

**Multiple Objectives in Planning School Performance Measurement:
Can the Diversity of Planners' Scholarship be Usefully Assessed
at the National Level?**

Paper prepared for
ANZAPS 2005
Adelaide, Australia
30 September – 2 October

Bruce Stiftel

Permanent contact info:

Professor
Department of Urban and Regional
Planning
Florida State University
Tallahassee, FL 32317-2280 USA
voice: +1.850/6448522
fax: +1.850/6454841
bruce.stiftel@fsu.edu

Temporary info (Aug-Dec 05):

Visiting Fellow
School of City and Regional
Planning
Cardiff University
Cardiff, CF10 3WA, Wales, UK
mobile: +44(0)7981/467718
fax: +44(0)2920/874845
bruce.stiftel@fsu.edu

Abstract:

Urban Planning School Performance Measurement can be useful to realistic self assessment by faculties, to assist schools in internal university competitions for resources, and to improvements in visibility of the profession. Yet, historically, U.S. planning schools have resisted efforts to compile and release school performance data. This analysis relays the history of unit performance measurement internationally and in the U.S. and describes efforts currently underway to develop a U.S. national system for collection and dissemination of comparative data on planning school performance.

Various national systems collect and openly report data on university programme performance in the U.S., but none of these include programmes in urban planning. These relatively decentralized efforts stand in contrast to the government-run university performance measurement systems in place or proposed in other English-speaking nations. Notable also is the distinction between performance measurement systems that conduct *de novo* assessments of quality and those that rely on measurements resulting from independent decision contexts. A current ACSP initiative proposes a Planning School Performance Measurement system that could capture the breadth of what planning schools do. The paper discusses principles for conduct of such a study and evaluates candidate measures in each of the broad areas of school activity: social science modelled research, design, outreach, teaching and reputation.

**Multiple Objectives in Planning School Performance Measurement:
Can the Diversity of Planners' Scholarship be Usefully Assessed
at The National Level?**

The past fifteen years have seen increased attention to university programme performance measurement in the interests of promoting quality in research and teaching, as well as cost efficiency. In some countries, including the UK, urban planning programmes have been directly affected by decisions resulting from these measurement systems; in others, including the US, planning programmes have been affected indirectly in terms of visibility and competition for resources. This paper reviews the development of university unit performance measurement internationally, examines the US record, and suggests a new US approach intended to benefit the development of planning education.

University performance measurement in general, and urban planning school performance measurement in particular, prompt wide disagreement. Institutions are quick to claim status positions from the results of performance studies. Perusal of university promotional materials quickly shows prominence given to the results of any ranking scheme that might be plausibly interpreted as showing the institution in question in a favourable light. Yet, by their nature, performance measurement programmes must focus on limited yardsticks, leading Thomas (2005, 241) to suggest that the British assessment exercise "poses great dangers of narrowing and distorting the purpose and scope of university life." When the ranking schemes in question are based on controversial performance measures, or where the performance measures used are not revealed fully, criticisms can be widespread and heated.

Central to the criticisms is the mismatch between the tendency to base unit performance assessments on a narrow list of measures, and the diversity of objectives served by urban planning education and scholarship. There are fears that while any performance measurement scheme must practically focus on a limited number of indicators, effectively assessing the quality of teaching and research at planning programs requires a very large number of variables, many of which are hard to measure. At the same time, there is recognition that university administrations and national governments are increasingly demanding unit performance measurements as the basis of resource allocation decisions, and that student and faculty recruitment is,

in part, a function of the publicity generated by unit performance measurement. There is also the expectation that good national comparative data will facilitate better internal policy decisions by faculties.

The Introduction and Use of University Unit Performance Measurement in Industrial Countries

The United Kingdom initiated a national program of unit performance measurement in the mid-1980s. This program is divided into two parallel activities: the Research Assessment Exercise (RAE) and the Teaching Quality Assessments (TQA), the latter of which is now a matter of individual university responsibility. Of central importance is that the RAE leads directly to re-allocations of funding to university programmes by the national government.

The system undergoes revision in each 5-7 year cycle. In the most recent RAE cycle in 2001, units were graded on a 6-point scale (1-5, and 5*) based on papers published, grants awarded, number of staff ("faculty" in U.S. usage) and of research students, the unit's research strategy, and measures of esteem including prizes, research roles, and advisory posts. (Jamrozik, Weller and Heller 2004; Punter 2001). Most important among the measures is the proportion of papers written by staff whose work is judged to be of "international or national quality" by a disciplinary peer panel who read up to four papers by each staff member at each school. Each school is allowed to "select", or prepare its own list of staff deemed "research active" to be included in the assessment, so there is some school discretion in the development of the faculty census. Larger staff is a positive in the ratings, but larger selection may lead to lower proportion of international or national quality papers.

The UK RAE is heavily dependent on direct peer review, leads to a single overall rating for each school, and directly affects unit funding. Proposals for changes to be implemented in 2008 call for greater reliance on independently-determined measures of performance, such as bibliometric measures, but the most important components of the ratings will still be based on direct peer review. The single overall scores assigned to each unit and the direct affect on funding will be retained. (Punter 2003).

In order to retain research funding, a unit must obtain a grade of at least 3a, and units earning 4 (virtually all work at national standard, with ten percent international), 5 (10-49 percent international) or 5* (majority of papers international) ratings enjoy increases in research funding, with some differences in proportional allocations as determined by national bodies in England, Scotland and Wales. Nationally in 2001, 80 percent of staff were in units earning ratings of 4 or higher. (Punter 2003).

In 2001, the Town and Country Planning peer panel included 13 people: 9 senior academics and four senior professionals. 28 British planning schools participated, submitting 1,440 pieces of research. The panel defined research quality as exhibiting, "substantive research content...quality of argument, and...contribution to the advance of theory and/or methodological development/policy development/good practice" (Punter 2003, 8-9).

Two schools (7%) earned grade "5*", 6 (21%) earned grade "5", 7 (25%) earned "4", 8 (29%) earned grade "3a", 3 (11%) earned grade "3b", 1 (4%) earned grade "2", and 1 (4%) earned grade "1" (Punter 2003, 29). So, 82 percent of the schools qualified for research funding, and 18 percent did not, but a subsequent decision by the English authorities led to reductions of funding for grade 3a and 4 schools. 11 of the planning schools improved their rating over the prior 1996 exercise. Planning had lower percentages of staff in the higher grade schools, 69 percent in schools with 4, 5 or 5* grades, compared with 80 percent across all fields nationally, and as a result, planning schools have suffered relative to those in other disciplines.

The RAE system is clearly intended to promote research accomplishment by concentrating university resources on those with the best research track records, and it has increased attention to research and publication in British universities. It has been the subject of considerable criticism, however, with the responsible national body concluding, "the amount of discrimination provided by the exercise is less than the length of the rating scale would suggest" (U.K. Higher Education Funding Councils 2003, 57). There is belief by some that faculty and schools now concentrate on grant getting and publishing to the detriment of teaching and professional relevance (Jamrozik, Weller and Heller 2004), and there is the fear that faculty cooperation and collegiality may have suffered (Thomas 2005).

The UK RAE has stimulated somewhat similar exercises in various countries including Australia, Canada, Belgium, Hong Kong, Ireland, New Zealand, the Netherlands, Poland, Slovakia, and Taiwan. Indeed, an OECD report characterizes evaluation of research as a "rapid growth industry" (Organization for Economic Cooperation and Development 1997). von Tunzelmann and Mbula (2003) find that Ireland and New Zealand are more oriented toward unit formulation of strategy, rather than summative evaluation of prior performance; that Taiwan is adopting the RAE model in order to promote publication by academics; that Belgium is adopting a system relying heavily on bibliometric measures; and that Poland, Slovakia, Hong Kong and Australia all use national research evaluations to distribute funding to institutions, while Ireland, France, Switzerland, Denmark, Japan and New Zealand rely more on self evaluation by units, within proscribed national structures and with oversight and concerns about "puffery". It appears that interest in intense models of national evaluation requiring direct review and assessment of scholarship by peer panels are limited to smaller countries and may be difficult to implement in larger ones (von Tunzelmann and Mbula 2003).

Netherlands undertook an RAE-like assessment of geography, planning, demography and cartography with results released in 2001. According to Voogd (2001), the exercise was intended to maintain and improve quality through feedback, rather than through funding reallocations, although it paralleled the UK system in producing a single measure of performance for each unit. Peer panellists in the Netherlands were required to read five research outputs per unit, selected by the unit director. There was some concern over selection of panellists exacerbated by the small size of the country, and as a result a high proportion of foreign panellists were appointed. The Dutch system was intended to base evaluations on the mission statements of individual research units, but according to Voogd (2001) there is little evidence that this took place. He is also critical of the direct peer review scheme, arguing that such reviews are not better than those of the best peer reviewed journals, only different.

New Zealand's Labour party promised greater accountability for research funding in 1999. The government's Performance-based Research Fund (PBRF) was designed to reward research excellence. Half of the nation's higher education institutions elected to participate in the initial round of Quality Evaluation (QE), some 22 institutions,

including 8 universities and 2 polytechnics. In all, 310 academic units were assessed. The QE process, modelled after the UK RAE, involved two phases: evidence portfolios prepared by staff and assessment by 12 peer review panels consisting of 165 leading researchers, 33 from abroad. Three principle sets of measures were used: number of research degree completions, external research income, and expert panel review of research quality. Results were complimentary of much about research in the nation, finding strengths broadly spread across many institutions. 5.7% of staff received the highest possible rating ("A"), while 39.9% were rated below the "C" level. Questions were raised about the extent of post-graduate education in institutions that did not perform well. In response to the study results, the government decided to increase research funding by NZ\$33 million over the subsequent four years. (New Zealand, Tertiary Education Commission 2004).

Eight New Zealand universities participated in the review of units in the subject area of Architecture, Design, Planning and Surveying with 175 eligible staff and 114 staff reviewed. Across these units, 3.1 percent of staff were rated "A", 20.7 percent "B", 39.4 percent "C", and 36.8 percent below C (referred to as "R")(New Zealand, Tertiary Education Commission 2004, 110). Two of the schools earned overall ratings of 6 on the systems' 10-point scale; five earned ratings between 2.3 and 3.5; and three earned ratings of 1.0 or less (New Zealand, Tertiary Education Commissions 2004, 111). Individual staff ratings are not released to protect confidentiality, but unit ratings are published.

Australia has been basing university research funding on measures of performance since the early 1990s, and has recently decided to create a national research assessment system, but implementation has not yet begun (Butler et al. n.d.). The country's Department of Education, Science and Technology convened a conference in June 2004 intended to advance discussion of the form of the system. Sir Gareth Roberts, a key actor in the UK RAE, keynoted that conference attributing increases in research quality and funding in the UK to the RAE. Australia's Chief Scientist warned of strategies that demand too much detail, but expressed the hope that demonstrating the accomplishments of Australian researchers would lead to greater funding for research from government and industry (Batterham 2005). Braithwaite (2005) argued that assessment of the policy sciences must be based on peer reviewed

accomplishments, primarily publication in refereed journals, rather than attempting to gauge impact on policy, which is difficult to fairly assess.

Iain McCalmann (2005), president of the Australian Academy of the Humanities discusses measurement of excellence in the humanities and creative arts. He identifies four broad methods for measuring research excellence in such fields: peer review, self-assessment, historical ratings and quantitative measures (citation indices, research grant income, numbers of post-graduate students, esteem measures including keynote lecture invitations, editorial board memberships and memberships on panels of learned societies and government bodies).

Outside the U.S., governments of industrial nations are increasingly using university unit performance measurement to provide feedback on the quality of research and to inform resource allocation decisions. These measurement exercises are most often based on a very limited set of measures, frequently utilize new judgements of quality made by peer review panels assembled for the purpose, and lead to single measures of unit quality. Preliminary evidence suggests that urban planning schools have not fared particularly well under these programmes, likely as a function of their multi-objective and interdisciplinary nature, which is hard to capture in a system that produce single overall scores of performance, leading Balducci (2005) to call for richer measurement schemes to be developed by planning school associations.

Unit Performance Measurement in the U.S.

In 1995, the (U.S.) National Research Council published results of a wide-ranging study of research-doctorate programs in the United States (Goldberger, Maher and Flattau 1995). The most recent of a series of such studies, it has widely been used as the basis of rankings claims by departments and universities. The NRC study included only disciplines in which there were more than fifty doctoral programs nationally, and as a result Urban Planning was not included.

The 1995 NRC study was the latest of five studies of performance in university departments published by the American Council on Education and the National Academy of Sciences Press beginning with the Cartter Report in 1966, and tracing routes to earlier national assessments that go back to 1927 (Ostriker and Koh 2003; Webster 1988). 29 variables were analyzed pertaining to 3634 academic programs in

41 disciplines at 274 universities at a cost of over one million dollars (Hargens 1996, 730). In contrast to university performance measurement programmes in other countries, the U.S. system is operated by an association of scholarly societies, independent of the government.

The variables used by Goldberger, Maher and Flattau (1995) were in three groups: (1) Reputation; (2) Students; and (3) Faculty. The **Reputation** variables were based on a survey of graduate faculty conducted by the NRC. The **Student** variables concerned graduate student headcounts and demographics, student support and time span of study. The **Faculty** variables differed somewhat for major disciplinary groupings, with all programs assessed on faculty size, seniority, and research support; and then Social and Behavioural Science programs assessed on publications and citations, while Arts and Humanities programs were assessed using honours and awards. This was the first of the NRC studies to report citation data made feasible by computerization of the *Current Contents* information on citations in journal publications (Hargens 1996, 732).

The results of the 1995 study have been intensely scrutinized, with many universities using the performance of their units in that study as the basis for internal decisions about resource allocations. In some institutions decisions about strategic investments have been tied to the potential to increase placement in the next NRC study, with the result that those disciplines not represented in the NRC study, including Urban Planning, have been ineligible for such investment.

A great deal of re-analysis has been done using the NRC data, including many studies that show reputational rank correlates with objective measures (e.g. Ehrenberg and Hurst 1998; Toutkoushian, Dundar and Becker 1998); studies that demonstrate concentration of publishing among small numbers of departments (e.g. Hodgson and Rothman 1999); and criticisms of behaviours directed to padding numbers of publications and citations without real intellectual merit (e.g. Berry 2000; Brunn 1996).

Among the NRC's own assessments of its work, the most comprehensive analysis is reported in Ostriker and Kuh (2003) who praise the effects of the 1995 study in terms of wide acceptance, comprehensiveness, transparency, and temporal continuity. They

go on, however, to find fault with the study's emphasis on exact numerical rankings, confounding of research reputation with educational quality, emphasis on reputational measures of scholarly quality, and inadequate review of data accuracy by schools. They also criticize the difficulty students face in accessing the data, the length of the ten-year interval between studies, and the groupings used to categorize fields.

Planning School Performance Measurement in the U.S.

Among U.S. planning educators there had been a long-standing reluctance to publication of comparative performance measurements. Results of a national reputational survey included in the first printing of the first edition of the *Guide to Graduate Education in Urban and Regional Planning* (Suskind 1974) were deleted from the second printing, and such a study has never been replicated. In the years since, when the Planning Accreditation Board and the Executive Committee (now Governing Board) of the Association of Collegiate Schools of Planning have considered school rankings, the weight of opinion has always been against undertaking such an endeavor¹.

Other fields closely related to planning have had various studies: *Design Intelligence* (Cramer 2004) has ranked programmes in Architecture, Landscape Architecture and Interior Design since 1998 using employer surveys. Architecture and Public Affairs are ranked bi-annually by *U.S. News and World Report* using surveys of academics². Groop and Schaeztl (1997) assessed geography departments based on teaching productivity indicators, placement of graduates, and publications counts that include books written and books edited. Strathman (1992) ranked 33 U.S. urban studies and urban affairs graduate programs based on a reputational survey and citation data.

Meanwhile the landscape of American higher education has changed. Disciplinary rankings have become widely used in the internal reward structures of universities as well as in the decisions of national bodies about such matters as invitational memberships and peer group identification (Hargens 1996, 730; Webster 1988).

¹ In a colloquy on the PLANET listserve (planet@listserve.buffalo.edu), former PAB chair Linda Dalton (24 April 2003) and former ACSP president Michael Tietz (28 April 2003) each recalled prior decisions against official ranking projects.

² Notably, the Public Affairs ranking of programmes in public policy and public administration (but not city planning) includes separate treatment of a City Management category.

Potential students and faculty often use publicized rankings in making decisions about institutions and about fields of study. Legislators and trustees have become accustomed to assessing accountability in significant part through national comparative studies of performance.

Urban planning programs may be losing visibility and resources because they do not participate in comparative performance measurement. Many students learn about graduate fields through rankings and their subsequent publicity, and many universities now base internal allocations and other decisions on results in national performance comparisons. These concerns were partly responsible for the creation of the A.C.S.P. Institutional Data Project in 1999 (see Rosenbloom 2002). Agreement to go forward on that project, however, required stipulation that data on individual schools would not be made public. So, the IDP allows schools to assess where they place in comparison to national averages, but individual school performance data is not available to persons outside of the institution in question.

In 2004, Stiftel, Rukaman and Alam (2004a) published an application of methods from the NRC study to the 84 U.S. urban and regional planning graduate programmes that were full members of ACSP and/or accredited by the U.S. Planning Accreditation Board. They expressed the hopes of: (1) advancing the debate among planning educators concerning appropriate performance measures; and (2) providing data to faculties concerning the relative performance of their school among planning schools generally. The study was limited to those faculty variables used by the NRC for which national data were readily available: principally faculty size, publication rates and citation rates, with the last two drawn from the *ISI Web of Science* database (Institute for Scientific Information 2003). There was no consideration of reputational data or student data. Nor was there consideration of honours or awards, since no national source existed for these data. The authors expressed the hope that others would undertake subsequent studies using other measures, cautioning that their data presented at best an imperfect partial picture of school performance biased toward social science models of planning scholarship and virtually ignoring design, outreach, and teaching.

Findings of the Stiftel, Rukmana and Alam (2004a) study showed that America's planning schools were most often at public research universities, were typically quite

small (mean faculty size = 10), and with seniority similar to that in other university fields. While scientific models were shown to substantially influence U.S. planning faculty work, only about one-half of faculty published an ISI-indexed article in the five-year study period. About two-thirds were cited during the study period. There was considerable concentration of activity among those who do publish, with fourteen schools and eight-six faculty (out of a total of 844) accounting for half of all publications, and five schools and nineteen faculty accounting for half of all citations. There were substantial differences among accredited and non-accredited schools, doctoral degree-granting schools and master's-only schools, publicly-supported schools and private schools.

Comments came quickly, including praise for the study's uses in promoting discussion of the field's paradigm and standards (Tietz 2004), and for the value offered schools in justifying themselves within universities (Myers 2004). Criticisms focused on the lack of measures of design scholarship, practice, teaching and non-English language materials (Forsyth 2004; Myers 2004; Albrechts 2004), as well as the methods of faculty census and the completeness of the publication outlets considered (Fainstein 2004).

Current U.S. Developments

In response to the discussions surrounding the Stiftel, Rukmana and Alam (2004a) study, ACSP assembled a Planning School Performance Measurement Working Group charged with proposing a programme of school assessment intended for implementation by the Association. Working Group membership consists of:

- Linda Dalton, Executive Vice Provost at California Polytechnic State University, San Luis Obispo;
- Ann Forsyth, Professor of Urban Design at the University of Minnesota, Twin Cities;
- Frederick Steiner, Dean of the School of Architecture at the University of Texas at Austin;
- Bruce Stiftel, Professor of Urban and Regional Planning at Florida State University;
- Dawn Terkla, Executive Director of Institutional Research at Tufts University; and
- Nohad Toulan, Dean Emeritus of the College of Urban and Public Affairs at Portland State University.

The Working Group, while cautious about predicting feasibility, quickly agreed to the desirability of assembling national data on school performance. We see three key reasons why U.S. planning schools would benefit from a national program of performance measurement:

1) *A national system of comparative data on school performance would provide faculty with realistic gauges of the relative quality of our work.*

It is easy to form opinions about how one's work fits into a peer group or a national comparison, but accurate opinions require good information. The same principles that lead us to recommend quality evaluation research for public planning programmes suggest that we should want good evaluation of our own efforts. Such evaluations would require cross-sectional and longitudinal comparisons.

2) *A national system of comparative data on school performance will allow schools to make believable strategic arguments to the administration of their universities.*

In the increasingly competitive resource-allocation decisions on campuses and within university systems, units that have believable comparative data are in stronger positions to make claims. It is inaccurate to think that only a small number of schools would benefit in this way, since a strong system of performance measurement will show many schools to be leaders in various areas and among different comparison groups, and will also show where resource allocations are tied to the quality of outcomes, allowing school administrators to make arguments for the benefits of increased resources. It is perhaps useful to know that 42 schools (50% of the schools studied) appear in the top 10 on at least one of the nine measures examined in the Stiftel, Rukmana and Alam (2004a) study.

3) *A national system of comparative data on school performance would improve the visibility of our profession and lead to stronger recruitment.*

As a small profession with limited public profile, city planning struggles to present itself to potential students and to help those students see potential in planning careers and in the schools that could prepare them for those careers. Especially when national data on performance are widely circulated for the fields of Architecture, Landscape Architecture and Public Affairs, the absence of data on Urban Planning schools reduces our comparative visibility. National school performance data would be promoted by universities and by our profession. They would garner media coverage

and web links, leading to greater profile for our schools and a stronger recruiting position for all planning schools.

While we easily agreed to the desirability of national performance measurement, the Working Group has had considerable challenge to develop operating principles for conducting the work. The most difficult issue we have grappled with concerns the treatment of design work and outreach.

The overlapping nature of planners' views for distinguishing research from practice on the one hand, and social science forms of scholarship from design-based forms of scholarship on the other, proved to be difficult to work through. Certain universities embrace artistic accomplishment as the equivalent of research and readily understand arguments that planners as designers need to engage in practice to vet their accomplishments and to influence the evolution of the art. This view is quite different, however from the ethos of community engagement often expressed as an effort, valued on its own, to bring the results of classroom and faculty work into the real world, and conversely, to bring the real world into teaching and research (Checkoway 1998). Certain universities prize community engagement and want to promote it without specific attention to whether the engaged work is research- or design-based. Measures that conflate design and outreach run the risk of not being persuasive in institutional environments which prize one but not the other.

We benefited from Crewe and Forsyth's (2004; 2003) analyses of scholarship in design in which they distinguish creative work that conforms to the standards, practices and sensibility of research, from creative work that stands above the typical through the production of prototypes, and from design practice that can have high artistic or technical merit. We have also drawn from Clay's (2003) discussion of the need to assess the impact of faculty work in moving a professional field, and from Steiner's (2005) enumeration of the mechanisms of judgment for design accomplishment, including design awards, publication of work, exhibitions, and competitions.

We also face considerable challenge in the need to create a system that will fairly measure all schools' performance despite differences in record keeping and tendencies to engage in gaming behaviour. The ACSP Institutional Data Project experience has

shown that data from national sources is generally more reliable, but a paucity of sources suggests that outreach and teaching data will most often have to come directly from the schools, and that certain design data will have to as well. But, we want to be careful not to burden the schools, especially small schools, with unrealistic data collection requirements, and we must design a system that ACSP can afford to implement.

Finally, we have to choose a method of faculty census that will capture the great preponderance of school accomplishments while responding effectively to differences in faculty appointment arrangements across institutions. The difficulties here include planning programmes that are housed in non-departmentally organized colleges, as well as large multi-disciplinary faculty groups that are affiliated with urban planning PhD programs but have no direct reporting lines within urban planning administrative units.

After considering these challenges, the Working Group agreed to the following six **principles**:

- The Planning School Performance Measurement (PSPM) system would be *broadly based*, intending to capture a wide range of planning school activities, including activities in research, design, outreach, and teaching, and including measures derived from both objective and reputational sources.
- Schools with *PAB-accredited bachelor's and master's degree programs* would be included in the PSPM system; non-accredited ACSP-member schools would not participate;
- Analyses and reports would *identify and report data by bachelor's degree and master's degree groupings, as well as by Carnegie category of institution*. Ph.D. programs would be identified, but would not be used as a separate category for display of results, since there is no clear method of determining what is, or what is not, a doctoral program in our field, and since determination of the faculty census for PhD programmes is especially problematic.
- The census of faculty to be included in analysis of faculty productivity measures would *include those faculty identified by the schools as "50% or greater in planning"* in the most recent submission of faculty lists to ACSP. This census would include faculty of any rank in which the word 'professor' appears in the title, so those with the job titles Assistant Professor of Practice or Research Professor would be included, while Instructor, Assistant Scientist, and Research Associate would not. Additional data counting "less than 50%

in planning" faculty will be collected and reported, again with the caveat that 'professor' appear in the job title, but such faculty will not be included in the main faculty productivity analyses.

- Wherever possible, the PSPM system should *rely on independently collected national sources of data*, in preference over data provided directly by the schools.
- The *number of indicators included in the PSPM system should be kept relatively small*, preferably less than 20.

The Working Group went on to envision **a six-year PSPM cycle, with three studies conducted during each cycle**, a first concerned with reputation; a second concerned with faculty scholarship, design and outreach; and a third concerned with teaching performance. Release of the three studies might be spaced two years apart, so that the reputational study might be released in 2007; the faculty scholarship, design and outreach study might be released in 2009; and the teaching performance study might be released in 2011.

Reputation

The reputational study would consist of results of a survey of faculty.

The population to be sampled for this survey would consist of persons included in the faculty census as outlined above. The sample would be sufficiently large so that, with each respondent rating 40 programmes, there would be 150 ratings requested for each programme. It is expected that a sample of about 315 faculty would facilitate this result. The sample would be constrained to ensure that the overall sample proportions approximately reflect the size of the various school faculties, and that at least one faculty member is chosen from each school.

Each rater will be assigned a randomly pre-selected group of 40 schools to rate, with the constraint that the rater's current school will not appear on the list. No faculty member will be permitted to rate his current school, nor any school affiliated with a university where s/he has previously worked or studied. Questions in the survey form would confirm which universities these are, and the returned questionnaires would be crosschecked to delete any answers reported for inappropriate schools. A list of census faculty for the schools will be included with the survey to facilitate recall by respondents. While we do not wish to over-specify the administration of the survey,

we believe a web-based survey may both lead to a high response rate and be cost efficient.

For each school, the raters will be presented with a 7-point Likert scale for assessing quality of the program's faculty in each of eight areas of study. The eight areas will include Planning Theory, Planning Methods, and six areas of specialization chosen based on frequency of listing in the current round of data collection for the *Guide To Undergraduate and Graduate Education in Urban and Regional Planning*. Results would be reported as interquartile ranges or as medians rather than as means in order to minimize the effects of outliers on results.

Faculty Scholarship, Design and Outreach

This study would be based on national data drawn from the Institute of Scientific Information, the Library of Congress, a group of foundations making fellowship awards, and a group of professional associations making design awards and invitational memberships, as well as from a survey of faculty, and a survey of schools.

A survey of census faculty would be conducted to ensure that national data collection properly distinguishes identities and recognizes former names and affiliations. Each faculty member identified in the most recent ACSP census will be asked their current institution, and rank, as well as the names of all former employers, and any prior names they used professionally.

Then, in a separate survey, each school will be asked to report key outreach, research and design data, for a five-year interval, such as:

- the number of times census faculty testified before local, state, national and international legislative bodies or investigative commissions;
- the number of projects initiated for public- or private-planning clients and the total dollar value of those projects;
- the number of faculty memberships on local, state, national, and international boards and commissions, and
- the number of exhibitions in which faculty work appeared, away from their home campus.

The intent is conduct web-based surveys and to stimulate high response rates through multiple reminders by letter, e-mail and phone.

Finally, name and employment histories provided in the faculty survey will be the basis for remaining data collection from the national sources. While final determination of variables to be included will be the responsibility of the study team, it is expected that variables would include measures like:

- density of ISI-listed publications,
- density of ISI-listed citations,
- new books authored or edited by faculty assigned ISBN numbers by the Library of Congress,
- total number of fellowships to faculty from a fixed list of granting institutions (such as: Fulbright program, Guggenheim Foundation, MacArthur Foundation, U.S. Presidential Fellowships, APA, etc.), and
- total number of national and regional awards to faculty from a fixed list of awarding institutions (such as: AICP Fellows induction, American Planning Association, American Institute of Architects, American Society of Landscape Architects, Environmental Design Research Association, etc.).

Outreach, research and design data will be reported by school (not by individual faculty member).

Teaching Performance

Data collection for the teaching performance study would be conducted in conjunction with data collection for the next edition of the *Guide to Undergraduate and Graduate Education in Urban and Regional Planning*. In addition data on degree completions would be collected from the U.S. Department of Education, National Center for Education Statistics, and data on AICP Exam attempts and pass rate will be collected from the Institute. While final determination of variables to be included would be done by the study team, the expectation is that data will include information on:

- the numbers of students admitted to each degree program in each of several recent years,

- the numbers of students first enrolling in each degree program in each of several recent years,
- the interquartile range of GRE verbal and quantitative scores of entering graduate students,
- the interquartile range of undergraduate Grade Point Averages for entering graduate students,
- the demographic makeup of entering students including race, gender and US v. foreign status,
- the percentage of full-time graduate students appointed to graduate assistantships, to non-service fellowships, and to tuition waivers,
- the number of degrees awarded in each of several recent years at bachelor's, master's, and doctoral levels,
- the number of attempts at the AICP Exam and the resulting pass rate by graduates of the school.

Logistics

The Planning School Performance Measurement system would be conducted under the oversight of an ACSP committee. Staff would be chosen through a national Request for Proposals. Different RFPs would be used for each of the three studies in the cycle, and different staff might well be chosen for each of the three studies envisioned.

Funding would be provided by ACSP sufficient to reimburse materials and student and contract labor, but faculty labor would be expected to be provided without cost to ACSP.

The Future of Planning School Performance Measurement

The internationally-growing practice of assessing unit performance through systems that lead to single overall scores is difficult for planning programmes that are designed to serve a wide range of multiple objectives and that draw from a wide range of disciplinary traditions. Especially when original panel assessments are used, there is the danger that the work of planning academics will not be evaluated by true peers, but rather by scholars who disagree with the approaches undertaken. Moreover,

planning education's need to be connected to practice through outreach and grounded research and teaching is not likely to be respected and rewarded.

To grow and prosper in an era of rising unit performance measurement, planning schools need to develop credible national or regional systems of assessing quality. These systems should utilize wide ranges of measures intended to capture social science modelled research, design work, outreach and teaching. To be credible, they will have to come from learned societies or professional associations, most promisingly, from associations of planning schools.

In nations that do not have a government-sponsored assessment scheme, such school association-based assessments could provide the evidence school administrators need to make credible claims for resources. In nations that base national-level funding decisions on the outcomes of narrower assessments, the more broadly-based assessments may offer units some ability to argue for discretionary funds at the institutional level. In all contexts, they will provide feedback to faculties about the perceived success of their own work, and they will be of recruitment value by increasing publicity for the field.

In the United States, the ACSP now has a proposal before it that intended to serve these goals. The ACSP Working Group on Planning School Performance Measurement has proposed a system of school assessment involving three separate studies, to be carried out on a six-year cycle, which measure school performance in teaching, social science-modelled research, design work, and outreach, utilizing 25 distinct measures, including eight measures of teaching performance, nine measures of scholarship and other creative work, and eight measures of reputation. This system would provide a great deal of useful information to faculties, would enrich student awareness as they choose fields of study and schools, and would allow unit administrators to argue more effectively for resources within the mission contexts of their individual schools.

REFERENCES

Albrechts, Louis. 2004. A challenge for the global planning community? *J Planning Education and Research*. 24(1):26-27.

- Australia, Department of Education Science and Training. 2004. A Measuring excellence in research and research training. @ Proceedings of a conference held at the Shine Dome, Canberra, 22 June. Canberra.
- Balducci, Alessandro. 2005. Collegiality within a mass university system: reflections from Italy. *Planning Theory and Practice*. 5(2):249-251.
- Batterham, Robin. 2005. Measuring excellence: a Chief Scientist perspective. Pp. 3-9 in *Measuring Excellence in Research and Research Training: Proceedings of a conference held at the Shine Dome, Canberra, on 22 June 2004*. ed. National Academies Forum and Australia, Department of Education, Science and Training. Canberra: Australian Academy of Science.
- Berry, Brian J.L. 2000. "Citations as measures of worth." *Urban Geography* 21:283-284.
- Braithwaite, Valerie. 2005. Excellence in social sciences in the context of impact on public policy. Pp. 44-50 in *Measuring Excellence in Research and Research Training: Proceedings of a conference held at the Shine Dome, Canberra, on 22 June 2004*. ed. National Academies Forum and Australia, Department of Education, Science and Training. Canberra: Australian Academy of Science.
- Brunn, Stanley D. 1996. "Quality research performance: issues of splitting, cloning, citing, significance and judging." *Professional Geographer*. 48: 103-105.
- Butler, Linda, Grit Laudel, Claire Donovan, Frank Jackson, David Siddle, and Ian Lucas. n.d. Strategic assessment of research performance indicators: an ARC linkage project. <http://repp.anu.edu/au/Linkage>. (accessed 4 May 2005).
- Checkoway, Barry. 1998. Professionally-related public service as applied scholarship: guidelines for the evaluation of planning faculty. *J Planning Education and Research*. 17(4):35-360.
- Clay, Philip. 2003. Scholarship reconsidered. Plenary address to the ACSP Administrators Conference, Amelia Island FL, 21 October 2003.
- Cramer, James P. 2004. America's Best Architecture and Design Schools. *Design Intelligence*. 10(11):2-39.
- Crewe, Katherine and Ann Forsyth. 2004. Research in design: definitions and limits. Manuscript. Minneapolis: University of Minnesota, Twin Cities.
- _____, and _____. 2003. LandSCAPES: a typology of approaches to landscape architecture. *Landscape J*. 22(1):37-53.
- Ehrenberg, Ronald G. and Peter J. Hurst. 1998. "The 1995 ratings of doctoral programs: a hedonic model." *Economics of Education Review*. 17:137-148.
- Fainstein, Susan. 2004. Letter to the editor. *J Planning Education and Research*. 24(2): .

- Forsyth, Ann. 2004. The view from design. *J Planning Education and Research*. 24(1):24-25.
- Goldberger Marvin L., Brendan A. Maher and Pamela Ebert Flattau, eds. 1995. *Research-doctorate Programs in the United States: Continuity and Change*. Report of the Committee for the Study of Research-doctorate Programs in the United States, sponsored by the Conference Board of Associated Research Councils, and conducted by the Office of Scientific and Engineering Personnel, National Research Council. Washington, DC: National Academy Press.
- Groop, Richard E. and Randall J. Schaetzl. 1997. "Productivity profiles of Ph.D.-granting geography departments in the United States: 1980-1994. *Professional Geographer*. 49:451-464.
- Hargens, Lowell L. 1996. Review of RESEARCH DOCTORATE PROGRAMS IN THE UNITED STATES: CONTINUITY AND CHANGE, ed. M.L. Goldberger, B Maher and P.E. Flattau. *Contemporary Sociology*. 25:729-732.
- Hodgson, G.M. and H. Rothman. 1999. "The editors and authors of economic journals: a case of institutional oligopoly?" *Economic Journal* 109:F165-186.
- Institute for Scientific Information. 2003. *Web of Science*.
<http://www.isi4.isiknowledge.com/portal.cgi/wos> (accessed January through March 2003).
- Jamrozik, Konrad, David P. Weller and Richard F. Heller. 2004. Research assessment: there must be an easier way. *Medical J of Australia*. 180(11):553-554.
- McCalmann, Iain. 2005. Excellence in the humanities, creative arts and media. Pp. 20-25 in *Measuring Excellence in Research and Research Training: Proceedings of a conference held at the Shine Dome, Canberra, on 22 June 2004*. ed. National Academies Forum and Australia, Department of Education, Science and Training. Canberra: Australian Academy of Science.
- Myers, Dowell. 2004. How can planning schools be usefully compared? *J Planning Education and Research*. 24(1):25-26.
- New Zealand, Tertiary Education Commission. 2004. A Performance-based research fund: evaluating research excellence: the 2003 assessment. @ Wellington, April.
- Organization for Economic Cooperation and Development. 1997. *The Evaluation of Scientific Research: Selected Experiences*. OECD/GD(97)194. Paris.
- Ostriker, Jeremiah P. and Charlotte V. Kuh, eds. 2003. *Assessing Research-doctorate Programs: a Methodology Study*. Report of the Committee to Examine the Methodology for the Assessment of Research-doctorate Programs, Policy and Global Affairs Division, National Research Council of the National Academies. Washington DC: The National Academies Press.

- Punter, John. 2003. "The UK research assessment exercise: the evaluation of planning research and its implications." Paper prepared for presentation at the Joint Congress of the Association of Collegiate Schools of Planning and the Association of European Schools of Planning, Leuven, Belgium, 8-11 July 2003. Cardiff, Wales CF10 3WA UK: Cardiff University. punterj@cardiff.
- _____. 2002a. "Reflections on the 2001 United Kingdom research assessment exercise for town and country planning." *Town Planning Review*. 73:333-349.
- Rosenbloom, Sandra, ed. 2002. "Report of the ACSP Institutional Data Project." Association of Collegiate Schools of Planning.
- Steiner, Frederick. 2005. Assessing design productivity. Presentation made to the ACSP Administrators Conference, Cincinnati OH, 2 April 2005.
- Stiftel, Bruce, Deden Rukmana and Bhuiyan Alam. 2004a. Faculty quality at U.S. graduate planning schools: a National Research Council-style study. *J Planning Education and Research*. 24(1):6-22.
- _____, _____, and _____. A reply to graduate planning school study responses. *J Planning Education and Research*. 24(2):
- Strathman, James G. 1992a. "A ranking of U.S. graduate programs in urban studies and urban affairs." *Journal of Urban Affairs*. 14:79-92.
- Susskind, Lawrence. 1974. *Guide to Graduate Education in Urban and Regional Planning*. 1st ed. Cambridge, MA: Association of Collegiate Schools of Planning.
- Thomas, Huw. 2005. Pressures, purpose and collegiality in UK planning education. *Planning Theory and Practice*. 5(2):238-247.
- Tietz, Michael B. 2004. Assessing the quality of U.S. planning schools: a comment. *J Planning Education and Research*. 24(1):23-24.
- Toutkoushian, Robert K., Halil Dunder and William E. Becker. 1998. "The National Research Council graduate program ratings: what are they measuring?" *The Review of Higher Education*. 21:427-443.
- von Tunzelmann, N. and E. Kraemer Mbula. 2003. Changes in research assessment practices in other countries since 1999. Working paper. University of Sussex, Science and Technology Policy Research.
- United Kingdom, Higher Education Funding Councils. 2003. *Review of Research Assessment*. Report by Sir Gareth Roberts to the UK Funding Bodies. Bristol. www.hefce.ac.uk.
- _____. 1999. *Research Assessment Exercise 2001: Assessment Panels Criteria and Working Methods*. Bristol.

Voogd, Henk. 2001. Comment. *Town Planning Review*. 73(3):350-352.

Webster, David S. 1988. "Reputational rankings of colleges, universities and individual disciplines and fields of study, from their beginnings to the present." Pp. 234-303 in *Higher Education: Handbook of Theory and Research*, vol. 8, ed. J. Smart. New York: Agathon.