7-29-2008

Assessing Research Performance: Implications for Selection and Motivation

Michel Avital
Case Western Reserve University, michel@avital.net

Fred Collopy
Case Western Reserve University

芽|Sprouts Working Papers on Information Systems ISSN 1535-6078

Assessing Research Performance: Implications for Selection and Motivation

Michel Avital
Case Western Reserve University, USA

Fred Collopy
Case Western Reserve University, USA

Abstract
An examination of a sample of 67 studies on research performance and productivity published between 1974 and 1998 identified two main perspectives named: evaluative approach and explanatory approach. The evaluative studies focus on assessing, comparing, and ranking researchers' performance. Though sometimes based on peer reviews, evaluative studies employ mainly bibliometric methods which stem from publication or citation counts. In contrary, the explanatory studies aim at the enablers of research productivity and evaluate their effect on researchers' performance. The explanatory studies examine determinants of research productivity, which have been classified as institutional, financial, collaborative, professional, personality, and demographic factors. A review of the literature found that most of the evaluative studies were conducted at the departmental or the institutional level, whereas explanatory studies were done at the individual level. These findings are counterintuitive because it is generally individuals who are evaluated and rewarded, whereas the controllable enablers of productivity are mainly institutional. Furthermore, although evaluative and explanatory studies are complementary in nature, using both perspectives in one study is surprisingly rare. Additional difficulties and perplexities related to measurement criteria and ranking schemes are identified and discussed.

Keywords: Research Performance Measures; Research Productivity; Research Success Factors

Permanent URL: http://sprouts.aisnet.org/1-14

Reference: Avital, M., Collopy, F. (2001). "Assessing Research Performance: Implications for Selection and Motivation," Case Western Reserve University, USA . *Sprouts: Working Papers on Information Systems*, 1(14). http://sprouts.aisnet.org/1-14

Assessing Research Performance: Implications for Selection and Motivation

Research performance has long been discussed in the scientific literature. Yet, the prevailing performance assessment measures are controversial and no criterion has been adopted as a ubiquitous standard. Although certain measures and ranking schemes have been institutionalized among both researchers and administrators, they are predominantly perceived as the lesser among evils rather than as an impartial yardstick of research performance. The following manuscript draws upon the pool of existing research of academic productivity to identify, reframe, and juxtapose factors and dimensions perceived as critical and significant. A special emphasis is given to research in social sciences and particularly to research in organizational and behavioral sciences.

Two main lines of inquiry have emerged in the development of research performance measures: an evaluative approach and an explanatory approach. The evaluative approach is rooted in 'rational models' and particularly in the school of scientific management. Common to evaluative studies is a retrospective perspective that focuses on the consequences of a scientific effort in an attempt to assess its outputs. The primary objective of evaluative studies is performance evaluation, which is mainly fueled by the need to compare and rank researchers, as well as programs and institutions. A few examples of evaluative studies are Cox & Catt (1977), Endler, Rushton, & Roediger (1978), and Kirkpatrick & Locke (1992). Although each of these studies is based a different approach to the operationalization of research performance, they share a common objective, which is to rank the main departments in a particular academic discipline.

Explanatory studies provide an alternative viewpoint on the examination of research performance. The explanatory approach may be attributed to 'natural models' such as the school of human-relations. Common to explanatory studies is a prospective perspective that concentrates on the antecedents of scientific pursuit in an attempt to evaluate their effect on research performance. The primary objective of explanatory studies is performance prediction that is based on the inputs and the context of research processes. The impetus of the explanatory approach stems from the desire to understand the factors associated with research performance in an attempt to better facilitate an environment that is conducive to creative and productive research. An example of an explanatory study is Jones & Preusz (1993) that examined circumstantial factors associated with research performance such as personal, attitudinal, and environmental factors. Another example is Hancock, et al. (1992) that identified differences between high and low performers based on individual and institutional factors associated with research performance. Both studies attempted to explore questions such as what are the characteristics of a productive researcher, or what conditions are conducive to productive research.

An a priori viewpoint taken on research performance, either evaluative or explanatory, plays a major role in the selection of research methodologies, has a direct effect on their consequent operationalization, and confines its possible insights and implications. The main features of the two approaches are contrasted in Table 1.

Approach:	**Evaluative**	**Explanatory**
Objective	Measure performance	Predict performance
Inquiry Approach	Retroactive	Proactive
Time line bearings	Past and Present	Future
Observations Focus	Outputs	Inputs
Indicators type	Consequential	Causal
Rationale	Accountability and Ranking	Productivity improvement
Ontology	Rational Choice	Human Relations

TABLE 1. OBSERVED CHARACTERISTICS OF THE TWO APPROACHES TO STUDYING AND ASSESSING RESEARCH PERFORMANCE

In addition to the different perspective taken, either evaluative or explanatory, studies vary in dimensionality and in unit of analysis. Dimensionality refers to how many evaluation criteria are used, and unit of analysis is either an individual researcher or a research group such as an academic program, a department, or a school.

To better understand how research productivity is perceived and treated, we conducted a literature review focusing primarily on the management related disciplines. The review was based on a sample of 67 studies dealing with research performance or research productivity, which appeared in the Sociofile 1974 to 1996, in the Social Science Citation Index 1983 to 1996, or in the Institute for Scientific Information Citation Database 1990-1998. Classification of the sample according to the three categories: approach, dimensionality, and unit of analysis, revealed the following ratios. The distribution between the two approaches was in favor of the explanatory studies with 72% explanatory studies versus 28% evaluative studies. Whereas about

85% of the evaluative studies dealt with departmental or institutional performance, almost all the explanatory studies examined individual performance. A single-criterion measure was used in about two-thirds of both the evaluative and the explanatory studies[1].

The evaluative studies and the explanatory studies are reviewed respectively in the next two sections. Then, observations and conclusions that can be drawn from the studies are presented and discussed.

EVALUATIVE STUDIES

Performance is associated with productivity, creativity, and achievement. In reviewing the historical development of research performance assessment, three main themes, or factors, emerge: reputation, yield, and influence. Although evaluation schemes often attempt to overcome biases inherent in earlier approaches, none has become the standard of research performance evaluation among scientists and administrators. Appendix A provides a summary of the evaluative studies.

Reputation-Based Factors

The traditional approach to research productivity assessment associates performance with reputation among peers. The first attempts to assess and rank the quality of academic programs were based on reputational ratings gathered by polls or peer reviews that aimed at "scientific eminence". For example, reputational ratings were the sole criterion of the two widely cited studies--Carter (1966) and Roose & Andersen (1970), which rated and ranked graduate programs in a broad range of academic disciplines. Both studies were based on reputational polls that required qualified and knowledgeable respondents from each field to stratify the available programs in their discipline. Another high profile reputational assessment is the controversial annual *Business Week's Guide to the Best Business Schools* that draws on an annual poll of alumni and recruiters.

Single-criterion reputation-based measures have been criticized by opponents that doubt both the theoretical validity of reputational polls and the reliability of the employed procedures. Jones, Lindzey, & Coggeshall (1982) criticized reputational measures as being subjective, biased, and insensitive to change . A study by Thoresen, Krauskopf, & Cox (1975) suggested that reputational polls are likely to be biased by a university's image that creates a A halo effect, which surrounds both academic programs and individuals attending the institution. Thus, they argued, reputational polls are likely to be positively skewed toward both large and long-established programs. Another study by Cox & Catt (1977) asserted that reputational polls may provide a global view or an overall assessment of an institution as a whole, but cannot reliably

[1] The sample consisted of 19 evaluative studies and 48 explanatory studies. Whereas 16 of the evaluative studies dealt with group's performance, 46 of the explanatory studies dealt with individual's performance. A single-criterion measure was used in 13 of the evaluative studies, and in 32 of the explanatory studies.

discriminate among the different programs within that institution. Furthermore, Webster (1981) concluded that reputational polls suffer from an anachronism because they stem from opinions about past performance, rather than current performance. Nevertheless, numerous studies find that ranking based on reputation is significantly correlated with ranking grounded in other, more objective, quantitative instruments (e.g., Cole & Cole, 1967; De Meuse, 1987; Kirkpatrick & Locke, 1992; Thomas & Watkins, 1998).

Yield-Based Factors

Among the single-criterion operationalizations, yield-based criteria were the most prevalent and constituted about two-thirds of our sample. Pointing to the drawbacks of reputational indicators, Cox & Catt (1977) suggested a ranking scheme that was based on the frequency of publication in the refereed journals of a discipline. According to their measure, for each paper published in a predefined set of journals during a given period, the affiliated parties get one unit of credit. Under the assumption that productivity per se, or yield, is correlated with quality, yield-based instruments seek to provide an objective index of quality and to deliver a separate rating for each researcher, program, or institution. By devising a quantitative measure that can provide separate ratings for each unit of analysis, yield-based indicators overcome one of the major deficiencies of the reputational indicators.

Some yield-based studies count the number of pages printed rather than simply the number of papers published. Proponents of this approach perceive the length of a manuscript as correlated with quality. Thus, they assert, a longer manuscript should be considered as more valuable than a shorter one. For example, Niemi (1988) ranked business schools based on a measurement technique that counts the number of standardized pages in two leading journals for each discipline[2]. A similar technique was implemented by Malhotra & Kher (1996) for ranking of both researchers and departments in production and operations management.

Further developments of yield-based performance measures refined the reliability and validity of the instruments. For example, the use of journal quality indices is a prevalent adjustment of the simple yield-based criterion. A journal quality index is a ranking scheme that seeks to sort journals in a discipline according to a quality criterion. In a simple yield-based criterion, one credit unit is attributed arbitrarily for each manuscript in a predefined set of journals, with no specific attention to the quality of the manuscript or the journal in which it was published. The implementation of a journal quality index undertakes to correct this by applying to each journal a weight that represents its relative quality in comparison to the other journals. Therefore, a yield-based criterion in conjunction with a journal quality index attributes a higher value to manuscripts that appear in the better journals. Using a journal quality index is demonstrated, for example, in Coe & Weinstock (1984) that management science journals based on a reputational poll among 188 chairs of AACSB-accredited management departments, which

[2] A "standard page" is a unit of analysis that is defined by a preset number of printed characters (e.g., one standard page = 1500 characters). The concept of standardized pages was devised in order to allow comparison among different journals while controlling for the different formats.

were asked to evaluate the quality and prestige of their journals. The study found that although there is a strong negative correlation (-.95) between the perceived acceptance rates and prestige ratings of the journals, the correlation with the actual acceptance rate was more modest (-.59). In another study, Vocino & Elliott (1984) ranked journals in the field of public administration based on intensity and extensity ratings. Intensity ratings measured the strength of feeling toward a journal and extensity ratings measured the breadth of recognition of a journal. Both studies found stability in journal rankings over time.

Most ranking schemes attribute one credit unit for each author with no regard to either the number of authors or their ordinal positions. An additional facet of yield-based criteria is a procedure for taking shared authorship and the consequent increased capacity into account. Based on citation analysis, Lindsey (1980) concluded that both a single-author article and a multi-author article have a similar impact. Therefore, he concluded, the failure to account for multiple authorship in the evaluation and ranking procedures is responsible for skewed rankings in favor of multiple-authors. He suggested that if the research effort is divided among more than one researcher, so should be the credit. Thus, offsetting coauthorship can be done by sharing equally one unit of credit among all the listed authors. Further adjustment, based on the premise that the authors ordinal position is correlated with their relative contribution to the paper, was suggested by Howard, Cole, & Maxwell (1987). Their study presented evidence showing that both the number of authors and their ordinal position should be accounted for. Hence, Howard, Cole, & Maxwell (1987) suggested that in the event of a multi-author article in which the alphabetical order rule was not applied, one unit of credit should be divided proportionately among the authors[3].

As far as departmental or institutional ranking is concerned, a study by Howard, Cole & Maxwell (1987) demonstrated that large departments appear more productive if the productivity ranking is based on a raw count of affiliated publications. Thus, in order to avoid discrimination against smaller departments, they ranked the departments based on the number of publications per capita. Small programs emerged when the size bias was corrected for.

The act of publishing is the fundamental criterion underling all yield-based measures, which inherently suffer from under-differentiation because they do not reflect the actual contribution of a specific work. While raw yield measures merely equate quantity with quality, quality-index adjusted measures rather identify the value of a manuscript with the perceived quality of the journal in which it appears. Thus, in the best case scenario, an adjusted yield-based measure implies that all the manuscripts in a particular journal have a similar quality or value, which, of course, is not so. Manuscripts printed in the same issue of a particular journal are likely to have different impacts on the field. Yet, neither yield-based indicators nor reputation-

[3] According to the study, authorship credit should be divided proportionately according to the following formula: credit = $(1.5^{n-1})/(\Sigma 1.5^{i-1})$, where *n* represents the number of authors and *i* represents the particular author's ordinal position.

based indicators can account directly for the quality or contribution of any specific manuscript. The implementation of influence-based criteria attempts to sharpen further the discrimination among papers by measuring the impact and influence of each manuscript on subsequent research in the field.

Influence-Based Factors

Influence-based criteria provide concrete and direct measures of quality. In relying on influence-based criteria, one assumes that the final judgement of a particular work's quality does not lie with journal editors and reviewers, but is a cumulative decision of the scientific community at large. Typically, the mechanics of influence-based criteria is based on gauging the impact of a particular manuscript or scientist on other related research.

In the social sciences, influence-based criteria often draw on citation counts based on records appearing in the *Social Sciences Citation Index* (SSCI) or in the *Science Citation Index* (SCI). For example, Endler, Rushton, & Roediger (1978) used the 1975 SSCI to evaluate and rank the impact of 4,977 manuscripts by 5,597 faculty members in 180 departments of psychology. In the same fashion, Garfield (1977, 1978) used both the SSCI and the SCI in a set of comprehensive studies that ranked the most-cited authors across disciplines. A similar approach, substituting textbook citation counts for journal citation counts, was taken by both Perlman (1984) and Gordon & Vicari (1992) in studies that ranked scholars in social psychology. Nederhof et al. (1993) made a further attempt to extend the coverage of the indices-based citation counts by including a wide variety of publications in addition to journal articles. Based on a survey of researchers from 70 departments in one university, they reported that for about third of the departments, publications not covered in citation indices accounted for at least 30% of the total citations.

Citation counts correlate with reputational rankings. Oppenheim (1997) examined scores of British academic departments in variety of disciplines and reported that citation counts for the period 1985-1992 are significantly correlated with the 1992 Research Assessment Exercise. Moreover, both So (1998) and Thomas & Watkins (1998) found high correlation between citation-based ranking and the more complex and expensive to produce peer review ranking.

Citation counts within the current research literature provide a quantitative assessment of one's recognition among peers. Nevertheless, in spite of their contribution to the arsenal of research quality assessment tools, influence-based measures suffer from under-differentiation. A citation is considered to be an indicator of an impact made on consequent research. Yet, although it is clear that the contributions of different research projects vary significantly, most of the current evaluation schemes treat the citations that represent these research projects as having an equal "influence" on a field. Under-differentiation is furtherly imposed by studies that report only on primary authors, and thereby, discount the contributions of coauthors. In addition, the SSCI and the SCI include primarily citations in journal articles disregarding the impact of much of the work that appears in books and conference proceedings.

Although influence-based criteria can provide an excellent measure of research quality, by themselves they fall short of assessing research productivity per se. Therefore, influence-based criteria are usually used in conjunction with yield-based instruments to create a superior quantitative measure of research performance. While the yield-based indicators provide measures of productivity, the influence-based indicators provide measures of impact and

effectiveness.

Multidimensional Studies

About third of the evaluative studies in our sample used a multi-criteria measure of performance. Multi-criteria instruments are designed to increase measurement accuracy and to overcome the pitfalls associated with the single-criterion measures. Typically, they are consisted of a combination of the previously identified factors--reputation, yield, and influence. Of course, in each study the instruments and the particular variables were crafted differently according to the objectives and the methodology of the study. Jones, Lindzey, & Coggeshall (1982), for example, used such a measure to rank 639 doctoral granting departments in seven major disciplines of the social and behavioral sciences. In the methodology section of their manuscript, they specifically noted that a multi-criteria approach was implemented in order to alleviate the criticism directed at the single-criterion measures.

Multi-criteria measures are also used to alleviate criticism of those who argue that the research performance construct extends beyond paper counts or citation counts. For example, in a study by Kirkpatrick & Locke (1992) research performance was perceived as scholarship, which was defined as a multidimensional construct consisting of three elements: productivity, influence, and reputation. The study, which ranked the departments of five disciplines in 32 business schools, operationalized productivity as the number of publications, influence as the number of citations, and reputation as peer ratings. The study found significant correlations among the three components of research performance, confirming similar findings in previous research (e.g., Cole & Cole, 1967; Jauch & Glueck, 1975). Yet, although the correlations among the components were significant, they were far from unity[4]. This observation of "convergent validity among the three measures" led Kirkpatrick & Locke to assert that each of the three dimensions focuses on a different facet of research performance, and that together they complement rather than substitute one for another. They argued, therefore, that combining all three components, instead of using any one alone, would provide "a more representative measure of faculty scholarship". Kirkpatrick & Locke did not, however, assess discriminant validity. Thus, the extent to which the components make unique contributions, rather than being redundant or overlapping, is not clear.

In a similar fashion, Jauch & Glueck (1975) asserted that although a single-criterion productivity measure is sufficient, researchers and evaluators simply do not accept it as a viable measure that accurately conveys research performance. Their study summarized the perceptions of psychology department chairs and researchers, who were asked to rank a set of indicators as to their suitability for the measurement of research performance. Overall, both chairpersons' and researchers' responses were similar. The study ranked the most desirable indicators in the following order: quality of publication, honors and awards, peer review, and influence on other research. Consequent research reaffirmed that both researchers and evaluators perceive the number of publications in top-ranked journals as the most desirable evaluation criterion (e.g.,

[4] For example, the correlations among the three components for the management departments, not adjusted to department size, were .53 for citations-articles, .85 for citations-peer ratings, and .80 for articles-peer ratings. The corresponding correlations after size adjustment were .41, .77, and .37.

Boyer, 1990; Zamarripa, 1995).

It is assumed that the most distinguished scholars publish extensively in top-rated journals, have significant impact on the field, and are perceived as eminent scholars among their colleagues. Furthermore, it has been demonstrated that these three components are significantly correlated. Yet, the available literature is ambiguous about the assertion that multi-criteria instruments actually provide superior measures to those achieved by more parsimonious single-criterion instruments.

EXPLANATORY STUDIES

Explanatory studies focus on the antecedents of scientific pursuit and attempt to evaluate their effect on research performance. Their primary objective is identification of conditions conducive to research productivity. In contrast to evaluative studies that concentrate on group performance, most explanatory studies deal with individuals. However, similarly to evaluative studies, explanatory studies tend to adopt a single-criterion strategy. About two-third of the explanatory studies in our sample used single-criterion measures centered mostly on demographic factors such as age, gender, and race. Additional factors, often used in the multi-criteria studies, included experience, personality, institutional, financial, and collaborative factors. Appendix B provides a summary of the explanatory studies.

Demographic Factors

A large body of research deals with the effect of demographics on research performance, particularly the correlation between gender or age and research performance. The study of demographic factors aims to predict the performance of individual researchers given the average performance of a comparable subset of scientists. Yet, in spite of their prevalence, the reliability of demographic factors alone in dealing with the complexities of predicting individual performance is open to question. Moreover, although discrimination based on demographics may contribute to bottom line efficiencies, it is often considered ethically or legally inappropriate.

A frequently used indicator is *age*. Levin & Stephan (1991) reported in a longitudinal study that life cycle effect varies significantly by field. They concluded that life cycle is correlated with publishing productivity and that generally scientists become less productive as they age. Underlying their study were the assumptions that research is as an investment-motivated activity that is associated with some future financial rewards, and that scientists are rational beings that engage in such activity predominantly for economical gain. Although life cycle effects are evident in communities of scientists, it is questionable whether they can be reliably applied for predicting individual productivity.

Another frequently investigated correlate of research performance is *gender*. Based on data from the 1980 UCLA HERI Faculty Survey, Rebne (1990) confirmed previous results concluding that women tend to produce less research than men across disciplines (e.g., Astin, 1978; Cole, 1979). The study found that aggregated production of journal articles yielded a per capita female/male ratio that ranges from .26 in management sciences to .74 in biological sciences. In another study, which surveyed 199 male and 200 female full-time academic members, Vasil (1996) also found that females exhibited lower levels of research productivity.

Efforts to explain gender differences in productivity have yielded some research on the effect of *marital status* on research performance. The hypothesis was that married women researchers would be less productive given their domestic responsibilities. A study by Reskin (1978) found no significant effect of marital status on women's productivity. In another study which controlled for differences in rank, Astin & Bayer (1979) found that marriage has some positive effect on women's performance. Furthermore, Cole & Zuckerman (1983), using both in-depth interviews and publication analysis, found that marriage does not impede the performance of women scientists. Not only did the study reject the hypothesis that domestic responsibility negatively impacts productivity, it found a slight productivity increase among women that have families with two or fewer children. In a similar study, Cole (1979) found that this effect extends to males as well. He concluded that married faculty were significantly more productive than unmarried, and speculated that being married probably fosters productivity by encouraging more stable work practices.

Ethnic minorities are under-represented in academic institutions relative to their respective share in the population at large. No conclusive evidence is suggested by the literature regarding the effect of *descent* or ethnicity on research productivity. Clemente (1974) found that non-black sociologists publish more frequently than blacks. Scott (1981) suggested that contextual factors may inhibit the research productivity of some black faculty. Rebne (1990) found that Asian and non-Asian minority groups differ significantly in research productivity across disciplines. In another preliminary study, which compared the mean two-year publication rate for non-Caucasians and Caucasians, Rebne (1990) found virtually identical performance for both groups.

Experience Factors

Research concerning personal factors attempts to identify the relationships between an individual's situational context and research performance. One frequently used indicator is *career age*. Career age represents professional experience or number of years in the field. A longitudinal study by Rebne (1990) suggested that publishing generally increases sharply during the initial years and reaches its peak performance within the first 10 years of work. Decline of the yield to levels below average performance tends to start after 25-30 years of activity. These findings were confirmed by Goodwin and Sauer (1995) in a study that examined the temporal distribution of publications for a sample of 140 tenured professors of economics. In addition, they found that the above pattern is valid for both high-rate publishers and low-rate publishers.

Another related indicator may be *rank* or professional standing as manifested by pre-tenure and post-tenure status. While it is reasonable to believe that tenure candidates are under pressure to produce, no research findings suggest a clear explanation of how this pressure affects actual research performance. Based on publication patterns for both the pre-tenure and the post-tenure periods of 97 sociologists from various institutions, Holley (1977) found substantial differences between pre-tenure output and post-tenure output, with pre-tenure output being of greater magnitude in most instances.

A significant relationship has been found between *education quality* and later research performance. Rebne (1990) found a correlation between productivity level of faculty and their

education quality at both the undergraduate and graduate levels. The education quality was measured in this study as the institutional selectivity in terms of the average cumulative SAT scores of entering students. Jones & Preusz (1993) surveyed 833 researchers and found a significant correlation between research productivity of faculty members and the perceived proficiencies in research methodology and techniques attained at the training stages. Somewhat contradictory findings are reported in a study by Long (1978) who found a strong effect of *doctoral program prestige* on the initial placement institution, but a weak effect of doctoral students' productivity on the initial position granted. Similarly, Fogarty & Ruhl (1996) found correlations between productivity and both the doctoral program=s prestige ratings and the initial placement institution=s reputation.

It has long been recognized that *trend* and *past performance* are excellent indicators of future performance. Following this rationale, Allison & Long (1990) examined the antecedents and consequences of 179 job changes across multiple disciplines. They concluded that performance record is consistent along one's career. Those who were upwardly mobile showed a substantial increase in their rate of publication and rate of citation, while those who were downwardly mobile showed a decrease in productivity.

Personality Factors

Other attempts to explore the antecedents of research productivity seek to identify personality determinants of highly productive researchers. For example, productive researchers emerge as highly motivated individuals who make deliberate choices about the type of research they undertook (Harris & Kaine, 1994), empowered by high self-efficacy, able to be engaged in multiple projects, (Taylor et al., 1984), set clear research goals (Locke et al., 1994; Taylor et al., 1984), and have excellent time management skills (Hancock et al., 1992).

Rushton, Murray, & Paunonen (1987) attempted to determine which *personality traits* are associated with academic excellence. Their literature review identified previously used personality traits that presented a successful scientist as "a person considerably less sociable than average, rather serious-minded, intelligent, aggressive, dominant, achievement-oriented, and independent. In addition, he or she is cognitively complex, has a radical imagination and a well-articulated self-concept. In short, the creative person is both introverted and bold." In a consequent study, Rushton et al. (1987) examined a group of 52 full-time psychology professors of varying ranks for correlations between the previously identified 29 personality traits and both research performance and teaching performance. The study suggested that "one might characterize the creative researcher as ambitious, enduring, seeking definiteness, dominant, showing leadership, aggressive, independent, not meek, and non-supportive."

Institutional Factors

Institutions set the stage for the research performance of faculty members. Institutional objectives and policies are the frame of reference for the development of performance evaluation principles that underlie the formulation of reward systems. Faculty members in most institutions are expected to contribute in three domains: research, teaching, and service. Yet, institutions differ significantly in their emphasis and expectations. Whereas research or doctorate-granting institutions tend to emphasize research as the most important domain, liberal-arts and community colleges tend to foster teaching at the expense of research. These institutional tendencies are

articulated in mission statements, implied by resource allocations, translated into promotion decisions, and rooted in the organizational culture. The implications of *institutional orientation* for the opportunity of one to engage in research and consequently to excel as a scholar are clear. Nonetheless, Gordon & Marquis (1966) found that academic freedom alone is not likely to generate scholarship. Academic freedom, they claimed, is conducive to scientific accomplishment only when the institutional settings in which research is conducted fuel the impetus to innovate. One such critical institutional enabler of the impetus to innovate is the *visibility* of research consequences in relation to the institutional objectives.

Many studies have reported that researchers perceive *teaching and administrative duties* as a hindrance to research productivity (e.g., Boyer, 1990; Hancock et al., 1992). The *competition* perspective and the *mutuality* perspective are two opposite theoretical views that attempt to deal with the relationships between research productivity and teaching. Proponents of the competition view suggest that research and teaching represent different and competing dimensions of academic investments, which do not have complementary relationships (Fox, 1992). Thus, aligned with the scarcity theory of role behavior, which posits a linear relationship between the time spent on a role and performance (Goode, 1960), they believe that an increase in time spent on other institutional activities is likely to cause a consequent decrease in research performance. In contrary, opponents of the competition view suggest that as long as they are being pursued in a low or moderate capacity, other institutional roles should have a positive effect on research performance (Rebne, 1990). This alternative hypothesis stem from the complementary role theory (Faia, 1980), which postulates a relationship of mutual enrichment among interrelated roles. The mutuality view is further supported by both Hicks (1974) and Hoyt & Spangler (1976), which found that moderate teaching activity enhances research performance and suggested that balanced activities enhance research productivity.

Another set of indicators is related to *institutional affluence*, or the ability of an institution to facilitate extensive research activity, to provide the necessary resources, and to attract top scholars. Frequently used indicators are total institutional or departmental expenditures (Jones, Lindzey, & Coggeshall, 1982); institution and department size measures (Rebne, 1990); and institutional prestige ratings (Long, 1978).

Financial Factors

Although spending does not guaranty results, limited resources are likely to prevent or hinder research activities. With this notion in mind, Jones, Lindzey, & Coggeshall (1982) used the amount of *direct expenditures on research* as an indicator of research performance. An alternative measure is *research grants awarded*. Grants provide an indication of past performance and a measure of reasonable future research plans that have been approved under peer review. The relative importance attributed to this indicator is reflected in the findings of studies that examined faculty perceptions about the relative importance of performance indicators. In a study by Zamarripa (1995), faculty members ranked the indicator "research grants received" as second, and in another study by Boyer (1990) they ranked grants as third after "number of publications" and "recommendations from peers". Nevertheless, the viability of this indicator is questionable. Gillett (1991) demonstrated that an index based on research grants awarded is not a reliable indicator of scientific performance. Furthermore, examining the effect of a faculty resource center on grant funding, Baldwin et al. (1994) reported that although grant awarded quadrupled over five years, the program failed to increase publication productivity.

Collaborative Factors

Mind sharing, mutual enrichment, and complementary relationships among researchers have long been recognized as major contributors to synergies resulting in collaborative projects. Prpic (1996) identified a significant intensive scientific collaboration among 385 eminent researchers. Jones & Preusz (1993) reported a significant correlation between research productivity and the extent of interactions with colleagues for discussions along with involvement in joint research projects. They suggested that personal relationships with colleagues are the basis for informal exchange of ideas that eventually may become joint research projects. One measure of collaborative relationships is the number and the extent of one's *ongoing contacts with peers*. Another measure is the number of active *joint research initiatives*, which is a more direct indicator of actual research efforts that may result in increased performance.

Two additional indicators are *participation in conferences* and extent of *involvement in consulting*. Both are in the same line as the previous indicators, but with more emphasis on level of exposure to potential sources of insights. Participation in conferences indicates a level of exposure to the field's community and current research. It provides one with the opportunity to attain direct impressions of current research, to meet with active members in the field, and to identify new partners for collaborations. The extent of consulting projects indicates a level of exposure to practitioners' agendas and to environmental trends concerning the discipline. Consulting projects enable one to become aware of current critical issues, to examine the implementation of theoretical solutions for practical problems, and to gain hands-on experience that cannot be obtained otherwise.

Both Hancock et al. (1992) and Zamarripa (1995) suggested that the number of *graduate students* supervised is correlated with research performance[5]. Hancock et al. (1992) found that high performers are more likely to supervise doctoral students. Consequently, Zamarripa (1995) reported that the "presence of graduate students working on research projects" was ranked among the first four indicators of research performance in a survey of 40 scientists and 49 research administrators in management schools. In the same vein, Pettigrew & Nicholls (1994) found that publication productivity is likely to be higher in doctorate granting universities.

DISCUSSION AND CONCLUSIONS

About 85% of the evaluative studies examined departmental and institutional performance. Given the nature and impetus of evaluative studies, this finding is counterintuitive because it is generally individuals who are evaluated, promoted, and rewarded. On the other hand, the explanatory studies examined individual productivity. This too is counterintuitive, considering that the controllable enablers of productivity are primarily institutional. In light of these observations, we propose that further evaluative studies should aim to understand individual performance and that further explanatory studies should include factors of group

[5] In some instances, the criterion "number of graduate students supervised" was perceived as a reputational indicator and not as collaborative-causal indicator (e.g., Jauch & Glueck, 1975).

performance.

The examination of research performance is essentially divided between evaluative studies and explanatory studies. Studies taking a wide perspective that includes both the development of performance measures and the examination of research performance correlates (e.g., Creswell, 1985) are rare. The failure to adopt a comprehensive look may be one of the main causes for the existing diversity and incomparability among the available instruments. Adopting an overarching approach, that successfully bridges between the explanatory and the evaluative perspectives, may ultimately shed light on the cause-effect relationship between faculty selection criteria, institutional enablers, and research productivity.

Some methodological issues, which are prevalent especially in evaluative studies, should be of concern. For example, all of the evaluative studies that focused on individuals used a single-criterion measure. The implementation of a single-criterion rather than a multi-criteria measure in these studies is in contradiction to our expectations. Relative to single-criterion instruments, multi-criteria measures are less prone to systematic measurement error and they are less sensitive to biases created by those individuals who modify their behavior in an attempt to gain favorable evaluations. For example, a yield-based criterion may hinder innovative but risky research and promote the writing of many trivial or conservative and paradigm-conforming articles. Therefore, while a parsimonious single-criterion measure appears to be sufficient in evaluative studies of groups, a multi-criteria measure seems warranted in assessments of individual's performance.

Furthermore, research performance is treated by the evaluative studies as a temporal event rather than as an ongoing dynamic process. Academic pursuit and the consequent outputs are often measured in these studies as a onetime cross sectional snapshot. No consideration, let alone representation, is given to either the direction or the pace of the research process. Therefore, two researchers or departments may appear similar although they are actually very different. For example, whereas two departments may have a similar score, one may have undergone a decline in the past few years, while the other may be experiencing growth at the same period of time. Treating both departments as having a similar scholarly capability overlooks the growth potential in the one department and the decay in the other.

Finally, evaluative studies that employ ordinal scales often suffer from over-differentiation. Given the precision of the available instruments, no reliable differentiation and consequently ranking can be attained among comparable institutions or comparable individuals. Although methodologically and culturally driven, a straightforward ordinal ranking of scientific excellence leads to non-sensible, arbitrary, and misleading classifications.

Although the general notion of academic excellence is quite clear, common measurement practices are plagued, and so far no single measure of research performance has emerged as a standard. The diversity of the current performance measures across disciplines and institutions is undoubtedly fueled by a combination of legitimate differences, individual perceptions, and political considerations. Nevertheless, we recommend that explicit attention be given in further studies to the effects, limitations, and biases inherent in the research methodologies employed.

REFERENCES

Allison, P.D. & Long, J.S. (1990). *Departmental effects on scientific productivity*. American Sociological Association, Philadelphia.

Astin, H.S. (1978). Factors affecting women's scholarly productivity. In H.S. Astin and W.Z. Hirsch (Eds.), *The higher education of women: Essays in honor of Rosemary Parks*, Praeger, New York.

Astin, H.S. & Bayer, A.E. (1979). Pervasive sex differences in the academic reward system: Scholarship, marriage, and what else? In D.R. Lewis & W.E. Becker (Eds.), *Academic rewards in higher education*, Ballinger, Cambridge, MA.

Baldwin, C.D., Goldblum, R.M., Rassin, D.K., & Levine, H.G., (1994). Facilitating faculty-development and research through critical-review of grant proposals and articles. (1994). *Academic Medicine*, 69 (1), 62-64.

Boyer, E.L. (1990). *Scholarship reconsidered: Priorities of the professorate*. The Carnegie Foundation, Princeton.

Cartter, A.M. (1966). *An assessment of quality in graduate education*. American Council on Education, Washington, DC.

Clemente, F. (1974). Race and research productivity. *Journal of Black Studies*, 5 (2), 157-166.

Coe, R.K. & Weinstock, I. (1984). Evaluating the management journals: A second look. *Academy of Management Journal*, 27 (3), 660-666.

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

Cole, S. & Cole, J.R. (1967). Scientific output and recognition: A study in the operation of the reward system in science. *American Sociological Review*, 32, 377-390.

Cole, J.R. & Zuckerman, H. (1983). *The structure of scientific careers: Are marriage and family incompatible with research productivity?* American Sociological Association.

Cox, W.M. & Catt, V. (1977). Productivity ratings of graduate programs in psychology based on publication in the journals of the American Psychological Association. *American Psychologist*, 10, 793-813.

Creswell, J.W. (1985). *Faculty research performance: Lessons from the sciences and the social sciences*. Association for the Study of Higher Education, ERIC Clearinghouse on Higher

Education, Washington, DC.

De Meuse, K.P. (1987). The relationship between research productivity and perceptions of doctoral program quality. *Professional Psychology*, 18 (1), 81-83.

Endler, N.S., Rushton, J.P., & Roediger, H.L. (1978). Productivity and scholarly impact (citations) of British, Canadian, and U.S. departments of psychology (1975). *American Psychologist*, 12, 1064-1082.

Faia, M.A. (1980). Teaching, research, and role theory. *Annals of the American Academy of Political Science*, 448, 36-45.

Fogarty, T.J. & Ruhl, J.M. (1996). The stratification of academic accounting in the U.S.: The role of doctoral funding grants. *Critical Perspectives on Accounting*, 7, 509-528.

Fox, M.F. (1992). Research, teaching, and publication productivity: Mutuality versus competition in academia. *Sociology of Education*, 65 (4), 293-305.

Garfield, E. (1977). The 250 most-cited primary authors, 1961-1975: Part 1. How the names were selected. *Current Contents*, 9 (49), 5-15.

Garfield, E. (1978). The 300 most-cited authors, 1961-1976, including co-authors at last: Part 1. How the names were selected. *Current Contents*, 10 (28), 5-17.

Gillett, R. (1991). Pitfalls in assessing research performance by grant income. *Scientometrics*, 22 (2), 253-263.

Goode, W.J. (1960). A theory of role strain. *American Sociological Review*, 25, 483-496.

Goodwin, T.H. & Sauer, R.D. (1995). Life cycle productivity in academic research: Evidence from cumulative publication histories of academic economists. *Southern Economics Journal*, 61 (3), 728-743.

Gordon, G. & Marquis, S. (1966). Freedom, visibility of consequences, and scientific innovation. *The American Journal of Sociology*, 72 (2), 195-202.

Gordon, R.A. & Smith, C.J. (1989). Research productivity in Social Psychology. *Personality and Social Psychology Bulletin*, 15 (3), 463-472.

Gordon, R.A. & Vicari, P.J. (1992). Eminence in social psychology: A comparison of textbook citation, Social Sciences Citation Index, and research productivity ranking. *Personality and Social Psychology Bulletin*, 18 (1), 26-38.

Hancock, T., Lane, J., Ray, R., & Glennon, D. (1992) Factors influencing academic research

productivity: A survey of management scientists. *Interfaces*, 22 (5), 26-38.

Harris, G. & Kaine, G. (1994). The determinants of research performance: A study of Australian university economists. *Journal of Higher Education*, 27 (2), 191-201.

Hicks, R.A. (1974). The relationship between publishing and teaching effectiveness. *California Journal of Educational Research*, 25, 14-46.

Holley, J.W. (1977). Tenure and research productivity. *Journal of Higher Education*, 6 (2) 181-192.

Howard, G.S., Cole, D.A., & Maxwell, S.E. (1987). Research productivity in psychology based on publication in the journals of the American Psychological Association. *American Psychologist*, 42 (11), 975-986.

Hoyt, D.P. & Spangler, R.K. (1976). Faculty research involvement and instructional outcomes. *Research in Higher Education*, 4, 113-122.

Jauch, L.R. & Glueck, W.F. (1975). Evaluation of university professors' research performance. *Management Science*, 22 (1), 66-75.

Jones, J.E. & Preusz, G.C. (1993). Attitudinal factors associated with individual faculty research productivity in academic dentistry. *Perceptual and Motor Skills*, 76, 1191-1198.

Jones, L.V., Lindzey, G., & Coggeshall, P.E. (1982). *An assessment of research-doctorate programs in the United States: Social sciences*. National Academy Press, Washington, DC.

Kirkpatrick, S.A. & Locke, E.A. (1992). The development of measure of faculty scholarship. *Group & Organization Management*, 17 (1), 5-23.

Levin, S.G. & Stephan, P.E. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review*, 81 (1), 114-132.

Lindsey, D. (1980). Production and citation measures in the sociology of science: The problem of multiple authorship. *Social Studies of Science*, 10, 145-162.

Locke, E.A., Smith, K.G., Erez, M., Chach, D.O., Schaffer, A. (1994). The effects of intra-individual goal conflict on performance. *Journal of Management*, 20 (1), 67-91.

Long, J.S. (1978). Productivity and academic position in the scientific career. *American Sociological Review*, 43 (12), 889-908.

Malhotra, M.K. & Kher, H.V. (1996). Institutional research productivity in production and

operation management. *Journal of Operations Management*, 14, 55-77.

Merton, R.K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22, 635-659.

Nederhof, A.J., Meijer, R.F., Moed, H.F., & Vanraan, A.F.J. (1993). Research performance indicators for university departments: A study of an agricultural university. *Scientometrics*, 27 (2), 157-178.

Niemi, A.W. (1988). Research productivity of American business schools, 1975-85. *Review of Business and Economic Research*, 23 (2), 1-17.

Oppenheim, C. (1997). The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British research in Genetics, Anatomy, and Archaeology. *Journal of Documentation*, 53 (5), 477-487.

Perlman, D. (1984). Recent developments in personality and social psychology: A citation analysis. *Personality and Social Psychology Bulletin*, 10, 493-501.

Pettigrew, K.E. & Nicholls, P.T. (1994). Publication patterns of LIS faculty from 1982-1992: Effects of doctoral programs. *Library & Information Science Research*, 16 (2), 139-156.

Prpic, K. (1996). Characteristics and determinants of eminent scientists' productivity. *Scientometrics*, 36 (2), 185-206.

Rebne, D. (1990). *Determinants of individual productivity: A study of academic researchers*. Institute of Industrial Relations Publications Center, UCLA.

Reskin, B.T. (1978). Scientific productivity, sex, and location in the institution of science. *American Sociological Review*, 83 (5), 1235-1243.

Roose, K.D. & Andersen, C.J. (1970). *A rating of graduate programs*. American Council on Education, Washington, DC.

Rushton, J.P., Murray, H.G., & Paunonen, S.V. (1987). Personality characteristics associated with high research productivity. In Jackson, D.N. & Rushton, J.P. (Eds.), *Scientific excellence*, 129-148, Sage, London.

Scott, R.R. (1981). Black faculty productivity and interpersonal academic contacts. *Journal of Negro Education*, 50 (3), 224-236.

Sharplin, A.D. & Marby, R.H. (1985). The relative importance of journal used in management research: An alternative ranking. *Human Relations*, 38 (2), 139-149.

So, C.Y.K. (1998). Citation ranking versus expert judgement in evaluating scholars: Effects of

research specialty size and individual prominence. *Scientometrics*, 41 (3), 325-333.

Stahl, M.J., Leap, T.L., & Wei, Z.Z. (1988). Publication in leading management journals as a measure of research productivity. *Academy of Management Journal*, 31 (3), 707-720.

Taylor, M.S., Locke, E.A., Lee, C., & Gist, M.E. (1984). Type A behavior and faculty research productivity: What are the mechanisms? *Organizational Behavior and Human Performance*, 34, 402-418.

Thomas, P.R. & Watkins, D.S. (1998). Institutional research ranking via bibliometric analysis and direct peer review: A comparative case study with policy implications. *Scientometrics*, 41 (3), 335-355.

Thoresen, R.W., Krauskopf, C.F., & Cox, J.D. (1975). Reputation, halo, and ratings of counseling programs. *Journal of Counseling Psychology*, 22, 446-450.

Vasil, L. (1996). Social process skills and career achievements among male and female academics. *Journal of Higher Education*, 67 (1), 103-114.

Vocino, T. & Elliott, R.H. (1984). Public administration journal prestige: A time series analysis. *Administrative Science Quarterly*, 29, 43-51.

Webster, D.S. (1981). Methods of assessing quality. *Change*, 10, 20-24.

Young, S.T., Baird, B.C., & Pullman, M.E. (1996). POM research productivity in U.S. business schools. *Journal of Operations Management*, 14, 41-53.

Zamarripa, E.J. (1995). Evaluating research productivity. *SRA Journal*, 26 (3-4), 17-27.

APPENDIX A. FACTORS AND MAJOR INDICATORS OF RESEARCH PERFORMANCE IN EVALUATIVE STUDIES

Factor	Indicator	Illustrative References
Reputation	Rank in a reputational poll	Carter, 1966 Roose & Andersen, 1970
	Honors and awards received	Jauch & Glueck, 1975
	Positions held in professional associations and journals	Jauch & Glueck, 1975
	Invitations to present papers	Jauch & Glueck, 1975
	Number of doctoral students supervised	Jauch & Glueck, 1975
Yield	Total publication yield (books, conference papers, etc.)	Jauch & Glueck, 1975
	Publication yield in a set of leading journals	Cox & Catt, 1977
	Publication yield in selected journals targeting a unique subset of members	Stahl, Leap, & Wei, 1988
	Publication yield in one top-ranked journal	Gordon & Smith, 1989
	Character yield (standardized page output)	Niemi, 1988 Malhotra & Kher, 1996
	Publication yield per capita	Howard, Cole, & Maxwell, 1987
	Publication yield adjusted to the number of coauthors	Young, Baired, & Pullman, 1996
	Publication yield adjusted to journal quality index	Coe & Weinstock, 1969, 1984 Vocino & Elliott, 1984
	Publication yield adjusted to journal citation impact efficiency index	Sharplin & Mabry, 1985
Influence	Journal citation counts (SSCI and/or SCI)	Garfield, 1977, 1978 Endler, Rushton, & Roediger, 1978
	Textbook citation counts	Perlman, 1984 Gordon & Vicari, 1992
	Citation counts of the three most frequently cited manuscripts	Cole & Cole, 1967

APPENDIX B. FACTORS AND MAJOR INDICATORS OF RESEARCH PERFORMANCE IN EXPLANATORY STUDIES

Factor	Indicator	Illustrative References
Demographic Factors	Age	Levin & Stephan, 1991 Rebne, 1990
	Gender	Astin, 1978 Cole, 1979
	Marital status	Asin & Bayer, 1979 Cole & Zuckerman, 1983
	Descent	Scott, 1981 Rebne, 1990
Experience Factors	Career Age	Rebne, 1990 Goodwin & Sauer, 1995
	Rank (tenure)	Holly, 1977 Hancock et al., 1992
	Past performance record	Allison & Long, 1990
	Research methodology knowledge	Jones & Preusz, 1993
	Education and training quality	Rebne, 1990
	Doctoral school prestige rating	Long, 1978
	Initial placement prestige rating	Long, 1978
Personality Factors	Personality traits	Rushton, Murray, & Paunonen, 1987
	Self-efficacy	Taylor et al., 1984
	Goals driven	Locke et al., 1994 Harris & Kaine, 1994
	Multiple project management skills	Taylor et al., 1984
	Time management skills	Hancock et al., 1992

more...

APPENDIX B (continue)

Institutional Factors	Institutional orientation	Boyer, 1990
	Visibility of research consequences	Gordon & Marquis, 1966
	Time spent on research	Hicks, 1974 Hoyt & Spangler, 1976
	Time spent on teaching and other institutional related roles	Hoyt & Spangler, 1976 Fox, 1992
	Institutional affluence	Jones, Lindzey, & Coggeshall, 1982 Zamarripa, 1995
	Size of institution/ department	Rebne, 1990
	Institutional prestige rating	Long, 1978
Financial Factors	Research grants awarded	Boyer, 1990 Gillett, 1991
	Direct expenditure on research	Jones, Lindzey, & Coggeshall, 1982
Collaborative Factors	Number of ongoing contacts with peers	Jones & Preusz, 1993 Zamarripa, 1995
	Number of joint-research initiatives	Jones & Preusz, 1993 Prpic, 1996
	Number of consulting projects	Hancock et al., 1992 Rebne, 1990
	Number of graduate students supervised	Hancock et al., 1992
	Presence of a doctorate program	Pettigrew & Nicholls, 1994

芽|Sprouts

Working Papers on Information Systems | ISSN 1535-6078

Office:
Sprouts
University of Amsterdam
Roetersstraat 11, Room E 2.74
1018 WB Amsterdam, Netherlands
Email: admin@sprouts.aisnet.org

芽|Sprouts