extent of disciplinary variations and tries to explain these variations through theoretical categorisation. The findings of the differences between disciplines would be more robust had many significant variables been controlled for (e.g. example level of teaching, academic rank, productivity of researchers). The introduction of such controls by Smeby (1998) in a paper of similar methodology makes that study superior. Also, since Moses only examines specific departments within disciplines the results may not be generalisable.

The analysis of scholarship by Moses is another point of note. As asserted before it has been argued that the interaction between research and teaching is mediated through scholarship. Given this likelihood it is important to attempt to quantify this. While the present study manages to compare the degree of scholarship between different disciplines, it does not provide an overall indication of how important scholarly activity really is for faculty, or how much time they spend in such activities. It could also be argued that of the variables chosen to describe scholarly activities some (e.g. delivered conference paper) are actually distinctly either a research or teaching activity and do not overlap both fields.

Further, for the purpose of generalisation a survey with a limited number of respondents, from a limited number of disciplines in one institution can only give limited information.

References:

Becher, T (1984) The cultural view, in B R Clark (ed.), *Perspectives on higher education*, Berkeley, University of California Press, pp. 165-98

Elton, L (1986) Research and teaching: symbiosis or conflict?, *Higher Education*, **15**, pp. 299-304

Smeby, J-C (1998) Knowledge production and knowledge transmission: the interaction between research and teaching at universities, *Teaching in Higher Education*, **3** (1), pp. 7-20

Whitley, R (1984) *The intellectual and social organization of the sciences,* Oxford, Clarendon Press

RESEARCH AND TEACHING AT A RESEARCH UNIVERSITY
*Higher Education, **40**, 2000*

Robert C. Serow

Serow reports findings from a case study of faculty role performance at a large research university in the United States. He examines the tension between the research and teaching components of the faculty role, by giving voice to those faculty members who see themselves, and are seen by others, as committed undergraduate teachers.

The sample used is limited to full-time faculty members in the natural, applied and behavioural sciences. Individuals with histories of active involvement in undergraduate education were identified through meeting several criteria including membership of committees dealing with undergraduate academic affairs and participation in one of the university's major teaching improvement initiatives. Serow identified 33 faculty members through the process, 29 of whom agreed to participate in the study. The sample were much more likely than their colleagues to be not just interested in instruction but also effective at delivering it. 25 of the 29 had been recipients of teaching awards. Additionally, they were more senior in rank, with 27 holding tenured associate or full professorships. In addition data was collected from publicly accessible archival material (faculty handbooks, records of contracts and grants etc.) and personal documents obtained from the participants including curricula vitae and course syllabi.

Kremer (1991) based on results obtained from a U.S. institution, hypothesises that full professors may not feel the same pressures as faculty members at earlier stages of their careers to excel in all academic pursuits. And thus may return to being more focused in a particular area, emphasising teaching or research, but not both. The sample is observed to be skewed heavily toward senior ranks.

It is evident that dependence on such a small sample at one institution restricts the generalisability of the conclusions. Additionally, through the structuring of the sample it is obviously not representative of the entire faculty population. It should not be taken to be representative of good teachers either (those who inhabit the high tail end of the teaching skill distribution). While through the measure of teaching awards, it is evident that most of the chosen sample are good teachers, they are distinct in that they see themselves, and are seen by others as committed teachers. Given this, and Kremer's assertion, the sample participants are not likely to perceive research in a positive light compared to teaching. Though the generalisability of the results of the study are restricted, our interest in the paper lies in the following two factors:

1. Since the case study format has the advantage of the researcher probing more tellingly into individuals' attitudes or behaviours than would be possible in a larger study, it is possible that light could be shed on the complexity of the interaction between research and teaching.
2. Despite the factors discussed previously if positive accounts of research are given by some of the sample this could evidence an association between research and teaching.

Serow observes that a major point of agreement among the interviewees was that research outranked teaching in the university's faculty reward system. This did not imply, however, that teaching and advising were ignored altogether. It was suggested by some respondents that by sponsoring teaching awards, cash grants for instructional innovations, and other incentives the administration's real aim was to

“avoid the ramifications of bad teaching” by minimising pressures from students, parents and legislators. This could be interpreted as a cynical view point.

Although all of the interviewees had established reputations as committed undergraduate teachers, significant differences were found in other dimensions of their faculty roles. 11 of the 29 faculty members were found to have attracted moderate to substantial amounts of external financing ($50,000 to almost $2 million). Serow classifies these as active researchers. The remaining 18 sample members (14 of whom had no funded research) were coded as less active.

In explaining the role of the undergraduate teaching faculty the accounts respondents gave could be divided between complementary role theory and competitive models. The complementary version considers the roles as mutually reinforcing. Proponents of the competitive model consider the two roles to be indirect competition with each other, with excellence in one coming at the price of the other. The patterns of the accounts are summarised in table 1.

Academics classified as active researchers are observed to generally support the assertion that teaching and research are complementary. In contrast, an overwhelming majority of less active researchers, view teaching and research as competitors. This finding is contrary to the generally held view among faculty that their research enhances their teaching, and that their teaching activities have a positive effect on their research (see Smeby: 1998 and Moses: 1990). But it is expected given the underlying undergraduate teaching commitment of the sample. It further supports Gottlieb and Keith’s (1997) finding for a sample of academics from eight industrialised countries that academic orientation (i.e. whether an academic’s primary emphasis is on research or teaching) is the most significant factor in determining how much time is spent on either activity.

Comments reflective of and describing the position in the study included:

I have 100 plus advisees of my own. This is 4 to 5 times the usual advising load.... It’s ultimately related to the reward system. If 70 percent of someone’s faculty appointment is (funded) by research, why should they spend time on advising? It’s a disincentive.

(Parents) expect their kids to get a quality education. They’re paying us to teach. I’m not doing my job if I use federal funds to buy my way out of teaching.

I’m not held in as high esteem in our department as some of our researchers. I don’t think that’s appropriate. I doubt that I’m paid as much. But am I doing as much good for society? Quite a bit more, actually.

References:

Gottlieb, E E and Keith, B (1997) The academic research-teaching nexus in eight advanced-industrialized countries, Higher Education, **34**, 397-420

Kremer, J F (1991) Identifying faculty types using peer ratings of teaching, research, and service, Research in Higher Education, **32** (4), pp 351-61

Moses, I (1990) Teaching, research and scholarship in different disciplines, Higher Education, **19**, pp 351-75

Smeby, J-C (1998) Knowledge production and knowledge transmission: the interaction between research and teaching at universities, Teaching in Higher Education, **3** (1), pp 7-20

KNOWLEDGE PRODUCTION AND KNOWLEDGE TRANSMISSION: THE INTERACTION BETWEEN RESEARCH AND TEACHING AT UNIVERSITIES

*Teaching in Higher Education, **3** (1), 1998*

Jens-Christian Smeby

In this article Smeby looks more closely at the ways in which there is interaction between research and teaching by asking university faculty members their opinions. The study is based on data from all fields of learning at Norway's four universities. Smeby examines the extent to which university teachers' think their research characterises their teaching and whether they think their teaching activities affect their teaching positively. The interaction between the two activities is examined in the light of variations between disciplines, teaching levels and the characteristics of the academics involved in the study.

The study is based upon both quantitative survey data and in-depth interviews. The survey data were generated by a 1992 questionnaire study among all faculty members of the rank of assistant professor or higher at Norway's four universities. The response rate of 69 percent is very high for this sort of study. Five disciplines, the humanities, the social sciences, the natural sciences, medicine and technology were distinguished. Data from a similar study in 1982 are also included. The survey did not include technology. The total sample size numbers 1592. 35 semi-structured, in-depth interviews were conducted with tenured faculty members in physics, Nordic languages and literature at the universities of Oslo and Bergen. Smeby defends his decision to use multiple data sources and collection methods by arguing that this is a way of overcoming weaknesses inherent in the singular use of either interviews or quantitative survey data.

He admits it may be questioned to what extent faculty members evaluation of the interaction between research and teaching are a good measure of the actual symbiosis between these activities. But argues that the nature of the symbiosis is complex and often indirect. And that focusing on faculty members' subjective experience is important towards understanding the nature of the relationship between teaching and research.

Norway's four universities are all research universities and there are only small differences in prestige between them. There are no differences in the working conditions for teaching loads between universities and positional groups, i.e. different ranking academics. So the result from the study may therefore not be generalised countries where working conditions are less homogeneous.

Faculty were asked the extent to which they thought their research influenced their teaching. More than 95 percent at PhD level and more than 90 percent at major subject level thought their teaching was affected by their research 'a great' or 'some extent', while there were few who thought this at a undergraduate level (about 50 percent). There were a higher percentage of teachers in the humanities and social sciences than in the natural sciences and medicine who thought that their teaching at a lower level was influenced by their research to 'a greater' or 'some extent' (see table 1). However, there were no significant differences between fields of learning in the percentages of those who thought that teaching at a doctoral degree level was influenced by their own research. Smeby also reports that there were only insignificant differences between professorial ranks, different age groups and academics with different levels of research productivity in their judgements about the extent to which research influenced their teachings.

Faculty were also asked if they thought teaching at a lower level should be more connected to ongoing research. Smeby reports that there was considerable scepticism amongst teachers concerning this. The tendency was that more teachers in humanities and social sciences agreed that undergraduate teaching should be connected more to ongoing research than those in the other fields. Once again Smeby reported insignificant differences for academics with different characteristics (age groups, levels of research productivity).

The in-depth interviews with teachers revealed the viewpoint that at a lower level it is difficult to have research experience in all the subjects one teaches. But at a postgraduate level there is consensus among disciplines about research experience and teaching. Teachers in physics mentioned three arguments about why teaching at a graduate level should be researched based.

- One, it gave a feeling of academic security. Developments within research area happen so quickly that it is difficult to keep up and follow them if one does not do research in the field.
- Two, it was easier to teach if one could use examples and anecdotes from ones own research.
- Three, doing research gave university teachers a critical attitude toward knowledge. 'This was very important because an independent critical attitude is one of the most important things we can communicate to our students'.

The survey also asked faculty members about the extent to which they thought teaching gave positive impulses to research. It was observed that the higher the teaching level higher the percentage of those who thought this was the case. However, faculty thought that teaching gave positive impulse to research to a lesser extent than they thought research enhanced teaching. A larger percentage of social scientists and faculty from the humanities thought that research at a lower level gave positive impulses to teaching than colleagues in other fields of learning (see table 2). Again Smeby reports insignificant differences by characteristics of academics (rank, age group, levels of productivity of research).

In-depth interviews with faculty revealed that they thought teaching had positive effects on their research in various ways.

- Teaching forced faculty be more thorough, i.e. helping to understand things actually never understood correctly, things learnt long ago but not used later.
- Graduate teaching could be stimulating for faculty's own research.
- Teaching can also be a way to become acquainted with a new field.
- It is a way to maintain subject breadth, which contributes to a better understanding of the whole discipline.
- It was an advantage to teach a subject which faculty were interested in and in some cases this contributed to bringing up problems which might result in research.
- Students asking questions could make teachers think about things they had not thought of themselves.
- Research would be meaningless if it could not be disseminated also amongst students.

Data from the surveys in the study shows that faculties think that the content of teaching is influenced by research and that teaching gives positive impulses to their own research. The interview material while illustrating this also shows that while there is a direct interaction there are also more indirect and complex relationships.

Interviews also showed that there was not only functional interaction between the two activities, that the activities are also in competition with each other. Interviewees revealed that teaching and administration took time which could have been used for research. While Smeby agrees that researchers experience a conflict between the time they use for research and teaching, he argues that interruptions need not be negative. Research is often lonely and psychologically stressing activity and therefore teaching can often be a positive interruption which gives the possibility of direct feedback. Also while several of the interviewees said that they would like fewer teaching duties none only wanted a pure research position. Smeby thinks that this indicates that there is more complementary than competitive interaction between research and teaching.

Several of the interviewees held that at lower levels there was little direct interaction because research is very specialised in relation to teaching. Smeby views this as showing that interaction between teaching levels is connected to both students academic and intellectual maturity. This is also supported by Lindsey *et al* (2002) who find that postgraduates consider themselves as stakeholders in research while undergraduate students did not perceive themselves as such.

The analysis shows that there are differences between disciplines concerning interaction between research and teaching at undergraduate level, but fewer differences at graduate level and no differences at postgraduate level. Interaction in the lower levels is stronger in the social sciences and the humanities than in other fields. Centra (1983) found for two large samples of US faculty members that the teachers of the social science courses were the only group for which there were consistent positive relationships between the number of published articles and student ratings of instructor effectiveness. Smeby considers this to be the result of different degree of academic specialisation. He hypothesises that it could be easier to use examples from teaching ones own research if the research is less specialised and teaching which takes up relatively general issues can positively influence research.

Another interesting aspect of the study is the insignificant connections between academic's age, his/her rank and research productivity and the degree of interaction between research and teaching. The strength of this study is in its display of the ways in which there is an association between research and teaching. By shedding light on the complexity of the interaction it advances the understanding of the linkage.

Apart from the difficulty to generalise on the basis of the study due to the homogeneity among the universities in Norway where there is no difference in the working condition of teaching loads between universities and positional group, a further weakness is observable which questions to what extent faculty members' evaluations of the interaction between research and teaching are a good measure of the actual relationship between these activities. It is observed that there is no connection between teachers' productivity and their views on the interaction between research and teaching. This means that teachers who thought that teaching gave positive impulses to their research were not scientifically more productive than those teachers who thought that teaching had a less positive influence. It is suggested that this indicates teaching does not have a clear and direct positive influence on teachers own research productivity. It could be inferred from this, though not concretely, that the belief amongst academics that teaching and research performance are related, and strongly so as this study shows may be stronger than the actual linkage.

References:

Centra, J A (1983) Research productivity and teaching effectiveness, *Research in Higher Education,* **18**(2), pp 379-89

Lindsay, R, Breen, R and Jenkins, A (2002) Academic research and teaching quality: the views of undergraduate and postgraduate students, *Studies in Higher Education*, **27** (3), pp 309-27

Leslie et al (1998): Table 1. Faculty size of respondents' institutions, reported amount of grant money for research, and presence of graduate programs by perception of effect of research on teaching effectiveness

	Response					
	Yes		No			
Variable	**%**	**n**	**%**	**n**	**N**	**p**
School faculty size					159	.003
Fewer than 101	83	44	17	9		
More than 101	97	102	3	4		
Grant money (typical per year)					157	.0001
$0-$10,000	70	19	30	8		
More than $10,000	97	126	3	4		
Graduate degrees					158	.02
Yes	95	119	5	6		
No	82	27	18	6		

Respondents replied to the question "Do you believe that faculty research activity enhances teaching effectiveness at your university?"

Moses (1990): Table 1. Percentage of staff agreeing and strongly agreeing

	Chemistry	*Engineering*	*English*	*Law*
My teaching is enhanced by my research	95	94	96	80
My research is enhanced by my teaching	75	52	82	60
Staff should only teach in their special field	5	33	36	20
Teaching in areas in which I am not a specialist is a welcome challenge	65	58	36	87
For my teaching my professional experience is important	90	92	64	87
One can be an excellent teacher without doing actively research	30	48	46	40

Moses (1990): Table 2. Scholarship analysis

	Chem.	Eng.	Engl.	Law
Occurrence: Scholarly activities				
Informal discussions with colleagues in the department about common research interests	+			
Joint research projects with colleagues in the department				
Reviewed articles for a professional journal	+			-
Participated in staff/postgraduate seminar	+			-
Developed a new way of dealing with a problem	+			-
Engaged in systematic study to gain new knowledge or acquire a new research technique			-	+
Developed a new course (subject)				-
Served as a guest lecturer in a colleague's class				
Delivered conference paper	+			-
Agreement with statement:				
I frequently discuss my teaching with colleagues in my department				--
When I revise a course I do library research to make the content up-to-date				++
I keep up-to-date with the developments in the subject area I teach			--	++
When I revise a course I examine teaching and assessment matters to see their appropriateness	--			++
I have a scholarly approach to my work				
I maintain professional contact with colleagues met at professional meetings	++			
I maintain professional contact with colleagues interstate and overseas	++		--	

+ = significantly higher occurrence of activity
- = significantly lower occurrence of activity
++ = significantly higher proportion of staff strongly agreed with statement
- - = significantly higher proportion of staff strongly disagreed with statement

Serow (2000): Table 1. Summary of characteristics of faculty participants

Field	Rank	Funding	Account
Applied sciences	Professor	> $1 million	Complementary
Applied sciences	Associate	> $1 million	Complementary
Natural Sciences	Associate	> $1 million	Complementary
Applied sciences	Professor	> $100,000	Neutral or Undecided
Natural Sciences	Professor	> $100,000	Competitive
Natural Sciences	Assistant	> $100,000	Competitive
Applied sciences	Professor	> $100,000	Neutral or Undecided
Applied sciences	Associate	> $50,000	Complementary
Behavioural sciences	Associate	> $50,000	Neutral or Undecided
Applied sciences	Associate	> $50,000	Neutral or Undecided
Natural Sciences	Professor	> $50,000	Complementary
Natural Sciences	Associate	> $1,000	Competitive
Natural Sciences	Professor	> $1,000	Competitive
Applied sciences	Professor	> $1,000	Neutral or Undecided
Natural Sciences	Associate	> $1,000	Competitive
Behavioural sciences	Associate	0	Competitive
Behavioural sciences	Associate	0	Competitive
Natural Sciences	Professor	0	Competitive
Applied sciences	Assistant	0	Complementary
Natural Sciences	Professor	0	Competitive
Behavioural sciences	Professor	0	Competitive
Natural Sciences	Associate	0	Competitive
Applied sciences	Professor	0	Competitive
Natural Sciences	Associate	0	Competitive
Applied sciences	Professor	0	Complementary
Natural Sciences	Associate	0	Competitive
Natural Sciences	Professor	0	Competitive
Natural Sciences	Professor	0	Neutral or Undecided
Behavioural sciences	Associate	0	Neutral or Undecided

Smeby (1998): Table 1. Percentage of faculty who thought their teaching was affected by their research 'a great' or 'some extent'. By teaching level and field of learning.

	Hum	Soc	Nat	Med	Teach
Undergraduate level	67	59	47	42	56
Major subject level	94	94	94	79	96
PhD level	95	97	98	97	99

Smeby (1998): Table 2. Percentage of faculty who thought their teaching to 'a great' or 'some extent' gave positive impulses to their research. By teaching level and field of learning.

	Hum	Soc	Nat	Med	Teach
Undergraduate level	45	42	26	22	19
Major subject level	85	80	50	72	60
PhD level	90	93	89	93	90

STUDENT PERCEPTIONS

TEACHING AND RESEARCH: STUDENT PERSPECTIVES AND POLICY IMPLICATIONS
Studies in Higher Education, ***23*** *(2), 1998*

Alan Jenkins
Tim Blackman
Roger Lindsay
Renee Paton-Saltzberg

ACADEMIC RESEARCH AND TEACHING QUALITY: THE VIEWS OF UNDERGRADUATE AND POSTGRADUATE STUDENTS
Studies in Higher Education, ***27*** *(3), 2002*

Roger Lindsay
Rosanna Breen
Alan Jenkins

The two studies evaluate student perceptions of staff research on learning. The authors use focus groups of undergraduate and postgraduate students from a range of disciplines at Oxford Brookes University. Student perceptions on the effects of lecturer research on learning are evidently relevant to the debate, but this is one of only a few studies which have investigated the issue by asking students directly.

The Lindsay *et al* study reports a quantitative analysis of the data discussed by Jenkins *et al* which focused on undergraduate students, and then reports both quantitative and qualitative data from a new investigation which replicates the original study using postgraduate students.

The method used was to tape-record focus group discussions with undergraduate students in eight subject groups, and post graduates in nine subject groups, with 4-6 randomly drawn students per focus group. The disciplines were chosen to ensure coverage of both academic and professional areas, and subjects which got different ratings in the Research Assessment Exercises (RAE) of 1992 and 1996. The focus group discussions were prompted by a common set of questions which explored whether and how students experienced staff research as part of the curriculum in that discipline.

A method of quantitative analysis was sought primarily to provide numerical scores that could be statistically related to RAE ratings which was used as a proxy for the quantity and quality of research activity in departments. Each sentence was categorised as positive towards research, negative towards research or irrelevant. Because the focus group context encouraged free discussion, most of the sentences produced were classified as irrelevant.

The undergraduate sample: Level of research activity was controlled by using two areas of study gaining relatively high ratings (3 or 4) in the 1992 RAE assessments, two areas gaining moderate ratings (2) and four areas earning relatively low ratings (1). Lindsay *et al* observe that the average percent of sentences categorised as positive with respect to research (19.5%) is almost double the average percentage of research-negative sentences (9.8%). A χ^2 test showed that the number of positive sentences was significantly greater than the negative sentence frequency for all disciplines. The ratio of positive (average 14.5%) to negative sentences (average 7.8%) about teaching is observed to be about the same. Table 1 shows the

relationship between RAE rating and the number of positive and negative *dialogue contributions* related to research and teaching. In this case a sentence was only included in the count if it was judged to make a new and substantive contribution to the dialogue. Observing values for research related comments it is seen that as RAE rating increases so do both the perceived benefits and the perceived disbenefits. While the association with RAE rating remained significant for teaching-related dialogue contributions, it was much weaker than for research. There seems to be no consistent relationship between positive dialogue contributions and RAE rating. Negative dialogue contributions, however increase in frequency as RAE ratings increase.

Postgraduate sample: Postgraduates produced 177 dialogue contributions, which were positive about research, and only 45, which were negative. When attention is restricted to comments about the impact of research and researcher attributes on teaching 107 comments were positive compared with only 37 negative comments. Table 2 shows that as RAE rating goes up, the percentage of positive comments about the effects of research on teaching tends to increase, and the percentage of negative comments tends to decrease. This pattern is different to the ones observed for undergraduate students.

Qualitative data: Jenkins *et al* (1998) reported that undergraduates perceived the following disadvantages associated with research of lecturers:

- They were less available to students
- They were sometime preoccupied with research at the expense of teaching and this could have too great an influence on the curriculum
- Research interests could distort the curriculum
- Undergraduate students did not perceive themselves as stakeholders in research

In the postgraduate sample the last stated issue did not arise. Postgraduates made it very clear that they did see themselves as stakeholders. An associated issue that arose from their perception of themselves as stakeholders led to a complaint not observed among undergraduates. They thought lecturer research should be useful/interesting/relevant, and not so academic that it was not related at all to what the students were doing.

The perceived benefits of research reported among undergraduates by Jenkins *et al* were:

- Knowledge currency. The need for research to be cutting-edge and not outdated.
- Competence in supervising project work. Motivation in using methodologies/ approaches used by supervisors in their research.
- Credibility enhancement. The desire to learn from people respected in their fields.
- Enthusiasm/motivation: The perceived enthusiasm, greater knowledge, and improved skills brought to teaching by lecturers actively involved in research

The first two stated categories together also accounted for the majority of positive comments among the postgraduate sample.

Students perceived clear benefits including staff enthusiasm and the credibility of staff. Both samples also associated some disadvantages with lecturer research, including reduced availability and curriculum distortion. Lindsay *et al* conclude that the linked studies they have carried out at undergraduate and postgraduate level appear to reinforce and confirm the conventional wisdom that research does confer benefits upon student learning. They point to the fact that both samples make a

considerably greater number of positive than negative statements about lecturer research activity. Whilst in general the two studies indicate positive effects and attitudes to research at undergraduate level, they also give particular importance of such research to student learning at postgraduate level. This seems to validate Smeby (1998) who analysed staff perspectives in four Norwegian universities and found that faculty thought their research influenced teaching to a greater extent at post-graduate level.

The overall results are not dissimilar to those found by Neumann (1994) who carried out an exploration at a research oriented Australian institution. She found that there were tangible benefits to students of staff research, mainly through students perceiving that their courses were up to date and that staff demonstrated interest in what they were studying. However, many students were also critical of subjects in which a teacher's individual research and research interests were seen to dominate, particularly at the expense of the aims of the course.

Given the nature of the study, the sample size for each discipline is too small. A focus group of 4-6 students can be overshadowed by one or two talkative individuals with strongly held views. The problem could be mitigated either by increasing the number of focus groups (and hence the sample size) in each discipline, or standardising (by a z score or similar method) the number of sentences or comments for each student. This would control for the skewness in the distribution of the number of sentences across the students. Given the length of the sessions, 1.5-2 hours, it is also possible that single issues were harped upon for long periods of time, skewing the number of positive or negative sentences in the undergraduate sample, where the results were not controlled for discourse contribution, except when linking with RAE ratings. The undergraduate analysis, where all sentences or dialogue contributions which referred to research or teaching were tabulated, also suffers in comparison to the postgraduate data where only sentences referring to the effects of research upon teaching were counted.

While the study accounts for variation between different disciplines it does not control for the disparity within each discipline. It has been demonstrated in other studies (Kremer: 1991) for example that being an assistant, associate, or full professor has an effect on research productivity.

The size of the faculty in each discipline is also likely to have an effect on the reliability of the views of students. It is likely that in disciplines where there are a fewer number of faculty, students in the focus group have taken courses with most of them and therefore can give an account which more accurately reflects the linkage between the research done in the discipline and student learning. The view in such cases would be more reliable than in disciplines where the faculty size is large and students may actually not have had any course with a significant number of academics.

Since none of the departments included in the study are rated 5 or 5* in the RAE ratings, researchers from the high end of the distribution are not even included in the analysis.

The actual link between teaching and research is not measured since the effect of teaching on research is ignored.

Since the results are from a limited number of disciplines, in one institution only, and based on a limited sample of students, the results cannot be generalised with any degree of confidence.

References:

Kremer, J F (1991) Identifying faculty types using peer ratings of teaching, research, and service, *Research in Higher Education*, **32** (4), pp 351-61

Neumann, R (1994) The teaching-research nexus: applying a framework to university students' learning experiences, *European Journal of Education*, **29** (3), pp 323-39

Smeby, J-C (1998) Knowledge production and knowledge transmission: the interaction between research and teaching at universities, *Teaching in Higher Education*, **3** (1), pp 7-20

THE TEACHING-RESEARCH NEXUS: APPLYING A FRAMEWORK TO UNIVERSITY STUDENTS' LEARNING EXPERIENCES

European Journal of Education, ***29*** *(3), 1994*

Ruth Neumann

Neumann interviews a range of students from first year undergraduates to postgraduates in a large Australian research university to gather their views about teaching and learning experiences. She argues in her review of past literature that no researchers have directly examined the teaching and research nexus with a focus on student views. Given that students are the recipients of university teaching, they are the most important group to consider in the examination of the teaching research nexus. Since this paper, studies replicating the methodology have been attempted (see Jenkins *et al:* 1988, Lindsay *et al:* 2002).

Neumann's student study comprised of eight first year students, three third year students, six students completing their fourth year and four students in their fifth or more year of study. Nine of the students had done or were doing honours and five were PhD students. The students participated in a 60 to 90 minute interview in which they discussed their various subjects and the teaching and learning experiences they had encountered.

Neumann's framework of analysis depicts the teaching research nexus as a multileveled relationship between teaching and research operating on three levels.

The tangible nexus: This refers to the communication of advanced knowledge and research skills to students. At this level teaching serves not only as a form of dissemination of the latest knowledge, especially in those areas where knowledge is developing too quickly for text books to keep pace, but also for academics to pass on the research skills and techniques necessary for study at an advanced level. Neumann observes that this level of connection was readily identified in nearly all interviews. Students across all disciplines and all levels of study discussed subjects they had enjoyed because the lecturer was at the forefront of knowledge. And the students were taught techniques which the lecturer had used in his/her research. This appeared to operate more strongly at the senior undergraduate levels particularly in the sciences and social science.

Neumann also reports however, that many students were critical of subjects in which teachers' individual research and research interests were seen to dominate particularly at the expense of the aims of the course. Yet there was also criticism of the subjects where the content was seen to be stale. References were made to dull and old textbooks, teachers with book wisdom but no first hand knowledge, and lecturers simply reading from textbooks.

The intangible nexus: Neumann refers to this as a more subtle interplay between research and teaching i.e. imparting to students a questioning critical approach to knowledge as well as a positive attitude to learning. She observes that this connection was also readily identified in virtually all interviews. Students' remarks reflective of this viewpoint included:

She was highly motivated, she absolutely loved her subject matter, so it really came out in her lectures and when it got to the practicals she used to bring in a lot of research grant stuff that she was using. And so you would see what was really going on and the techniques they were using in their researching at the time. So it really

made you see what was going on at that level, you were not just learning the course for the sake of obtaining a degree.

He was conveying the present day excitement of it, mostly he was conveying his interest and excitement in his field which was good, it made a lot of people think more about what we were being told, it was not just like getting another telephone book full of facts.

In most cases in the interviews the academics talked about were mentioned by name, hence it was possible for Neumann to collect information about their research activities independently of the interviews. She found that more than three quarters of the students discussed subjects which they had enjoyed where their teacher was described as being very good and was also an active researcher. In the discussions of the influence of individual teachers on students' subjects selection and their enjoyment of learning only one student selected among his good teachers an academic who was known to be a teacher only.

Neumann notes though it can not necessarily be inferred from this that all active researchers are good teachers. For in discussing subjects which were found to be unenjoyable and where the teaching was considered poor, about a quarter of the students referred to academics who either let their own research dominate the curriculum or were uninterested in spending time with their students because they were perceived to be more interested in spending time researching.

The global nexus: This refers to the direction given to course offerings by departmental research activity. Neumann discovers that few examples of this form of connection were found among the undergraduate student interviews, but it was well illustrated by the doctoral students and those considering honours. E.g. She quotes the case of a student who had been fascinated by the excitement of science, however decided against honours in the subject because the research direction and style of academics in that department were primarily taxonomic and therefore largely descriptive. The student ended up selecting a subject where the academics research styles that suited her more analytical interests.

An important issue Neumann finds arising from the interviews is the importance to students of the opportunity of seeing their teachers as real people and to be able to glimpse what they do, how and why. Nearly all of the more advanced students commented on how much more enjoyable third year study was compared with first year. The opportunity to see teachers as people, which was considered important, was not really possible in large classes.

As Neumann concludes it would appear that the opportunity for interaction with teachers is an important component of the teaching research nexus. This is a very important point because it suggest that lower levels of associations found between the activities of research and teaching at lower levels in several studies do not prove that there is not a complementary role between the two activities. Rather it could be the structure of the university system where classes at lower levels are often very large in size which prevents interaction between teachers and students. Thus, particularly inhibiting more intangible tacit aspects of knowledge and learning.

Neumann's findings are backed up by Jenkins *et al* (1998) and Lindsey *et al* (2002) who also demonstrate strong positive student perceptions of staff research. These studies also observed similar disadvantages in that research oriented teachers tended to be less available to students and were often preoccupied with their research at the expense of their teaching.

It should be borne in mind that the number of participants involved (28 student case studies) does not allow generalisations to be made.

References:

Jenkins, A J, Blackman, T, Lindsay, R O and Paton-Saltzberg, R (1998) Teaching and research: student perceptions and policy implications, *Studies in Higher Education*, **23** (2), pp 127-41

Lindsay, R, Breen, R and Jenkins, A (2002) Academic research and teaching quality: the views of undergraduate and postgraduate students, *Studies in Higher Education*, **27** (3), pp 309-27

Lindsay *et al* (2002): Table 1. Percentage of dialogue contributions classified as positive (+), negative (-) or irrelevant (I) with respect to teaching and research used by undergraduate students in focus group discussions and categorised according to the 1992 and 1996 Research Assessment Exercises.

RAE rating	Research +		Research -		Research I		Teaching +		Teaching -		Teaching I	
	1992	1996	1992	1996	1992	1996	1992	1996	1992	1996	1992	1996
High (%)	34	40	21	13	45	47	20	30	14	16	66	54
N	2	4	2	4	2	4	2	4	2	4	2	4
Medium (%)	25	31	9	9	65	60	17	22	10	11	73	67
N	2	3	2	3	2	3	2	3	2	3	2	3
Low (%)	18	9	8	3	74	88	15	29	7	7	77	64
N	4	1	4	1	4	1	4	1	4	1	4	1

Lindsay *et al* (2002): Table 2. Percentage of dialogue contributions produced by postgraduate students in focus group discussions categorised as positive (+), negative (-) and irrelevant (I) with respect to teaching and research and grouped by outcome of the 1992 and 1996 Research Assessment Exercises (percentages in brackets).

RAE Rating	Impact of research on teaching +		Impact of research on teaching -		Impact of research on teaching I	
	1992	1996	1992	1996	1992	1996
High	38	26	8	4	1994	925
(%)	(1.9)	(2.7)	(0.4)	(0.4)	(97.7)	(96.9)
N	2	4	2	4	2	4
Medium	58	34	17	6	2447	1506
(%)	(2.3)	(2.2)	(0.7)	(0.4)	(97.0)	(97.4)
N	2	3	2	3	2	3
Low	1	37	12	27	502	2512
(%)	(0.2)	(1.4)	(2.3)	(1.0)	(97.5)	(97.5)
N	3	1	3	1	3	1

EXAMINING THE TIME SPENT ON RESEARCH & TEACHING AND OTHER STUDIES

A PRELIMINARY INVESTIGATION INTO THE CORRELATION BETWEEN TEACHING AND RESEARCH IN THE FLORIDA STATE UNIVERSITY SYSTEM
Management Research News, ***21*** *(9), 1998*

H. Eugene Baker III
Homer Bates
J. Victoria Garbacik-Kopman
John McEldowney

THE RELATIONSHIP BETWEEN RESEARCH PRODUCTIVITY AND TEACHING EFFECTIVENESS OF FINANCE FACULTY
Journal of Business Education, ***1****, Proceedings 2000*

Homer Bates
Cheryl Frohlich

The Florida Legislature enacted a State University System Teaching and Departmental Incentive Program (TIP) in 1993 to "recognise, promote and stimulate high quality and productive teaching". The program was designed to reward faculty members for undergraduate teaching effectiveness by raising their base salary by $5000. Faculty were eligible to apply for these awards if they had taught undergraduate classes for each of the previous three years, and if they were expected to continue having undergraduate teaching responsibilities in the future. Applications for the awards were considered by a committee comprised of faculty and university administrators. This committee made recommendations to the president of each university who selected the final recipients of the TIP awards.

The program presents an opportunity to examine the research records of faculty identified as outstanding teachers and to contrast these with the research records of comparable non-recipients. Baker *et al.* (1998) match the publication record of all management faculty members who received these awards with non-recipients. Bates & Frohlich (2000) attempt the same for finance faculty members.

There were 797 recipients of the TIP awards in the nine State Universities of Florida, based on their teaching quality during the three academic years from 1990-91 through 1992-93. Baker *et al.* (1998) identify 17 management faculty members in the State University System as recipients of the award. Five are Professors, eight are Associate Professors and four are Assistant Professors. All of these were PhD holders. Each award recipient is matched, by the authors, with a non-recipient at the same university, who possessed the same rank (i.e. Assistant, Associate or Professor), and who had received his/her doctoral degree within three years of the recipient. 15 pairs are consequently established. Matches are not possible for two of the recipients because at their University no one of the same rank possessed the same degree.

Bates & Frohlich (2000) identify 13 finance faculty members as recipients of the TIP awards. Five of these are Professors, five are Associate Professors and three are Assistant Professors. Employing similar matching criteria to Baker *et al*, the authors manage to establish 12 pairs. In this case the year of attaining their doctoral degree between recipients varies by no more than five years (instead of 3 in the case before).

The TIP award was based on teaching effectiveness for three academic years (1990-91, 1991-92, 1992-93). The authors examine research publications for the same time period. Both papers define published research in two ways:

- The total number of articles published during the 37-month period (August 1, 1990 through August 31, 1993)
- The total number of articles published during the 37-month period in "major" journals

Baker *et al.* examine publications in management journals. They define "major" journals as one of the top 20 as ranked by Extejt and Smith (1990). Bates & Frohlich examine publications in finance journals. Their definition of "major" journals is based on the 16 principal journals identified by Borokhovich *et al.* (1995).

Both papers use a non-parametric sign test to test two hypotheses:

H1: Overall research record of recipients = Overall research record of non-recipients
H2: Major publication record of recipients = Major publication record of non-recipients

If the number of research publications of the recipient exceeded the matched non-recipient a (+) is given. If the reverse is true a (-) is given. If the research records are the same the observation is excluded. The test statistic is the number of (+) signs.

Baker *et al*. find for management faculty that the overall research record of the recipients ranged from 0 publications for six faculty members to a high of 19. For the non-recipients the research record ranged from 0 for nine faculty members to a high of 6. The publication records in major publications ranged from 0 to 1 for both groupings.

Using the sign test, the overall number of publications resulted in seven instances where the recipient's publications exceeded the non-recipient's; four cases in which the non-recipient's publications exceeded the recipient's; and four occurrences of equality. The value of the test statistic shows no significant difference between the number of (+) and (-) signs.

Similarly for the second hypothesis, Baker *et al.* find 2 instances where the recipient's major publication record exceeds the non-recipient's and 13 cases where there is a tie. They conclude that the hypothesis that the research records are the same cannot be rejected.

Bates and Frohlich discover for finance faculty that the overall research record for the recipients ranged from 0 publications for seven faculty members to a high of 3. For the non-recipients the research record ranged from 0 for five faculty members to a high of 7. The publication records in major journals ranged from 0 to 3 for recipients and from 0 to 6 for non-recipients.

Employing the sign test, the overall number of publications result in 3 (+)s, 7 (-)s and 2 exclusions. The authors employ a Wilcoxon Signed-Rank test at the 5% confidence level and conclude they cannot reject H1.

For the second hypothesis, Bates and Frohlich find 2 (+)s, 5 (-)s and 5 ties. Once again they determine that the hypothesis that the research records of the two groupings are the same cannot be rejected.

There are several failings in these two studies which need to be considered when examining these results. Firstly the sample sizes are very small. There are only 15 matched pairs in the first study, and 12 in the second. It would be very difficult to generalise these results given the limited number of observations. If all 797 recipients of the TIP awards were subjects of study, the results would be far more robust.

Secondly research productivity is defined through quantity (the number of publications) and not quality. This is a common feature of many studies trying to quantify the link between research and teaching. Ratings of research productivity by citation counts come closest to a measure of research quality.

Third, by identifying teaching quality solely as receiving a TIP award, the authors examine only the high end of the teaching quality distribution. Generalising the results to be applicable for all or most faculty, through examining the tail end of (what is presumably) a normal distribution is not advised.

Also it is possible that rank and year of degree (with a five year variation) may not be adequate enough controls for correct matching; it is not obvious what measure of good teaching was employed by the committees/presidents to hand out teaching awards; there is no examination of causality in either paper; generalising the results of the link between research and teaching obtained from examining just finance and management faculty is not advised.

References:

Borokhovich, K A, Bricker, R J, Brunarski, K R, and Simkins, B J (1995), Finance research productivity and influence, *The Journal of Finance*, **50** (5), pp 1691-1717

Extejt, M and Smith, J E (1990), The behavioural sciences and management: an evaluation of relevant journals, *Journal of Management*, **16**, pp 539-551

WHAT MATTERS MOST: TEACHING OR RESEARCH? EMPIRICAL EVIDENCE ON THE REMUNERATION OF BRITISH ACADEMICS

Centre for Economic Policy Research
Discussion Paper Series no. 2628, November 2000

Rob Euwals
Melanie Ward

The paper examines the impact of productivity on pay within academia, drawing upon a detailed data set of academics from five old, established universities. It aims to provide some insight into the determinants of academic remuneration, with a view to highlighting the skills currently most valued within the profession. The analysis also investigates the relationship between teaching and research skill.

The data used in the paper comes from a postal questionnaire sent out in 1995/96 to academics in a cross section of five distinguished universities: Aberdeen, Dundee, Glasgow, Heriot-Watt and St. Andrews. The response rate achieved was 30%, reasonably high for this sort of study. Of the 878 academics who responded, part time faculty and those who were paid on the clinical scale were dropped. This left 635 observations for the analysis. Data were weighted for non-response at a faculty level by sex. We would not expect this to be an adequate enough control to eliminate non-response bias.

The data set contains information on an individual's gender, nationality, length of working experience, length of current tenure, length of time out of the labour force, job rank, university, discipline of faculty (arts, social science, engineering, science, and medicine), research publications, grants and teaching skills.

Such detailed data is not available in other existing literature for the United Kingdom. At face value, the list of variables seems adequately comprehensive. There are problems though with the measures of both research and teaching. The split of publications into books, chapters and papers is valuable. A point at times ignored in other studies. The measure of number of publications does not, however, allow us to answer the question "are good researchers also good teachers". To measure the quality of research, a superior variable would be citation count. This is a common blemish in studies in the area. The bigger problem is the measure of teaching. Euwals and Ward describe their variable as:

Teaching skill = 1 if skilled teacher (based on student's evaluations), = 0 otherwise

No more elaboration is given. Apart from it not being obvious what the measure of teaching skill actually is, a further problem arises from a large number of academics in the sample (80.2%) reporting they are not skilled teachers according to the criterion. By restricting the range this may depress the existing relationship in the results.

In the first step of their analysis Euwals and Ward investigate the hypothesis that productive researchers are also the best teachers. They run a probit regression of teaching skill against individual, job and research productivity characteristics to test this relationship. Their results are presented in table 1. The omitted or base category is female, non-UK citizen, Glasgow University, science faculty, non-doctorate holder. Euwals and Ward find no strong support for their hypothesis. Among the measures of research, only the number of chapters has a positive and significant effect on the probability of being a skilled teacher. Experience is revealed to have a significantly positive effect, and there are some differences in teaching skill by faculty.

In the second part of their analysis Euwals and Ward investigate the determinants of academic salary with a view to highlighting the skills currently most valued within the profession. Is it the case for example, that teaching skills are highly rewarded, or is success purely determined by publication record? This may not seem directly consequential for the purposes of the current review. However, it has been often argued that promotion for faculty on the basis of research alone sends a signal to young academics to reduce the time and effort spent on teaching to a minimum so that they can get on with churning out publications. It provides a clear incentive for faculty to neglect teaching in favour of research.

Euwals and Ward's simulations reveal teaching skill to have a sizeable impact on pay. While they find an additional paper increases the expected wage by 0.3 to 0.4 percent, remarkably changing from an unskilled to skilled teacher increases the reward to between 12 and 15 times this effect. Also the probability of getting a professorship is seen to be affected substantially by teaching skill. These results are surprising in the light of the commonly held view that rewards are tied to research and not significantly to teaching.

The biggest problem with this analysis, and especially that of the linkage between research and teaching is the measurement of teaching skill. Apart from the meaning of the measure, and the likelihood that it could depress the results, as discussed before, there is a further issue. It is most likely that teaching skill follows a normal distribution over the population of academics. To regard it as a dichotomous variable, as in this study, is an over-simplification. With measures of both roles, especially teaching, being flawed it is apparent the results of the linkage found in he study will be imperfect.

THE ACADEMIC RESEARCH-TEACHING NEXUS IN EIGHT ADVANCED-INDUSTRIALIZED COUNTRIES
Higher Education, **34**, 1997

Esther E. Gottlieb
Bruce Keith

The study uses the Carnegie International Survey (CIS) of the academic profession to compare research and teaching activities in eight countries: the former West Germany, United Kingdom, Sweden, United States, Australia, Israel, Japan, and South Korea. The data collected is for the time period 1991-93. The CIS collected data on another six countries also. In order to make the teaching-research relation the central comparison of the study, and not the country specific system, Gottlieb and Keith elected to drop Brazil, Chile, Mexico and Russia. These countries had the lowest orientation toward research among the sample of 14 countries. Hong Kong was dropped for being a city-country and therefore having a different system. The Netherlands was dropped because some essential variables were found missing from their data set.

The data set for the 8 countries analysed in the study covers a total of 13,984 academics. Gottlieb and Keith group the respondents into research oriented and teaching oriented categories according to their self-reported primary preference. More than 90% of research oriented (RO) respondents in each country declared that they were required to do research. Almost all the RO respondents in the UK (96%), Israel (99%), US (99%), Australia (98%), and Sweden (96%) were currently active in researching. The proportions were lower for West Germany (87%), South Korea (74%) and Japan (65%). Interestingly the United Kingdom is the only country in the sample where the teaching oriented cadre is greater in size than the research oriented group. Almost 79% of all teaching oriented (TO) faculty in the sample claim that they are required to do research, and 73% are currently active.

The major part of the study is spent on examining time investments, the work load, of academics. Time investments in the study are represented in the study by two variables:

1) the mean weekly hours spent on teaching, which is defined as “preparation, classroom instruction, advising students, reading and evaluation of students’ work”
2) the mean weekly hours spent on research which is defined as “reading literature, writing, conducting experiments, and field work”

when classes are in session.

Gottlieb and Keith regress weekly hours spent of research activities on selected characteristics. And then do the same for weekly hours of teaching activities. The results of their regressions are included as table 1 and table 2. Both sets of estimations are linear standard multiple regression models which express the number of hours of either activity as a function of country, sex, rank, academic orientation, course load, class size, number of publications during the previous 3 years, and hours spent on alternate activity. In both sets of regressions the omitted or base category is teaching oriented, American, female faculty member. This is what the Constant term indicates.

Interestingly it can be observed in the first set of regressions that faculty members in the United Kingdom spend significantly fewer hours on research than their

counterparts in the United States or in fact any of the other countries in the sample. Those maintaining an academic orientation that leans toward research are likely to spend 5 more hours per week on research than their teaching oriented counterparts. Similarly, as the number of articles published during the previous three years increases, the amount of hours spent on research also increases. For every five articles published, time spent on weekly research activity increases by about one hour. This substantiates Hattie & Marsh (1996) who find in a review of literature that time on research is positively related to the number of articles published. Males spend an average of almost 2 more hours per week on research than their female counterparts.

Observing the second set of regressions faculty in Australia, South Korea, West Germany, and the UK spend significantly more time on teaching related activities than the faculty in the US. As expected faculty members who are oriented toward research spend considerable less time on teaching activities than their counterparts of teaching orientation (about 2.4 hours less per week). Similar to the previous set of results, for every five articles published during the previous three years time devoted weekly to teaching activities goes down by approximately an hour. Finally when knowledge of faculty members' sex is added to the equation, Gottlieb and Keith find that males spend significantly less time on teaching than do their female counterparts, with an average of 1.75 fewer hours per week.

The relationship between time on research and time on teaching is confusing. Hattie & Marsh (1996) in their review of literature located 14 correlations between time on teaching and time on research, with values ranging from -.46 to .19, and an overall correlation -.17. Gottlieb and Keith, however, find a positive relationship between time spent on either activity. This appears surprising. Time is a finite resource. It would be expected that as an academic spent more time on teaching, he/she would have to reduce the time spent on research (and vice versa). The contrary finding in this study could be explained in several ways. First it is possible this indicates the sample can be grouped into two categories: hard workers who spend more time on both research and teaching and have a longer average working week; and lazy workers who work less overall hours per week and less hours on each activity also. A second explanation could arise from the complementary nature of the two activities. It could be optimal for academics' research productivity to spend time on non-research roles (teaching). Only increasing time spent on research would not allow an academic's research to benefit through other particular avenues which would only be possible through teaching (example: student feedback, clarification of thinking etc.).

A third (and perhaps most plausible) explanation is the presence of omitted variable bias. The consequences of omitting time spent on service/administration activities (found to be highly significant by Oliveras *et al*: 2003) could lead to the coefficient of time spent on research or teaching (whichever is the independent variable given the regression) being biased as well as inconsistent.

The coefficient of determination, R^2, is not particularly high in either sets of regressions. About 25% of the total proportion of variance is explained in research hours by the selected variables the results of which are shown in table 1. The explained variance is lower, 15%, in the second model.

Gottlieb and Keith further analyse how respondents perceived the relationship between teaching and research obligations by examining the responses to three questions. They analyse the competing and complementary aspects of teaching and research separated for RO and TO faculty. They observe the following responses:

1) *teaching is influenced by my research commitments:* 43% of RO faculty think their research obligations have complementary influence on teaching compared to 29% of TO faculty. 48% of TO faculty think research has no influence on their teaching, compared to 30% RO.
2) *research is influenced by the number of course I am assigned to teach – course load:* the highest proportions of both orientations perceive course load as having a negative influence on research
3) *research is influenced by the number of students in my course – student load:* the highest proportions of both groups (48% TO and 56% RO) see no influence on research.

As Gottlieb and Keith themselves state the measures used do not capture all aspects of teaching or research performance, nor do they report on quality of teaching, research or publications. Also it should be noted when looking at findings of the aggregate model that there is likely to be greater variations within countries than is found to exist between them because Gottlieb and Keith have chosen to analyse systems which are more similar than different.

The data set includes information on research publications. It would be of value to extend the analysis conducted to examine how time spent on teaching affects research productivity. Support for or refutation of Mitchell and Rebne's (1995) results could be provided. That study finds that the relationship between research and time spent on teaching is non-linear, with teaching loads of up to one course per semester being facilitative of research productivity. This aspect of the relationship has not been examined elsewhere.

References:

Mitchell, J E and Rebne, D S (1995) Nonlinear effects of teaching and consulting on academic research productivity, *Socio-Economic Planning Sciences*, **29** (1), pp 47-57

Oliveras, E, Blake, J and Dowds, J (2003) The interaction between teaching and research: perceptions of accounting academics in Spain and the United Kingdom, *Working Paper Series Universitat Pompeu Fabra*

NONLINEAR EFFECTS OF TEACHING AND CONSULTING ON ACADEMIC RESEARCH PRODUCTIVITY

*Socio-Economic Planning Sciences, **29** (1), 1995*

John E Mitchell
Douglas S Rebne

The study tests the proposition that moderate amounts of faculty times spent on consulting and teaching are facilitative of research productivity and establishes explicit values for the point at which constancy and teaching time cease to have this facilitating affect. This is done by fitting continuous piecewise-linear regression models to 1980 data, from a sample of 5605 U.S. faculty.

Mitchell and Rebne consider two ways of explaining the effects of time spent in alternative roles such as teaching and consulting on research productivity.

1. *Complementary roles:* This is the notion that roles may be mutually reinforcing. In this case time spent on a given academic role would be predictive of success in that role as well as other occupational roles associated with it. This suggests non-linear relationships between time spent in various academic roles and individual performance with moderate activity in alternative roles being associated with highest levels if performance.
2. *Time scarcity theory:* The scarcity theory of role behaviour assumes that commitment of time and energy to one role must come at the expense of success in another. This position suggests that time spent in any role except research would be negatively in a linear related to research performance.

Mitchell and Rebne argue that light to moderate teaching duties as well as time spent in preparation might be expected to enhance research performance insofar as keeping abreast of developments in field would serve both roles. However, a heavy emphasis on teaching should have a more negative effect as faculty find themselves without sufficient time to devote to research. With regard to consulting activity they hypothesise positive effects to arise with new research problems being found through the applied problem solving process. Again however, high levels of activity in this role should detract from research performance.

Mitchell and Rebne's position seems to be that complementary role theory holds but that it inherently assumes that at higher levels of activity time scarcity inevitably emerges as a negative factor. Their focus then is on the effects of low to moderate alternative activities and the identification of the point at which further such efforts become dysfunctional to performance.

The data they use to test their hypothesis was constructed in 1980 by the UCLA higher education research institute. It covered a nationally representative sample of 98 US colleges and universities. While the authors admit that the data set is old, they argue that it is the most comprehensive of its kind. The sample for the purpose of this study covers a total of 5605 academics.

Mitchell and Rebne measure research by the self-reported publishing activity. The respondents in the survey had been asked to report the total number of works, refereed articles, books, monographs published or accepted for publication in the past two years. Using these the authors construct a basic productivity index using the following categories:

1 = 0 publications in past two years

2 = 1 to 2
3 = 3 to 4
4 = 5 to 10
5 = more than 10 articles, books or monographs

The response codes were converted to mid-point values in order to approximate actual output. Brew and Boud (1995) have argued that ratings of research productivity by citation counts come closest to a measure of research quality. Mitchell and Rebne understand that their measure is a purely quantitative one. However, they argue that such measures have been found to correlate well with qualitative measures such as citation scores and assessment of general quality (Cole:1979 and Long:1978). A shortcoming of the basic productivity index is that it provides no way of assigning different weights to articles, books, edited volumes etc.

Activity in various academic roles was measured by reports of average weekly hours spent during the current academic term. Response categories consisted of four-hour intervals. Later while fitting their model the authors convert these intervals into their midpoints.

One limitation of their analysis that Mitchell and Rebne pick up is their need to assume that levels of activity were fairly consistent for the years preceding the survey. Given the often lengthy period between the time research is initiated and the time it reaches print, the effect being sought is more properly that of a lagged variable reflecting activity in the mid 1970s. Consideration of the lagged period of research has not been picked up in other studies.

The general forms of relationship between time spent on academic roles and research productivity are shown in Table 1. As should be expected, the highest level of productivity was achieved by those spending the most time on research and writing. The data also indicates the relationships with time spent on teaching, administrative work and consulting are non-linear and in keeping with the common complementary role argument that roles may be mutually reinforcing, except at levels of high activity. The fact that only so much time and energy is available to any one person and commitment to one role prevents the development of excellence in others.

Mitchell and Rebne use a piecewise-linear regression technique in the context of the non-linear relationship. This approach consists of two linear pieces or segments where the function changes its slope at the threshold value. More specifically it is assumed that research productivity increases linearly with time spent on teaching until a certain threshold level after which it either decreases linearly with teaching, or increases at a much lower rate. The authors consider that the preliminary analysis of the data (Table 1) indicates that the three best predictors of research productivity are time spent on research, time spent consulting and time spent teaching. They restrict focus to these variables fitting a model in which these are the independent variables and the number of papers published in the last two years is the dependent variable.

Oliveras *et a*l (2003) demonstrate using data for accounting academics in the UK and Spain that in fact service activities have a high and significant impact on the available time for research activities. While this is not obvious in Mitchell and Rebne's data set, it is still possible their model suffers from excluded variable specification bias.

The proportion of explained variance in Mitchell and Rebne's final model is 27.4 percent. The results indicate that a researcher can typically expect to produce about one additional paper every two years or each additional six hours of research per

week. For less than 8 hours per week or around one course per semester of teaching the increase in productivity from an hour of teaching is greater than that from an hour of research, with one extra paper per two years resulting for each 3 hours of teaching per week. Teaching more than 8 hours per week has a slightly detrimental effect on research productivity. For small amounts of consulting (less than 4 hours per week) the increase in productivity from an hour of consulting is greater than that of an hour of research, resulting in one extra paper every two year for each hour of consulting per week. Additional consulting does not hurt research productivity but the positive effect is very marginal.

The analysis supports the view that it is inappropriate to regard academic job content in zero sum terms when research productivity is the outcome of interest. Time spent on non-research roles of research and consulting is seen as not negatively related to productivity at all levels of activity. The results indicate that up to four hours per week of consulting and up to eight hours per week of teaching are indeed facilitated of research productivity.

As the authors themselves suggest replication of the present study is advisable since the activity patterns have likely changed since 1980. Apart from likely excluded variable bias and the outdatedness of the data in the study another factor should also be noted. In the study the average work week is approximately 44 hours. The results pertain to this mean. If the average work week for academics is different in the UK we would expect the thresholds points to be different as well.

References:

Brew, A and Boud, D (1995) Teaching and research: establishing the vital link with learning, *Higher Education*, **29**, pp 261-73

Cole, J R (1979) *Free science: women in the scientific community*, Free Press, New York

Long, J S (1978) Productivity and academic position in the scientific career, *American Sociological Review*, **43**, pp 889-908

Oliveras, E, Blake, J and Dowds, J (2003) The interaction between teaching and research: perceptions of accounting academics in Spain and the United Kingdom, *Working Paper Series Universitat Pompeu Fabra*

THE INTERACTION BETWEEN TEACHING AND RESEARCH: PERCEPTIONS OF ACCOUNTING ACADEMICS IN SPAIN AND IN THE UNITED KINGDOM

Working Paper, Universitat Pompeu Fabra, 2003

Ester Oliveras
John Blake
Jack Dowds

This working paper reports a survey of the experience and opinions of accounting academics in two countries, Spain and the United Kingdom. Specifically the study focuses on the following issues: i) the impact of teaching and service on time available for research; ii) the integration of teaching and research and; iii) the perceived value of teaching and research for career success.

Data was collected for the study through an e-mail questionnaire survey. The survey collected information on individual characteristics, the division of time between activities, and perceptions on several statements related to teaching and research roles. In Spain the questionnaire was sent to 750 members of the Spanish Association of Accounting Lecturers. The response rate was 12% (90 usable replies). In the United Kingdom lecturers were identified through accounting department web pages for universities. 1135 questionnaires were sent out, and the response rate was again 12% (136 usable replies).

Given the low response rates, the conclusions drawn from the study should be treated carefully and mainly regarded as orientative. Also it is quite possible that results obtained may be specific to this particular discipline.

The respondents had been asked to distribute all of their working time among the categories of research, teaching, administration and other activities. The responses were analysed using regression analysis. Oliveras *et al* ran linear standard multiple regressions which expressed the proportion of time spent on a role as a function of the proportion of time spent on all other activities. Due to the formation of the data in this way all their models have very high explanatory power. Oliveras *et al* find that the addition of qualitative variables such as gender, nationality or age group were found to be insignificant. This surely is to be expected, given once again the way their models are formulated. It should definitely not be inferred from this that these individual characteristics are not relevant in explaining the amount of time academics spend in various roles.

Interestingly, Oliveras *et al* observe that time used in service activities has a higher impact on the available time for research activities than teaching. The percentage of research time is negatively affected by 0.484 times the percentage of teaching time, whereas it is negatively affected by 0.678 times the percentage of service time. The authors explain this by arguing that teaching and research enrich each other, whereas administrative/managerial responsibilities hardly have any input towards improving teaching or enhancing research.

This finding leads to an important consideration. Even in a situation where complementarity might otherwise be expected to exist, it could be vitiated by time-scarcity. If time spent on administrative duties/paper work has been growing, time spent on research (and therefore research) would suffer significantly. It would be interesting to view a longitudinal series of time spent on services in this light.

Alternatively it could be argued that the interpretation of Oliveras et al is erroneous. The findings could actually be revealing that if teaching load rises, time spent on

research is affected significantly but this effect is mitigated somewhat by academics spending less time on service activities (hence the lower value of the teaching coefficient in absolute terms). However, if administrative load is increased, research activity suffers greatly because time spent on teaching cannot easily be reduced.

The value of the analysis could be improved greatly if data were also compiled on the publication record of the respondents. By examining how time spent on alternative activities affects research (weighted publication scores, or citation counts) support for or refutation of Mitchell and Rebne's (1995) results could be provided. That study finds that the relationship between research and time spent on teaching is non-linear, with teaching loads of up to one course per semester being facilitative of research productivity.

It should be noted that all results reported from the regressions in the discussion above are preliminary, given the study is a working paper. Also regression results discussed in this review are limited to findings quoted in the main text of the paper. The authors were unwilling at this stage to share their full set of regression findings.

The rest of Oliveras *et al's* paper deals with academics' perceptions of the integration of teaching and research, and the relative importance of the roles for career success. The main findings on perceptions of the link are listed below:

- 51% of UK accounting academics claim that students at their university welcome teaching which presents the lecturer's own research. For accounting academics working in Spain the number is 37%.
- In response to the statement "Meeting the curriculum needs of students sometimes leads to interesting research", 49% of Spanish and 34% of UK academics either disagreed or strongly disagreed.
- 70% of UK respondents agreed that research should be shared through teaching. This number is only 29% for Spanish respondents.
- More than 50% of both Spanish and UK accounting academics agreed that all lecturers should also undertake researchers. Less than 20% in both samples disagreed.

While the all encompassing perception (bullet point 4) among accounting academics from both countries seems to indicate that they believe research ensures the lecturer maintains updated knowledge in their field, their are significant differences between academics from either country on the other responses. The perception of the integration of research and teaching is observed to be higher in the UK than Spain. Oliveras *et al* explain that Spain has a tradition of magisterial lectures in which student participation is limited. On the other hand, the teaching system in the UK encourages a higher degree of participation from the student in their own learning process, which might lead to a richer exchange between lecturers and students, and therefore to a wider acceptance of research.

With regard to the importance of teaching and research for career success in their institution:

- 57% of UK respondents and 69% of Spanish respondents agreed that research was more important than teaching for career success
- Validating this response, 60% of UK accounting academics and 89% of their Spanish colleagues disagreed with statement "Good teaching can lead to promotion, even with a weak research record".
- Meanwhile 80% in the UK and 89% in Spain agreed that good research could lead to promotion even with a weak record.

- 53% of UK respondents and an overwhelming 91% of Spanish respondents felt good teaching was not adequately rewarded.

In sum, there is a strong perception that promotion is easier to achieve with a good research record even if the teaching record is weak. The opinions are more dominant in Spain. This is indicative that there may be less reward disparities in the UK than in Spain. Still there is dissatisfaction, among UK academics, with the reward obtained as a result of good teaching. This implies that though the disparities in the reward structure between teaching and research may be less in the UK than Spain it is still significant.

References:

Mitchell, J E and Rebne, D S (1995) Nonlinear effects of teaching and consulting on academic research productivity, *Socio-Economic Planning Sciences*, 29 (1), pp 47-57

THE TEACHING AND RESEARCH RELATIONSHIP WITHIN AN INSTITUTIONAL EVALUATION

*Higher Education, **40**, 2000*

Javier Vidal
Miguel A. Quintanilla

The main issues addressed in the paper are a) whether research should be evaluated from an institutional point of view, and if so, b) which specific aspects of research activities can be evaluated within each institution. The authors' related aim is to determine to what extent the relationship between research and teaching should be analysed together from an institutional point of view. To achieve this, they present an analysis of the relationship between research and teaching. This will be the primary focus of this review. The context is Spanish universities.

Vidal & Quintanilla propose a transference-interference model of analysis for the teaching-research relationship. In simplest terms this means that in some circumstances the relationship would be positive (transference) while in other circumstances it would be negative (interference). To explain:

- *Transference*: If something is good for research then it is also good for teaching, or if something is bad for research, then it is also bad for teaching
- *Interference*: If something is good for teaching it is bad for research, and if something is bad for teaching it is good for research

The concept of transference seems identical to complementary role theory, and the notion of interference is analogous to competitive theory which are terms more widely used in literature to explain the relationship between research and teaching (see Faia:1976 and Mitchell & Rebne: 1995).

Vidal & Quintanilla report that academics at Spanish universities spend 46% of their time on teaching activities, 41% on research, and 13% on administration and other activities (1991 National Institute of Statistics Survey). They also note two other issues:

- First, that the structure of the academic staff at the Spanish universities is designed to fit in with the university's teaching needs, and
- Second that research achievements are more valued than teaching qualifications in the selection and promotion of academics.

The authors use two sources of data to perform their analysis. Firstly they interviewed 36 researchers at a medium-sized research oriented Spanish university. The sample covered a broad range of disciplines, and also a broad variety of research experience. An open interview format was chosen and individual reports were made for each session.

The second source of information was reports from the Spanish National Program for the Evaluation of University Quality. The complete reports include self-assessment and external assessment. In the self-assessment phase, researchers have to consider to what extent their teaching role affects their research activities, and how teaching activities are affected by high quality research projects in the department. Vidal & Quintanilla selected 20 self-assessment reports from 10 different public universities. The sample was screened to cover a broad variety of university types, and a range of departments (different sizes, research quality, and disciplines).

From the qualitative information gathered from both sources, all possible connections between teaching and research were extracted. These were then categorised according to the concepts of transference and interference. The transferences pointed out in the study are as follows:

- Research activity leads to an improvement in teaching quality (but not vice-versa). Moreover, an academic cannot be good without doing research, although a good researcher can be a poor teacher.
- Certain infrastructural means obtained through research projects are also used in teaching activities.
- Research activities contribute to updating the curriculum, positively affecting the most specialised courses.
- If courses are related to research profiles of the teachers, the relationship is favourable.

This last stated idea is observed by Vidal & Quintanilla to be the most common transference. This notion that it is easier to do research if teaching is oriented towards research interests, indicates the positive relation between specialised courses and research. It could further suggest a positive link between research and teaching at higher levels of teaching where courses tend to be more specialised in comparison to more modest association at undergraduate level. Smeby's (1998) finding that a greater percentage of faculty perceive the two roles to be mutually reinforcing at postgraduate than undergraduate level could partly be explained by this.

Also interesting to note is the consensus among Spanish faculty that research activity improves quality of teaching, but that the enhancing relationship does not work the other way. The adoption of this position seems to be unique in the literature of this field.

The interferences observed in the study are listed below:

- Some aspects involved in teaching activities hinder good research. For instance, having to teach several different courses, huge groups of students, having many hours of teaching and also having an unfavourable teaching schedule, reduces the possibilities for research.
- The setting-up of new programs increases the time required for teaching and in consequence decreases research activity.
- Research collaboration with external institutions usually requires travelling, and this affects teaching activities.
- Research (the most specialised) affects the most general and basic courses negatively.

The most common interference found is that it is difficult to carry out research if there is a lot of teaching to be done (bullet point 1). This issue that there is only so much time and energy available to any one person and that commitment to either role prevents the development of excellence in the other, has often been discussed in past literature. And yet as Mitchell and Rebne (1995) demonstrate time-scarcity only starts becoming a factor at high levels of activity. And that the relationship is mutually reinforcing till then. As the Spanish faculty in the sample state: it is teaching "several courses" or having to handle "huge groups" of students that is problematic. A zero-sum relationship between research and activity should therefore not be inferred from the comments in the study.

Vidal & Quintanilla observe that for Spanish academics, almost nobody agrees with the idea that, there is no link between teaching and research in the university context. The analysis fails to resolve however, if this unavoidable relationship is positive or negative. It is possible this indicates that correlation studies, which find no association between research and teaching, are not indicative that there is in fact no linkage, but that the incompatibilities and complementarities tend to cancel each other out. It is also possible, as Mitchell and Rebne (1995) discover (and this would not be picked up by either correlation studies, or proved by qualitative studies of this nature), that the competing nature of the activities (mainly through time scarcity) starts affecting the complementary nature of the roles only when activity levels cross certain thresholds.

Vidal & Quintanilla's study could have helped provide greater empirical evidence by constructing a method of quantitative analysis based on the number of positive and negative discourse contributions, similar to the technique adopted by Lindsay *et al* (2002). While the procedure has its weaknesses it does allow an empirical base to quantify the linkage, its direction, and its extent.

References:

Faia, M A (1976) Teaching and research: rapport or mesalliance, *Research in Higher Education*, **4**, pp 235-46

Lindsay, R, Breen, R and Jenkins, A (2002) Academic research and teaching quality: the views of undergraduate and postgraduate students, *Studies in Higher Education*, **27** (3), pp 309-27

Mitchell, J E and Rebne, D S (1995) Nonlinear effects of teaching and consulting on academic research productivity, *Socio-Economic Planning Sciences*, **29** (1), pp 47-57

Smeby, J-C (1998) Knowledge production and knowledge transmission: the interaction between research and teaching at universities, *Teaching in Higher Education,* **3** (1), pp 7-20

Euwals & Ward (2000): Table 1: Probit for Teaching Staff

Dependent variable	**Teaching skill = 1 if skilled teacher (based on student's evaluations) = 0 otherwise**
Number of Obs.	635
Individual characteristics	
Intercept	*-2.387 (0.340)*
Male	*0.048 (0.166)*
UK Citizen	*-0.087 (0.209)*
Job characteristics	
Experience	*0.110 (0.031)*
Experience2/10	*-0.029 (0.008)*
Tenure	*0.032 (0.027)*
Tenure2/10	*-0.001 (0.009)*
Time-out	*-0.090 (0.061)*
University	
Aberdeen	*-0.155 (0.186)*
Dundee	*-0.158 (0.181)*
Heriot-Watt	*0.296 (0.250)*
St. Andrews	*0.118 (0.190)*
Faculty	
Arts	*0.484 (0.186)*
Engineer	*0.158 (0.224)*
Medicine	*0.403 (0.215)*
Social Sciences	*0.756 (0.190)*
Publications	
Books (weighted)	*0.003 (0.031)*
Chapters (weighted)	*0.076 (0.028)*
Papers (weighted)	*-0.080 (0.059)*
Other	
Grants (weighted)	*0.034 (0.044)*
Having PhD	*0.286 (0.163)*

Figures in parentheses are standard errors

Gottlieb & Keith (1997): Table 1. Regressing weekly hours of research activities on selected characteristics

Variable	
Constant	5.672 (.482)
Australia	-4.081 (.355)
Israel	2.183 (.517)
South Korea	-0.014 (.400)
Sweden	-4.207 (.372)
West Germany	-4.225 (.362)
UK	-5.246 (.354)
Japan	0.895 (.328)
Published articles	0.206 (0.013)
Total enrolment	0.005 (.001)
Total courses	-0.186 (.041)
Teaching/research orientation	5.188 (.124)
Academic rank	0.378 (0.088)
Hours spent on teaching	0.378 (0.009)
Sex of respondent	1.874 (0.249)
R^2	*0.253*

While the study does not explicitly say so, presumably the numbers in the parenthesis are standard deviations. It is indicated repeatedly in the main text of the article that almost all of the variables are found to be statistically significant.

Gottlieb & Keith (1997): Table 2. Regressing weekly hours of teaching activities on selected characteristics

Variable	
Constant	18.614 (.484)
Australia	1.778 (.378)
Israel	-1.671 (.550)
South Korea	3.032 (.424)
Sweden	-3.048 (.397)
West Germany	0.989 (.386)
UK	1.733 (.379)
Japan	-1.343 (.348)
Published articles	-0.199 (.014)
Total enrolment	0.005 (.001)
Total courses	1.046 (.043)
Teaching/research orientation	-2.362 (.140)
Academic rank	0.443 (.094)
Hours spent on research	.043 (.010)
Sex of respondent	-1.75 (.264)
R^2	*0.153*

Mitchell & Rebne (1995): Table 1. Productivity by time spent in academic roles

	Publications, 1978-80			
Time spent in roles:	*None* *0*	*Low* *1-8*	*Medium* *9-20*	*High* *20-45+*
Scheduled teaching	4.01	4.41	2.14	1.66
Teaching preparation	3.89	4.04	2.84	1.70
Advising students	3.45	3.30	3.22	1.22
All teaching (1-3) above	3.16	5.13	4.31	2.46
Consulting	3.17	3.63	3.83	2.93
Administration	2.64	3.54	3.45	_
Committee work	3.11	3.30	3.48	_
Research and writing	0.77	2.06	4.27	6.36

Copies of this publication can be obtained from:

DfES Publications
P.O. Box 5050
Sherwood Park
Annesley
Nottingham
NG15 0DJ

Tel: 0845 60 222 60
Fax: 0845 60 333 60
Minicom: 0845 60 555 60

Produced by the Department for Education and Skills

ISBN: 1 84478 162 3
Ref No: RR506

www.dfes.go.uk/research