

EVALUATING RESEARCH PRODUCTIVITY

A compilation of papers originally presented at a conference
sponsored by
The Merrill Advanced Studies Center
June 2001

Editor – Mabel L. Rice
Technical Editing – Joy Simpson

MASC Report No. 105
University of Kansas

TABLE OF CONTENTS
MASC Report No. 105

INTRODUCTION

Mabel L. Rice 1
Director, Merrill Advanced Studies Center

EXECUTIVE SUMMARY 3

KEYNOTE ADDRESS

Joan F. Lorden 11
Associate Provost for Research and Dean of the Graduate School
University of Alabama at Birmingham
*Measuring Success: Lessons from the NRC Study of the
Research Doctorate*

RESPONSE to the Keynote Address

Robert E. Barnhill 23
Vice Chancellor for Research and Public Service, University of Kansas

PANEL OF RESEARCHERS

Carol Shanklin 43
Professor of Institutional Management & Dietetics, and Assistant Dean
of the Graduate School, Kansas State University
*A Faculty Perspective: Intrinsic Research Rewards that Make a
Successful Faculty Member Tick*

Michael Podgursky 47
Professor of Economics, University of Missouri – Columbia
Statistical Analyses of Economics NRC Scores

Susan Kemper 51
Distinguished Professor of Psychology and Gerontology, University of Kansas
A Passion for Reputation, A Taste for Originality

FIRST PANEL OF RESEARCH ADMINISTRATORS

Thomas H. Rosenquist 57
Vice Chancellor for Research, University of Nebraska Medical Center
*Evaluating Research Productivity of Individual Investigators:
Chairmen's Perspectives in a Medical Center*

Suzanne Ortega 61
Dean of the Graduate School and Vice Provost for Advanced Studies
University of Missouri – Columbia
Evaluating the Quality and Quantity of Graduate Student Research

K. Michael Welch	67
Vice Chancellor for Research and Senior Associate Dean University of Kansas Medical School <i>Achieving the “Gold Standard”: A Reductionism Directive</i>	
SECOND PANEL OF RESEARCH ADMINISTRATORS	
R.W. Trewyn	71
Vice Provost for Research and Dean of the Graduate School Kansas State University and President, KSU Research Foundation <i>Evaluating University Research Productivity: What’s the ROI ... and Who Cares?</i>	
Jack O. Burns, Vice Provost for Research.....	77
and Mary M. Licklider, Director, Grant Writing and Publications University of Missouri – Columbia <i>The Road to Success: MU’s Master Plan for Research and Technology Development</i>	
James R. Bloedel.....	81
Vice Provost for Research and Advanced Studies, Iowa State University <i>Judging Research Productivity on an Entrepreneurial Campus</i>	
A REFLECTION ON A DAY SPENT DISCUSSING EVALUATION	
David Shulenburg, Provost, University of Kansas.....	85
HIGHER EDUCATION ADVOCACY: THE INTERFACE OF TWO CULTURES	
Kim A. Wilcox.....	89
Executive Director, Kansas Board of Regents	
ADMINISTRATIVE PERSPECTIVES	
Rollin C. Richmond.....	93
Provost, Iowa State University <i>Keeping a Broad Perspective on Research</i>	
James R. Coffman	101
Provost, Kansas State University <i>Research and Institutional Mission and Niche</i>	
LIST OF PARTICIPANTS.....	105

INTRODUCTION

Mabel L. Rice
University Distinguished Professor
Director, Merrill Advanced Studies Center
University of Kansas

This year marked the fifth annual conference on research policy hosted by the Merrill Center in Valley Falls, Kansas. Our topic, "Evaluating Research Productivity," generated lively discussion, as you will see in this collection. Our group included twenty-two administrators and senior faculty from the four-state region of Kansas, Nebraska, Iowa and Missouri. Keith Yehle from the office of Senator Pat Roberts also joined us. Discussions centered on methods of ranking universities and evaluating faculty in the context of graduate education.

The 2001 topic built on discussions at the four previous regional conferences. The inaugural conference in 1997 focused on pressures that hinder the research mission of higher education, with special consideration of public research universities. In 1998, we turned our attention to competing for new resources, and ways to enhance individual and collective productivity. In particular, our keynote speaker of that year, Michael Crow, encouraged us to identify niche areas for research focus, under the premise that it was most promising to do selective areas of investigation at the highest levels of excellence. In 1999, we examined in more depth cross-university alliances. Keynote speaker Luis Proenza encouraged participants to think in terms of "strategic intent" and he highlighted important precedents in university-industry cooperation as well as links between institutions. In 2000, we focused on making research a part of the public agenda. We heard from George Walker who encouraged us to meet the needs of our state citizens, business leaders and students who are quite able to "carry our water" and champion the cause of research as a valuable state resource. This cutting-edge topic included an exploration of the dynamic interface between research initiatives at public universities and the response of public constituencies in light of actual and potential research outcomes in science.

This year's keynote speaker was Joan Lorden from the University of Alabama at Birmingham. Dr. Lorden brought to the conference her experience with the topic of evaluating research productivity. She drew upon the position paper she co-authored in February 2000 for the National Association of State Universities and Land Grant Colleges. For our group, she provided a valuable overview of key elements to consider when selecting measures for evaluating performance, with a focus on the very important National Research Council (NRC) study, which appeared in 1995. With this starting point, our conference participants elaborated and expanded on issues of research evaluation from the

perspectives of research administrators, faculty researchers, provosts, and the Executive Director of the Kansas Board of Regents.

The dynamic interactions of the conference are reflected in this collection of papers, which conveys a sense of the multiple perspectives and the complex issues that bear on ways to measure an institution's research productivity. One axiom emerged throughout our discussions: We become what we measure. With this in mind, participants emphasized the role of public universities in the national dialogue and stressed the importance of developing measurement systems that capture the strengths of our public institutions.

It is with pleasure that I encourage you to read each of the following collection of papers. I wish to express my appreciation to Joy Simpson, for her assistance with the organization of the conference, her careful note taking, and her editorial assistance with the collection of papers. Patsy Woods provided help with the budget and fiscal arrangements. We are appreciative of the gracious hospitality provided by the owners and staff of The Barn Bed and Breakfast Inn, and the relaxed and pleasant atmosphere provided for discussion. We are especially grateful for the support and generosity of Virginia and Fred Merrill.

EXECUTIVE SUMMARY

KEYNOTE ADDRESS

Joan F. Lorden

Associate Provost for Research and Dean of the Graduate School
University of Alabama at Birmingham

- Evaluation is useful to set benchmarks, to recognize excellence and to promote improvement. We become what we measure, so it is important to choose wisely.
- The study released by the National Research Council in 1995 on research doctoral programs in the U.S. achieved some important goals. It provided broad coverage in terms of the number of universities. The report was derived from national datasets using a process of uniform data collection. Unlike the *U.S. News and World Report* survey, it provided in-depth analysis, and the rankings were done by scholars in the field of study. The primary features of the NRC study were: rankings, a reputational survey, longitudinal comparison, institutional information and objective measures of performance.
- The Council on Research Policy and Graduate Education provided feedback to the NRC and suggested several changes for the next study, critiquing in particular the undue emphasis on ranking programs based on reputation. The CRPGE suggested that reputational rankings don't reflect the tremendous change some fields of study have experienced in the last 20 years because reputations are slow to change. The CRPGE is a group established by the National Association of State Universities and Land Grant Colleges.
- The survey portion of the study exemplifies many of the problems encountered when measuring research quality, but it successfully demonstrated the link between graduate education and research.
- Other questions to consider about rankings include: Is it valid that the top-ranked programs consistently have a larger faculty and more students? How can we evaluate niche programs?
- According to the survey, high-ranked programs: have a large faculty; are well funded; publish successfully; and give their graduate students research assistantships. The count of per capita publications tends to correlate with ranks, but awards and honors are the marker of significance for the arts and humanities in the top quarter. In science and engineering, federal funding is the highest in the top quarter rankings.
- In general, the rankings don't tell us much about the experience of students or outcomes of graduate education. We do see that students from the higher ranked programs are supported more often on research assistantships, whereas students from lower-ranked programs are supported more often on teaching assistantships. We also see that time

- to degree increased more in the lower-ranked schools. We would benefit from knowing what happens with graduate students while on campus and the outcomes of their education.
- The University of Alabama at Birmingham determined that funding is an important measure. Most areas that have experienced NIH funding success have been interdisciplinary. To provide incentives, UAB established an umbrella operation for interdisciplinary centers and guidelines, and also invests in targeted areas.
 - When choosing measures for the future, we should ask: What are the goals? Who is the audience? Do the measures reflect our values? Do we understand their limitations? How will the measures be used?

RESPONSE TO THE KEYNOTE ADDRESS

Robert E. Barnhill

Vice Chancellor for Research and Public Service, University of Kansas

- We should select and promote measures that reflect the values we think are important.
- The National Science Foundation annually collects data on the federal R&D expenditures in science and engineering. This information has become our “gold standard” for national comparisons. Rankings of this type also provide a surrogate for market share in terms of the percentage of the federal R&D funds obtained by a given university. Although federal expenditures in R&D measure national research competitiveness, this statistic underestimates the local impact of research. The University of Kansas (KU) uses the same methodology, but extends it to include fields outside of science and engineering and to include research training grant expenditures. This is a measure of RD&T—research, development, and training expenditures. KU’s expenditures for RD&T rose 15% from fiscal year 1999 to 2000.
- The U.S. Department of Commerce estimates that in Kansas each \$1 million in R&D funding creates 40.6 jobs. The three Kansas research universities had \$335.2 million in RD&T expenditures in fiscal year 2000, which implies that more than 13,600 jobs are due to this source of funding. The average salary in these jobs exceeds the average salary in our state.
- Graduates are the largest form of technology transfer from research universities. The annual income of the alumni of the three Kansas research universities, who currently reside in the state, is \$9 billion. About 1/3 of this total, or \$3 billion, is due to the increased salaries they earn due to their degrees. The state tax paid by these graduates is \$700 million annually, a figure that exceeds the annual state appropriation to the three universities of \$400 million.
- To maximize research productivity, we must minimize internal competition between academic departments and research centers. KU uses a multiple credit algorithm to accomplish this; expenditures are recorded in two lists, one according to departments and one according to centers.

PANEL OF RESEARCHERS

Carol Shanklin

Professor of Institutional Management and Dietetics, and Assistant Dean of the Graduate School, Kansas State University

Michael Podgursky

Professor of Economics, University of Missouri - Columbia

Susan Kemper

Distinguished Professor of Psychology and Gerontology, University of Kansas

- To maintain research productivity, it is important to encourage participation at all levels: faculty at all stages of their development; pre-doctoral and postdoctoral training programs; and mentoring and support of the most productive faculty so we don't lose them.
- Some research supports the notion that productivity is greater in larger institutions and departments because of the "intellectual synergy." Other factors to consider at the departmental level are: workloads, availability of leave-time and travel funds, the number of students on research support, availability of non-governmental funds, and availability of star faculty.
- Senior faculty are motivated to remain active as scholars by the intrinsic rewards of mentoring their graduate students. They also thrive on public recognition of their contributions to the profession. Interdisciplinary teams can energize faculty by creating opportunities and stimulating new research.
- The National Research Council rankings in economics are strongly associated with objective measures of productivity such as total citations or total pages in refereed journals. NRC data also establishes a link between size of department and rank. The large departments in the top 50 tend to have faculty in a variety of fields, which would seem to discredit the strategy of building a "unique niche."
- The individual faculty member is motivated to be productive because of his/her "passion for reputation" and "taste for originality." The challenge is to find how these attributes then lead to publications, citations and impact assessments at the unit-level. "Bibliometrics" is not helpful in actually fostering productivity.
- The researcher who is productive over a long career may experience multiple peaks and valleys as he/she invests additional time in acquiring new skills and competencies in order to develop new lines of investigation. The system for evaluating research productivity at the unit-level must reflect this non-linear career trajectory at the individual level.

FIRST PANEL OF RESEARCH ADMINISTRATORS

Thomas H. Rosenquist

Vice Chancellor for Research, University of Nebraska Medical Center

Suzanne Ortega

Dean of the Graduate School and Vice Provost for Advanced Studies,
University of Missouri – Columbia

K. Michael Welch

Vice Chancellor for Research and Senior Associate Dean for Research
and Graduate Studies, University of Kansas Medical School

- One departmental model for faculty evaluation is, in a sense, “Darwinian.” It quantifies annual research productivity for all investigators; derives an average; compares each investigator with the average investigator; and then distributes rewards accordingly. Investigators who do poorly over time are weeded out, whereas the strongest, most adaptable in the department, thrive. Whether the evaluation system is “Darwinian” or “Egalitarian,” faculty members are not completely satisfied; however, it seems that faculty in departments with a major Darwinian component are more satisfied than they are in circumstances where the chair gives a highly subjective, or no, evaluation.
- We haven’t determined how to recognize the quality or productivity of activities that are not for an academic audience. In evaluating the outcomes of graduate education, how do we judge the placement of chemistry students who go into non-academic institutions? Should the standard be placement in Fortune 500 companies? Do we count the number of students who start their own companies? Do we count patents? Is a number an adequate indicator of productivity, or do we attach a dollar value? As universities move in the direction of increased collaboration with industry, with increased public accountability, and respect for the wide range of career opportunities for our doctoral degree recipients, it will become more important to develop assessment and evaluation strategies that align with the values and goals of our non-academic audiences.
- We must be careful to develop appropriate measures of quality and impact in the arts and humanities or we may erode the position of these disciplines at our institutions, especially when measures of “impact” drive resource allocation models in the future. By intention or happenstance, our support of the arts and humanities will be an important statement about our institutional values.
- It is difficult to convince faculty that they should be interested in assessment as a strategy for improving the things they care about, i.e., the preparation of the next generation of scholars and researchers; faculty often believe that administrators actually want a quick and efficient way of allocating—or more frightening still, reallocating—resources.

- For biomedical institutions, total NIH award is a measure that meets key characteristics: it can be measured in a simple, easily understood and goal-directed manner. It has a strong association with other markers of research productivity and it is a clear outcome. However, the use of a productivity index must not be confused with the goals and values of the institution, which include scholarship, clinical care, education and service.
- To make NIH funding the gold standard in an institution, each school must have its own mandate to increase NIH funding and create a strategic research plan for a 5-7 year period, with award targets as their goal. Administrators who set the goals and oversee the process should be held accountable, using NIH awards as the productivity measure of the programs in their area of responsibility.

SECOND PANEL OF RESEARCH ADMINISTRATORS

R.W. Trewyn

Vice Provost for Research and Dean of the Graduate School, Kansas State University

Jack O. Burns

Vice Provost for Research, University of Missouri – Columbia

James R. Bloedel

Vice Provost for Research & Advanced Studies, Iowa State University

- Those who are concerned about the University's return on investment are: governing boards, accrediting bodies, funding agencies, state legislators, taxpayers, prospective students, employers of students and bill-paying parents. Education is a value-added product. It creates a significant difference in income for the student, particularly when comparing the salaries of high school graduates with those who earned a masters degree. It also makes a difference to society; we have estimated that Kansas State University alumni paid \$250 million in taxes from the earnings garnered from a college education. In agriculture, the value of research can be measured in dollars; for example, KSU high-yield wheat has been shown to generate \$64 million more in income for farmers.
- Increasingly we must look at technology transfer and create new productivity measures: licensing income; licensing-linked research funding; companies launched and jobs created. In a document published by Kansas State University in 1998 on the economic impact of research and teaching, we estimated that the return on investment is \$17 for every dollar.
- The master campus plan at the University of Missouri – Columbia involves these goals: maximizing internal resources and communications; enhancing research compliance; providing grant assistance; nurturing technology development; expanding external partnerships; and fostering governmental relations. To make better use of internal resources, MU has established a campus network of 55 grant writers and a grant information system. Through its office of Technology and Special Projects, MU

provides mentoring on technology transfer and develops relationships with economic development entities as it encourages entrepreneurship on campus.

- The scholarship of the scientific faculty is now more diverse. Many individuals are interested in the wide-range of experiences that result from entrepreneurial activities, not as a substitute for their more traditional scholarship activities, but rather as a complement to their professional experiences while serving our institutions. Their contributions not only add to the research culture on our campuses, they also provide unique training opportunities for our undergraduate and graduate students. These training opportunities support current trends in graduate education that emphasize the importance of meeting the needs of students interested in careers in industry.
- If we are to attract and retain faculty who are entrepreneurial, a broader definition of productivity is needed; “objective-driven scholarship” can apply to educational initiatives as well as extension activities. Evaluation would then be based on “impact on the field.” To meet this standard, the faculty must demonstrate a set of contributions that has impacted a field in a way that modified thinking and/or trends among other scholars in the same area. For Promotion and Tenure, I suggest setting up external study sections with experts chosen on the basis of their capacity to assess the impact of research. This process would parallel the one established by NIH and NSF for evaluating grants and contracts.

A REFLECTION ON A DAY SPENT DISCUSSING EVALUATION

David Shulenburg, Provost, University of Kansas

- A market model of evaluating our productivity does not work unless universities can demonstrate that they are covering the full cost—and yet all our activities are subsidized. This is why we fall short in using measures such as the quantity of external funding.
- Our arguments about state funding and higher education may not succeed because the public knows our contributions are not unique. Higher education does improve an individual’s income, but if our state did not support universities, students would be able to seek an education elsewhere. Likewise, the public may not accept the argument that our institutions give a good rate of return on monies invested because we cannot say that the leverage we provide for investment in higher education is really better than the benefits derived from money spent on traffic safety or early childhood education, etc.
- The public relies on *U.S. News and World Report* for information on college rankings, not the National Research Council. The value of the process is further jeopardized if our evaluation schemes do not measure up to our ideals, and our faculty don’t believe in the process.

HIGHER EDUCATION ADVOCACY: THE INTERFACE OF TWO CULTURES

Kim A. Wilcox, Executive Director, Kansas Board of Regents

- There are many differences between the culture of academe and that of the legislature. It is difficult to understand the way compromise is reached in the legislature, but it is critical to accomplishing their objectives.
- We need to spend more time thinking about what it is we are doing in academia and how our work can be cast into an appropriate form.
- The universities need a focused message in communicating with the legislature. Too often, we find ourselves espousing our own individual needs and positions. It is important to unify our voices and if we do this, it makes it easier for the press to espouse our position and for the legislature to accomplish our objectives.
- Our long-term work with legislators should combine “friend raising” with fund raising.

ADMINISTRATIVE PERSPECTIVES

Rollin C. Richmond, Provost, Iowa State University

James Coffman, Provost, Kansas State University

- Academics in the developed world have contented themselves for many centuries with the same approaches to education as they themselves experienced. Few scholars have read the literature on learning styles and best practices for teaching.
- We are likely to experience increasing difficulties in attracting public and private support for our institutions unless we change the way that we reward academic scholarship. Scholarship can be viewed from a broad perspective. It is integral to all three components of higher education: learning, discovery and engagement. Scholarship can be communicated through: teaching materials and methods, classes and curricula as well as publications, presentations, exhibits, performances, and also patents, copyrights and the web. We need to place scholarship in the context of the institution it serves, not just the discipline it supports.
- Applied science is what John Maddox says has dramatically changed and improved the lives of people in the 20th century. Iowa State University has invested in applied science via the Plant Sciences Institute.
- Funding patterns for higher education in Kansas reflect more than 100 years of decisions to foster a high participation rate via community colleges and technical schools as well as the regional and research universities. In an economic development context, we see a difference in funding priorities in Nebraska, Iowa, and Missouri where more investment is being made in research universities.
- While every research university works to the limits of its ability to expand research and development, this happens in a context in which education retains primacy. The federal agenda for academia is focused on research,

- and the state agenda is founded and evaluated primarily on the basis of undergraduate education. These two forces frequently are in conflict.
- The standard model of the complete scholar is too constraining to be affordable. Not everyone is able to maintain a research output that is nationally competitive, and even fewer can establish and maintain a national reputation. While it is in everyone's best interest to celebrate and capitalize upon those who can produce optimally in teaching, research and service, we should recognize that not everyone can do this over the entire course of a career. It is most effective to create flexibility in roles and rewards so that work can be allocated according to an individual's strengths, especially during the post-tenure period.

MEASURING SUCCESS:
LESSONS FROM THE NRC STUDY OF THE
RESEARCH DOCTORATE

Joan F. Lorden
Associate Provost for Research and Dean of the Graduate School
University of Alabama at Birmingham

Why Measure Performance?

For a complex organization like a university, it is a formidable challenge to communicate what we do and its value. The task is made more difficult by the wide variety of audiences that must be addressed. The 2000 Merrill conference focused on making university research part of the public agenda. The public, at this conference, was broadly defined to include not just public officials, potential donors, and industry but also those that we might think of as an internal audience: alumni, students and their parents. Whether the audience is internal or external, we need tools for communication.

Universities have two basic ways of expressing what it is we do and why it is important: numbers and stories. By stories, I mean the narrative explanations that we offer about the significance of the work we perform. Our publications highlight student achievements, faculty discoveries, and the services the university provides for the community. These narratives provide the context for the statistical data that we assemble. As instruments of persuasion, the narratives are compelling and effective, but they are unwieldy when our need is to: set benchmarks, recognize excellence, and promote improvement. Quantitative performance measures can be tools for persuasion but ideally, they are also objective indices that we can use to help set and meet our goals.

The kinds of questions we are called upon to answer imply comparison, either with our own past performance or with that of others. How do we compare? Do we meet the standard? Are we getting better? What are the best practices? To be able to answer these questions, we need to measure the right things in the right way. We need to be sensitive to the limitations of what we are measuring. And when we start comparing ourselves to others, we need to recognize the very real possibility that if we tie the measures to rewards, we will change behavior and become what we measure.

I would argue that there are four basic elements to consider when we select measures for evaluating performance: goals, audience, values, and the practicality of the measures themselves. We need to be able to answer the questions:

- What are we trying to achieve?
- Who are we trying to reach? Do the questions we pose address the concerns of the audience?
- What do we consider important? What behaviors do we want to promote?
- What are the practical limits of what we are trying to do? If we say we are measuring quality or effectiveness, do the measures we have at hand really allow us to do it?

For the remainder of my talk, I would like to discuss *Research Doctorate Programs in the US: Continuity and Change*, published in 1995 as a case study. This work is usually referred to as the NRC study, since it was conducted by the National Research Council (NRC). The NRC appointed a Committee for the Study of Research Doctorate Programs that actually undertook the study. This study was the subject of a position paper by the Council on Research Policy and Graduate Education (Lorden & Martin, 1999). Some of my comments will be drawn from that paper. Following my critique, I would like to provide a few examples of the ways we have tried to measure performance at the University of Alabama at Birmingham (UAB) in research and graduate education.

Why Examine the NRC Study?

The NRC study is interesting for a number of reasons, not the least of which is that it produced rankings based on program reputation. Let me mention a few arguments for taking time to look carefully at the NRC study:

- Any rankings tend to be quoted and used. (Take a look at your campus website; See also Webster & Massey, 1992.) For these reasons, it is worth knowing where they come from and what they might or might not mean.
- There are many criticisms of the NRC study, but the reputational rankings, unlike other published ranks, were done by people in the field and were presented alongside objective measures of performance. The choice of the measures reveals the values of the academy.
- Intended as a study of graduate education, the NRC study is almost as much a study of research. Graduate education and research are so

closely intertwined that many of the measures selected are those that one might choose if evaluating the research programs of a university.

- Although the study was thoughtfully designed, its measures have limitations. It is one thing to measure quantity; it is another to measure the quality of what is produced. Because the NRC study is the major national study of graduate education and is soon to be repeated, its measures deserve scrutiny.

As background, let me mention a few of the main features of the NRC study. First and foremost, there was a reputational survey from which rankings of programs within disciplines were derived. The reputational survey was supplemented with objective measures. Similarities with earlier studies allowed for longitudinal comparisons within most disciplines. Basic institutional information such as enrollment, library holdings, and level of research funding was collected and reported.

Evaluating the NRC Study

How does the NRC study look if we apply the four questions raised above about goals, audience, values, and practicality of measures?

- One of the declared goals of the NRC study was to assess graduate education in the United States at a time when doctoral enrollment had reached an all time high and more institutions were offering the doctoral degree. There was a perceived need for data to guide policy decisions. The study was undertaken to address the quality and quantity of research doctorates. The objectives for the study were to: update data last collected in 1982; collect new information; analyze components of the new database; and make the data available for further analysis. In addition to having people in the field rate the merits of different programs, the NRC collected quantitative measures of faculty productivity covering publications, citations, and funding. These measures were presented so that one can get a sense of their distribution across the faculty in a program. The information reported included: the percent of faculty publishing, the number of publications/faculty members, and the gini coefficient that measured the degree to which publications were concentrated in a small number of program faculty. Whatever the initial intent, once published, the study became a ranking of graduate programs and universities, and there was little discussion of other measures (e.g., Webster & Skinner, 1996).
- The study was aimed at multiple audiences. The committee that guided the study expected that it would be useful to students and their advisors making choices about graduate programs and that it would

inform the judgment of university administrators and other decision makers at the state and national levels, including those in funding agencies. Finally, the data were presumed to be of interest to scholars in graduate education and to those involved in the research enterprise. One can reasonably ask if it is possible to address an audience this broad with a single document when only limited interpretation is provided. Individual institutions have used the rankings as publicity and to argue for investment (Lorden & Martin, 1999; Mervis, 2000), but there has been no systematic “study of the study.” We do not know what the overall impact has been in terms of student choice, or in terms of institutional and other investments. Much of the public debate has revolved around what the rankings mean and how much weight to give them.

- The NRC study is an example of self-examination by the academy. The committee that guided the study was drawn from the academy and the product must be assumed to reflect its values. The choice of variables was based on assumptions about the features of doctoral training environments that facilitate the preparation of research scholars and scientists, including: the reputation of the faculty, their publications, and their funding and awards. Measures related to the subjects of graduate education, the students, were limited to: the number enrolled, the number of women, and the number of degrees reported. The information secured on graduates was: sex, minority status, the percent supported as research or teaching assistants, and the time to degree.
- Because the NRC study was national in scope, it included a broad coverage of disciplines and institutions and a uniform method of data collection. The breadth of coverage was one of the features that made the study interesting and useful, but it also meant that some measures were too costly to undertake, particularly those related to students and alumni. As a practical matter, the study was also too big to be done very often. Data points ten years apart will not track the movement of faculty in and out of programs that can result in significant changes in program profiles.

Given the broad goals and wide audience that the NRC study aimed to reach, it is not difficult to enumerate omissions or aspects of the study that could have been done differently. As a study of graduate education, the most obvious omission was a valid measure of program effectiveness. Program effectiveness was presented in a measure called 93E and was obtained through the same survey of reputation that produced 93Q, the measure of program quality used for the rankings. These two measures, 93Q and 93E, were highly correlated, leading to the criticism that the raters had no real knowledge of program

effectiveness and so made the assumption that reputation and effectiveness were the same. The measure was inadequate to the task.

The study also lacked any measure of student outcomes. This is an admittedly difficult area to tackle on a national scale, but we would all agree that it is an important measure of the success of a program. For prospective students and their advisors, it may be **the** measure of interest. The absence of measures for student outcomes does not imply that the study committee disregarded this valuable information, but it does illustrate that we tend to measure what we can and not necessarily what we need or value.

A similar point can be made about other aspects of graduate education that have become increasingly important. Interdisciplinary programs and research have grown, and many emerging areas are inherently interdisciplinary. Others, such as the disciplines encompassed by the biomedical sciences, have evolved into interdisciplinary fields. It is not a simple task to measure how interdisciplinary a program is and whether it produces students with breadth of training. If we value interdisciplinarity and want to promote it, we need to find a way to capture it in studies of graduate education.

Without question the variable that has been the most frequent object of discussion is 93Q. Defined as the mean scholarly quality of program faculty, 93Q was the variable on which programs were ranked and was intended to capture the perceived intellectual resources in a general field. The measure does not, however, tell us about the experience of students. Nor does it capture the quality of faculty performance or mentoring in graduate education. In fact, we know little about what 93Q actually measures. Reputations are slow to change and may be influenced by a variety of factors, including halo effects of the institution or one prominent faculty member. We could ask whether older programs or larger programs with many graduates are more likely to be familiar to the raters. Translated into rankings, 93Q was reported to two decimal places, and many programs differed only in the third decimal place. The confidence intervals presented in the appendices revealed that the quality of programs as established by this measure could be distinguished only in rather broad terms.

What Did We Learn?

Given the limitations of the NRC study, we still learned several things. There were interesting findings about change in programs over time. We also gained information about: the impact of program size; faculty involvement in research and its relationship to rankings; and the relationship of program ranking to student variables. For example, when comparisons with the 1982 study were possible, they indicated that there is great stability at the top and the bottom of the rankings. Differences in fields over time could also be discerned. Biology underwent radical change during the 1980's and the taxonomy of programs in the 1995 study bore little resemblance to the earlier study. In the sciences and engineering, the greatest growth in programs occurred at the top of the rankings.

In the arts and humanities, the largest decreases in program size occurred in the top quarter. In general, new programs tended to be ranked low.

Looking at the measures of research and scholarship, several points stand out:

- Federal funding in science and engineering was highest in programs in the top quarter.
- The percentage of faculty who are publishing varied little across the sciences and engineering.
- Per capita publications correlated significantly with ranks.
- Citations/faculty were much higher in the top quarter in the sciences.
- Numbers of awards and honors were much higher in programs ranked in the top quarter in the arts and humanities.
- Citations tend to be more concentrated in a few faculty members in the top ranked programs, rather than being broadly distributed, indicating the presence of a few highly influential individuals.

Examining the relationship of program size to ranking reveals that top ranked programs in all fields tended to have larger numbers of faculty and larger numbers of students than lower ranked programs, although as noted above, the top ranked programs in the humanities tended to decrease in size during the 1980's. The relationship of size to ranking has raised questions about how we evaluate the overall quality of niche programs. Do programs that are more narrowly defined necessarily offer a weaker intellectual experience? The NRC data also allow one to ask how a particular program does, given its size.

The NRC study confirmed what most institutions know—that time to degree increased across all ranks and all disciplines when compared with the 1982 data—but we also learned that the increase was greatest in the lower ranked programs. This may be related to another observation: students from lower ranked programs were more often supported by teaching assistantships. Those in more highly ranked programs were more often supported as research assistants, coincident with the greater availability of research funding in those programs.

Stepping back from the data, one can build an image of what it would take to construct a top ranked program. Some of the elements would certainly be: “star” faculty, a wide range of faculty representing all aspects of the discipline, and resources sufficient to support a large number of students without heavy reliance on teaching assistantships. Having done that, it is worth stopping to reflect on the study by Cerny and Nerad, known as the “Ten Years Later” study,

in which the emphasis was on student outcomes. In a presentation to the National Association of State Universities and Land Colleges (November 1999), Cerny presented data from a national dataset in which graduates who had received their degrees ten years before and who were employed, were asked about their graduate experience (see also Mervis, 2000). The response to questions about whether or not a person would repeat the degree program and, if so, if he or she would do so at the same institution, was remarkable when put in the context of the NRC study. A high rank was clearly no guarantee of a positive experience. The relationship between program rank and the willingness of graduates to repeat the experience at the same institution was weak to nonexistent. This is not to say that faculty reputation and productivity are unimportant in graduate education, but clearly, we need to know more about what it means to the education and experience of the student. If our goal is to sustain and improve graduate education, the question of how to rate a student's experience or how to define the effectiveness of a program deserves an answer. These are topics that we must tackle both at an institutional and a national level.

Measurement of Performance in an Institutional Context

Moving from a national study of research and graduate education to assessment at the institutional level allows one to revisit the question of goals, audience, values, and practicality on familiar turf. I would like to give you a few examples of our efforts to assess the performance of programs at UAB. As background to this discussion, let me point out that UAB is a relatively new institution. We have only existed as an independent institution since 1970. The growth of the research enterprise and graduate education during this short history has been substantial. Although it offers a comprehensive education with undergraduate, graduate, and professional degrees in a wide variety of areas, the university has focused its development around its strengths in medicine and health.

The easiest measure to present when discussing research is funding. As we have seen in the NRC study, the ease with which data can be collected and presented is not necessarily an indication of its importance or relevance. Just as the rankings of the NRC study were the easiest piece to communicate, an institution's extramural research funding is the easiest way to express the level of research activity.

In communicating with both internal and external audiences, all institutions produce graphs that show the changes in research dollars over time. One that we use frequently at UAB is a graph that shows extramural grants and contracts awarded over the past decade. This clearly puts UAB in a positive light, since we have grown from about \$100 million to just over \$300 million in awards. Like other institutions, we compare ourselves to various peer groups, depending on the audience we are trying to reach. A table that I have often shared with trustees and visitors shows how UAB compares in federal research and

development expenditures with other institutions, nationally and in the south. Nationally, UAB ranked 27th in 1999, and among southern institutions, we ranked third, just behind Duke and the University of North Carolina at Chapel Hill. In funding from the National Institutes of Health, our single largest sponsor, UAB ranked 20th nationally in 2000 and fourth in the south, following North Carolina, Duke, and Baylor. This is a good showing for a relatively new institution and the achievement has garnered attention nationally, but it is not the whole story of our research enterprise.

It is clear from examining research rankings based on funding that if you evaluate your institution by funding alone, it is difficult to move up and keep up. In 1997, UAB ranked 30th in federal research and development expenditures with \$125 million. To move up to 27th by 1999, we had to increase our federal expenditures by \$40 million. Although our numbers were increasing, we were not alone. This level of growth is difficult to sustain. Over the same time period, Duke had an increase of \$36 million in funding and moved up from 24th to 23rd. The University of North Carolina at Chapel Hill had a \$25 million dollar increase and decreased in the rankings from 20th to 25th. Do these changes in rank translate into significant changes in the quality of the institutions or the worth of the research supported? I would never argue that, but federal research funding is important because it is awarded based on what most investigators would agree is a fairly rigorous peer review process. While the process is not without its faults, the award of millions of dollars of research support annually from a variety of agencies represents a significant consensus on the value of the work produced by the faculty of the institution. Ultimately, the merits will be assessed by the impact of the findings on problems the research addresses as they are published and translated into products.

Another measure that indicates the success of the faculty at UAB is the percent of proposals to federal agencies that are funded. The fact that we exceed the national average is an indication of the quality of the proposals that are subjected to the peer review process. As a public institution, we also consider in various ways the return on the state's investment in the institution. Appropriations to support public universities differ substantially from state to state. Compared with other research-intensive universities, UAB's state appropriation is modest. The current appropriation is about equal to the university's expenditures in federal research and development dollars. It is important to note the impact of the federal dollars in jobs, income to local government, and spending for goods and services in the city and state because this can be compared to the impact of other state investments and industries.

Limited state funding has led UAB to focus on biomedical research with the goal of being preeminent in this area. The rest of the institution does not flounder as a result, because our emphasis is on building depth and making connections in related fields. We can measure the impact of the investment by looking at the number of students in the undergraduate programs doing research

or by the number of participants in the research enterprise from schools outside medicine. Our approach has been to stay focused, set priorities, fund according to opportunities for leveraging university funds, and then monitor success. The questions we ask are: do we have the basic resources to make an impact? Will we be able to muster enough external resources to sustain a program?

As part of a strategic planning exercise that the university undertook in 1994-95, we affirmed the importance of collaborative, interdisciplinary research to the institution. This was an important factor in the success of the university in research, and it was agreed that as the university matured, we needed to have structures in place to sustain and foster a collaborative environment. As part of this effort, we developed a mechanism for funding our major University-wide Interdisciplinary Research Centers (UWIRC) based on evaluations conducted every three years.

The success of a center is judged on the extramural research funding associated with it, but a number of other qualitative and quantitative criteria are also important in determining whether a center will receive funding and how much it will receive. A defining characteristic of a UWIRC is substantive involvement of members from at least three of the university's twelve schools. Some, like the Center for Aging, have participation from all twelve. The other factors on which these centers are judged include: the resources that they provide for the campus; their outreach efforts; their contribution to education and the intellectual environment through courses and seminars; their contribution to translational research; and the extent to which they serve as a resource for the state and region. Evaluating the centers on these criteria has led to changes in the way they operate. Many have started outreach programs that they would not otherwise have had. Others have initiated post baccalaureate certificate programs or specialized courses. Over the six years that the program has been in operation, we have experienced increased participation as the centers develop pilot grant programs and encourage new interdisciplinary activities.

In graduate education, programs are evaluated on the basis of internal training grants for the award of fellowship and assistantship positions. The programs define their mission, describe the curriculum and degree requirements, and then provide data on the following measures, some of which overlap with those in the NRC study:

- Applicant population (e.g., size, number of international and minority applicants);
- Characteristics of students matriculating (e.g., grades, test scores, percentage of minority students, percentage of international students);
- Funding levels and sources for research and student support;
- Training experience of mentors;
- Publications of mentors and students;
- Intellectual environment, facilities;

- Time to degree;
- Attrition;
- Placement of graduates.

As is the case with the UWIRC funding program, we evaluate graduate programs on the basis of defined criteria and tie resources to changes in behavior that are demonstrated by performance. For example, enrollment of African American doctoral students has more than tripled since 1989, increasing from 53 to 178. Programs with high attrition rates have been forced to examine the reasons and take corrective actions. In both cases, putting a spotlight on an issue made a difference in the attention that programs gave it.

Conclusions

Historically, universities have been among the most stable institutions in our society and their contributions are numerous. The performance of universities as the purveyors of knowledge and the creators of new knowledge has been a great success story. University research has been the basis for improvement in the lives of all citizens and has been an economic engine for the nation. Universities have served as the entryway for new citizens seeking full participation in the cultural and economic life of the country. Graduate education draws students to the U.S. from around the world. Despite these and other successes, universities, particularly public universities, are increasingly being held accountable by a multiplicity of audiences. In this environment, measurement of performance is inevitable.

As members of the academy, we need to develop ways to capture accurately the performance of our institutions and their programs. In choosing measures for the future, we need to bear in mind our goals. Why are we engaged in a measurement process? Are we asking how to move up in the ranks? Or, do we want to know how we have served the state or advanced our mission? We need to ask whom we are trying to influence: faculty and administration, prospective students, donors, the legislature? These groups have different questions for us, and measures designed to answer those questions will be more effective. Most importantly, we need to acknowledge the power that performance measures have to change the behavior of individuals. We will not always agree on what to measure and not everything that we value will be easily captured in quantitative measurements. But as members of the academy, we are in the best position to develop valid measures that will promote our values and apply them in ways that sustain and enhance our mission.

References

Goldberger, M.L., Maher, B.A., and Flattau, P.E. (Eds.) (1995). *Research Doctorate Programs in the United States: Continuity and Change*. Washington D.C.: National Academy Press.

Lorden, J.F., and Martin, L.B. Towards a Better Way to Rate Graduate Programs. (www.nasulgc.org/councils_research.htm)

Mervis, J. (2000). Graduate Educators Struggle to Grade Themselves. *Science*, 287, 268-570.

Webster, D.S., and Massey, S.W. (1992). Exclusive: The Complete Rankings from the *U.S. News & World Report* 1992 Survey of Doctoral Programs in Six Liberal Arts Disciplines. *Change*, 24, 20-45.

Webster, D.S., and Skinner, T. (1996). Rating PhD Programs. What the NRC Report Says...and Doesn't Say. *Change*, 28, 22-44.

RESPONSE TO THE KEYNOTE ADDRESS

Robert E. Barnhill
Vice Chancellor for Research and Public Service
President, KU Center for Research
University of Kansas

Prologue: Research Competitiveness

Research productivity is important to the nation. Since "we become what we measure," we need good methods of evaluating this productivity. Joan Lorden has introduced some important questions for our topic today, "Evaluating Research Productivity."

It is my experience that **leadership at every level is essential for institutional research competitiveness.** This was the principal conclusion from a 1995 conference on Research Competitiveness. The American Association for the Advancement of Science (AAAS) convened some forty people at Kiawah Island, South Carolina. Included were

experts in research policy such as Roger Geiger, Irwin Feller, Susan Cozzens, and Harry Lambricht. The purpose of the meeting was to help EPSCoR states become more competitive in research. The AAAS invited two "outliers," that is, two people who had been successful in non-EPSCoR states, to pass around their secrets of success. Those two people were George Walker from Indiana University and me, representing Arizona State University. This was my first meeting with George Walker and also with the national research policy experts. Subsequently, my institutions, Arizona State University (ASU) and the University of Kansas (KU), and I personally have profited from meeting George Walker and the other research policy gurus.

We prepared manuscripts prior to the 1995 AAAS meeting, which then became a published book (see references). Roger Geiger's pre-meeting manuscript described the overall research scene, focusing on federal expenditures. He mentioned that only five universities had made a considerable

Prologue: Research Competitiveness

- AAAS conference on research competitiveness, 1995
- Necessary for research competitiveness:
 - Presidential backing
 - Research VP leadership

Leadership at every level is essential for institutional research competitiveness.

improvement in research competitiveness in the 1980s and early 1990s; ASU was one of the five. Geiger went on to say, “Presidential backing for strengthening research is a virtual prerequisite. In some cases, presidents have identified themselves with ambitious research goals; in others, presidents have more quietly backed the efforts of provosts or vice presidents for research (ASU).” Geiger also said, “An institutional commitment to research almost presupposes the organization of research administration under a single office. The office of the vice president for research does far more than standardize research accounting and offer administration support. It should become the initiator of and advocate for proactive policies.” The Kiawah Island conferees agreed that leadership is essential for institutional research competitiveness. This includes not only the president/chancellor, but also the faculty and the rest of the university research community.

Prior Merrill Center Research Policy Meetings

Let me review for you the last three Merrill Center conferences on research policy. The keynote speakers were Michael Crow, Columbia University, Luis Proenza, University of Akron, and George Walker, Indiana University.

Michael Crow, now the Executive Vice President for Research at Columbia, emphasized

the “niche” strategy of highlighting a few areas of institutional expertise. Luis Proenza, formerly Vice President for Research at Purdue University and now President of the University of Akron, discussed “strategic intent” and its ramifications in collaborative efforts. George Walker, Vice President for Research and Graduate School Dean at Indiana University, has discussed the Indiana story of mobilizing “the public” to support research. I will take up each of these three themes in addition to our topic for today.

Prior Merrill Center Research Policy Meetings

- Mobilizing for Research Opportunities in the Next Century
 - Michael Crow, Columbia University, 1998
- Building Cross-University Alliances that Enhance Research
 - Luis Proenza, Purdue/University of Akron, 1999
- Making Research Part of the Public Agenda
 - George Walker, Indiana University, 2000

Current Merrill Center Research Policy Meeting

Our current topic, evaluating research productivity, involves devising ways to measure what we think should be measured. Joan Lorden has initiated this discussion for us, drawing on her work with Lawrence Martin, SUNY at Stony Brook, and others in the Council on Research Policy and Graduate Education. This council of chief administrative officers in charge of research policy, administration and graduate education is drawn from the membership of the National Association of State Universities and Land Grant Colleges. The NASULGC position paper co-authored by Dr. Lorden focuses on the U.S. National Research Council ratings of graduate programs. "Towards a better way to rate research doctoral programs" is posted on the NASULGC web site: www.nasulgc.org/councils_research.htm

Current Merrill Center Research Policy Meeting

- Evaluating Research Productivity
 - Joan Lorden
 - University of Alabama at Birmingham, 2001

Setting the Stage for Success

The new millennium is an exciting time for research in general and science in particular. It is an excellent time for organizations to take stock of their goals, resources and impacts.

Most universities have great stability and long and honorable histories. But these days it is also important that institutions

develop a certain level of flexibility so that they can move as quickly as possible when necessary. Each university must answer these questions:

- Can we remain relevant to today's fast moving world? Or will we be relegated to a genteel backwater role in American society?
- If we wish to remain, or become, relevant, how can we do it?
- What are reasonable goals and how can we achieve them?
- How can we measure our progress toward these goals?

Setting the Stage for Success

- Setting the agenda for 21st century science:
Only the flexible will thrive.
 - Andy Grove, Michael Crow
- Strategic Intent (Competing for the Future)
 - Direction
 - Discovery
 - Destiny

Michael Crow estimates that in the near future there will be about 75 significant research universities in the U.S. These select universities will obtain almost all the competitive federal funding.

Andrew Grove, CEO of Intel Corporation, has written the book, *Only the Paranoid Survive*. Let me paraphrase the book's title to: *Only the Flexible Will Thrive*. Only those universities which are flexible in their approach and which have clear goals and expectations will do well or even have the chance of being among Crow's 75 universities.

Two years ago, Luis Proenza introduced us to the key concept of "strategic intent," as examined in the book *Competing for the Future*. Strategic intent has the attributes of direction, discovery and destiny.

Direction: "Most companies are over-managed and under-led." That is, "more effort goes into the exercise of control than into the provision of direction."

Discovery: "Strategic intent should offer employees the enticing spectacle of a new destination or at least new routes to well-known destinations."

Destiny: "Only extraordinary goals provoke extraordinary efforts." Thus, numerical goals are less energizing to employees (or researchers) than goals such as being "the best" in defined competitive areas.

Strategic intent goes beyond strategic planning. Strategic planning is a "feasibility sieve." Strategic intent goes beyond the feasible to what is barely possible, e.g., President Kennedy's vision of a space landing on the moon, or our efforts today to find a cure for cancer.

Arrival at Destination

There are several ways to tell that an institution has arrived at a suitable research destination. Examples are shown here.

Arrival at Destination

- High institutional rankings
- World class research areas
- Cash
- Fullest utilization of university community
- Value added to society

Successful Examples

Let me take you through four examples of strategic intent applied to public universities.

1. Arizona became a state, the 48th, in 1912. It still feels like a frontier. A few of you know that I spent eleven exciting years at Arizona State. I want to discuss the example of ASU's friendly rival down the road, the University of Arizona. From Roger Geiger's book, *Research and Relevant Knowledge*, "the same factors that have been identified in

Successful Examples

From the 1960s:

- University of Arizona
 - Roger Geiger, Research & Relevant Knowledge
- University of Utah

From the 1980s:

- Arizona State University
- University of Alabama at Birmingham

the advancement of other research universities—establishing centers of research excellence, academic leadership, and the availability of resources—were vital to Arizona as well." In 1959, President Richard A. Harvill stated that "Arizona's role in the expanding research economy would be to concentrate on fields in which it possessed some natural advantage." (Clark Kerr has also used this phrase "natural advantages.") At the time, just after Sputnik in 1957, the University of Arizona had only \$1 million in federal funding and no nationally recognized departments. In the years that followed, two centers emerged, one in astronomy and one in anthropology. Each relied on natural advantages: astronomy on Arizona's clear skies and nearby mountains for observatories; and anthropology on the presence of a large number of Native American tribal nations. (There are twenty-one tribal nations in the state.) In 1966, the two corresponding departments became the first University of Arizona departments to receive national recognition in reputational rankings.

Geiger discerns a pattern to establishing these university centers of research excellence:

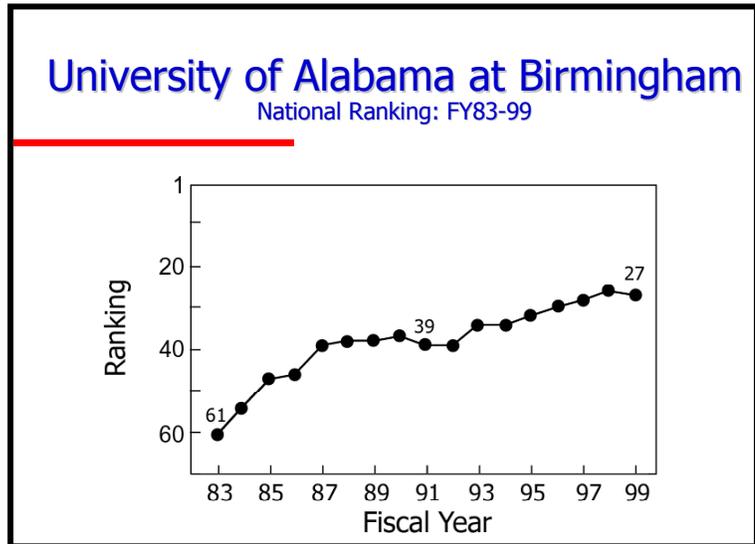
- a natural advantage
- topics a little off the beaten academic path
- areas of excellence that have far-reaching effects on the rest of the university.

Note especially Geiger's expansion on his third point: "...achieving these pockets of research excellence...overcame a kind of defeatist attitude that was prevalent on the campus."

2. Before going to Arizona State University, I spent twenty-two pleasant years at the University of Utah, in Salt Lake City. Technology sectors in Salt Lake City account for some \$10 billion in annual revenues. Five of the six key factors in the city's development as a technology center hinge on the University of Utah. One spin-off company—the Evans & Sutherland Corporation—has helped create more than 150 computer and software companies. In 1965, David Evans came to the University of Utah to chair the Computer Science Department. In the 1970s, he brought Ivan Sutherland to the university with the strategic intent of forming the premiere computer graphics group in the country. Evans & Sutherland formed their company in the university's new research park. (Many of my own students in the mathematics department worked for the new company.) The University of Utah Research Park was itself a product of strategic intent. Wayne Brown, Dean of Engineering, worked with President David Gardner to inaugurate the research park. Their strategic intent was to develop a place where local entrepreneurship and expertise could flower. Evans & Sutherland became the anchor tenant of the new park. The three elements of direction, discovery and destiny prevailed for all of these people relative to their respective goals.

3. I now turn to a more recent example, Arizona State University, where I served from 1986-1997. ASU is a large university in the Phoenix metropolitan area, a location with considerable high tech industry. However, ASU only formally adopted a research mission in 1980. At about the same time, C. Roland Haden, the new Dean of Engineering, met with local business people who wanted ASU to become a significant research university with the goal of stimulating economic development. "Engineering Excellence" was born from these meetings and sold to Governor Bruce Babbitt and other political and business leaders. Unlike many universities in the early 1980s, ASU was growing and thus received new science faculty positions to which excellent people were hired. This combination of Engineering Excellence and the emphasis on science hiring lifted the entire university (cf. Geiger's remarks above). At ASU, I served for five years as Chair of Computer Science and Engineering and thus worked within Engineering Excellence on the front lines. I then served for six years as the University's second Vice President for Research. During that time, ASU's external funding doubled and, in 1994, ASU became a Research I university for the first time in its history.

4. Finally, Joan Lorden's school, the University of Alabama at Birmingham (UAB), has recently become a major success story. Although it is located in an EPSCOR state, UAB has made great strides in research productivity. Dr. Lorden shared a little about this leap forward in her presentation. Here is a longitudinal delineation of the university's federal expenditures over the past twenty years.



Lessons Learned

Lessons Learned

*Strategic intent by top leadership,
coupled with natural advantages
and local expertise,
can lead to research enhancement
that lifts the entire institution.*

Strategic intent by top leadership, coupled with natural advantages and local expertise, can lead to research enhancement that lifts the entire institution. Lifting the entire institution is a phenomenon that occurred at all four of the public universities I've mentioned when all the critical elements were in place.

Action Agenda

In his book, *Only the Paranoid Survive*, Andy Grove of Intel includes a variety of useful advice. As noted earlier, I've modified the title of his book for application to our universities: *Only the Flexible Will Thrive*. Grove discusses "strategic inflection points," which demarcate times of strategic changes in the performance of a company. These changes can be either positive or negative.

Positive strategic inflection points are reached more often if we apply strategic intent to our universities. Having goals that reflect our institutional missions can affect major changes in the output of our university research communities.

Performance Measures

Performance metrics are important because we will become what we measure. Thus we should select and promote measures that reflect values we think are important.

Joan Lorden has played a leading role in bringing these issues to the forefront. I believe her work with the Council on Research Policy and Graduate Education of NASULGC will have national influence.

Performance measures are used to rank and rate universities nationally, as well as to provide accountability locally. Well-known rankings are published by *U.S. News and World Report*, the National Research Council on Graduate Education, the Carnegie Foundation, and in the book by Graham and Diamond, *The Rise of American Research Universities*. The numbers collected by the

Action Agenda

- Only the Paranoid Survive →
Only the Flexible will Thrive
- Strategic inflection points
- Academic performance measures

Performance Measures

Performance metrics are important because we will become what we measure.

- Road map
- Goals: THINK BIG!
- Research performance measures
 - "NSF numbers," federal and total expenditures
 - Graham and Diamond, *The Rise of American Research Universities*
 - The Center (University of Florida) study
 - NRC rankings of graduate programs
 - US News and World Report rankings
 - AAU membership criteria

National Science Foundation ("NSF numbers") provide rankings based on both federal research expenditures and all research expenditures. There are recent interesting studies by The Center at the University of Florida and by the Association of American Universities that use multiple dimensions of quantitative measurements.

A "road map" can be a useful guide. By "road map" I mean a well thought out formal "action agenda" document. This concept is adopted from the Japanese semi-conductor industry where it has been used since the early 1980s when Japan became a threat to American dominance in that field.

Universities often do not set research goals or, if they do, the goals don't have quantitative measures. My counsel is to encourage setting goals that are both ambitious and multidimensional. I will return to this topic in connection with my present institution, the University of Kansas, in a moment.

If we would like to enlist the public in support of research, it is essential to have quantitative goals that are easily understood by the public. This is another important reason to collect accurate performance measures.

Tactics: Intra- and Inter-institutional

The University of Kansas (KU) provides an interesting case study for us today. When I returned to my undergraduate alma mater in 1997, KU had reached a research equilibrium, wherein its national research ranking was fairly static, and, **at the institutional level**, little change had occurred within memory. State support of the university had apparently been mediocre for some time and, consequently, support for research was sparse. However, the faculty and the university appeared to be better than was indicated by the institutional ranking. In particular, KU had a group of entrepreneurial research centers with faculty eager to step up the pace.

Tactics: Intra-, Inter-institutional

University of Kansas:

- One university (Megathemes)
- One state (Strategic initiatives)
- Four corners (Merrill, Heartland VPRs)
- EPSCoR states (AAAS)
- National level (NASULGC, CSSP)

We decided to inventory our intellectual capital on the four KU campuses. We did this by means of a call to the Deans and Center Directors to elicit faculty proposals for research attention. This was not a formal call for financial proposals, but rather a call for feasibility of "world class" research. Forty-seven

proposals were submitted and a steering committee of Deans, Directors, and others looked for "mega themes," that is, for topics that met three major criteria: at least 50 faculty working in areas that have demonstrated, peer-reviewed strength, that are also of significance to our public. The steering committee was unanimous in selecting four mega themes: information technology, human biosciences, the human condition, and environmental science and engineering.

What is "world class" research? In my opinion, a group is doing world-class research if every international meeting in their area must invite a member of that group to participate.

Next, we inventoried the three research universities of Kansas: the University of Kansas, Kansas State University and Wichita State University. Partners in this process included the AAAS, KTEC (Kansas Technology Enterprise Corporation), EPSCoR, the Senator Pat Roberts Advisory Committee on Science, Technology and the Future, and KU's Merrill Advanced Studies Center. In due course, we determined four strategic initiatives in science and technology for the state: information technology, human biosciences, agricultural biotechnology, and aviation. We are working at the state, regional and national levels to promote these initiatives.

Example of Performance Measures: KU

Performance measures that follow national norms are best for national comparisons. The federal research and development expenditures in science and engineering (i.e., those areas with significant external funding) are collected annually by the National Science Foundation (NSF) and are freely available. So this has become our "gold standard" for national comparisons. Such rankings also provide a surrogate for market share:

the percentage of the federal R&D funds obtained by a given university. Market share corrects for variation in federal R&D funds available; since such variation has been considerable over the years, this is an important consideration.

KU Measures

- Federal R&D expenditures in S & E
 - All universities, public universities
 - Research competitiveness
 - Rankings, market share
- RD&T expenditures (Research, Development and Training)
 - The public face of research:
 - State economic development
 - Direct economic impact, jobs from research
 - Graduates: best form of "tech transfer"

Although federal expenditures in R&D are the best available measure of national research competitiveness, this statistic underestimates the local impact

of research. Thus, in Kansas, we use the same methodology as the NSF, but have extended it to include: (1) fields outside of science and engineering, and (2) research training grant expenditures. We call this statistic "research, development and training expenditures" (RD&T). These figures are particularly helpful in discussing local economic development impacts of university research.

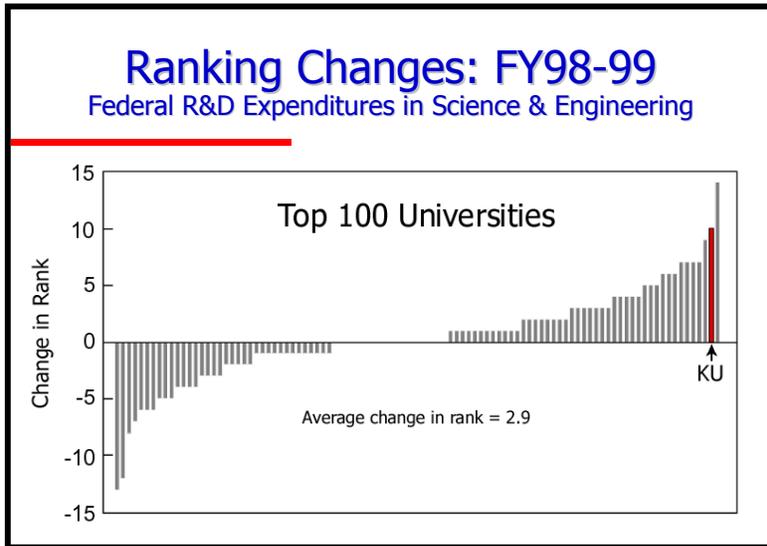
KU's FY99 Rankings

KU advanced in the rankings according to federal R&D in science and engineering, among all universities, by ten positions between fiscal years 1998 and 1999. The corresponding KU rise among public universities was seven positions.

KU FY99 Rankings Federal R&D Expenditures in Science & Engineering		
	1998	1999
All Universities	93 rd	83 rd
Public Universities	60 th	53 rd

Source: National Science Foundation Survey

Rankings Changes: FY98-99



This ranking change was the second largest among the top 100 universities, ASU being the university with the largest change. The average change in ranking, positive or negative, was about three positions. KU's ranking change within **public** universities was also the second largest nationally.

KU's RD&T Expenditures – FY00

KU's federal R&D expenditures in science and engineering rose 20% between fiscal years 1999 and 2000. National rankings for 2000 will be available later from the NSF. KU's total research, development and training (RD&T) expenditures rose 15% from fiscal year 1999 to 2000. KU uses NSF methodology to determine its total RD&T number.

**University of Kansas
Research, Development & Training Expenditures - FY00**

Source of funds	Science and Engineering Research	Training and Non – S&E Research	Total
FY00 Expenditures (dollars in thousands)			
Grants and Contracts	101,916	31,719	133,635
Federal Government	68,950 ^a	22,914	91,864
State and local governments	5,486	5,719	11,205
Industry	15,712	242	15,954
All other sources	11,768	2,844	14,612
Institution Funds	46,754	12,769	59,523
Institutionally financed organized research	31,494	2,138	33,632
Unreimbursed IDC and related sponsored research	15,260	10,631	25,891
TOTAL	148,670	44,488	193,158^b

^{a,b} Represent 20% and 15% increases from FY99, respectively

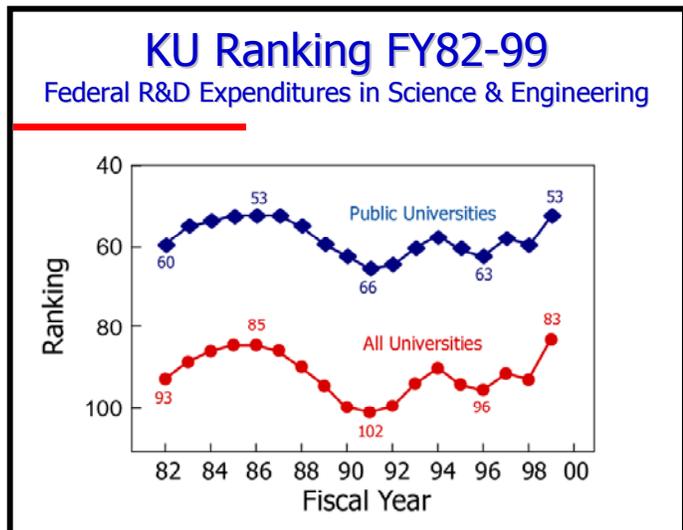
KU Longitudinal Studies

KU Longitudinal Studies

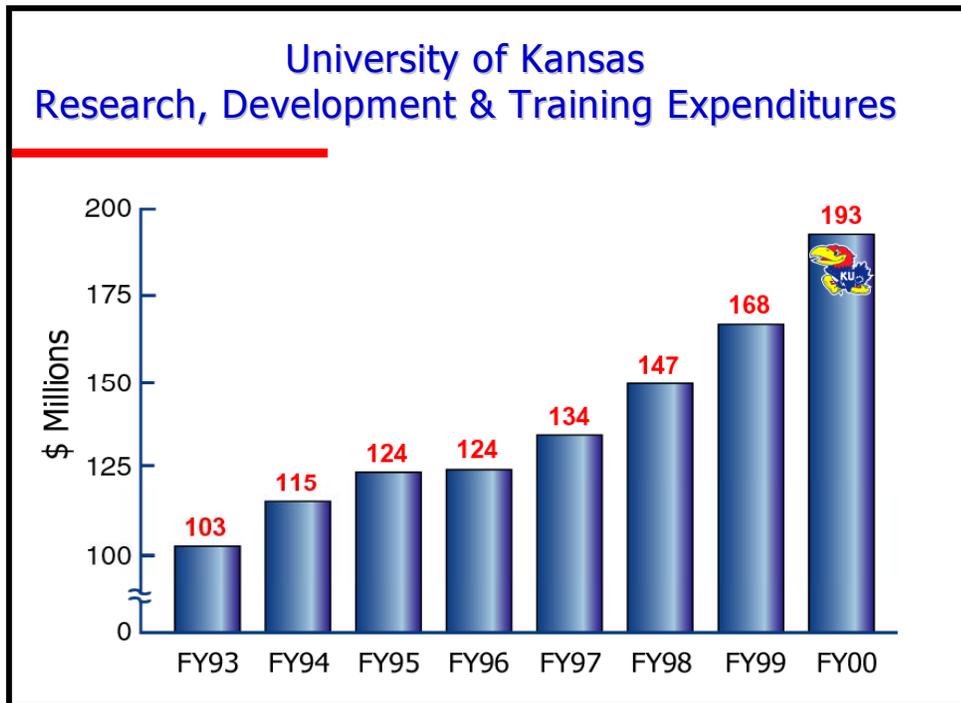
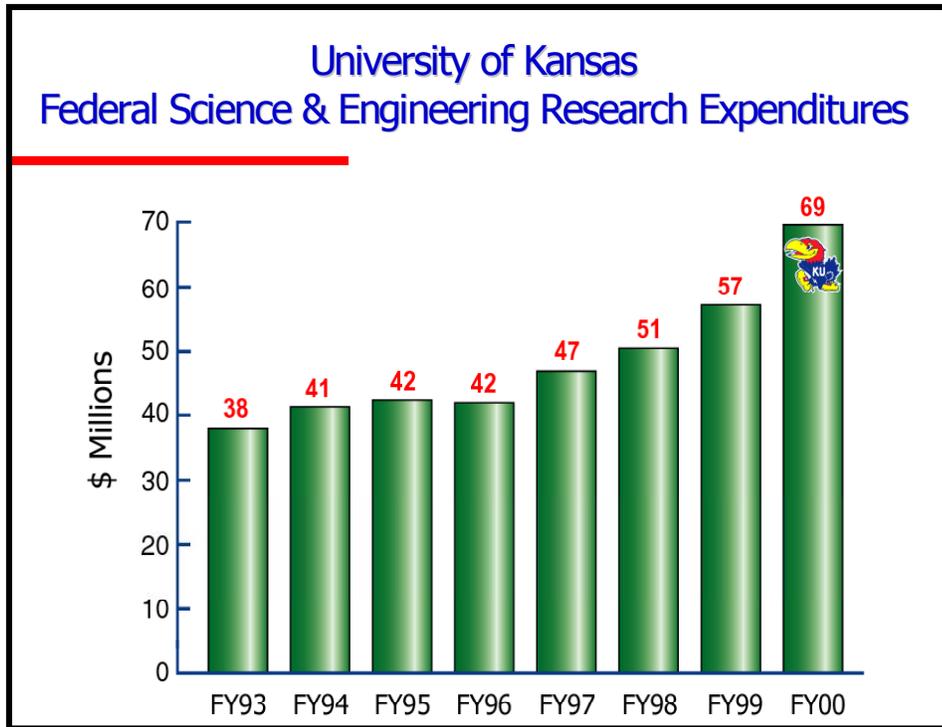
- **Rankings**
 - All universities, public universities
- **Expenditures**
 - RD&T
 - Federal R&D in Science and Engineering

It is always wise to study an institution over several years, as exemplified on the longitudinal graphs that follow.

In order to provide historical context, here are KU's research productivity rankings among all universities and among all public universities for the last 20 years. The criterion is federal expenditures in science and engineering.



In the next two graphs, you will see the dollar amounts producing the last few years' rankings, as well as the RD&T numbers.



KUCR Research Strategy

- Promote research as a topic on the campus
- Stimulate bigger projects: collaborations
- Provide much better assistance to faculty
- Make strategic investments of time and money

*A necessary condition:
the ability to operate relatively autonomously*
(cf. *The Innovator's Dilemma*)

KU's research strategy—exemplified by its research administration arm, the KU Center for Research, Inc. (KUCR)—involves many aspects, of which a few are listed here. Since research administration is a “disruptive technology” (in the sense of Clayton Christensen's book, *The Innovator's Dilemma*), it must operate relatively autonomously.

A necessary condition to maximize research productivity is to minimize internal competition between academic departments and research centers. We have devised a multiple credit algorithm to accomplish this task. Expenditures are recorded in two lists, one according to departments and one according to centers. This simple expedient has helped reduce competition between departments and centers.

Research Productivity on a Campus

- Multiple Credit Algorithm
 - Credit for all
 - Reduce internal competition
 - Stimulate research activity

National R&D Environment

- Lester Thurow: "Building Wealth"

- Private rate of return: 24%
- Public rate of return: 66%

"Put simply, the payoff from social investment in basic research is as clear as anything is ever going to be in economics."

Some background is necessary and helpful in understanding what is nationally possible. Lester Thurow, MIT professor of management and economics, wrote the lead article in the June 1999, *Atlantic Monthly*, "Building Wealth: The New Rules for Individuals, Companies and Nations."

Thurow writes, "A successful knowledge-based economy requires large public investments in education, infrastructure, and research and development." He quotes rates of return on R&D as: 24% for private investment; 66% for public investment. ("Public" rates of return mean that the benefits accrue to the whole society.) "Put simply," Thurow continues, "the payoff from social investment in basic research is as clear as anything is ever going to be in economics."

Some sound bites:

- 50% of economic progress since World War II is due to technology. This includes the fact that almost 3/4 of patents issued depend at least in part on publicly funded research.
- Alan Greenspan has stated that: "...the unexpected leap in technology is primarily responsible for the nation's phenomenal economic performance."

National R&D (continued)

- 50% of economic progress since WW II due to technology
- Alan Greenspan: "...unexpected leap in technology is primarily responsible for the nation's phenomenal economic performance."
- Internet economy: \$300 billion, 1.2 million jobs
- Information Technology (PITAC):
 - 1/3 of economic growth
 - 1/3 of all corporate R&D
 - 55% of all venture capital
 - New startup every hour
 - 7.4 million jobs
 - 80% higher salaries

- Internet economy: \$300 billion with 1.2 million jobs
- Information Technology bits from the PITAC report (see slide, above)

My own scientific career in Numerical Analysis and then in Computer Aided Geometric Design causes me to think that information technology advances during the next few years will dwarf what has come before, in terms of ubiquitous computing and visualization possibilities. These advances will include such visionary topics as molecular level, fault tolerant computer architectures that resemble biological systems, as well as advances in brain imaging and gene therapy due to virtual reality and computational power. For an institution to become a research leader, it must possess leadership that is aware of and can utilize national trends on the local level.

State Rationale for Research

Research universities provide unique cultural and economic advantages to society in general and to local communities in particular. Cultural opportunities include the advantages of a liberal education and all its corollaries. Economic impacts include the value added to graduates' incomes, as well as the economic ripple effect due to R&D dollars.

State Rationale for Research

- Cultural quality of life
 - Graduates
 - Society
- Economic impact
 - Graduates value-added
 - Jobs due to research university
 - R&D jobs
 - Companies
 - Spin-offs

Economic Impact of Research I University

Graduates are the largest form of technology transfer from research universities. We have quantified the economic impact of this important asset for our state of Kansas: the annual income of the alumni of our three research universities, who currently reside in Kansas, is \$9 billion. About 1/3 of this total, or \$3 billion, is due to the increased salaries they earn

Economic Impact of Research I University

- Graduates – Best form of tech transfer
- R&D jobs
 - RD&T statistics (“enhanced NSF numbers”)
 - AAU / US Commerce Department indicators: In Kansas, 40.6 jobs are created per \$ million of R&D funding
 - In FY00, KU’s RD&T expenditures resulted in >7,800 jobs
 - RD&T expenditures at the 3 Kansas research institutions in FY00 resulted in >13,600 jobs

because of their degrees from our three universities. The state tax paid by these graduates is \$700 million annually, a figure that exceeds the annual state appropriation to the three universities of \$400 million.

What is the ripple effect of R&D funding in Kansas?

The U.S. Department of Commerce estimates that, in Kansas, each \$1 million in R&D funding creates 40.6 jobs. The three Kansas research universities had \$335.2 million in RD&T expenditures in fiscal year 2000, which implies that more than 13,600 jobs are due to this source of funding. Moreover, the average salary in these jobs exceeds the average salary in our state. This type of economic information is what truly catches the attention of state legislators.

A Poll of the Public

Everyone knows that the National Institutes of Health (NIH) have received significant appropriations in recent years. Research!America has made many of the persuasive arguments that have promoted the NIH's budget. In 1999, I met with Mary Woolley, head of this group, and learned that they do state surveys and want to move beyond a focus on

biomedical science to support of science in general. Ms. Woolley also confirmed that Kansans' attitudes toward scientific research are of great interest to Research!America because of recent decisions on the topic of evolution.

Thus, I called together my counterparts from the KU Medical Center, Kansas State University, Wichita State University, and the Kansas Technology Enterprise Corporation (KTEC) to meet Mary Woolley. The result of this meeting was a poll of the Kansas citizenry. My favorite

A Poll of the Public

- Research!America → NIH budget
- Kansas, the evolution state
- Forming a state partnership
 - KU, KSU, WSU, KTEC
 - Senator Pat Roberts Advisory Committee on Science, Technology and the Future

Research!America Poll of Kansas

- State support of university research
 - In Kansas: Favored by 93%
 - National Average: Favored by 82%
- Publicity: Press conference in the state capitol featuring Sen. Pat Roberts and CEOs of the three research universities

among the statistics gathered by Ms. Woolley's firm is: 93% of Kansans favor state support of university research, whereas the comparable national average is only 82%. Senator Pat Roberts announced the results of the survey at a press conference with the CEOs of the three Kansas research universities by his side.

A Tipping Point

A Tipping Point

*Changes made by a few people
can have large impacts.*

- Malcolm Gladwell book, The Tipping Point
 - **Law of the Few:** Paul Revere and William Dawes
 - **Stickiness Factor:** "Selling the Endless Frontier"
 - **Power of Context:** Environmental tipping point

Malcolm Gladwell's book, *The Tipping Point*, indicates that changes by relatively few people can have large impacts. There are three rules for a tipping point: the Law of the Few, the Stickiness Factor, and the Power of Context.

My considerable oversimplification of the book is the following:

The Law of the Few: The example of Paul Revere illustrates that some people have exactly the right connections for making a significant impact, while others in the same situation cannot because they do not have these resources.

The Stickiness Factor: Successful projects frequently have some feature, say, a snappy title or phrase, which makes people remember them favorably. The image "sticks" in their mind. My own advocacy example is "Selling the Endless Frontier." This echoes "Science, the Endless Frontier" from Vannevar Bush's letter to Franklin D. Roosevelt encouraging federal support of research after World War II.

The Power of Context: "Environmental tipping points are things that we can change." Gladwell gives the example of fixing up a small portion of a run-down neighborhood. By this example, the neighborhood as a whole improves itself. My hopeful example would be the historical indifference of a legislature to university research.

Black Elk, Oglala Sioux

In the research arena, change is a necessity. Sometimes it is tempting to think that we have invented everything. I am always brought back to Earth when I turn to this late 19th century saying by Black Elk, an Oglala Sioux elder.

Black Elk, Oglala Sioux

"Little else but weather ever happened in that country -- other than the sun and moon and stars going over -- and there was little for the old men to do but wait for yesterday."

Let us not be like Black Elk's "old men." Rather, let us embrace change and use it to advance science and society in the 21st century.

References

Bush, Vannevar (1945). *Science—The Endless Frontier*. National Science Foundation.

Geiger, Roger (1993). *Research and Relevant Knowledge: American Research Universities since World War II*. Oxford University Press.

Gladwell, Malcolm (2000). *The Tipping Point: How Little Things Can Make a Big Difference*. Little, Brown and Company.

Graham, Hugh Davis and Diamond, Nancy (1997). *The Rise of American Research Universities: Elites and Challengers in the Postwar Era*. Johns Hopkins University Press.

Grove, Andrew S. (1996). *Only the Paranoid Survive: How to Exploit the Crisis Points that Challenge Every Company and Career*. Doubleday Publishing Company.

Hamel, Gary and Prahalad, C.K. (1994). *Competing for the Future*, Harvard Business School Press.

Stokes, Donald E. (1997). *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press.

Teich, Albert H. (Ed.) (1996). *Competitiveness in Academic Research*, American Association for the Advancement of Science.

Thurow, Lester (June 1999). Building Wealth: the New Rules for Individuals, Companies, and Nations, *Atlantic Monthly*.

A FACULTY PERSPECTIVE: INTRINSIC RESEARCH REWARDS THAT MAKE A SUCCESSFUL FACULTY MEMBER TICK

Carol Shanklin

Professor of Institutional Management & Dietetics
and Assistant Dean of the Graduate School
Kansas State University

Many factors influence faculty productivity. My remarks will focus on the intrinsic rewards that motivate senior faculty members to remain actively engaged in research. To illustrate my points, I will use personal examples or examples that colleagues have shared with me. Before I discuss my own experiences, I will present research I have reviewed on productivity in higher education.

Dundar and Lewis (1998) summarized results of several studies on research productivity. Individual and organizational attributes that have been found to affect research productivity include: individual attributes, institutional and departmental attributes, and departmental culture and working conditions. Individual attributes include: innate abilities—such as IQ, personality, gender and age—and personal environmental influences. Inconsistent results were reported when these variables were investigated. Dundar and Lewis (1998) identified as personal environmental factors the quality and culture of graduate training and the culture of the employing department. Most research studies have found a positive correlation between departmental culture and research productivity. The departmental culture refers to shared values and attitudes within the academic unit. Faculty and administrators who learned to place a high value on research as graduate students tend to foster a research-oriented culture throughout their professional lives. Faculty members who work in a research-oriented culture maintain dialogue with other researchers through internal and external communication, and seek opportunities for collaborative research projects. Departments with a research culture recruit and hire new faculty with strong research credentials or the potential to be successful researchers, and provide the faculty with development opportunities.

Institutional and departmental attributes influence research productivity. Dunbar and Lewis (1998) cited several research studies that found a relationship between productivity and the organizational size, including the number of faculty. In general, productivity is greater in larger institutions and in departments with a critical mass of faculty. “Intellectual synergy” appears to increase dialogue and cooperation among faculty members. Other research studies cited by Dunbar and Lewis (1997) did not find a positive relationship between size and

productivity. Grunig (1997) reported that other variables associated with increased research productivity are: university funding, availability of technology and computing facilities, and the number of books and journals in the library. Departmental working conditions that influence productivity include: workload policies; availability of leave-time, travel, and institutional funds for research; number of students on research support; availability of “star” faculty; and availability of nongovernmental research funds (Dunbar and Lewis, 1997). Limited institutional-level studies have been conducted on many of these variables.

What else motivates a senior faculty member? I believe there are many intrinsic research rewards that influence productivity. The privilege of mentoring graduate students has been one of the most important factors in my own productivity. Observing the professional growth that occurs during graduate school, especially during the research phase of a student’s program, is rewarding personally and professionally. Most of you can probably still recall the emotions you felt when you hooded your first doctoral student. Hopefully this same feeling of accomplishment motivates you to be an effective mentor to your graduate students and postdoctoral fellows today. An effective mentor must remain current in the discipline and actively engaged in research. Mentoring does not end with the awarding of the degree; it continues during the initial phase of the student’s early research career that spans the time beyond the publication of the research. I find it very rewarding to work collaboratively on funded multi-university research projects with individuals whom I mentored during their doctoral program. Another intrinsic reward is to receive requests from prospective students who are seeking an advisor or mentor.

The opportunity to work on interdisciplinary research teams with colleagues in other departments and universities is another intrinsic reward that stimulates my desire to remain an active researcher. Involvement in interdisciplinary teams has broadened my approaches to investigating research questions. Through this kind of research, I am often introduced to resources that were unfamiliar until I began working with colleagues from other disciplines. I also value our interaction in team meetings, our work together in the laboratory, and presenting papers together at professional meetings.

Professional recognition is another important intrinsic reward. This includes: being invited to present papers at national and international meetings; being selected to author or co-author position papers for professional associations; serving as chair of peer review panels for government agencies, foundations, and professional associations; and serving as editor of a journal or a member of an editorial review board. Accepting one of these responsibilities involves more work; however, I believe the intrinsic benefits outweigh the additional time demands since faculty gain public acknowledgement of their contributions to the discipline and recognition of their expertise. These opportunities also increase the network of colleagues with whom faculty can

consult on different topics, and can lead to multi-university or industry partnerships.

I can recall the excitement I experienced when I was asked to present one of the keynote addresses at the Asian Congress of Dietetics in Seoul, Korea. Since I was one of the few dietetic professionals in the U.S. conducting research in environmental issues, I was invited to speak on the topic: "Dietitians' Roles in Protecting the Environment." Not only did this experience provide me with an opportunity to visit Korea and speak at an international meeting, but it also gave me an opportunity to network with other researchers in my discipline and to visit with alumni who had graduated from our department. Networking at the meeting led to a visiting professor from Andong University completing her sabbatical at Kansas State University. Her goal was to learn the waste characterization methodology that we were using. Our research resulted in two publications and a collaborative research project. As another example, a professional association asked me to be the lead author on a position paper regarding natural resource conservation and management. I found it very rewarding to use my research in this way.

Other valuable achievements for faculty involve selection for a distinguished faculty award at their university or alma mater, and professional recognition by a society or association. Administrators should join in the public acknowledgement of contributions their faculty make to the profession and to their disciplines. These accomplishments should be communicated at the university level and through alumni publications.

These are some of the intrinsic rewards that have influenced my continued involvement in research and graduate education. Bland and Bergquist (1997) identified other intrinsic factors that influence a faculty member's vitality and productivity. Examples are: socialization, subject knowledge and skills, past mentors, work habits, adult career development, a vital network of colleagues, simultaneous projects in progress, sufficient time, and morale. Bland and Bergquist suggest that institutions can enhance faculty members' productivity by: establishing clear, coordinated goals that emphasize faculty members' core functions of research and teaching; providing a supportive academic culture where intellectual inquiry is valued; fostering a positive group climate that is essential for interdisciplinary research, as well as sufficient and accessible resources with frequent communication; and providing professional growth opportunities. They also recommend targeted recruitment and selection that supports the mission of the academic unit and the university.

In summary, many factors influence a senior faculty member who strives to maintain research productivity while also making contributions to the discipline and to the university. University administrators should not undervalue the benefits of intrinsic rewards, as they seek to increase the overall productivity of faculty members at their institutions.

References

Bergquist, W.H. (1997). *The Vitality of Senior Faculty Members. Snow on the Roof-fire in the Furnace*. ASHE-ERIC Higher Education Report vol. 25, no. 7. George Washington University.

Dundar, H. & Lewis, D.R. (1998). Determinants of Research Productivity in Higher Education. *Research in Higher Education*, 39, 607-630.

Grunig, S.D. (1997). Research, Reputation, and Resources: The Effect of Research Activity on Perception of Undergraduate Education and Institutional Resource Acquisition. *Research in Higher Education*, 68, 17-52.

STATISTICAL ANALYSES OF ECONOMICS NRC SCORES

Michael Podgursky
Professor of Economics
University of Missouri - Columbia

Economists are obsessive about measuring things. Thus it was inevitable that when the rankings by the National Research Council (NRC) were released, particularly given their "high stakes" character, economists would start doing statistical analyses of them. There have been several good studies of these rankings. Today I am going to focus on what I consider the most thorough of these—a recently published study by a well-known econometrician at Purdue (Thursby, 2000). This study was published in the *Journal of Economic Literature*, one of the "flagship" professional journals.

Thursby actually had access to the underlying individual survey responses. He did quite a bit of reliability checking and explored a variety of hypotheses concerning the rankings. I will cover some of the highpoints of his study. Here are a couple of the questions that were addressed.

1. Are the NRC rankings measuring something real?

The answer to this seems to be "yes."

As all of you are aware, the NRC rankings are based on a subjective survey of faculty concerning the quality of departments and Ph.D. program effectiveness. On the whole, economists are wary of subjective survey data. We usually prefer "hard" data like prices and quantities. And skepticism is not just limited to economists. When these rankings were first released I heard grumbles from my fellow department chairs in the College of Arts and Sciences that this was a "beauty contest" and similar disparaging remarks.

However, these NRC rankings in economics are **strongly** associated with objective measures of productivity such as total citations or total pages in refereed journals—measures that economists take seriously. If the NRC score is the dependent variable, publications and citations—current and lagged—explain about 90 percent of the variation in NRC scores.

There are a couple of things to note about the statistical model that generates this good fit. First, the effect of citations and publications is non-linear. Specifically, these variables demonstrate diminishing returns. That is, it takes progressively more and more citations or journal pages for your NRC score to move up as you get closer to the top. For example, it might take 30 journal pages to move you from 98 to 97th rank, but 250 pages to move a department from 26 to 25th rank.

In addition, the effect of citations and publication is different depending on their age. Current journal pages have a bigger effect than older pages. Citations to older articles count more than citations to current articles.

So, these subjective NRC rankings are strongly associated with objective measures of departmental productivity in ways that make sense.

Let me note, however, that this strong association holds for **total** department productivity. The relationship is much weaker when we consider **per capita** productivity. When per capita measures of journal pages and citations are used in the model instead of total pages and citations, the percent of variation explained falls from 90 percent to just 60 percent. So, size clearly matters in the NRC rankings.

2. What factors produce higher rankings?

If we think of NRC scores as outputs, what inputs produce a yield with higher output? Thursby again fits a non-linear statistical model to these data. He found that total faculty size matters. No surprise here. The proportion of full professors, external grants, library expenditures, research assistants per faculty member, and the number of faculty per undergraduate student all mattered. The latter variable, by the way, fully explained the public-private difference in rank. The private schools had far more faculty per undergraduate student. As with the former analysis, most of these inputs displayed diminishing returns.

The link between size of department and rank is clearly visible in even casual inspection of the NRC data. The average department size in the lowest quartile of departments was 17.4 whereas the average size in the highest quartile was 36.1.

More recent evidence on the effect of size may be found in the *U.S. News and World Report* top 50 rankings. We counted the number of regular faculty based on information on the web sites of these departments. The attached chart shows the number of regular faculty in the top 50 economics Ph.D. programs according to the most recent *U.S. News* ranking (see figure 1). [I omitted two departments: CalTech and Claremont-McKenna. Caltech doesn't have an economics department; it has a quantitative social science department. Claremont-McKenna draws on faculty from a variety of private colleges in the area.] For background, let me indicate the size of economics departments for some of the Midwestern public universities represented at this conference: the University of Missouri-Columbia 17, the University of Nebraska 17, the University of Kansas 17, and Kansas State University, 16. (None of these is in the top 50.) The department at Iowa State is in the top 50, and has 29 faculty.

There are a couple of things to note here. First, the average size in the top 50 is 32. Furthermore, as you can see from the ranking, there is no department in the top 50 that is as small as my department.

If these were data on firms in an industry, the interpretation would be clear. There are "economies of scale" in this market and if you want to be competitive you need to be larger than a minimum size threshold.

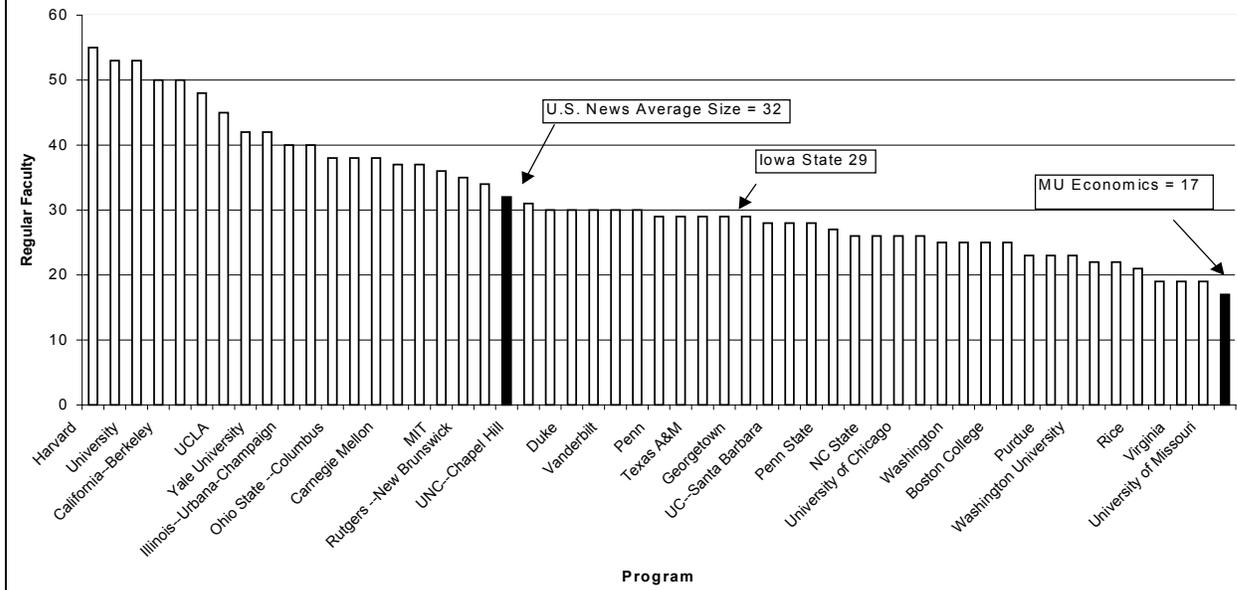
Several years ago, a previous Chancellor at the University of Missouri urged departments to pursue what he termed a "unique niche" strategy. He believed that a department could pull itself up and gain high professional rank by "putting all its eggs in one basket," finding a niche, and becoming the best around in that niche area. Aside from the fact that this is a very risky strategy (suppose a physics department picked "cold fusion"), there is no evidence in these data to suggest it works in raising NRC rankings. So far as I'm aware, none of these departments is highly specialized. On the contrary, they tend to be fairly diversified, with faculty in a variety of fields. To be sure, some departments stand out in some fields more than others, but all of them have high quality productive faculty in a variety of fields. They all produce dissertations in all or most of the major fields in the profession.

In short, I see no quick, inexpensive gimmick that will move a department up in these rankings. You need to be big and you need to be good.

References

Thursby, Jerry G. (2000, November). What Do We Say about Ourselves and What Does It Mean? Yet Another Look at Economics Department Research. *Journal of Economic Literature*, 33, 383-404.

Figure 1: Size of Economics Programs in U.S. News Top 50 Compared to MU Economics Department



A PASSION FOR REPUTATION, A TASTE FOR ORIGINALITY

Susan Kemper

Roy A. Roberts Distinguished Professor
Psychology and Gerontology
University of Kansas

When Mabel Rice first invited me to this retreat on “Evaluating Research Productivity,” I dashed off to look up the topic, where else, on the Internet. I found lists of rankings, lists of productivity indicators, and policy statements designed to increase research productivity. I even read some of this stuff, learning much more than I previously knew about citation indices and their limitations, per capita publication counts, the research productivity ranking of Hong Kong University’s economics department, and biases of impact and reputation rankings. I acquired a few, new buzzwords and phrases. I particularly like the notion of “streamlining the administrative process” [note: not streamlining the administration] as one key to increasing research productivity and the distinction between measuring the “vitality “ of the institution, not its “quality of research.”

I found little of this sort of thing to be helpful to me, professionally, and certainly not something I could pass along to my students as sage advice. It seems to me that the researcher has gotten lost in all these discussions of evaluating research productivity—the researcher, the one who writes the books, referees the articles, and delivers the keynote addresses. It seems to me that what is missing is a bridge between the formal metrics for evaluating research productivity (citation counts, peer evaluations, dollar cost/benefit analyses), applied at the unit level, and the activities of individual researchers.

I take it that the goal of evaluating research productivity is to better help the university manage its resources: resources flow to units with greater productivity, and improvements in productivity reduce costs and increase quality. Measuring research productivity involves establishing appropriate benchmarks or indicators, such as counts of publication, citations, and patents, and scaling each unit relative to appropriate professional, national or international standards.

While there appears to be an extensive literature on “bibliometrics” and “scientometric,” there is little that relates to the activities of the individual researcher—with one exception. The only discussions I uncovered that relate such measures of research productivity to individual faculty members concerned tenure disputes; I found citation counts evoked by those denied tenure as an objective measure of impact. This to me signals a disturbing trend, one of

substituting unit-level aggregations of quantifiable data—this new science of bibliometrics—for more qualitative reviews of individual researchers' productivity.

I next turned to my collection of “faculty survival guides” and found much advice regarding that “pesky tenure problem” as well as networking, socialization, dealing with sexist comments, preparing course syllabi, devoting time to writing, and choosing committee assignments. But again, there is surprisingly little advice on the evaluation of research productivity at the level of the individual researcher.

After logging a lot of time web surfing and even a few hours in the library, I have found one source that I think is helpful: *Advice for a Young Investigator* by Santiago Ramón y Cajal, the pioneering Spanish neuroanatomist who is credited with the discovery of the synapse and shared the 1906 Nobel Prize in Physiology. Writing in 1897, he offered a wide range of advice on many matters pertaining to evaluating research productivity.

I'll spare you his advice on choosing a wife—one key to research productivity—and turn to his typology of “diseases of the will” that I found illuminating:

We have all seen teachers who are wonderfully talented and full of energy and initiative—with ample facilities at their disposal—who never produce any original work and almost never write anything. Their students and admirers wait anxiously for the masterpiece worthy of the lofty opinion they have formed of the teacher. But the great work is never written, and the teacher remains silent.

Let us not be deceived by optimism and good intentions. Despite their exceptional merit, and the zeal and energy they display in the classroom, such teachers suffer from a disease of the will...their students and friends may nevertheless consider them abnormal and suggest some adequate form of spiritual therapy, with all due respect to their fine intellectual abilities. (p. 75)

Cajal classifies these diseases of the will as “the dilettantes or contemplators; the erudite or bibliophiles; the instrument addicts; the megalomaniacs; the misfits; and the theory builders” (p. 75). He is most dismissive of the contemplators as “likeable for their juvenile enthusiasm and piquant and winning speech as they are ineffective in making any real scientific progress” (p. 77) and he recognizes that “cold-hearted instrumental addicts cannot make themselves useful” (p. 82) and he notes that the misfits, who occupy a professorship “simply to collect the salary, and to enjoy the incidental pleasure of excluding the competent,” are “hopelessly ill” (pp. 82-83).

For the rest, Cajal has some recommendations regarding evaluating research productivity at the individual level. Cajal's recommendations ring as

true today as they did in 1897. Cajal reminds the bibliophile that “We render a tribute of respect to those who add original work to a library, and withhold it from those who carry a library around in their head” (p. 78). He advises the megalomaniac to “tackle a small problem first ...[an approach which] may not always lead to fame but [to] the esteem of the learned and the respect and consideration of our colleagues” (p. 80). He notes that rather than bemoaning the lack of able assistants, or laboratory equipment, or government funding, that “dreamers do not work hard enough” (p. 80). And he reminds the theorist that “Theories desert us, while data defend us” (p. 86).

For all, Cajal cautions that independent judgment, intellectual curiosity, perseverance, and concentration are required. He also recognizes that excellence in research and teaching are complementary, requiring similar intellectual traits. He notes that time constraints and competing time demands erode the scientist’s devotion to research.

Cajal foresaw the current debates over basic versus applied science as a “false distinction” since “nothing in nature is useless” and applications “will always come, whether in years or perhaps even in centuries” although he does bemoan that “the struggle for patents, and the fever of industrial competition, have disturbed the august calm in the temple of Minerva.”

And he called for public support of science, pointing out that “today’s statesmen undoubtedly have limitations, one of which is not realizing...that the greatness and might of nations are the products of science, and that justice, order, and good laws are important but secondary factors in prosperity” (p. 91).

Beyond these prerequisites, Cajal emphasizes that research productivity results from a “passion for reputation, for approval and applause,” and a “taste for originality, the gratification associated with the act of discovery itself.” So in evaluating research productivity, our challenge is to evaluate this “passion for reputation” and this “taste for originality” that then lead to publication lists, citations counts, and impact assessments for unit-level aggregations.

I’d like to now introduce a bit of gerontology into our discussion. In the book *Age and Achievement*, Lehman (1953) evaluated research productivity in a number of different fields. He identified, using standard sources, the most significant achievements in those fields, often recognized years later by Nobel Prizes or other awards. And he determined the age of the individual at the time these contributions were made. He published a series of graphs showing striking agreement: research productivity peaks in the 30s, and rapidly declines thereafter—in medicine, in mathematics, or in German grand opera.

Recently Simonton (1990) has reanalyzed Lehman’s data, drawing a different conclusion. Simonton has shown that research productivity is a function, not of chronological age, but of time-in-profession. Following a 4 to 8 year period

of training and apprenticeship, research productivity peaks and then rapidly declines. What Simonton's analysis makes clear, is that one is "over the hill" professionally not by age 40 or so but approximately 15 years after entering the profession. Simonton also makes clear, and illustrates with numerous examples, that one can experience many productivity peaks by shifting careers, frameworks, and methodologies at the price of investing 8 to 12 years in a new period of training and apprenticeship.

So my point is that our system of evaluating research productivity assumes that productivity is a monotonically increasing function. It is not, and the successful researcher who is able to sustain productivity over a long career may experience multiple peaks (and valleys) as he or she invests additional time in acquiring new skills and competencies in order to develop new lines of investigation.

So my survey has identified two dissociations between the unit-level measures of research productivity and the efforts of individual researchers. First, productive researchers are distinguished from less productive dilettantes, bibliophiles, instrument addicts, megalomaniacs, misfits, and theory builders by their "passion for reputation" and "taste for originality." These traits do not appear among the five factors of personality modeled by Costa and McCrae (1998), nor do they map only McClelland's (1953) achievement motive, and they certainly don't correspond to any of the profiles on the MMPI (Hathaway & Meehl, 1967). My concern is that these qualities are also overlooked by bibliometrics. Counting publications, estimating impact, and assessing costs/benefit ratios may be appropriate for unit-level aggregations of researchers but not for the evaluation of research productivity of individual researchers. Individual researchers are motivated, not by their count of publications or listing of citations, but by their "passions" and "tastes."

Second, research productivity is not linear but is best described by Simonton's exponential function of "investment time." Our system for evaluating research productivity at the unit-level must reflect this non-linear career trajectory at the individual level. We must encourage and allow individual researchers, who have demonstrated their productivity, to seek out new venues. We must grant them the time to invest in new skills and competences if we wish to sustain their productivity over a long career. Bibliometric practices that assume productivity is monotonic will discourage those very investments that promote research productivity.

Thus, I think no discussion of "evaluating research productivity" at the institutional level can be complete without a discussion of "evaluating research productivity" at the level of individual researchers. Whenever we aggregate data to examine trends in research productivity or to look at the leveraging of state funds, we are in danger of overlooking the variability of individual research careers. The unit level aggregations will average over nonproductive dilettantes,

bibliophiles, instrument addicts, megalomaniacs, misfits, and theory builders. Aggregations at the unit-level will also mask the peaks and valleys of individual research careers. We need to question how bibliometrics affects not only institutional research planning but how it may impact the very “passion for reputation” and “taste for originality” that motivate researchers to be productive and to sustain their productivity over a long career.

References

Ramón y Cajal, S. (2000). *Advice for a Young Investigator* (N. Swanson & L. W. Swanson, Trans.). Cambridge, MA: MIT Press.

Costa, P. T., & R. R. McCrae (1998). *Emerging Lives; Enduring Dispositions: Personality in Adulthood*. Boston: Little, Brown, & Co.

McClelland, D. C. (1953). *The Achievement Motive*. New York: Appleton-Century-Crofts, 1953.

Hathaway, S. R., & Meehl, P. E. (1967). *Atlas for the Clinical Use of the MMPI*. Minneapolis, MN: University of Minnesota Press.

Lehman, H. C. (1953). *Age and Achievement*. Princeton, NJ: Princeton University Press.

Simonton, D. K. (1990). Creativity and Wisdom in Aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging* (pp. 320-329). New York: Academic Press.

EVALUATING RESEARCH PRODUCTIVITY OF
INDIVIDUAL INVESTIGATORS:
CHAIRMEN'S PERSPECTIVES IN A MEDICAL CENTER

Thomas H. Rosenquist
Vice Chancellor for Research
University of Nebraska Medical Center

The overall success of the research enterprise of any institution depends critically upon the quality and success of its individual investigators. Thus, each institution must develop and implement a well-defined system for recruitment, retention and reward of excellent, well-funded scientists ("the 3 R's"). Both recruitment and retention are dependent upon the adequacy of the reward system.

In the typical model for a reward system in an academic medical center, "reward" means tangible resources that include salary, space, and discretionary dollars. The assignment of these resources is invariably based upon an evaluation of research productivity, according to department chairs in academic medical centers. However, the nature of the evaluation process is highly variable among different medical centers, and among various departments within a given medical center. At the University of Nebraska Medical Center, this variation is manifested by a range of evaluation methods that may be characterized as "Darwinian" on the one extreme, to "Egalitarian" on the other.

The ultimate "Egalitarian" method holds that all investigators are equal in both need and merit, and therefore all receive the same consideration for distribution of resources. Egalitarian methods are by definition more subjective and qualitative; e.g., a person may be judged on "collegiality" or "leadership." Darwinian methods tend to be more objective and quantitative.

Purely "Darwinian" evaluation is one that compares each investigator with all other relevant investigators, and then bases the assignment of resources (the "Reward") upon this comparison. Investigators who do poorly over time will be weeded out, whereas the strongest, most adaptable, will thrive; hence, "Darwinian."

Most methods for evaluation of individual investigators involve some combination of Egalitarian and Darwinian. In general, basic sciences departments are more on the Darwinian side, whereas clinical departments tend to be more Egalitarian; colleges of medicine typically are more Darwinian than other colleges in a given medical center.

Faculty members, of course, are not completely satisfied in any case. However, it seems that faculty in departments with a major Darwinian component are generally more satisfied with evaluative criteria than they are in circumstances where the chair or another evaluator gives a highly subjective, or no, evaluation. Differences in evaluative criteria among similar departments is a source of discontent; for example, high-performing researchers in a center where **some**, but not **their**, department use Darwinian evaluation methods, typically are less satisfied and are more likely to be susceptible to recruitment by other centers.

A model that may be used for the Darwinian method quantifies annual research productivity for all investigators; derives an average; compares each investigator with the average investigator; and then distributes rewards accordingly.

The following formula has been applied to one or more successful applications of the Darwinian model:

Annual Productivity (A) = Publications (P) + Funding (F), where

P = journal “power” X author position;
F = total grant dollars as PI + FTE % paid on grants

“A” is calculated for each investigator (A_{inv}). Then, all A_{inv} are used to derive an average of A (A_{avg}) for all relevant investigators.

From the above data, an “Annual Productivity Quotient” (A_q) is derived for each investigator, where

$$A_q = A_{inv} / A_{avg}.$$

By this quotient, it is possible to determine how any investigator compares with the average investigator, and to apply this datum to the merit-based distribution of resources. For example, Investigator Smith experiences an outstanding year, with an A_q that is 2.7 times the average for the comparison group (typically, members of a single academic department); Investigator Smith’s annual salary increase, for example, could be calculated as 2.7 times the average raise for members of this comparison group. All others in the group are judged and rewarded by the same criteria, yielding 100% distribution based upon merit.

One of the principal criteria in the typical evaluation plan is the level of NIH funding; but the most Darwinian plan of all (“Darwinus Maximus”) holds that the **only** criterion that needs to be evaluated is the level of NIH funding of an individual. In order to achieve NIH funding, it may be reasoned, an investigator must be well-published and have a solid national reputation. To achieve more than one major grant, or to be principal investigator on larger program-type

grants, the investigator must have good knowledge and experience with the NIH process; a solid national reputation in research; understanding of the politics of his or her funding institute; and demonstrate a high level of both leadership and collegiality. In a word, success in NIH funding is both necessary and sufficient for judging research merit in an academic health sciences center, according to this plan of evaluation.

In general, even departments that use a highly Darwinian model for **annual** evaluations (objective, quantitative, limited), also permit a more subjective and qualitative and comprehensive evaluation as part of the kind of **episodic** evaluation (two or three career episodes) that may lead to promotion or tenure. In this case, objective criteria of research success may be combined with other criteria, such as leadership, professionalism, character, or collegiality; peer review and editorial activity; research awards; or election to national office. Obviously, virtually all of these criteria either reflect a history of strong annual evaluations, or are essential characteristics for obtaining strong annual evaluations. Therefore the annual and the episodic evaluations are part of the same fabric.

There are two recommendations that arise from the above thesis:

First, the most effective tool for evaluating the annual research productivity of an individual investigator is objective, quantitative, and limited to a few specific criteria that are research-sensitive—especially success in publication and in obtaining grant funding, with emphasis upon NIH funding.

Second, within a given health sciences center, it would be best for similar units to use the same evaluation criteria (e.g., all basic sciences departments; all clinical departments). Although it is necessary for such standardized criteria to be imposed by a college- or university-level authority, it is unusual for this to occur. Research administrators should do an evaluation of criteria for evaluation of individuals within their respective institutions, and make recommendations for standardization.

Acknowledgements: The above thesis is based in part upon data obtained from department chairs at the University of Nebraska Medical Center, and from discussions of these topics at meetings of the Association of Anatomy, Cell Biology and Neurobiology Chairs (AACBNC). The research development experience shared by Dr. Roger Markwald was especially productive and useful.

EVALUATING THE QUALITY AND QUANTITY OF GRADUATE STUDENT RESEARCH

Suzanne Ortega

Dean of the Graduate School and Vice Provost for Advanced Studies
University of Missouri – Columbia

Two fundamental assumptions from sociology shape my thinking about the evaluation of graduate student research productivity. First, the meaning of any social process, including research and evaluation, is shaped by the context within which it occurs. Second, the whole is always larger (or at least different) than the sum of its parts.

The most important analytic elements of a social context include: the intended uses of the process outcome and the various audiences to receive the “message” about the outcome. As a sociologist, I see two forces involved in shaping the context for graduate student evaluation. Accreditation requires that institutions be able to use objective indicators of student learning outcomes to map their planning efforts. Also, senior administrators and university communities want to utilize resources to improve the quality of graduate education (a proxy for which is often an increase in the number of graduate programs ranked in the top 10 or in the top quartile of the next NRC study of doctoral programs). In terms of audience, there are two that overlap: academic and non-academic, the latter including industry and governmental leaders, and the public at large. The academic audience is internal to an institution and is composed primarily of university administrators, faculty, and current graduate students. Each of these audiences will have a slightly different use for evaluation and assessment data on graduate research productivity. As a result, each will have a differential stake in the efficiency and/or comprehensiveness of the assessment process and each will be interested in a somewhat different set of outcome measures.

Let’s look at how our thinking about assessment correlates with the premise that the whole is always something larger (or at least different) than the sum of its parts. Universities are more than the sum of the departments that comprise them. Graduate programs are more than a simple sum of individual student learning outcomes. This also leads us to believe that the quality of graduate research is more than a simple aggregation of the number of graduate student papers presented or published.

Joan Lorden’s model for measuring research productivity is an excellent framework for the remainder of my remarks. I will focus on: goals, audiences, values, and practicality.

Aside from satisfying the mandates of accrediting agencies such as the NCA, we might first ask what our **goal** is in graduate program assessment (and by implication graduate research). I would submit that there are two major goals. The first goal is to provide the information necessary to create self-directed plans for improvement in graduate student learning outcomes and in the overall quality of graduate programs. The second major goal is to provide the data necessary to guide resource allocation decisions. Resources can, of course, range from the very tangible to the much less tangible. This can include hiring new graduate faculty, providing enhanced graduate stipends and benefits, and assisting with professional travel and development opportunities, as well as the more intangible aspects of acquiring prestige as a result of achieving some highly desired and difficult to attain outcome.

There are two **audiences** for assessment data and each is differentially interested in one of the two goals above. Internal audiences include all stakeholders within the university. Certainly students and staff are impacted by the perceived prestige of an institutions' graduate research profile and by the relative proportion of resources that flow to it. However, I would like to focus on the two internal audiences that seem to have the most impact on the way research is assessed and used. Administrators—Chancellors, Provosts, Research Administrators, and Graduate Deans, for example—primarily look at research assessment as a tool for making strategic decisions about the use of existing resources and as a platform from which to argue for more external resources, whether that be prestige or funding. Of course, administrators are also interested in improved educational outcomes, but on a day-to-day basis, I believe most are willing to trust another internal audience—the faculty—to make sure that improvements in graduate research training are taking place and that those improvements are reflected in the assessment data they produce.

Faculty, of course, have a major stake in evaluating the quality and the quantity of graduate research. Yet, I would have to say that I have used up more of my reserve of good-will capital with faculty on the assessment issue than on anything else. Even though I keep telling faculty that they should be interested in assessment as a strategy for improving the things they care about, i.e., the preparation of the next generation of scholars and researchers, they believe I actually want a quick and efficient way of allocating—or more frightening still, reallocating—resources. In fact, I suspect most faculty end up going along with our standard graduate research assessment procedures only because they are worried that if they don't comply, they might lose funds. They are skeptical at best that any new resources or opportunities will be forthcoming as a result of an honest evaluation of either graduate research quantity or quality. The ambivalence of faculty may be attributed to the competing and in some ways contradictory use of data to: (1) make resource allocation decisions (an approach that many faculty fear and resent) and (2) make informed, self-directed decisions about program improvement strategies.

Externally, we can divide audiences into two subtypes, other academics and non-academics. When we speak in the language of graduate program rankings and prestige, I would submit that our primary, but not our exclusive, audience is composed of other members of the academic community. This is particularly true with reference to the National Research Council (NRC), where the primary indicators of graduate faculty quality—number of publications in refereed journals, proportion of faculty supported on extramural funds, or even number of degrees conferred—reflect the standard academic **values** of peer review and publication as the appropriate measure of research productivity. To the extent that some portion of our external non-academic audience is composed of aspirants to the academic roles, values, and community (i.e., prospective graduate students), the language of rankings and prestige will be compelling and influential for them, as well.

Although our non-academic external audiences probably share in the same general goals for assessment, i.e. resource allocation and program improvement, it is quite likely that business, governmental, and non-profit leaders will have different performance standards. We are in a situation where appropriate indicators of productivity and quality are still contested within the academy, and we have yet to consider how we might develop productivity and quality measures that address the core values of industry or the public at large. If we consider job placement a measure of student learning, how do we apply this to non-academic placements? By the size of the firm? By the firm's profitability measures? By dollars spent by the firm on R&D? Would placement in a federal agency be ranked higher in quality than placement within a county or municipal social service agency? If the productivity of our graduates is an indicator of graduate program and individual graduate student research quality, what kinds of non-academic research productivity measures speak to the core values of our non-academic audiences? Are patents valid indicators of research productivity? Is quality then measured by patents that lead to the development of start-up companies and by the profits they derive? Is there a metric by which we can gauge the impact and quality of scholarship that leads to new public policy or law? As universities move in the direction of increased collaboration with industry, with increased public accountability, and respect for the wide range of career opportunities for our doctoral degree recipients, it will become more important to develop assessment and evaluation strategies that align with the values and goals of our non-academic audiences. To date I have heard no considered and sustained discussion of the measures we should use.

More than anything else, I believe, the **practicality** factor has led most institutional research offices, accrediting agencies, and organizations involved in educational ranking to use productivity indicators as their best, and often only, measures of research quality. Whether measured in absolute numbers or per capita, indices based on the number of refereed publications, the number of awards, and the amount of extramural research funding have the real advantages of being routinely collected as part of other faculty and student

evaluation processes. Because they are numeric, they have the added advantage of being standardized and easily summarized. In its last iteration, for example, the NRC basically relied on the faculty productivity measures identified above to measure graduate faculty quality. Faculty quality, in turn, was used as the indicator of graduate program quality. Although the research protocol is not yet set, a shift toward inclusion of more student outcomes in the next NRC study will likely parallel indicators of research productivity for the faculty.

Clearly, research productivity bears an important relationship to research quality. At the individual level, however, productivity is a necessary but not a sufficient condition for research and graduate program quality. Here, I would simply reiterate that we must begin turning our attention toward the development of easily collected quality measures, appropriate at both the individual and the program/institutional levels, and pertinent to both academia and the broader community.

The practicality of an assessment and evaluation strategy depends as much on the process we use to collect data as it does upon the simplicity, reliability, and validity of the indicators we choose to collect. We will be well served, then, if we can embed the assessment of research quality into a common data collection process that has the capacity to address a variety of institutional needs. This process should recognize the differing values and priorities of our various audiences. At the University of Missouri - Columbia, for example, we are trying to create an integrated assessment process that inputs data from the annual reviews of individual graduate students and merges it with routinely collected institutional measures. Institutional measures typically include: proportion of students supported on assistantships or fellowships; part-time/full-time enrollments; number of degree recipients, and average time to degree. Where there are sufficient numbers of graduate degree recipients to do so, we also utilize summary reports from the NSF Survey of Earned Doctorates. This database, in turn, will provide much of the information about graduate education necessary for state- and institutionally-mandated five-year academic program reviews. By reducing the number of unique reports that departmental chairs and directors of graduate studies must provide, we are optimistic that one of the big stumbling blocks to meaningful assessment will be removed. I would caution, however, that efforts to use student learning outcomes and graduate program quality for thoughtful self-improvement often run at cross purposes to the academic review process, which is fundamentally about resource allocation. We will have to continue to monitor whether the savings in faculty time and the possibilities of creating a truly useful body of information for program development can offset this potential “danger.”

In the end, our core **values** should guide assessment, and not simply issues related to expediency or audience. It seems to me that one of the core values we need to resolve is the question of measuring quality or productivity, per se. If it is quality we want to measure, we must determine how to differentiate

it from productivity. In general, I suspect, we are talking about the impact of research when we assess quality. How, then, do we measure impact? Once again, we may find that different audiences will be convinced by different measures. It is also important to keep in mind the distinction between the impact of individuals and the impact of programs.

Within the academy there may already be a fair amount of consensus about how to measure the impact/quality of an individual's research and scholarship. One standard indicator is publication in peer-reviewed high visibility journals with high rejection rates. Citations and the amount of extramural support for research are other standard measures. I would note, however, that these typical measures of scholarly impact work much better for the sciences than they do for the arts and the humanities. Earlier today, Dr. Lorden mentioned the adage: "We are what we measure." We must be careful to develop appropriate measures of quality and impact in the arts and humanities or we may erode the position of these disciplines at our institutions, especially when measures of "impact" drive resource allocation models in the future. By intention or happenstance, our support of the arts and humanities will be an important statement about our institutional values.

If ambiguities remain in the assessment of individual research, this is even more true of efforts to assess quality at the program level. The "value-added" dimensions of a high quality graduate research program will likely be its defining characteristics. Although we may not yet have the measures, I suspect that two important value-added indicators will be: the capacity of programs to foster interdisciplinary research skills and agendas, and the capacity to provide professional development opportunities for the next generation of scientists and scholars, namely, teamwork, sensitivity to issues of diversity and internationalization, communication skills, etc.

In summary, I find that we do not have a measure of impact that is relevant to audiences outside of academia. To do this successfully may entail tackling prejudices about applied research. We almost certainly will have to move beyond a hierarchy that gives preference to basic over applied research. We may need a separate metric appropriate to each kind of research. However, practicality will likely force us to compare the two and ask questions such as: how many refereed articles in what tier of journal does it take to equal the impact of one patent or five technical reports?

Whether we tackle these questions effectively, or indeed at all, will reflect on another basic value issue that universities are now facing—the extent to which we choose to be internally or externally focused. In the next several years we will learn something very important if our measures of research quality remain simply new and improved measures of traditional academic productivity rather than evolving to meet the challenges we have before us.

ACHIEVING THE “GOLD STANDARD”: A REDUCTIONISM DIRECTIVE

K. Michael Welch
Vice Chancellor for Research
and Senior Associate Dean for Graduate Studies
University of Kansas Medical School

Introduction

This paper takes the position that, from an administrative point of view, research productivity must be measured in a simple, easily understood, well rationalized and goal directed manner. I will put the case that, for biomedical institutions, total NIH award (from the National Institutes of Health) is the measure that meets these key characteristics. This simple index is a compound of all other quality measures, each of which correlate strongly with the overall index. The paper uses the example of biomedical research conducted in institutions that comprise a medical center.

Productivity Considerations in Medical Institutions

Administrators must first set the **goal**, define the **audience**, enunciate the **values** and consider the **practicalities** that may determine the process of achieving the goals. The most commonly articulated goal of most biomedical institutions is to become a leader among peers in biomedical research. Their audience is not only the community of medical science, but also an expanded interest sphere that includes health care delivery systems whose viability may depend on recognition of research excellence being inextricably entwined with the highest standard of health care. More and more the public has joined the audience as it uses science in everyday life, thirsts for biomedical information and demands more of its health care systems. All biomedical institutions should strive for the values of delivering the highest standard of health care, educating the best providers and creating new knowledge. Any institution designated as a research university must achieve its goals while dealing with contemporary challenges in providing healthcare and educating health professionals. Practical issues of size, funding environment and budget must be taken into account.

Biomedical Research Productivity: A Top-down Approach

When taking a top down approach, always involve faculty in decision-making and communication of goals and process. Where can you get a better, already paid for, group of consultants?

First set the goal. If the University of Kansas wishes to move into the top twenty-five public universities for example, it must understand what it takes to get there and define the productivity measures that best fit the goal. Arguably, the top biomedical research universities are defined by NIH award ranking. Therefore, the best measure of productivity is the amount NIH awards to an institution through its granting programs. Accordingly, all productivity should be directed to the goal of achieving NIH funding.

The rationale for using NIH awards as the sole productivity measure includes the simplicity of having one clearly understood outcome. NIH funding is generally accepted as the “gold standard” of quality and the measure used by most major ranking authorities. Further, no matter what the measures used for research productivity, they all flow towards the NIH award. For example, there are strong associations with other markers of research productivity such as alternative funding, publications and markers of national recognition. Thus, total NIH award is a compound or global index that reduces complexity of measurements. As a marker of new knowledge, it also correlates with clinical excellence by enabling the acquisition of cutting-edge scientific information to enhance clinical practice. In and of itself, it is also a marker of research infrastructure growth through the indirect cost mechanism. Since the ranking of medical establishments by NSF (the National Science Foundation) using NIH funding is very similar to their ranking by the popular magazine *U.S. News and World Report*, the public audience that we must engage for approval of research funding may have a closer understanding of this measure than more complex academic measures of productivity.

Implementing the Process

Each school in the health center should be given a separate mandate to increase NIH funding as its foremost goal. Each should complete a strategic research plan for a five to seven year period. Realistic dollar award targets must be set, knowing also that movement up the ranks by institutions competing for the same fund pool demands planning, not to just keep pace but to move ahead. Individual departments or centers must buy in to the mandate that NIH awards are the ultimate measure of productivity. The departments/centers should be structured to provide an appropriate balance of activities that should support the overall values of the school, emphasizing integration and collaboration. For example, a clinical department might be designated to focus on clinical care, which will assist the school in its research agenda through the residual income it generates. Promotions and Tenure Committee policies and procedures must take into account the essential value that an individual faculty member brings to the research mission of the school when not engaged directly in research, but rather in clinical, educational and service activities.

Department/Center Strategy

The research strategy is accomplished at the level of department/center. The Chair/Director must be given autonomy in setting up and maintaining the process. The value of every faculty member must be recognized as contributing to research. Roles and responsibilities must be clearly defined for each individual, with the prime understanding of the contribution to overall research productivity. Accordingly, all incentive programs should be directed to the department faculty as a team.

Administrator Productivity

Administrators who set the goals, generate strategic plans and oversee the process must be held accountable, using NIH awards as the productivity measure of programs in their area of responsibility.

Caveat

The use of a productivity index must not be confused with the goals and values of the institution. Such an index is simply a chosen measure of how the mission of the institution is best monitored. A mission of scholarship and balance in clinical care, education and service for example, is not directly articulated by such an index. A productivity index is part of process, although it should reflect how well philosophy, mission and policy are accomplished.

Conclusion

Vice Chancellor Barnhill, in an address to the conference on productivity measures, expressed his preference to replace the attitude of “only the paranoid survive” with “only the flexible thrive,” to which in turn I would add “with a like-minded inflexible goal.” If one goal of the institution becomes achieving top university status, and this is equated with NIH award ranking, the level of NIH funding received would seem a logical comprehensive measure of research productivity.

EVALUATING UNIVERSITY RESEARCH PRODUCTIVITY: WHAT'S THE ROI ... AND WHO CARES?

R. W. Trewyn
Vice Provost for Research and Dean of the Graduate School
Kansas State University
President, KSU Research Foundation

Public research universities face many challenges in the 21st century, not the least of which involves documenting the value-added outcomes that derive from the teaching, research, and public service missions of the institution. Governing boards, accrediting bodies, funding agencies, state legislators, taxpayers, and the American citizenry in general want to know. Prospective students and their parents want to know: what sort of return on investment (ROI) can they expect? In the new millennium, ROI is a concern of more than just investment bankers and stockbrokers. And universities—just like other entities seeking monetary investments—will be well served if they can provide compelling answers to questions about the ROI they generate in fulfilling their missions.

Teaching Value-Added

The annual teaching/education value-added for an institution can be estimated from the number of alumni residing in the state by using U.S. Census Bureau data for mean annual earnings by level of education. For example, there are 45,564 alumni in Kansas who received bachelor's degrees from Kansas State University (KSU). An individual with a bachelor's degree earns \$19,114 more annually (on average) than someone with a high school diploma, so one can estimate the economic value gained from an undergraduate education by multiplying $45,564 \times \$19,114 = \$870,910,296$.

A recent summary of the value-added earnings of KSU graduates in Kansas for all levels of education projected an annual value-added impact of more than \$1 billion, with nearly \$3 billion in total annual earnings (see attachment). The ROI to the state of Kansas was estimated to be almost seven-fold, that is, for every \$1 the state invested annually, it was estimated that \$6.80 is added to annual earnings. Moreover, KSU graduates are likely contributing significantly more in the way of tax revenues to the state than the state is providing to the university each year.

Public Service Value-Added

The monetary value of university public service activities can also be estimated if appropriate tracking systems are in place to monitor the time spent by faculty, staff, and students in providing such services. At land-grant institutions, extension service activities can be quantified, so this offers one

specific measure. Unfortunately, most other university units don't track public service activities particularly well, if at all.

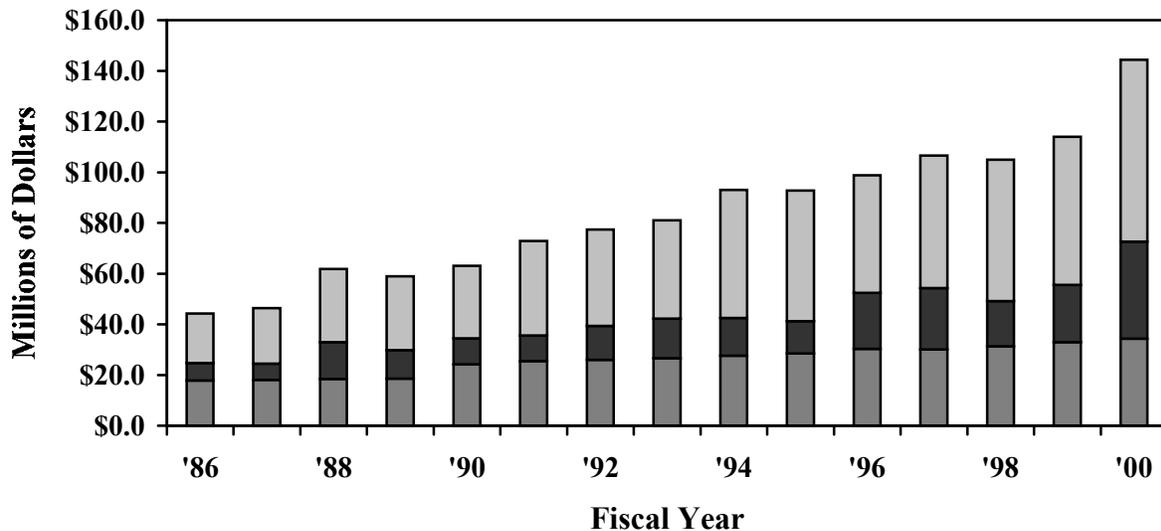
This difficulty became apparent in a recent assessment of the economic impact of teaching, research, and service at KSU, where it was found that quantifiable public service outcomes were small compared to those generated by teaching or research. As a result, these service end-points were merged with teaching or research in the final report.

Research Value-Added

Determining the ROI for university research is unquestionably the most difficult parameter to establish. Consequently, other parameters are sometimes used to assess the impacts of research. Job creation, which is an important consideration from an economic development standpoint, is one such indicator that is commonly utilized. And funding for university research and scholarly activity creates jobs in the same way that R&D funding does in the private sector.

As illustrated below in the fifteen-year summary of funding in support of research and scholarly activity at KSU, the base for FY 2000 approached \$150 million. These funds came from a variety of sources: competitive (extramural awards), donated (transfers from the KSU Foundation to the university), and appropriated (predominantly legislated land-grant support). A research and scholarly activity funding base at this level would create or support about 6,000 jobs as these funds are expended, based on a Kansas-specific multiplier of 40.6 jobs created per \$1 million in research support (derived by the AAU using Commerce Department statistics).

KSU Research/Scholarly Activity Funding Base



While the ROI associated with basic research cannot be quantified in economic terms, we can make reasonable estimates from the outcomes of applied research. Land-grant universities have been doing this for years in the agricultural realm where agricultural economists provide relevant data on a regular basis.

The 1998 publication entitled, *VALUE ADDED: The Economic Impact of Teaching and Research at Kansas State University*, estimated an annual impact of nearly \$1 billion from research, and as already noted, this undoubtedly is a significant underestimate of the total impact because the outcomes from basic research cannot be quantified. Moreover, many non-agriculture departments failed to track the impacts of their applied research.

The insert at the right illustrates how the value-added benefits of applied research can be calculated. In this example from 1994, the specific contribution from KSU to the \$1.5 billion Kansas wheat harvest amounted to \$64 million.

Example: In 1994, wheat harvested from 11.4 million acres in Kansas produced a total wheat crop of 433.2 million bushels. The average yield was 38 bushels per acre. KSU-developed wheat varieties were planted on 36.2 percent of the total acres, with Karl and Karl 92 accounting for 23.6 percent; Larned, Newton, Eagle, Arkan, and Sage made up the remaining 12.6 percent. Compared to non-KSU varieties, Karl and Karl 92 offered a yield advantage of 6 bushels per acre, while the other KSU varieties provided an advantage of 1.5 bushels per acre. With wheat priced \$3.50 per bushel, the value of the total wheat crop exceeded \$1.5 billion, and based on the relative yield advantage of the KSU varieties, \$64 million of that was due to the value-added benefits of KSU research.

The 1998 value-added report established an economic impact in Kansas of more than \$2.4 billion from teaching and research combined. Based on a state allocation to KSU of \$145 million the previous year, a return on investment of seventeen-fold was established, that is, the university returned \$17 to the state for every \$1 of state funding.

Research productivity can also be evaluated by assessing the outcomes of university technology transfer efforts. The return on investment in this instance can come in many forms, including signing fees and royalty payments from traditional patent licensing arrangements, and related research funding that goes back to the institution. Additionally, launching start-up companies can lead to job creation, equity acquisitions, and a variety of revenue streams.

Communicating Value-Added Outcomes

All institutions—public and private—have customers and stakeholders that need information about ROI and other productivity information packaged in an institution-specific manner. Every effort should be made to provide such information, tailored to the needs of the particular customers or stakeholders, whenever possible.

On the other hand, when a single governing board serves multiple universities and/or when multiple universities are located in a single state, it may be advantageous to provide pooled productivity data.

For example, the chief research officers at the University of Kansas, Wichita State University, and Kansas State University presented a combined *Research Infrastructure Report* to the Kansas Board of Regents on March 16, 2000. This report contained economic impact estimates for alumni of the three research universities who live in Kansas, and joint R&D expenditure data; no institution-specific information was included. Combined ROI information was also provided for the three universities, which included return on state investment in education and jobs created by R&D activities. The collective outcomes were substantial and impressive.

Clearly, the most effective way of communicating value-added outcomes is to make them audience specific. One size does not fit all. And we need to listen to those customers and stakeholders. They care, and the ROI data we provide can return huge dividends.

Attachment

Kansas State University Estimated Economic Impact of Graduates in Kansas

Degree Earned	Alumni in Kansas ⁽¹⁾	Aggregate Earnings ⁽²⁾	Education Value-added ⁽³⁾	Aggregate Value-added
Associate's	1,432	\$55,107,656	\$7,615	\$10,904,680
Bachelor's	45,564	2,277,379,848	19,114	870,910,296
Master's	8,253	496,566,504	10,186	84,065,058
Doctoral	881	60,954,628	19,206	16,920,486
Professional ⁽⁴⁾	1,108	100,443,524	40,671	45,063,468
Total Impact		\$2,990,452,160⁽⁵⁾		\$1,027,863,988⁽⁶⁾

- (1) The number of KSU graduates residing in Kansas. Data provided by the Alumni Association as of November 1999.
- (2) Based upon U.S. Census Bureau data regarding the mean annual earnings by level of education for those in the workforce aged 18 and over. For 1998, the mean annual earnings by degree/diploma level were: high school, \$30,868; associate's, \$38,483; bachelor's, \$49,982; master's, \$60,168; doctoral, \$69,188; professional, \$90,653.
- (3) The value-added impact of the education for alumni living in Kansas was calculated as: associate's, \$38,483 - \$30,868 = \$7,615; bachelor's, \$49,982 - \$30,868 = \$19,114; master's, \$60,168 - \$49,982 = \$10,186; doctoral, \$69,188 - \$49,982 = \$19,206; professional, \$90,653 - \$49,982 = \$40,671. To reduce the potential for double counting and overestimating the impact, the value-added differential was calculated on the mean earnings for an associate's or bachelor's degree minus those for a high school diploma and on the mean earnings for a master's, doctoral, or professional degree minus those for a bachelor's degree. As a result, individuals who received their bachelor's degree from one university and their master's, doctoral, or professional degree from another (a common occurrence), would have the appropriate portion of the education value-added earnings credited to the appropriate institution.
- (4) Professional degrees include only DVM degrees.
- (5) According to the Kansas Department of Revenue's Office of Policy and Research, the average for taxes paid in Kansas (State income tax, sales tax, and property tax) amounts to approximately 8.3% of personal income. **Based on \$2.99 billion in estimated earnings, KSU graduates would be contributing \$250 million annually to the tax rolls in Kansas, which is \$100 million more than the state appropriates to KSU.**
- (6) In FY99, the State of Kansas appropriated \$151.2 million to the general operating budget of KSU (Kansas Board of Regents Data Book, February, 1999). Using this current-day financial subsidy, **the state's annual return on investment (ROI) just for the education value-added portion of alumni earnings would amount to \$6.8 million (\$1,027.9 million ÷ \$151.2 million), i.e., for every \$1 invested by the state, \$6.80 is added to annual earnings in Kansas.** However, that 7-fold return doesn't take into account the value-added benefits of research and public service activities at KSU, which increase the total ROI to 17-fold or more based on the results of a recent research value-added study.

THE ROAD TO SUCCESS: MU'S MASTER PLAN FOR RESEARCH AND TECHNOLOGY DEVELOPMENT

Jack O. Burns
Vice Provost for Research

Mary M. Licklider
Director, Grant Writing and Publications
University of Missouri-Columbia

“If you don’t know where you’re going, you might end up somewhere else.” It’s a truism, but it has been an apt one for too many university research endeavors over the years. So it was something of a break with the individualistic traditions of academia in 1998 when the University of Missouri-Columbia’s Research Division began building on the campus-wide strategic planning process to create a Master Plan for Research and Technology Development.

We had a strong starting point in campus planning. The first goal of the campus strategic plan was (and still is) to “strengthen research, graduate, and professional programs and improve our stature among public AAU (the Association of American Universities) and Research I institutions.” Campus priorities under this goal included increasing the diversity of the campus community, involving students in research activities, and maintaining strong lines of communication with constituencies both internal and external to the campus. In addition, several other initiatives had resulted in identification and targeting of several broad areas of research priority: special state funding for mission enhancement, the campus plan’s emphasis on encouraging disciplinary strengths, and campus-level encouragement for the formation of interdisciplinary research teams.

These campus-wide goals and priorities suggested the outlines of a first Master Plan for Research and Technology Development, and this first research plan was organized around fourteen action steps. In 1999, we updated and refined that first plan, adopting six goals for research at the University of Missouri (MU):

- Maximize internal resources and communications,
- Enhance research compliance,
- Provide grant assistance,
- Nurture technology development,
- Expand external partnerships, and
- Foster governmental relations.

These goals were further articulated in objectives and action steps, and we held ourselves accountable to these by establishing measures and assessments against which we could gauge our progress. Finally, we reported on accomplishments that had been achieved toward the action steps established in the first plan.

It is probably worth noting that the original 14 action steps had been organized under five headings: technology transfer, external partnerships, grant assistance, federal relations, and state relations. The overlap of these headings with the six goals of the last two years reflects stability in our larger priorities but not a static planning document. The original five headings saw the merger of state and federal concerns under “governmental relations”; compliance emerged as a separate goal; and more explicit attention came to be focused on internal processes. The objectives, action steps, and measures operate at a level of detail that allow the plan to be responsive to changing conditions, emerging concerns, and completed objectives. Examples from three of our goals will help illustrate how this works.

The first goal in last year’s update of the Master Plan was to maximize internal resources and communications. Seven objectives supported this goal, including plans to enhance internal funding programs, improve internal and external communications, increase support for scholarship in the humanities, enhance MU’s national stature, and sustain a nurturing environment for women and minorities. To track our progress toward these objectives, we established several measures: increasing funding for small internal research grants by \$50,000; increasing the number of local, state and national media stories about MU by 10 percent; adding at least one National Academy or similarly prestigious faculty member per year; and helping with campus strategies to overcome barriers to recruitment and retention of a diverse faculty. We were able to report, too, that during 1999-2000:

- Federal research expenditures had risen 20 percent over the previous year (65 percent in 3 years);
- NIH funding had been secured to establish the MU Center for Phytonutrient and Phytochemical Studies;
- Awards in two of our most popular internal funding programs had risen nearly 70 percent over the previous year;
- The Research Division had added a Faculty Fellow position to lead initiatives in the humanities;
- A strategic and quality improvement plan had been developed and implemented for the Office of Sponsored Program Administration; and
- “Electronic” improvements included a complete redesign of the Research Division Web site, creation of an electronic Grant Data Form for internal processing of grant proposals, and integration of the local pre- and post-award grants databases.

Eight objectives defined the goal of providing grant assistance: developing a campus network of grant writers, increasing private foundation funding, expanding participation in the Community of Science, providing grantsmanship seminars, disseminating information about external funding opportunities, establishing externally funded centers, improving processes and services of the Office of Sponsored Program Administration, and developing and involving student teams in research. Here again, we established empirical measures that were quantitative wherever possible. We targeted a 10 percent increase in dollars requested through proposals receiving grant writer support, 95 percent faculty participation in the Community of Science; 20 percent reduction in instances of awards preceding proposals; and reduction in the mean proposal review time to less than 3 days, clinical trial implementation time to less than 45 days, and grant award implementation time to less than 2 weeks.

Accomplishments related to this goal included quantitative results and the development of a number of tools designed to clarify, speed and/or simplify external funding processes. Quantitative results included a drop in proposal review time from 21 days to 5 days and in award implementation time from 54 days to 37 days. Some of the new tools were documentary, such as template agreements and budget templates. Some were related to information access, such as creating appropriate access to the local grant database for faculty and departmental personnel. And some were training and support “tools”: the enhancement of a Grants and Contracts Support Group as a means for substantive policy flow, and the addition of 2 more grant writers to the growing campus network.

Our fourth goal, nurturing technology development, is an area receiving attention on campuses across the country. For MU, this is a relatively new priority. This novelty is reflected in objectives having to do with increasing the visibility of the new Office of Technology and Special Projects, enhancing its infrastructure, providing mentoring on technology transfer, and stimulating entrepreneurship. Other objectives are more likely to be ongoing even as the Office becomes more established: enhancing intellectual property protection and processes, promoting university research and technology to the corporate sector and economic development entities, and involving student teams in research and technology transfer. Benchmarks for this goal are also cast in quantitative terms where possible. Besides decreasing processing time for intellectual property agreements by 20 percent, we sought to increase numbers of invention disclosures by 15 percent, applications for Small Business Innovation Research and Small Business Technology Transfer by 10 percent, and cooperative agreements and licenses each by 10 percent.

Although the Office of Technology and Special Projects was only a year old, we could still report on accomplishments. During 1999-2000, besides getting the Office established and functioning, seminars were held on all four University of Missouri campuses, the organizational structure was established for the

Missouri Seed Capital Fund, and three new companies were formed around University of Missouri technologies.

By including progress reporting and accountability measures, our plan is a very useable document for guiding day-to-day decisions of senior staff within the Research Division. We monitor our progress more formally at semi-annual, day-long retreats. The mid-year retreat is an important opportunity to monitor progress and adjust our course. The summer retreat focuses on updates and revisions. The agenda for this second retreat also includes discussion of the larger issues of how appropriate the goals and objectives are and whether any should be deleted or replaced. The Master Plan is revised and updated each year following the summer retreat. To ensure that revisions continue to accurately reflect evolving campus priorities, we seek comments and revision suggestions on draft documents from two faculty advisory committees, the Deans, Faculty Council, Provost, and Chancellor.

The resulting document presents a credible statement of purpose and action. But the Plan cannot be actualized by Research Division personnel alone. We post it on the Web. In addition, realizing that the Web is a “passive” medium that requires the reader to seek out a particular piece of information, we print hard copies and distribute them widely. All faculty receive copies, as do Missouri’s Congressional delegation, local legislators, and media outlets. We use it to communicate institutional priorities to faculty recruits and others visiting campus.

By seeking input from faculty and campus administration in the development of the Master Plan but not allowing it to be bogged down in endless committee meetings, and by communicating our goals and strategies widely among the University’s constituencies, we can be reasonably sure that everyone knows where we are going—which certainly raises the odds that we will get there.

JUDGING RESEARCH PRODUCTIVITY ON AN ENTREPRENEURIAL CAMPUS

James R. Bloedel
Vice Provost for Research and Advanced Studies
Iowa State University of Science and Technology

One of the key benchmarks in judging the academic achievements and excellence of faculty is the quality and quantity of their scholarship. For individuals in the social, biological, engineering, physical, and mathematical sciences, research productivity often serves as one of the primary criteria for making this judgment. At Iowa State University this criterion is related to the broader area of discovery, one of the three major tenets of the University's strategic plan.

Traditionally, the assessment of research productivity has been based on the number of research publications in high quality journals as well as the level and consistency of research funding acquired from competitive sources. Despite this well-established practice on many campuses, it is becoming progressively more difficult to utilize only these norms for judging research productivity. The faculty are becoming appreciably more diverse in the exercise of their scholarship. In addition, there is an evolving emphasis on fostering the scholarship of teaching among faculty. Perhaps most important, campuses are becoming much more entrepreneurial. Universities are seeking partnerships with industry, and, as an integral part of those partnerships, entrepreneurial activity among its faculty is encouraged. As a consequence, many are actively engaged in research and the development of intellectual property that can lead to patent applications and the execution of licenses and options. The more aggressive faculty also are becoming involved in the establishment of start-up companies that utilize the intellectual property they developed. Given the time constraints under which all faculty operate, this type of entrepreneurial activity may at least partially displace the more traditional scientific pursuits, such as publishing in scientific journals and acquiring grants from foundations and agencies.

Since professional diversity is now encouraged on our campuses, it may be time to reassess how the scientific productivity of our faculty is to be judged. This presentation focuses on the development of a format for evaluating scientific faculty for promotion and tenure at an entrepreneurial institution. It should be emphasized, however, that the framework suggested here could be modified for use in the annual assessment of faculty achievement.

Two faculty profiles illustrate the type of scholarship activity that could characterize the research portfolio of young scientists who have been particularly active in entrepreneurial endeavors. The first is an assistant professor of

electrical engineering who has served on the faculty for six years. Although having published only eleven manuscripts in reputable journals, this faculty person also had four patent applications. Furthermore, the applied research of this individual was particularly excellent, one project leading to an R&D 100 Award. Grant activity consisted of one grant from a private company that resulted in four years of continuous funding at a level that adequately funded the laboratory's entire research program. Letters from five references were very good, and the teaching assessment of this faculty person indicated an above-average, but not exceptional, performance. The second faculty profile characterizes an assistant professor in mathematics who also has served on the faculty for six years. Although only seven publications in refereed journals were produced, three widely acclaimed web-based courses were planned, organized, formulated and executed under the leadership of this individual. Furthermore, all of the courses are under consideration for copyrights. This faculty member was judged to be an excellent teacher, and the letters of support were very good. Grant activity consisted of one three-year grant supporting the generation of web-based educational materials in mathematics.

To assess these two faculty and their somewhat diverse areas of scholarship, we should begin by establishing a definition of research that can be useful in their applications for promotion and tenure. Clearly a broader definition is required. In general, I favor defining research as objective-driven scholarship. This definition is applicable to research activity across many disciplines, including the arts and humanities as well as the sciences. Furthermore, when appropriate, objective-driven scholarship can apply to educational initiatives as well as extension activities. Consequently, it is feasible to use this definition for evaluating faculty whose emphasis has been in areas outside of traditional research, namely learning and engagement. At Iowa State University, these areas receive significant emphasis in the promotions and tenure process. This broader definition of research certainly includes the more entrepreneurial activities of the faculty mentioned above.

Once a working definition of research is established, it is necessary to develop an approach to **evaluating** faculty with a diversity of achievements and contributions. Based on the examples I have given, this evaluation should recognize some degree of parity among research contributions resulting in journal articles, patent applications and disclosures, and/or intellectual property related to the educational mission of the university. Similarly, a broader view in evaluating laboratory funding may be necessary. For example, although a more classic research career may utilize funding from the National Institutes of Health (NIH) and the National Science Foundation (NSF) predominantly, successful entrepreneurial faculty may derive their funding from private sources very interested in the intellectual property being developed. In some instances, this can be highly competitive, at least at the higher funding levels. Based on this view, the critical issue is the adequacy of funding for supporting both the quality

and the quantity of the laboratory's scholarship, not the specific source of research support.

If one accepts the premise that a significant degree of heterogeneity exists among the scholarship activities of scientists on many campuses, it is necessary to derive a common denominator by which research productivity can be judged fairly across the diverse research programs of faculty competing for promotion and tenure in a Research I institution. One criterion capable of meeting this objective is the **impact** of the faculty person on his/her field. This criterion, **impact on the field**, can be applied to any discipline and any area of scholarship. To meet this standard, the faculty must demonstrate a set of contributions that has impacted a field in a way that modified thinking and/or trends among other scholars in the same area. Almost by definition, implementing this criterion requires the utilization of external peer reviewers; a traditional promotion and tenure committee could not adequately assess the faculty based on "impact on the field" because it would not have the required expertise. Impact assessment is best judged by individuals who are working in the same area and have had multiple years of experience assessing the impact of new ideas and new findings on their field of expertise. External experts could provide an unbiased evaluation of a specific person in the context of other faculty at the same institution, and in comparison with individuals throughout the discipline.

Given the importance of external evaluations in formulating judgments regarding the impact of faculty on their field, I strongly advocate a method of assessment similar to that used in the evaluation of grants and contracts for funding agencies such as NIH and NSF. This new system utilizes an approach analogous to the study section/research council system with which we are all familiar. Iowa State University currently considers applications for promotion and tenure from approximately 70-80 faculty each year. Given the areas of scholarship represented across these faculty, the initial reviews could be performed by study sections in seven areas: engineering; chemistry and physics; math, statistics, bioinformatics and related disciplines; language and literature; fine and performing arts; humanities and social sciences; and medical/veterinary sciences. Study sections could be comprised of two to three invited external experts approved by colleges for assessing their faculty. The exact number would depend on the number of applications that a given panel was going to consider.

Invited participants would be provided with a \$300-500 honorarium for their services. The members could be invited to the University for deliberations over one day, or the members could discuss each of the applicants using a conference call format. The latter approach clearly would result in minimizing the expenses of the overall process. If visits to the campus were considered preferable, assuming a \$1000 travel allotment for each person and an average number of members per section of 2.5 across the seven areas to be reviewed,

the maximum cost per year would total \$25,850. The use of conference calls could accomplish the review for a fraction of this cost (\$8350).

Following the assessment of the “study sections,” a “council” consisting of the Provost’s team (consisting of the associate provosts and vice provosts) would then review the pending recommendations for consistency and fairness. The “council” would have the responsibility for modifying recommendations should it be necessary.

In summary, this presentation illustrates a progressively more common profile of scientific faculty at an entrepreneurial institution and provides a practical suggestion for fairly and adequately addressing the evaluations required for their tenure and promotion. Although it could be argued that the faculty profiled above are not worthy of tenure and promotion based on any criteria, there is no question that the scholarship of the scientific faculty at our institutions has become much more diverse. Since this trend often parallels a related change in the priorities of the university, there is little doubt that a broader perspective is required for faculty evaluation than has been employed in the past. In fact, on some campuses, chairs and deans have viewed entrepreneurial activity negatively. This must change if campuses are to attract and retain the best of the new breed of faculty. Many of these individuals are interested in the wide-range of experiences that result from entrepreneurial activities, not as a substitute for their more traditional scholarship activities, but rather as a complement to their professional experiences while serving our institutions. Their contributions not only add to the research culture on our campuses, they also provide unique training opportunities for our undergraduate and graduate students. These training opportunities support current trends in graduate education that emphasize the importance of meeting the needs of students interested in careers in industry. The best programs and the most progressive campuses will be those that can accommodate to the new trends and maintain excellence in programs that continue to train students for academic careers.

A REFLECTION ON A DAY SPENT DISCUSSING EVALUATION

David Shulenburger
Provost
University of Kansas

Evaluation, our topic for this retreat, is part of a much larger activity in which universities are engaged. Initially, the focus was on assessment of graduate and undergraduate programs. Now essentially everything in university life is subject to extensive evaluation.

There was a time when the academy was a more self-contained environment and informal evaluation of its activities was sufficient. This began to change with the advent of significant federal funding as society began to look to universities as vehicles for improving the world. Once this expectation was placed upon us, we were obliged to demonstrate what we were accomplishing.

Because the missions of universities are amorphous, no clear measures have emerged. Our situation contrasts sharply with the business world where evaluation is easy. The standard there is whether or not you make a profit. There are many who quest for a simple, straightforward evaluative model like that of the business world for universities. They yearn for an evaluation of university effectiveness so simple and clear that it cannot be argued or debated.

A storm of controversy engulfed academia as a result of the release of the National Research Council (NRC) evaluations early in the 1990s. In spite of all the adjustment and tinkering that will occur before NRC measures us again, I suspect that the uproar will be no less this time than last. What does that tell us? We don't agree among ourselves about what quality is. We don't agree about the objective function of evaluation. I don't mean to criticize the NRC process. What I suggest is that we are a long way from having an absolute criterion for reaching the gold standard of university evaluation, a simple non-debatable measure like the business world's profit measure.

We've talked about market models of evaluation at this conference and they surely should be considered as they begin to approximate the accepted model of the business world. Examples of "market measures" are the quantity of external funding the university attracts or the number of students who choose to attend. These models are often toyed with, but the simple problem is that the market model doesn't work unless you can demonstrate that you are covering the full cost. Essentially all of our activities are subsidized. Unless we can

assure the public that we are not “buying” our market with our subsidy, the market measures we develop will not be accorded respect by the public.

We should be aware that some of the measures we use are not persuasive to the public and for good reason. It does not persuade the public when we say that our graduates have higher incomes because we educated them and they consequently pay sufficiently higher taxes such that the state is now earning a surplus because it funded higher education. That argument fails not because the public refuses to believe that higher education raises income; it fails because the public knows that if we did not provide higher education to citizens of our state someone else would. For decades New Jersey chose to have almost all its students go to private universities or go out of state for higher education because of the dominance of this logic. Similarly, a claim that we have significantly increased revenues for farmers by developing a high yield variety of wheat does not work unless we are measuring only the difference in yield between our variety and that of the next best producing variety. The public is sophisticated enough to credit us only with the unique gains that we produce.

Leverage is another kind of argument that we make in support of our impact. We argue that if a certain amount is allocated to the University of Kansas, the institution will return to the state’s economy a multiple number of dollars. Is leverage not persuasive? It is, but the public and legislators know that the leverage we create ought to be compared with leverage that might be derived if the funds were spent on some other function of government. Is our leverage really better than that derived from money spent on traffic safety? On early childhood education? On expansion of the lottery?

Suzanne Ortega said that she lost more personal capital talking with faculty about assessment than any other activity. Why is this? I think it is because faculty do not really believe that assessment measures what education produces. The faculty believe assessment is something we have to do to satisfy those who accredit us, but they don’t believe that assessment measures whether the education we provide makes a difference. They see assessment as spending lots of time with activities that aren’t of value and which produce measures that won’t be convincing to others.

Our discussions have included the great importance we accord to evaluation of faculty based on publications in refereed journals. This is an important issue to me. I’ve talked with many academics, librarians and commercial publishers about this topic. Some believe that universities created the problem of rapidly increasing journal costs by rewarding faculty based on the number of articles they publish in refereed journals. Not everyone is convinced, as I am, that journals contain important research findings and new insights. Much of the rest of the world thinks we’re playing games with the articles we write, and the commercial publishers believe that as long as we are stupid enough to

continue the game, they are going to make money from it. Cynicism is raised to exponential levels by this cycle.

We risk being self-congratulatory by putting time into evaluations of our own construction that the rest of the world does not accept. How many times have you heard someone outside academe or the federal government talk about NRC rankings? The silence on this issue from the rest of the world reflects the fact that NRC rankings have zero meaning to them. What ranking does mean something to the public? For the people on the street, the most credible ranking is by *U.S. News and World Report*. Interestingly, academics tend to think it is meaningless, but the public demonstrates belief by continuing to make the rankings issues of *U.S. News and World Report* best sellers.

“Aunt Emma” stories are also effective. The credible information most people use to evaluate our institutions involves the stories about specific students whose lives were changed because they came to our universities.

If our evaluation schemes do not measure up to the high ideal of quality held by academe, we are endangering ourselves by engaging in them. The public will see us as phonies if we use measures of our effectiveness that even we cannot accept as legitimate. The public is willing to tell us that we are wearing no clothes if we insist on parading about naked.

There are reasons for measuring external funding—to encourage the faculty to seek it, and because we don’t have enough money to fund their research unless they seek it. There are good reasons to fund research competitively—because awarding funds objectively puts scarce resources to the most valuable use. If we use external funding as a measure, the reason should be that we want to encourage our faculty to put effort into successfully acquiring external funds. We should be very careful about arguing that we are a better university because we bring in more external funds or that society is better off because we are successful in the external funding game.

Susan Kemper expressed a healthy skepticism about the process of evaluation. It is important to keep her message in front of us. Evaluation can be a useful activity, but we must keep our perspective clear and claim no more for the activity than it merits.

HIGHER EDUCATION ADVOCACY: THE INTERFACE OF TWO CULTURES

Kim A. Wilcox
Executive Director
Kansas Board of Regents

After spending nearly three years in state government, I still think of myself as an academic. Sometimes, in fact, being in Topeka I feel like Gulliver in a strange land.

The differences and similarities in culture between higher education and the legislature can be seen in many different arenas. For example, in academe, we develop independent lines of research, but we all believe in the value of collaboration in pursuing our research goals. Legislators value independence, as well. In fact their very success in politics is predicated on convincing the electorate that they are independent thinkers and are unique from all other candidates for the position. At the same time, legislators can't get anything done without collaboration. They must be able to build coalitions and implement group goals. Like ours, the legislative merit system is highly individualized (elections) but the day-to-day working mode is collaborative. Their method of collaboration, however, is different from ours. We try to preserve the collegial environment at all cost. For us, compromise means allowing people to follow their own agendas even if we don't agree with what they are doing. In the legislature, compromise means developing a coalition, negotiating a middle ground, and closing a deal to accomplish a short-term goal; the long-term environment is less important

For scientists, peer review is the highest level of accountability. From the outside, this can look self-serving, in that we are essentially reviewing ourselves. Moreover, while we realize that some politics are in play in any human system—including scientific peer review—we trust that the basic process is an objective one. For others, and especially for politicians, it is easy to assume that the politics of the process are much more important in determining the outcome.

Legislators think academics are smart, and they value our knowledge. At the same time, they don't think we're grounded in the real world. Our work is often driven by a variety of goals and motivations that may not be apparent to the public and particularly to members of the public who are focused on local or regional concerns. Legislators also think we don't listen very well. From their perspective, after long conversations about our work and their aspirations for the state and for our science, we go back to our labs and do what we've been doing for years. For us, those conversations were more background and context for a career of science; for them, the same conversations were a request (or perhaps even a demand) for specific activity.

One of the most fundamental differences between the research community and the legislature is the role of data. Much of the decision-making in the legislature is driven by anecdotes. An anecdote that is replicated a couple of times may prompt us to begin a series of experiments. By contrast, it may prompt the legislature to invoke a law. This difference is driven in large measure by the different timelines on which we operate. In Kansas, 125 of our legislators are elected every two years. As a result, they need “to do something” in Year One, in order to run for re-election in Year Two. They can’t wait for several years to test every hypothesis and divine the preferred strategy for achieving particular goals.

Leadership is another area where the legislative and academic cultures differ. On campus, committee chairs are seen as facilitators (and report writers). They are not empowered to determine the agenda, or the direction of the deliberations. For example, there would be an uprising if a faculty committee chair failed to hold a meeting of her/his committee. By contrast, in the legislature, if a chair doesn’t want to hold a meeting, the committee doesn’t meet. Legislative leadership is crucial in every sense of the word: how an item is discussed, whether it is discussed, and whether there is action.

While not claiming to be an expert, I would offer some advice to the academy in working with the legislature. First we need to provide a focused message. As academics, we tend to seek out complexity. We must realize, however, that the press doesn’t deal well with complexity. As a result, our public message too often is lost or confused. We need to spend more time thinking about what it is we are doing and how our work can be cast into an appropriate form. When I was full-time at the university and someone would ask what I did, I would go on at great length trying to describe my applied work in experimental phonetics. My wife would frequently remind me to find a better way to describe my research, because most listeners were yawning before I got to the good part. Generally, people want to know what it is we are going to fix and how long it will take to fix it. This is not how we’re used to talking in academia, but we can become better at it.

We can also do a better job of unifying our voices. Too often, we find ourselves espousing our own individual needs and positions. Legislators are not going to give money to one professor versus another or to one department versus another. Instead, they are going to allocate a lump sum to higher education. The extent to which higher education delivers a single message makes it easier for the press to espouse it and easier for the legislature to accomplish our objectives. Too often faculty members are talking about one thing and administrators another. Or, a department advocates one objective, and another department advocates something else. This is healthy in an academic sense, but the more consistent we are in our message, the better we will ultimately fare in the legislative process.

Our long-term work with legislators should combine “friend-raising” with fund-raising. The faculty can play an important role in this, especially in the “friend-raising” area. Professors can put a human face on what we do, not as advocates, but in telling legislators about their work. What we do is interesting, and it is a source of pride to legislators. We should capitalize on that opportunity. By contrast, almost anything faculty do in asking for funding looks self-serving. That message generally is better carried by the administration, with the faculty’s role being to stay in tune with the administration’s message.

Anthropologists will tell you that any cultural generalization is just that, a generalization, and shouldn’t be construed as definitive particularly in analyzing a given situation. My comments should be taken in that light. They are generalizations derived from my observations over a couple of years. Nonetheless, I hope that they have some utility for you.

KEEPING A BROAD PERSPECTIVE ON RESEARCH

Rollin C. Richmond
Provost
Iowa State University of Science and Technology

Introduction

The background and interests of administrators inevitably influence how we perceive the research contributions of our colleagues. I have spent the majority of my professional career engaged in the study of evolutionary processes mostly at the molecular level. It is an exciting time to be an evolutionary biologist since the tools of molecular biology have enabled biologists to ask and often answer fundamental questions about the origins and evolution of organic diversity. It is now a commonly, although not universally, accepted fact that evolutionary processes have created the organic world as we see it including the human species and the full array of its culture.

E.O. Wilson (1998) in his sweeping volume, *Consilience*, propounds the view that human culture can be unified under the banner of the scientific paradigm. Wilson envisions a time when scholars in the humanities and the social sciences will recognize that they study phenomena that are as subject to scientific investigation as is the analysis of human disease or the fundamental properties of matter. I wish to make a similar, though much more limited point, in this paper. As academic administrators who evaluate the research contributions of our colleagues, we need to take a broader perspective on what we consider high quality research. Moreover, we need to realize that the academic disciplines, including that of education, share more in common, as Wilson would teach us, than is the general view.

Much of the academic world accords higher prestige to research that seeks to elucidate the fundamental properties of our world. Research whose goal is to advance the daily lives of people or to improve existing processes often is not accorded the kind of recognition academics reserve for the study of **fundamental** questions. Given the often-biting criticism leveled at scholarship in the academic world, I believe it is time for us to reconsider the values we attach to research productivity in the hope that our research culture will respond to the real and often un-addressed needs of human society.

The Critics View of Academic Research

Martin Anderson (1996) in his book, *Imposters in the Temple*, writes, "The dirty secret of academic intellectuals is that much of what they write and hold up to themselves and to the rest of the world as the highest expression of what they

do is inconsequential and trifling. Taken as a whole, academic research and writing is the greatest intellectual fraud of the twentieth century.” Unfortunately, there is an element of truth in this accusation that cuts to the core of what the academic community is about. I do not subscribe to the notion that it is the whole truth, however. I imagine that the postmodernist trend in literary analysis is a, if not the, motivation for Anderson’s assertion. This approach to research so prevalent in literary and historical scholarship has made apologists of a broad range of our colleagues in the humanities and social sciences. Indeed some humanists have gone so far as to assert that the findings of science, like art, are totally subjective. Most of us have done little to advance the contrary view and thus counter the perspective among the lay public that much of research is a waste of time amounting to expensive navel gazing. The topic for this meeting, Evaluating Research Productivity, requires a broader perspective that must take into account the ways academic research is viewed by the society that provides the resources for our efforts.

The management guru Peter Drucker in an interview in *Science* (July 18, 1997) provides a second critical view of academic research. Drucker is reported to have opined, “I consider the American research university of the last forty years to be a failure. The great educational needs of tomorrow are not on the research side but on the learning side.” This is a devastating criticism that academics worldwide must address. Despite the fact that many academics believe teaching and learning are important—indeed the most important missions of their institution—few academics have engaged in scholarship designed to investigate the most effective ways of helping students learn the disciplines they care for so deeply. Few scholars have even read the existing literature that provides valuable information on learning styles, including the factors of age and experience, and the most effective processes for learning.

So much is at stake in our world that ultimately must be addressed by teaching people how to analyze their situation and respond to it in an effective manner. The great problems of the developing world—overpopulation, poverty, disease, illiteracy, and corruption—all must be confronted by the people who experience them. However, the academic world could do much to help if we provided models of the best way to teach people and do it quickly. Academics in the developed world have contented themselves for many centuries with the same approaches to education as they themselves experienced. How is it that we have largely failed to apply the tools of science to learning?

A part of the answer to this question lies in the way in which we evaluate research and the values we espouse in making these judgments. Let’s begin by enumerating the characteristics of a scholar and the values that scholarship invokes.

The Qualities of a Scholar

There are three general characteristics that good scholars exhibit (see Glassick, C., Huber, M., & Maeroff, G., 1997). First, a scholar must have integrity. By this is meant truthfulness, fairness, the absence of fraud and dissembling. A scholar of integrity should use her talents to advance the general good of her society. A scholar of integrity gives credit to those whose work she builds upon and to those who have assisted her. Second, a scholar must persevere in his work and disseminate his scholarly findings to others. Without the critical review of colleagues, scholarship may well lack rigor. Finally, a scholar must show courage, sometimes at significant personal risk, in her search for truth. The system of providing academic tenure is a means of encouraging scholars to show courage and take risks with their scholarship so long as such efforts advance the search for truth.

The Criteria for Good Scholarship at Research Universities

Ernst Boyer of the Carnegie Institution for the Advancement of Learning argued in *Scholarship Reconsidered* (1997) that the evaluation of faculty performance in research universities should be changed to include the expectation that faculty will be both good teachers and good scholars. He did not mean that faculty should place less emphasis on research. Indeed he argued that "...original research and publication should remain the basic expectations and be the considered the key criteria..." for promotion and tenure. Boyer's work drew the attention of scholars to the research opportunities that analysis of learning entails. He did not intend this message for a subgroup of faculty in the schools of education, but rather he hoped to change how academia views scholarship in all the missions of research institutions—learning, engagement, and research. While research, and often fundamental research, is the core mission of research universities, Boyer proposed that teaching undergraduates is as important for advancement as the discoveries made in fundamental mathematics. Research on learning would advance our ability to help our students, and would demonstrate to the citizens who support our research that we are concerned with issues of direct relevance to them. Thus, Boyer would have us use a significantly broader definition of research in making judgments about the academic worth of our colleagues. He would have us apply the standards of good scholarship across all the missions of the research university, not just its research mission.

One of the potentially important results of good teaching and effective learning among undergraduates at research universities is that it will ultimately help advance the disciplines themselves. Academics often forget that among the students they teach are those who will themselves become academics. Jaroslav Pelikan (1992) in his book, *The Idea of the University*, makes a point similar to E. O. Wilson's in his discussion of consilience. Pelikan argues that great scholars are often much influenced in their research by what they learned in other

disciplines as undergraduates. If we are willing to accord prestige and value to high quality undergraduate instruction, we not only may enhance the likelihood that research universities will continue to be supported by society, but we may even be contributing to the development of knowledge that will help take academic disciplines to new levels of understanding.

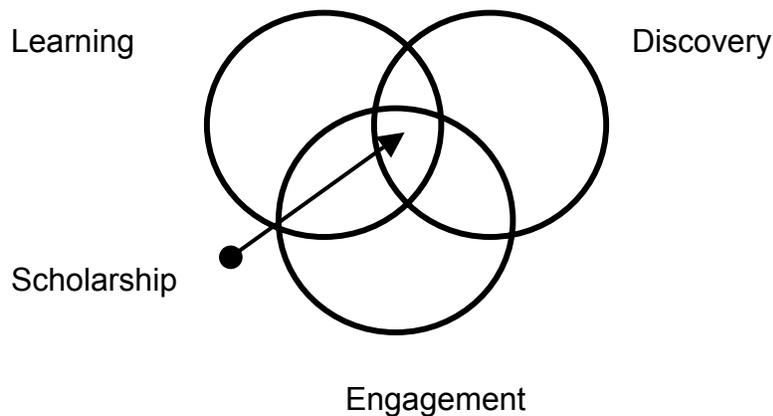
The land-grant institutions that this country so wisely created in 1862 are held responsible for providing education to students who might not otherwise have the opportunity or resources to obtain a university education. The land-grants also carry responsibility for applied research that is designed to help people in their daily lives. Applied research has not been accorded much prestige in American academia, and it is time for academic administrators to reconsider the significance of applied scholarship. John Maddox (1998), a theoretical physicist and former editor of the journal *Nature*, wrote in his book *What Remains to be Discovered*: "...the science that has dramatically changed and improved the lives of people in the past century is applied science." I turn now to the efforts Iowa State University has made in encouraging research across the spectrum of its missions, and especially in the areas of applied scholarship in the plant sciences, a discipline that is critical to the economic future of an agricultural state.

The Plant Sciences Institute at Iowa State University

The State of Iowa and Iowa State University undertook a joint public-private effort to develop a new institute devoted to the disciplines that are critical for the continuing development of agriculture in Iowa. In 1998, the Legislature set aside the first installment of state funds to allow the university to create an institute that would bring together existing faculty and draw new faculty to the university. The Plant Sciences Institute (PSI) takes as its mission "enriching agriculture, the environment and our lives through science." The institute now consists of nine centers covering a broad spectrum of the plant sciences disciplines and includes work on plant transformation and gene expression, designer crops, plant responses to environmental stresses, seed science, plant genomics, bioinformatics and statistics, plant breeding, crop utilization research and a center for designing foods to improve nutrition. The PSI is actively recruiting faculty and also engages existing faculty from our colleges of Agriculture, Liberal Arts and Sciences, Family and Consumer Sciences, and Engineering. Its public funding is increasing every year, and it receives wide bipartisan support from the State Legislature. In addition, the university has received substantial private gifts in support of the institute and allied disciplines. It is clear that the notion of an interdisciplinary group of faculty devoted to improving the lives of Iowans through science is attractive to both public and private organizations. The PSI and the values it projects underscore a wider effort at Iowa State University to reconsider how scholarship is valued and rewarded within the academic world.

How Iowa State University Interprets and Values Scholarship

Over the last three years, the university worked to create a strategic plan that would support and encourage the development of scholarship across our missions in a manner that draws from the ideas advocated by Ernst Boyer. The following diagram represents the interplay of our missions at the university and the scholarship of our faculty.



The three circles represent the missions of the university. The identifying words differ from the standard missions of teaching, research and service because we wish to emphasize that each of these missions relate fully to faculty and students. A faculty member engaged in teaching a course to her students is as likely to learn from organizing a course for students, as the students are to learn from her. Moreover, students are likely to teach faculty not only about the efficacy of their pedagogical methods but sometimes about the discipline being investigated. In a similar way, discovery and engagement are behaviors that both faculty and students will do. The intersection of the three missions of the university is where we encourage most of our scholarship to occur. If we are able to change the culture of reward in our institution to favor faculty whose scholarship informs and supports each of the university's missions, we will have made Iowa State University a better place for all who work there.

The university's strategic plan builds upon a revision of the university's tenure and promotion guidelines adopted by the Faculty Senate and the full faculty in 1999 (see: www.provost.iastate.edu/handbook/99toc.html). This remarkable document builds on the ideas first espoused by Boyer and encourages faculty to use the principles and values of good scholarship in every aspect of their work. Iowa State University characterizes good scholarship in five ways:

- Scholarship develops and communicates new understanding, new knowledge, insights, technologies, materials, uses, and beauty.

- The audiences for scholarship are: peers, undergraduate and graduate students, postdoctoral associates, users, patrons, and the public generally.
- Scholarship can be communicated to others through: teaching materials and methods, classes, curricula, publications, presentations, exhibits, performances, patents, copyrights, and the web.
- Scholarship is validated if it is: original, significant, accurate, replicable, of substantial scope, applicable, of depth and duration of influence, adopted by peers, and has impact or public benefits.
- Scholarship can be documented by: being validated by peers, communicated to peers and broader audiences, recognized, accepted, cited and adopted by others.

Potential Difficulties in the Evaluation of Research

The traditional means of evaluating research has many advantages. It is relatively easy to count publications, determine the quality of journals used, and count the number of invited presentations at conferences or papers in invited volumes. It is even easier to assess the amount of research money an investigator brings to her institution. However, these measures will not accurately assess the worth of research contributions or identify individuals who are truly making efforts to use the tools of good scholarship across the missions of their institutions. We need a broader means of measuring the significance of scholarship, and we need to place it in the context of the institution it serves, not just the discipline it supports. Administrators need to examine their reward structures and ask if these rewards bring about the changes we seek. Often administrators espouse an ethic of collaboration and breadth, but reward individuals whose scholarship is narrow and of real significance only to a limited academic audience. Finally we need to recognize and reward faculty who attempt to improve the quality of learning, our principle goal. We are likely to experience increasing difficulties in attracting public and private support for our institutions unless we change the way that we reward academic scholarship. The adoption of a broader perspective on research is perhaps the best means we have to ensure the continuation of one of the most productive, responsible and truthful institutions in our society—the research university.

References

Anderson, Martin (1996). *Imposters in the Temple*. Hoover Institution Press.

Boyer, Ernst (1997). *Scholarship Reconsidered*. Baltimore: Carnegie Foundation.

Drucker, Peter (1997, July 18). Interview in *Science*.

Glassick, Charles E., Huber, Mary T., and Maeroff, Gene I. (1997). *Scholarship Assessed*. New York: Jossey-Bass.

Maddox, John (1998). *What Remains to be Discovered*. Free Press.

Pelikan, Jaroslav (1992). *The Idea of the University*. New Haven: Yale University Press.

Wilson, Edward O. (1998). *Consilience*. New York: Vintage Books.

RESEARCH AND INSTITUTIONAL MISSION AND NICHE

James R. Coffman
Provost
Kansas State University

Every university operates within a context, which, at its center, is relatively constant, although its expression changes very rapidly at times. This context is comprised of charge (mission), culture, and institutional type. These three aspects combine to render every institution unique.

The university is, in the first instance, a place to educate students; everything else is added on or happens in support of the central mission of teaching and learning. Without the teaching and learning mission, universities would be research institutes. Having said that, we also acknowledge that research and graduate education grew rapidly in the post-World War II era, and in recent times, our expectations for research have increased as the country depends more and more on its universities for research and development. During the last fifteen years or so the research mission has expanded to include a technology transfer component (read economic development). While every research university works to the limits of its ability to expand the research and development effort and secure funding to do so, this happens in a context in which education retains primacy.

Viewed from this perspective, research and creative work at a university enhances the undergraduate experience in important ways, from the nature of the faculty to exposure to the processes of inquiry and creation. Research and creative work are the very foundation of graduate education.

It should be noted that all this happens in the presence of a state and federal agenda. The federal part of the agenda is focused, for the most part, on research and research funding, and the state level agenda is founded and evaluated primarily on the basis of undergraduate education. These two forces frequently are in conflict. The degree of conflict is dictated to a significant extent by the degree to which state government comes to understand the relationship between the potency of research in its universities and the long-term enhancement of the state's economic base. On any given day, however, most legislators are driven by the most immediate concerns of their constituents, which stem primarily from issues of undergraduate education.

The four-state region of Kansas, Missouri, Iowa and Nebraska provides an interesting long-term study in the ways universities are viewed as economic development engines. Driving through the various states, one is reminded that, by comparison, Kansas spends a lot of money on highways (judging by the

consistency of the driving surface). Leafing through data on comparative funding of research universities, it is equally evident that the other three states place far more emphasis on funding their universities. Time will tell which model yields the highest outcome in socioeconomic terms.

These funding patterns reflect more than 100 years of the political decision-making process, which, today, defines each state's approach to such matters. The outcomes therefore, have very, very deep roots and, in fact, reflect the history and culture of the individual state. Reflecting their intensely populist ethos, the citizens of Kansas have created a system that maximizes post secondary educational opportunity for traditional age students. This has resulted in a high participation rate via community colleges, technical schools and colleges, regional and research universities. In terms of resources, it has resulted in a comparatively low per student funding level, and a comparatively high per capita funding of post secondary education. Thus the resource issues, which affect the research universities in Kansas, are not a result of penury on the part of the taxpayer, but rather an unusually high degree of dilution resulting from emphasis on participation rate.

A central point is that these circumstances are not an accident. It is how Kansans have wanted it to be for well over a century, and changing the pattern is not going to happen overnight. In fact, one could argue that Kansas has it right. That certainly is so if a high participation rate is the right priority, the current fiscal straits of some community colleges notwithstanding.

Within this milieu, each university has a mission, or charge, and a culture of its own, based upon which it must establish a niche of optimal competitiveness for the future. At Kansas State University (KSU), for example, within the traditional mission of a land grant university, our strongest position, within and without, is that of a truly student-centered research university. A very decentralized, college-oriented administrative system and culture, which honors teaching and learning as well as research (and extension), has evolved over more than a hundred years. Today, our university has both strengths and weaknesses as a result of a long, incremental process of academic evolution. The trick is to emphasize the strengths and minimize the weaknesses in today's environment. The descriptor "student-centered research university" can best express our strengths at KSU. A long, gradual acquisition of traditional strengths and weaknesses characterizes every research university in the country, although the descriptors differ, at least in terms of emphasis. Having said all this, every research university must pursue the enhancement of its research base, within its unique context of culture and fiscal circumstances, as aggressively as possible. Reasons to pursue research include: creating an atmosphere of inquiry throughout the university; providing an appropriate research base to support doctoral programs; furthering economic development; and obtaining (grant and contract) money to fund graduate student stipends and expand the scientific equipment infrastructure.

In order to accomplish this at a competitive level in today's extraordinarily under-funded environment, it is, in my view, necessary to recognize that the standard model of a complete scholar (in which each faculty member is expected to produce optimally in teaching, research and service over the entire course of a career) is too constraining to be affordable. Faculty who fall short in research over an extended period of years, underutilize other talents and tend to burn out. The resources attached to their research time are utilized poorly or not at all. Every person is different to begin with, and circumstances change on an individual basis over one's career. Ernest Boyer (1990) understood this as the basis of his landmark publication *Scholarship Reconsidered*, in which he introduced a new vernacular under the terms scholarship of teaching, discovery, integration and application.

A framework is thus provided in which, over the course of a career, each person's strengths and passions can, to a greater extent, be capitalized upon, thereby enhancing the collective productivity of any academic unit. The truth is that, over the course of a career, not everyone is equally able to maintain a nationally competitive level of research (or other creative) output. An even smaller number are able to establish and maintain a national reputation, and an even smaller percentage are able to consistently frame the right questions to the right funding agency to bring in resources at a level that can allow competitiveness in research. So, while it is in everyone's best interest to celebrate and capitalize upon those who can "do it all," we should recognize that not everyone can, and the university has many different kinds of important work to do. Within the scholarly milieu, then, it is most effective to engender enough flexibility in the system of roles and rewards to allocate work according to individual strengths, to the extent possible—especially during the long post-tenure period.

The collective productivity of an academic unit, be it a department, college or university, is, of course, guided by its mission and molded along the contours of its culture by its existing and potential strengths and by the expectations of taxpayers and tuition payers. Research (and other creative work) is an essential tool in continuously creating the most aggressive and innovative advances in the service of knowledge for the sake of knowledge, and it also serves the people who pay the bills.

Reference

Boyer, Ernest (1990). *Scholarship Reconsidered: Priorities of the Professoriat*. Carnegie Foundation for the Advancement of Teaching.

CONFERENCE PARTICIPANTS 2001

Keynote Speaker

Joan F. Lorden, Associate Provost for Research and Dean of the Graduate School, University of Alabama at Birmingham

Kansas State University

James R. Coffman, Provost

James Guikema, Associate Vice Provost for Graduate Research

Carol Shanklin, Professor, Institutional Management and Dietetics, and Assistant Dean of the Graduate School

R.W. Trewyn, Vice Provost for Research and Dean of the Graduate School and President, KSU Research Foundation

University of Kansas

Robert Barnhill, Vice Chancellor for Research and Public Service and President, KU Center for Research

Don Hagen, Executive Vice Chancellor of the Medical Center

Susan Kemper, Distinguished Professor of Psychology and Gerontology

Kathleen McCluskey-Fawcett, Interim Dean, College of Liberal Arts and Sciences

Elias Michaelis, University Distinguished Professor and Chair of Pharmacology & Toxicology, and Director, Higuchi Biosciences Center and Center for Neurology & Immunology Research

Mabel Rice, Director, Merrill Center, and University Distinguished Professor, Speech-Language-Hearing

Dick Schiefelbusch, Director Emeritus, Institute for Life Span Studies

David Shulenburg, Provost

Steven F. Warren, Director, Schiefelbusch Institute for Life Span Studies and Director, Mental Retardation & Developmental Disabilities Research Center

K. Michael Welch, Vice Chancellor for Research and Senior Associate Dean for Research & Graduate Studies at the Medical School, and President of the Research Institute

University of Missouri - Columbia

Jack O. Burns, Vice Provost for Research

Suzanne Ortega, Dean of the Graduate School & Vice Provost for Advanced Studies

Michael Podgursky, Professor of Economics

University of Nebraska

Thomas H. Rosenquist, Vice Chancellor for Research, Medical Center

Iowa State University of Science and Technology

James R. Bloedel, Vice Provost for Research & Advanced Studies

Marit Nilsen-Hamilton, Professor, Department of Biochemistry, Biophysics and
Molecular Biology

Rollin C. Richmond, Provost

Kansas Board of Regents

Kim A. Wilcox, Executive Director

Other Participants

Jon Josserand, Governmental Relations, University of Kansas

Marlin Rein, University Director of Budget & Governmental Relations,
University of Kansas

Keith Yehle, Legislative Assistant to U.S. Senator Pat Roberts