

UC Berkeley

Working Papers

Title

Academic reviews and the culture of excellence

Permalink

<https://escholarship.org/uc/item/0jt1s2jp>

Author

Trow, Martin

Publication Date

1994

A1458
no. 94-16

✓ 1/10/95

Academic Reviews and The Culture of
Excellence

Martin Trow
Graduate School of Public Policy
University of California, Berkeley

Working Paper 94-16

INSTITUTE OF GOVERNMENTAL
STUDIES LIBRARY

DEC 20 1994

UNIVERSITY OF CALIFORNIA



UNIVERSITY OF CALIFORNIA AT BERKELEY

3

Academic Reviews and The Culture of
Excellence

Martin/Trow
Graduate School of Public Policy
University of California, Berkeley

Working Paper 94-16

Working Papers published by the Institute of Governmental Studies provide quick dissemination of draft reports and papers, preliminary analysis, and papers with a limited audience. The objective is to assist authors in refining their ideas by circulating research results and to stimulate discussion about public policy. Working Papers are reproduced unedited directly from the author's pages.

Academic Reviews and the Culture of Excellence*

Martin Trow
Graduate School of Public Policy
University of California, Berkeley

The assessment and evaluation of academic programs in colleges and universities has become a major industry in Europe and the United States. The control and maintenance of academic "quality" engenders a steady flow of conferences, books and articles, and governmental policies and programs.¹ The evaluation of university departments, their research, programs of instruction, courses and whole institutions² is the main business of a growing number of academics, university administrators, civil servants and external consultants. Two questions about this new industry come to mind:

How can we account for the current, relatively sudden upsurge of interest in the subject of "quality assurance" in higher education?³ And how

* Revised and expanded version of a paper read at a seminar in the Office of the the Chancellor of the Swedish Universities, Stockholm, March 1994. My thanks to Michael Burrage, Oliver Fulton and Herbert Kells for their comments on an earlier draft.

¹ The literature is very large, and growing. But as illustrations of quite different kinds of contributions to the discussion, see Roger Ellis, ed., Quality Assurance for University Teaching, London, The Society for Research into Higher Education, 1993; P.T. Knight, ed., The Audit and Assessment of Teaching Quality, Birmingham, The Standing Conference on Educational Development, 1993; and Research Funding and Quality Assurance, Stockholm, The Bank of Sweden Tercentenary Foundation, 1993.

² There are quite fundamental differences in the problems associated with the assessment of "teaching" and "research," problems made more difficult since the two activities are often so closely intertwined in research universities. But the discussions about assessing research and teaching, and the actual procedures put in place to do so, are commonly separate and distinct. Without accepting the legitimacy of that distinction, this paper discusses "academic reviews" without distinguishing between those addressing teaching and those addressing research.

³ In a recent essay reviewing the state of 'quality assurance' in Western Europe, Diana Green observes that "...there is no universally agreed definition of what is encompassed by 'quality assurance.' It is particularly difficult to determine whether, and in what ways, 'quality assurance' differs from 'quality

did the great universities of the world come to be so productive of research and scholarship, of trained and educated people, and of wide services to their societies over the past century and a half without much or any formal external assessment of their quality?

An answer to both questions might be found in this direction: The growth of concern about the quality of work in colleges and universities, and the need for stronger assessments of that quality from outside the institutions themselves, arise out of fundamental changes in the system of higher education as it moves in all countries toward the provision of broader access. The growth of mass higher education, and its increasingly visible differences from the traditional "elite" forms of higher education,⁴ raise questions in government, industry and in the universities themselves about the adequacy of traditional forms of quality control. Questions are asked (and in some countries firmly answered by government) about the efficacy of applying the traditional forms to the new mass institutions, or to the mass education elements of elite institutions, or to the elite institutions themselves as they become more like mass institutions.⁵

To a high degree, the control over their own performance by the elite universities in most Western societies lay in the trust those societies have placed in the academics, trust in their competence, in their intrinsic motivations to maintain the quality of their work and its products, and in the institutionalized arrangements the academy and its disciplines have created over the years for the control and maintenance of quality. Before we dismiss them, let us look briefly at some of those arrangements and quality control mechanisms created by the academics themselves. Not all of these are found in any given country or university. But they all serve similar functions: to maintain, and to demonstrate a concern for, the quality of teaching and/or research in the university by the academic community itself. Among these are the following:

control', 'quality assessment', 'quality management', 'quality improvement' and the term used more frequently in the USA, 'accreditation'. The concepts are frequently used interchangeably." (Diane Green, "Quality Assurance in Western Europe," Quality Assurance in Education, Vol. 1 No. 3 1993, pp. 4-14.) To avoid this confusion of nomenclature, and the more serious issue of assuming the function in the name, I refer to all these activities generically as "reviews," and then try to distinguish them by their nature and function.

⁴ See Trow, "Problems in the Transition from Elite to Mass Higher Education," in Policies for Higher Education, from the General Report on the Conference on Future Structures of Post-Secondary Education, 55-101. Paris: Organization for Economic Co-operation and Development, 1974.

⁵ Whatever the sources of the growth of the "quality assurance" industry, "in all countries, accountability is the keynote of government/HE relations." Green, op. cit., p.6.

the quality of university teachers

1. The single most important force for the maintenance of high quality in academic work lies in the close scrutiny and competitive review of candidates for appointment to university. This competitive review before appointment may be carried out by a department, a dean, a rector or president, an academic senate, or by some combination of those actors. It may involve ad hoc committees wholly composed of qualified academics from within the university, or committees which include members from other institutions. The quality of an academic community is almost wholly determined by two factors: the attractiveness of the institution to the ablest scholars and scientists, and the care and rigor with which it recruits and appoints people.

2. Many institutions appoint new staff to a probationary period before gaining a permanent appointment -- a period of anywhere from one year to six (as in the US) before the review leading to a permanent appointment ("tenure"). This review can be a significant element in the maintenance of the quality of a department or university only if it is not pro forma, and can lead to the dismissal of a candidate. This period, if extended, may include interim reviews, like the "progress" review of the candidate's performance within the department after three years found in many American institutions.

3. Promotions and other rewards in elite colleges and universities are in principle, and often also in fact, based on meritocratic assessments of an academic's achievement: among these are permanent appointment; promotion in rank, especially to the higher ranks; salary levels linked to rank or achievement, and the honors of all kinds awarded by the academic community. These are all motivations for high academic performance.

the quality of students

4. Selective admissions, and requirements for admission of all kinds maintain the academic quality of students, and directly affect the nature and level of instruction. This selectivity in admissions is also an important element in the attractiveness of the institution to scholars and scientists, and thus affects the quality of academic performance indirectly.

5. Examinations of all kinds serve to motivate students, and to monitor their performance. In some countries a special concern for examinations leads to their administration by academics who have not been the student's teachers. In other places objectivity is preserved by replacing the name by a code number on examination scripts. In addition, the quality and

objectivity of examinations and the standard and integrity of a course of study are sometimes reinforced by involving academics from other universities, as in the use of external examiners for all honors examinations in British universities. Another illustration of this is the presence of academics from other departments and other universities at doctoral examinations in almost all countries and universities.

6. The value of examinations as monitoring instruments is sustained in some institutions and systems by the recurrent analysis of patterns of grades awarded, with special concern for variations in the awarding of grades (including the class of honors) between departments and universities, and of patterns of grade inflation over time.

7. The awarding of fellowships, scholarships and other honors on the basis of academic performance rewards student performance and sustains motivation for work of high quality.

the quality of research and scholarship

8. Research grants are commonly made on the basis of competitive excellence by peer review.⁶

9. Reference is commonly made to citation indices as indicators of the quality and significance of published research, of the academics who publish it, and of their departments and institutions. These indices can serve as symptoms of a decline in performance, as well as of distinction.

10. Publication in refereed journals and by legitimate publishers of academic and scholarly books is also an indicator of scholarly quality.

11. Reviews of books in respected academic journals are also used as indicators of their scholarly quality.

12. A powerful indicator of the scholarly reputation of a department, closely watched by many leading American research universities, is the "take-rate" -- the acceptance of first offers to candidates for appointment, both as academic staff and as students. A decline in these acceptances is taken as an indicator of declining quality, calling for remedial measures.

⁶ For a description of current peer review procedures (and changes in those procedures) in a major federal research agency in the USA, the National Institutes of Health, see Eliot Marshall, "NIH Tunes Up Peer Review," Science, vol. 263, 4 March 1994, pp. 1212-1213.

13. The quality of departments in research universities is monitored through the use of periodic external competitive assessments by other academics. This is now common throughout Western Europe, but the best example is still the periodic assessments of the quality of a wide range of graduate departments in leading research universities by the National Academy of Science in the U.S. These assessments, carried out every ten years, are diagnostic rather than budgeting tools, and are not keyed to funding. Even more important than the level of quality of a department is its direction of movement, which can be seen by looking at the trends in its ratings from one assessment to the next.

14. Universities in many countries have developed a variety of internal reviews of departments, clusters of departments, or whole institutions. Some of these reviews are regular and periodic, others are ad hoc. These are ordinarily designed to help the units they review, to be supportive and not directly determine their funding, though they may influence administrative decisions.

the quality of the curriculum, courses, and instruction

15. Almost everywhere, what is required to earn a degree in a given institution is periodically reviewed by academic departments, faculties, colleges and universities, and sometimes by outside bodies (ministries, professional bodies) depending on the nature of the degree, the character of proposed changes, and where authority over such changes lies. Universities in different countries differ in these respects, but regardless of where formal authority for such changes lies, the academics who have the special knowledge in given areas and who actually teach to the degree, play a major role in the determination of its requirements. The key point here is that the nature and variety of work required of students (and indirectly, of their teachers) is subject to periodic critical review and discussion within the academic community.

16. Specific courses within a curriculum are ordinarily in the discretion of smaller units than is the curriculum: a college, or a department, or a single instructor. Courses are added to or subtracted from a curriculum in response to changes in the map of knowledge, or student demand, or to changes in the interests of teachers. But ordinarily, the decision will be made by some body which represents the department or the faculty or university, not on the whim of an instructor. And the discussions and criteria entering such a decision are among the ways through which an institution seeks to maintain or enhance the quality of its work. Any given decision may not contribute to the "quality" of education offered by the institution, any more

than may any specific new appointment to the staff. But the internal discussions and decision about a course are ordinarily rooted in a wish to maintain or improve the quality of the education offered by the institution, however "quality" may be defined in a given case.

17. Student assessments of teaching are now almost universally used as a way of monitoring or assessing the quality of instruction in universities. It is debatable as to how well students can judge the quality of the teaching to which they are currently being exposed, but there can be little doubt that student assessments can identify egregious ineptitude, or the violation or evasion of the common norms of academic life: that instructors be competent in their subjects, meet their classes on time, treat their students fairly and with respect, and not exploit their vulnerability through political or ideological indoctrination. If students cannot assess the quality of teaching or know its long-term effects on them, they can certainly identify the phenomena of "non-teaching" by teachers, and bring those incidents to the attention of others who may be able to take action. And that is certainly part of an institution's own process of quality control and maintenance.

18. Most universities now have units and programs of staff development. Under various names, these units employ groups of specialists whose primary skill and task is to help teachers improve the quality of their work in the classroom, and to help them better organize their course work.

the coordination and monitoring of the mechanisms of quality control

19. Many of the activities described above are largely in the hands of the academic staff, whether in their departments or in broader faculty or university wide committees. But increasingly, in many countries, more power and authority is flowing to the senior academic administrative officer, the rector or vice-chancellor, who thereby become more like the presidents of American colleges and universities who have had this authority since their founding. These senior administrative officers and their staffs have a wide variety of functions, but among them is the task of monitoring and overseeing the various quality control activities sketched above. Unlike deans and heads of department, these administrative officers have responsibility for the quality (and financial soundness) of the whole institution, not just a part of it. And if, as is usually the case, they have an interest in its quality and reputation, their efforts toward strengthening and monitoring the quality control mechanisms already in place do much to keep those mechanisms alive and functioning. Indeed, to exercise academic leadership is in large part to ensure that these mechanisms are functioning as they ought, and are not captured by the communities that they nominally regulate.

The Culture of Excellence, Trust, and Quality Assessment

All of these mechanisms of internal quality control -- of staff, students, research, curriculum and courses, coordination and monitoring -- taken together comprise (and are reflections of) a culture of excellence, a set of values and attitudes shared in varying degrees by the members of an institution which reflect their commitment to the maintenance and enhancement of the quality of their academic work, and of the quality (and reputation) of the institution as a whole. I mentioned earlier that not all these mechanisms can be found in all research universities, sometimes by the accident of history, sometimes due to the weakness of the norms themselves. The effectiveness of the quality control mechanisms in any particular institution does not depend on the number of these mechanisms present, but on the norms and values of the academic community, on the strength of the culture of excellence that underlies and animates those that are in place. And the strength of that culture is currently quite variable among institutions that call themselves universities. Nor are the mechanisms themselves of equal weight in the maintenance of quality; a university can be of high quality though committed to broad access. But it is difficult to imagine an institution of high quality that does not attend carefully to the appointment and promotion of academic staff, or whose academic leadership does not monitor the effectiveness of the mechanisms for quality control that are in place.

There are surely other mechanisms of quality control that I have not mentioned, but these are the most common, and most (but not all) are present in the leading research universities of the world. But in many countries, and most dramatically in recent years in the United Kingdom, the development of new forms of external evaluative assessment by agencies of government arises out of the growing mistrust by those governments of the intrinsic motivations of academics to maintain the quality of their professional and scholarly work, and in the effectiveness of their quality control mechanisms. What has not been demonstrated -- or even asserted -- is a decline in the quality of academic work. In the UK, the rapid transformation of the university system over the past decade and a half, ironically driven by central government policy, has led to the withdrawal of trust in that system by central government, and the creation of a parallel machinery of quality assessment managed by agencies of government, reporting to those agencies, and directly keyed to the funding of the departments and institutions being assessed.⁷

⁷ The special development of central government policy in higher education is discussed in my paper "Managerialism and the Academic Profession: The Case of England," in Thorsten Nybom, ed., Studies of Higher Education and

However many and however effective are these mechanisms of quality maintenance in universities, they require a large measure of trust on the part of the supporting society. These arrangements and mechanisms are largely invisible to outsiders. Even when the rules and procedures are public knowledge, the operation of those procedures often rest on professional or expert judgements which by their nature are necessarily arcane or obscure. For example, however clear are the procedures for the appointment and promotion of academic staff, the decisions finally rest on judgements of scholarly or scientific achievement and potential which can only be made by the professional community, and are always subject to doubts about the intervention of "particularist" criteria arising out of biases of the electors -- scholarly, personal, ideological -- or even of racial and gender prejudice. A good deal of time and energy is spent in the best institutions in trying to insulate the crucial acts of judgement -- with respect both to staff and students -- from these biases, which cumulatively must affect the academic quality of teaching and research. But not all universities, not all departments, not all academics maintain the highest standards in their performance, or in the way they operate the mechanisms of quality control. And that fact partly accounts for the sheer number and variety of these mechanisms. But it also helps explain how the rise of mass higher education, and the growth of institutions in which these internal mechanisms of quality control are weaker, has weakened the trust that has traditionally been accorded the leading universities in almost every country, and has underpinned their autonomy and freedom from external evaluation and management.

The tension between university quality and external trust has a number of sources. First, the monopoly of knowledge and expertise in universities leaves their "clients" -- their students, and ultimately the larger society which supports the universities -- vulnerable to exploitation. "Professional privileges -- their autonomy and monopoly over services -- are granted by the society in exchange for the guarantee that practitioners will not use these privileges to exploit clients."⁸ In this respect academics are similar to other professions, for whom the ordinary controls of the market -- caveat emptor-- are defeated by the imbalance of knowledge between professional and client. This vulnerability of clients is sharpened by the evidence that professional groups (including academics) do in fact sometimes defend their own interests at the expense of their clients; for example, even in leading research universities some members of the academic staff enjoy a degree of

Research, The Council for Studies of Higher Education, Stockholm, 1993:94. pp. 2-23.

⁸ R.K. Merton, Social Research and the Practicing Professions, Cambridge, Mass.: Abt Books, 1982, p. 121. See also the essays in B.R. Clark, ed., The Academic Profession, Berkeley, The University of California Press, 1987.

control over their own time and effort that is necessary and justified by the nature of scholarly and scientific work even if they do very little of it. Similarly, relatively weak universities claim the autonomy and levels of support that are granted the leading universities.

In addition, the products of universities are so diverse, and so hard to measure, that they defeat ordinary efforts to link support to external measures of success and "output." This makes the assessment of higher education's efficiency, its "value for money," a slogan rather than a reality, the political rhetoric of governments which want to be able to evaluate the effectiveness and efficiency of universities in the same way that investors assess the efficiency and success of private business firms.

The lack of trust in the internal mechanisms of quality control listed above is found also in most of the burgeoning literature on "quality assessment," perhaps not surprising since that literature arises out that absence of trust, and offers alternatives to it. There is an assumption in much of the "quality assessment" literature that traditional (i.e., "elitist") conceptions of quality control do not embody any specific agencies of quality assurance. One student of the subject puts the matter this way:

"The traditional concept [of quality] is associated with the notion of distinctiveness, of something special or 'high class'. This approach to quality underpins the elitist view of the high quality of an Oxbridge education in the United Kingdom, or in the Grandes Ecoles in France. This is not quality to be judged against a set of criteria. It is exclusive, and apodictic (one instinctively knows it). The traditional view in education is that universities embody quality. They do not need to demonstrate it. This traditional notion of quality is self-evident. There are no agencies external to the university, or agents within the university, specifically charged with the task of quality assurance." (Green, op. cit., p. 8, emphasis supplied.)

The author completely ignores the variety of quality control mechanisms present in research universities sketched in above, and apparently has little direct familiarity with how leading research university become and remain "leading research universities." But the greater part of the literature on "quality assurance" and "quality control" is contributed by academics from institutions of mass higher education where those mechanisms are weaker or wholly absent, and whose recommendations are designed to supply from outside the institutions a system of reviews and regulations for quality control as a substitute for the weak culture of quality control inside the institutions. In the UK this same machinery is now applied to research universities, where it introduces constraints and

incentives irrelevant to and at odds with the work of those universities. And tendencies in this direction can be found elsewhere.

Why this has happened this way in the UK is a difficult question to answer. One explanation which we have heard from a number of sources, both in government and in academia, is that the merger of the polytechnics with the universities in the UK has created a large number of "universities" whose quality as *universities* cannot be assumed, but must be imposed from outside. This explanation would suggest that the high measure of trust (and autonomy) granted to the British universities before 1979 was placed in elite institutions whose members' motivations could be assumed to be intrinsic and professional; those assumptions cannot be made with respect to at least some of the "new" universities which were recently polytechnics. Their academic staff do not all hold higher degrees; many are part-timers and are simply not fully socialized to traditional academic values relating to the pursuit and maintenance of quality.

Indeed, if one considers the inventory of quality control mechanisms that we find in most elite universities listed above, many of them are not present in institutions of mass higher education marked by weak academic communities. Many of those institutions do little research, and thus do not possess all the mechanisms for quality control associated with research and publication. In addition, many have modest admissions standards. And finally, many institutions of mass higher education advance staff more on the basis of seniority than merit. Even where some of these elements of quality control are present (as, for example, student selectivity in British polytechnics), many others are not, and the whole cluster is weaker. In the UK the most intrusive forms of external assessment and evaluation have been introduced since the merger of the polytechnics with the universities was achieved, or was on the political agenda. And the argument is made that the new forms of assessment, especially of "teaching," are imposed on the old universities because they are really needed for the new universities. The parallel justification for the external research assessments is that this is really a device to create or preserve a differentiated university system, a way of preserving the differential funding necessary to support research universities as opposed to the primarily teaching institutions (i.e., the old polytechnics), and beyond that, of institutionalizing a hierarchy of institutions within the old set of universities. And the leading British research universities have accepted the external research assessments linked to funding with little complaint, presumably in the (perhaps unwarranted) belief that this would preserve their incremental funding and institutional status.

Those are the sub rosa "official" explanations which can hardly be given wide circulation since they contradict the official doctrine that all universities are equal, or at least are assessed by common standards, and are

on the way to becoming more equal. But there are other ways of explaining these developments in the UK which are also consistent with the historical evidence. One is that the very deep cuts in funding for the universities over the decade of the '80s created anxieties in government that they would lead to real declines in the quality of research and instruction, and the enormous concern –indeed the apparent obsession—with issues of "quality" in the White Paper of 1991⁹ arose out of fear that the cuts had had, or would have, the negative effects on universities that many academics were claiming. From this perspective, the new machinery of external assessment was created to provide evidence that would persuade both Government policy-makers and the academic community itself that the cuts had not had these effects, that there was an enormous amount of fat in the old universities, and that they could achieve prodigies of "productivity gains" without loss to students or knowledge.

In addition to these possible motivations is the evidence that these devices for "quality control" also greatly strengthen the power of central government over the universities, institutions which leading politicians in Thatcher governments believed badly needed fundamental change and reform if they were to contribute properly to national wealth and strength.¹⁰ The merger allowed the application to the old elite universities of mechanisms of external assessment and control that were more familiar to the institutions of mass higher education, the former polytechnics. From this perspective, the merger was enormously convenient to a government committed to reducing the autonomy of the elite sector. So the evaluative assessments, and most especially the criteria against which the assessments were to be made, were also a way of shaping and steering the universities at the same time as they served the reassuring role of appearing to defend a central value of British higher education, the maintenance of high and common standards across the whole range of universities.¹¹

I opened with two questions: first, why the increase of interest in evaluation and quality control by governments, especially over the past five years? And second, how did the great universities of the world come to be so great and productive over the past 150 years without any external evaluations? I have answered the second question by reference to the emergence in these universities of a culture of excellence, reflected in and sustained by a variety of institutional mechanisms for quality assessment and

⁹ See my "Thoughts on the White Paper of 1991," Higher Education Quarterly, 46, No. 3 (Summer 1992), pp. 213-226.

¹⁰ See Trow, "Managerialism and the Academic Profession, op. cit.

¹¹ See my "Academic Standards and Mass Higher Education," Higher Education Quarterly, 41, No. 3 (Summer 1987), pp. 268-292.

control.¹² I approached the first question chiefly by reflecting on the recent history of British higher education. But the currents we see there in exaggerated form are also present in other countries. Everywhere that mass higher education grows, it exerts pressure on the values and resources of the elite universities, and this threatens their own internal forms of quality control.¹³ And that in turn weakens the confidence of governments and other institutions in the universities' capacity to sustain their own quality of work. We see this also in some parts of the United States. But we cannot assume that developments in other places will mirror those in the US or UK; conditions differ, as do different societies' responses to similar trends.

The policies and politicians currently shaping British higher education do not put great weight on the historical evidence that the greatest research universities in the world developed without much external pressure or evaluation; those conditions, they might suggest, hardly exist any more. Nevertheless, everywhere the leading research universities -- in Japan and the United States, in Brazil and Canada and Australia, in the UK and Western Europe -- are still the freest of governmental control and interference. But while many distinguished elite institutions around the world defend their autonomy successfully, it is a posture of defense; in many places the autonomy of universities is under pressure, and in some places under attack.

A Typology of Academic Reviews

Since external evaluations are a central instrument for increasing governmental influence over universities, let us look at a range of reviews of academic programs, and see how these external evaluations differ from other forms of academic assessments and reviews. Distinguishing and combining two dimensions of such reviews may help our discussion. One dimension distinguishes reviews which are initiated within a college or university vs. those initiated by outside authorities. The other dimension distinguishes reviews which are supportive of the unit under review as contrasted with reviews which are evaluative. If we look at these two dimensions together,

¹² The development of the assumptions and norms associated with competitive excellence varied greatly among the great universities of the West; it is at the heart of their histories. All universities continue to change along this dimension -- and not only in one direction!

¹³ Mass higher education introduces other bases of legitimacy of higher education, most notably egalitarianism and social mobility through broader access. These values are in some respects incompatible with traditional conceptions of "academic quality," which accounts for some of the confusion and controversy in the literature of quality assessment and evaluation.

they generate four familiar kinds of review: Internal supportive (Type I); Internal evaluative (Type II); External supportive (Type III), and External evaluative (Type IV).

A Typology of Academic Reviews

<u>Origin of the Review</u>	<u>Function of the Review</u>	
	Supportive	Evaluative
Internal	I	II
External	III	IV

Let us consider each of these types of academic reviews, their characteristic strengths and limitations. It is important to stress here that in all cases these are reviews of academic units whose quality is achieved and maintained (if at all) through the quality control mechanisms discussed above. The enormous stress currently being placed on these reviews, especially when they are called "evaluations," should not obscure the fact that reviews are no substitute for the institution's own quality control mechanisms and procedures. At their best reviews are only able to encourage and then monitor the effectiveness of those procedures, and provide clues as to how they might be made more effective. At their worst they hinder the emergence of a culture of excellence within a university or weaken it where it exists, and encourage a variety of patterns of behavior at odds with it.

Some Characteristics and Functions of the Types of Academic Review

Type I: Internal supportive reviews

Examples: Various kinds of reviews of academic units carried out by many colleges and universities in modern societies.

The first, and in some places the most familiar, form of academic review is that carried out within a university by its own staff in support of the work of the unit under review. These reviews may be regular and recurrent, or "special;" they may or may not involve outside academics; they may be reviews of a department, a "program", a college, or the institution as a whole. They may be focussed on the undergraduate program and instruction, or on graduate education and research, or on both simultaneously. They may

involve a one or two day visit by a review committee, or be extended over weeks or months. They may involve discussions with students as well as academic staff, or even with "alumni" and support staff. In other words, they may vary quite widely in character and procedures. But they have in common that they are oriented primarily toward helping the unit under review to identify its strengths and weaknesses, and by making recommendations to senior academic and administrative officers about ways in which the unit can be helped. These recommendations may involve suggestions about the allocation of resources to the unit, and thus engender some tension between their supportive and evaluative roles. But they are usually not directly linked to funding, and may not affect the unit's funding at all. What are some of the special characteristics and functions of these kinds of "internal-supportive" reviews?

1. They are done by peers who are in every way close to the unit under review, and who are best qualified to learn about its character: the "quality" of its work (along several dimensions), and its trajectory, or direction of development. If done properly, such a review can learn what the unit needs to improve its work: does it need advice, or leadership, or new blood, or additional resources, or more careful planning, or something else? How well will the unit make use of new resources of ideas or money or people? So such a review can be both diagnostic and advisory, and not just "evaluative" of the unit's current performance. A review has to know a good deal more about a unit to be helpful than to give it grades.

2. Since such a review is supportive, and done by colleagues, it is more likely to have the trust of the unit's members, and that in turn makes it more likely for the latter to be truthful and candid in what they perceive to be a non-adversarial relationship.

3. Since such reviews operate closely and consultatively with the unit, their errors can be identified and corrected through discussion. The possibility of correcting errors in reviews in turn gives the unit the courage to be more open and candid about its problems; it is less inclined to fear that such candor will be used against it.

4. Reviews in depth by colleagues can identify real problems early so that remedial action can be taken before they adversely affect the quality of the unit or its work. By contrast, "external evaluative" reviews can only identify problems when they have already done their damage, enough damage to show up on crude indicators of performance. But serious problems in an academic unit (including the negative effects of external reviews themselves) are often long delayed in showing themselves to outsiders.

5. Finally, internal reviews by peers are educative in a broad sense: they teach all participants what it means to be a self-governing and self-regulating institution. And they reinforce the academic values that are implicit in internal reviews.

But such internal supportive reviews have problems and limitations of their own. Among these:

1. They can cause strains among colleagues, especially when reviewers are critical of a unit in ways that the unit doesn't accept, rightly or wrongly. This strain is likely to be sharper when critical comments touch on personnel or resources -- for example, on the "non-productive" members of the unit, or the proliferation of small seminars to the detriment of the core program, or some broader pattern of mediocrity or poor performance.

2. The other side of the same coin is the possibility that the review will suffer from the reviewers' concern to avoid just those strains, and to be reluctant to make tough critical criticisms and observations about colleagues with whom they may have personal relationships, and almost certainly ongoing professional relationships. The operation of this kind of "senatorial courtesy" by members of a guild (the local academic community) can reduce the value of the review for decision-makers or the unit, or both.

A compromise is to add competent outside peers to the supportive review. These can be expensive in both time and money. But there are obvious trade-offs in advantages and disadvantages: the outsiders have less intimate knowledge of the unit, but they are also less constrained in what they say by personal ties and continuing relationships. But their presence, even in token numbers, may help keep a predominantly internal committee "honest"-- ie., less likely to allow personal ties shape the review in obvious ways.

The internal supportive review has much to recommend it: it is central to the continuing quality of many great colleges and universities. But the success of such reviews depends in large part on the presence of an academic culture which a) accepts the responsibilities of maintaining and improving the quality of its work, despite the considerable time, effort and personal strains that may entail; and b) creates the necessary institutions and procedures for such self-assessments -- for example, periodic reviews. All this requires that academics be motivated by professional pride and ambition for status and wider recognition, not just personally, but also for their departments and universities. This in turn assumes a certain competitive environment, competitive at least for comparative institutional status and prestige, an environment which is not welcome in all European university systems which have emphasized the formal equality of all their universities.

Ironically, the great transformation of British higher education has created just such a competitive climate for universities, while forcing their competitive efforts to be oriented toward the external criteria and evaluative reviews of the Government's own funding councils. By contrast, Swedish universities are currently trying to create the academic culture appropriate to the new autonomy recently granted to them. One element of this requires them to face the problems of creating internal review procedures required for quality maintenance in the absence of management by central government.

Type II: Internal evaluative reviews

Examples: Institutional assessments initiated by the rectors and principals of Swedish universities and colleges, as required by the University Chancellor.

Institutional reviews initiated by the vice-chancellors of British universities.

"Self-studies" carried out by American colleges and universities in preparation for an "accreditation" review.

Reviews initiated by American university presidents in response to some problem or crisis: for example, the review of biology departments at UC Berkeley, mentioned below.

Here we see the evidence of strong institutional leadership and strategic planning. These internal evaluative reviews are most commonly initiated by institutional decision-makers (presidents, deans etc.) who are forced to make cuts in budgets and want to set priorities rather than distribute the cuts equally across the board. The evaluations, carried out usually by groups responsible to the president/rector, may well be based on a wide variety of indicators of quality and effectiveness, and thus may use the results of Type I reviews, though this carries the danger of contaminating those reviews, and undermining the candor and trust that makes them most effective. Nevertheless, in their efforts to allocate scarce funds in the service of high quality (whether to pursue or to retain it), institutional leaders naturally use all the information available to them, including Type I reviews, supplemented by such other indicators as student demand and retention, time to degree, research productivity, quality of students recruited, etc etc.. All of this can be put in a broad context that can distinguish between the quality of a department of comparative literature and a department of chemistry. With enough of this kind of information, a rector or president faced with budgetary problems, or a decline in demand in certain areas, or paralyzing internal conflicts, or the seizure of a department by true believers, can act firmly, even taking a department into "receivership" and appointing a

chairman from outside its number with clear powers from above and instructions for making reforms.

Type II reviews can emerge from Type I reviews, especially when the latter give early warning of serious problems. An example occurred at Berkeley in the early 1980s when a number of independent indicators pointed to serious declines in the quality of some areas of biological science at the university. Several of the biology departments (there were a large number scattered throughout the university) earned poor national rankings in a then-recent national ranking of graduate departments carried out by the National Academy of Science. (That would be a special kind of Type IV review, as we will see.) In addition, not enough people who were the first choice for appointment to Berkeley departments of biology were accepting their offers, in competition with offers from other leading research universities. In the face of these grave warnings of problems which if not corrected could lead to a serious loss of status of the University as a whole, the Chancellor appointed a special review committee of the most distinguished biologists within the university. Their sweeping report was supplemented by the parallel appointment of a review committee of distinguished biologists from other universities which came to similar conclusions. These reviews were both supportive and evaluative: evaluations whose chief functions were diagnoses and prescriptions for strengthening the whole provision of biological studies at Berkeley, an activity which at the time involved some 20 departments and about 250 regular academics. Out of these reports and further consultations with the biological community at Berkeley, a broad plan was developed to respond to the problems, a plan involving major expenditures on new buildings and research facilities, and major reorganizations of the departments of biology, affecting how the various kinds of biologists related to one another both in teaching and research.¹⁴

The main differences between Type I and Type II reviews are that the latter are initiated by central university administrators, are almost always ad hoc rather than recurrent and routine, and are directly linked to action, including the allocation of resources. But that action is not formulaic, as are evaluative reviews (Type IV) initiated by central government; and very often a finding of weakness in a department or sector leads to the allocation of larger resources in support of new leadership and other reforms, organizational or curricular, as we saw in the example of biology at Berkeley.

¹⁴ This example is developed more fully in Trow, "Leadership and Organization: The Case of Biology at Berkeley," in Higher Education Organization: Conditions for Policy Implementation, edited by Rune Premfors, Stockholm: Almqvist and Wiksell International, 1984, pp. 148-178.

Type III: External supportive reviews

Examples: Activities of the Office of the Chancellor in Sweden
Audits conducted by the Higher Education Quality
Council in the United Kingdom, an autonomous body sponsored by the
Committee of Vice-Chancellors and Principals.

The traditional reviews of colleges and universities by
committees of the regional and professional accrediting bodies in the US.

The reviews formerly carried out by the Council for
National Academic Awards in and for the British polytechnics (no longer
functioning).

This type of review is relatively rare, perhaps because "external" usually means "government" or "the state," and because those institutions are more likely to be concerned with the "efficient" use of public funds than with the more delicate business of helping universities or departments strengthen themselves. Moreover, the power of the state over the universities is potentially so great that their involvement in assessing the strengths and weaknesses of academic units can hardly be separated in the minds of academics from that potential power. And that natural connection naturally affects the way in which units respond to the visits and requests for information on which external reviews are based.

But in the current climate of relations between universities and the state in many countries, governments are requiring that some body external to the universities assess their work and report back to the central funding agencies; that is what "accountability" has come to mean. And the choice comes down to external reviews which are primarily devices for evaluating the "quality" of academic units, and tying funding to those assessments, (Type IV reviews), and similar reviews which have as their primary function the identification of academic strength and weakness with an eye to confirming the former and amending the latter (Type III).

External supportive reviews have several advantages in performing their functions. For governments and the larger society they have the legitimacy of being external to the units reviewed, and thus presumably free from the constraints of collegiality that are inherent in internal reviews. This greater objectivity gives them a certain authority in certifying the quality, or quality seeking efforts, of the units in question. Put simply, outside agencies, whether governments, foundations, or business firms, are more likely to believe the assessments of external agencies than they are the outcome of an institution's self-studies. But this very fact presses bodies making Type III reviews to become more evaluative, even when they start with the avowed mission to be helpful and supportive of the units under review rather than evaluative. But insofar as they can avoid being pressed into the evaluative

function, Type III reviews are more likely than Type IV reviews to gain the trust of the units being reviewed, and of having their inquiries answered truthfully and with candor.

More compatible with their self-defined mission, Type III reviews have the large advantage over all others of being able to carry news and information about good practice from one institution to another. In the process of doing reviews in different institutions, a central reviewing body becomes a repository of special knowledge and expertise about educational reform and innovation -- about what characterizes strong academic units, and how weaker ones can become stronger. This capacity to learn and then to teach about academic quality across institutional boundaries, free from the constraints of guild loyalties and jealousies within universities, is surely the greatest contribution Type III reviews can make to academic life.

The chief drawbacks of Type III reviews are inherent in their being external. They are necessarily further removed from the units under review, and even if based on periodic visits their assessments of the quality of units must be more superficial than internal reviews. This drawback is much reduced if what they are assessing is the quality of self-assessment procedures -- ie., if they are conducting audits of procedures rather than evaluations of the quality of academic units.

Type III reviews are inherently unstable activities; unless firmly institutionalized and defended, they are vulnerable to pressures from government to become more evaluative, more "useful" to central governmental funding agencies which want to distribute their funds in ways that create -- and appear to create -- incentives for better performance.¹⁵ And from a government's point of view, the best way to do that is to get objective

¹⁵ "Management consultants in Britain have told the Higher Education Quality Council that its quality audits should be more explicit in their criticisms of institutions, and that sanctions should ensure that advice is heeded." [The management consultants involved are Coopers and Lybrand, who are the chief consultants on quality assessment and control issues to central government and the Higher Education Funding Council, its chief administrative arm.] "The consultants say that while the reports contain a great deal of information, they are 'rather light on judgement'." As a long-term option "they call for reports to contain explanation, judgement and sanction and they recommend the development of a national code of practice for quality control." The HEQC resisted the advice, claiming that a harder tone would "be a move towards an adversarial (if not confrontational) style for audit, and would be likely to lead to fewer substantive critical comments being made than at present." Having made these objections, the Council "does envisage reports with a 'firmer, more judgmental style' in the future." Moreover, "the HEQC intends to conduct a feasibility study on this proposal." (The Times Higher, Feb. 18, 1994, p. 48.)

evaluations by external agencies, and link funding to those evaluations. This pressure creates a drift from Type III toward Type IV reviews.

The pressures on the Higher Education Quality Council (HEQC) in the UK to move in that direction are paralleled by similar tendencies among regional accreditation agencies in the United States. These bodies, nominally voluntary associations of colleges and universities within large regional areas of the country, have traditionally sent visiting teams to their member institutions with two distinct functions: one is to certify to the federal Department of Education that the institution in question is a genuine college or university, to "accredit" it so that federal funds can go to its students. And certain other privileges often accrue to "accredited" institutions which are set forth in the legislation of the various states.

But this "evaluative" function of these accreditation visits until recently have been chiefly ceremonial; the criteria for accreditation are so modest as to be met easily by almost all non-profit institutions of further and higher education. "Accreditation" in the U.S. has been historically a "threshold" function, to determine that an institution is really a college and not a secondary school or a fraud.¹⁶ It is almost unheard of for a regional accreditation agency to withhold accreditation from any responsible institution. The second function of visiting committees is to provide helpful advice and counsel to the visited institution, based on what the committee has learned from the institution's own self-study and during its three or four day visit. However useful the reports of these visiting committees are, and they vary from marginally useful to useless, they are kept in confidence between the visiting committee and the institution, and are intended to be wholly "supportive." The chief effect of a visit from a regional accreditation body in the US is to require the institution to carry out a self-study, which often has real value, but also substantial costs in staff and faculty time. Thus, these Type III reviews have as their chief contribution the Type I reviews that are ordinarily part of a self-study.¹⁷

But recently, several regional bodies in the US have attempted to expand their functions beyond the preparation of a supportive report and the ceremonial accreditation of the institution to public bodies, and taken on a

¹⁶ See Harold Orlans, Private Accreditation and Public Eligibility, Lexington, Mass., D.C. Heath and Co., 1975; and his "Accreditation in American Higher Education: The Issue of 'Diversity'," Minerva, Vol. XXX, No. 4, Winter 1992.

¹⁷ This point was made to me by H.R. Kells in private communication.

normative and regulatory function.¹⁸ They begin to prescribe policies for the institutions they review - for student admissions, staff appointments and the curriculum. They require institutions to report on such matters as the racial and ethnic distribution of their students, staff and governing boards, and to describe what efforts they are making to "improve" those figures. Moreover, they threaten to withhold accreditation from colleges or universities for not meeting criteria that go far beyond the traditional criteria that define a legitimate degree-granting institution in the United States. And they begin to look very much like Type IV reviews.

Type IV: External evaluative reviews

Examples: Directly linked to funding: the reviews of the Higher Education Funding Councils in the UK., every three years, involving assessments of research and teaching of every department in every university in the country, linked directly to government funding of those units.

Not directly linked to funding: the ten year reviews of graduate departments in leading research universities carried out by the National Academy of Science in the U.S.; and the reviews of leading four year colleges, conducted annually by U.S. News and World Report in the U.S., and rankings by other magazines of colleges and universities in other Western nations. (Such independent agencies have been ranking American universities and colleges since 1914).¹⁹

Type IV reviews are of two sub-types: Type IV(1), those operated by government agencies that are linked directly to funding; and Type IV(2), those operated by an independent agency, and not linked to funding. What are some of the characteristics of these types of reviews?

1. Type IV(1) reviews convey the sharpest expressions of the lack of trust by a government in the intrinsic motivations of academics or in their own internal processes of quality control.

2. The criteria of these reviews operate as instruments of management and control over constituent institutions. Linked directly to funding, they are powerfully coercive. They tend to reward short-term over long-term achievement. They tend to reward quantity of research over quality. (No

¹⁸ See, for example, Gerhard Caspar, "Government and the University," The Newton D. Baker Lecture at George Mason University School of Law, April 19, 1994, pp. 11-19. Gerhard Caspar is President of Stanford University.

¹⁹ See Martin Trow, "The Analysis of Status," in Burton R. Clark, ed., Perspectives on Higher Education, Berkeley and Los Angeles, The University of California Press, 1984, pp. 132-164.

agency can really assess competitively the quality of research done by thousands of scholars in hundreds of departments in a given year.) They thus encourage a lot of poor research, and the diversion of energies of academics who are not oriented toward research from useful teaching and service into poor research. They encourage research of the kinds favored by the review committees, and inhibit innovation and ideas which are not yet widely accepted. Perhaps most costly, they reduce diversity in academic life by applying common criteria of research and teaching across diverse institutions and subjects, and even between different approaches to the same subject.²⁰

3. Type IV(1) reviews encourage an adversarial relation between governments and universities, and the rehearsed presentation of a department's virtues. They focus the attention of academics not on the quality of their teaching or research, but on the preparation of a department's presentations and documents; ultimately, those are the things that will be assessed and evaluated, not the work itself. Science and scholarship becomes subordinate to rhetoric, and the preparation of these arguments comes to be the full-time work of professionals in the universities.

4. These reviews are remote from the life of a department and cannot really diagnose its work. Reviewers cannot know whether the work of the department is really the product of a few stars or is more widely distributed; or of its younger or older members; or whether the quality of work is on the rise or falling, or what if anything might be right or wrong with it. The evaluations simple reward "strength" and punish "weakness," though a closer analysis might call for a quite different strategy.

5. By introducing powerful and coercive criteria from outside, Type IV(1) reviews weaken and undermine the institution's own internal procedures for quality control. (Why serve as an external examiner when review committees will pay no attention to the quality of examinations)? They convert non-researcher academics into second class citizens, despite the manifest value many of them have for active research communities. They undermine the informal division of intellectual labor within departments, and make it more difficult for scholars to move their fields of interest in mid-career if that entails a period of low productivity. These brief and superficial reviews cannot take into account that most good departments train their students for competence in their subjects, while at the same time they also seek to encourage creativity, and that different teachers play different roles in

²⁰ These observations are based on an on-going study of the impact of central government policy on English universities currently (1993/94) being conducted by Oliver Fulton and myself. See Trow, "Managerialism," op. cit.

those quite different (and sometimes antagonistic) activities.²¹ They cannot respond to the quite different demands placed on teachers by students with low and with high levels of motivation, and the different styles of teaching those different kinds of students require. They discourage forms of instruction (eg., tutorials) for which they have no sympathy or positive criteria. They introduce a constant set of concerns and anxieties into the academic enterprise which have nothing to do with research and teaching of high quality.

6. They introduce into intellectual life in the universities a pattern of judgement which is essentially mindless, arbitrary and unresponsive to that complex activity and its environment. They thus demean the activity itself, both to its practitioners and to the society at large. (This is the opposite of the educative role of Type I reviews discussed above.)

7. The chief redeeming feature of Type IV(1) reviews is that they satisfy the government of the day, at least for the day.

Type IV(2) reviews, external and evaluative but not linked to funding, largely address the comparative reputation of departments and affect those reputations. Originating in the U.S., such broad evaluations have in recent years become much more widespread, and are often conducted by weekly magazines where they purport to provide consumer education and protection in an area not governed by the market. These judgements from afar cannot be diagnostic or supportive, but they avoid the worst of the effects of governmentally sponsored evaluations by not having such powerful coercive sanctions and rewards. Nevertheless, they do affect the reputation of departments and colleges, and their attractiveness to students and staff. And there is a danger of their perpetuating an institution's reputation (for better or worse), even if undeserved. "Give a dog a bad name ..." is the generic problem of status by reputation.

Moreover, insofar as the assessors use similar criteria over time, what may be significant is not the department's ranking at any one time, but its direction of movement between two assessments. This can be indirectly (and

²¹ All assessments, all evaluations are efforts to rationalize efforts to support academic quality, and thus at their best they serve competence rather than creativity. Creativity escapes the categories of assessment, almost by definition; when present, it creates new categories, new criteria of excellence. It is a challenge to all forms of assessment in higher education that they serve the pursuit of competence without hindering or penalizing creativity. Reviews are better able to meet that challenge if they can learn about the department and its life in detail and in depth. That they do not and cannot is the strongest argument against Type IV reviews.

crudely) diagnostic, and lead to Type I and Type II internal reviews for finer diagnostics and analysis, as at Berkeley in connection with its biology departments. One might imagine that this could also be a virtue of Type IV(1) reviews, except that they exert their power through funding rather than through reputation, with all the attendant problems and pathologies of that arrangement already described.

The Dynamics of Review

The model of academic reviews presented in this paper is designed not only to clarify the differences among them, but also to help us see their influences on one another, and of their tendency to move or change over time. Some tendencies of this kind are immediately apparent.

1. Type I reviews (internal supportive) tend to be captured by the guild, by the academic communities they review, and lose their sharp critical edge unless carefully monitored and reinforced by senior academic or administrative officers. Academics, both reviewed and reviewing, need to be reminded periodically that a weak system of internal reviews invites reviews, evaluation and regulation from outside the university. Put in the form of propositions: the stronger and more common are internal reviews, either Type I or Type II, the weaker will be the pressures for external reviews, and vice versa. Similarly, the stronger and more common are supportive reviews, either Type I or Type III, the less need will be seen for evaluative reviews, and vice versa.

2. There is a tendency for all external supportive bodies (the Chancellor's Office in Sweden, the Higher Education Quality Council in Britain, the regional accreditation bodies in the US) to be pushed by government agencies, and perhaps also by their own inclinations, into evaluative roles, toward the exercise of power and not just influence through persuasion, from friendly advice toward regulation. This is a drift of Type III toward Type IV review. It is most clearly under way in the HEQC audits and in some regional and professional accrediting bodies in the US.

3. This drift also can be seen in internal supportive reviews. Good administrators want to know what is going on, and seize on internal reviews to help them make decisions and allocate resources. As that practice becomes more transparent, it tends to reduce the level of trust between reviewers and reviewed, increasing the adversarial element in the relationship, and thus paradoxically (a) weakening the reviewer's capacity to be supportive, and (b) reducing the level of candor by the unit under review on the basis of which help can be given or policy decisions made. Therefore, wise administrators

try to insulate internal supportive reviews from the appearance of direct involvement in resource allocation.

Thus, Type I reviews, by far the most important for the quality of academic work of all kinds, are vulnerable to capture by the academic guild, and are also vulnerable to becoming part of the funding mechanisms of university administrators. Both tendencies undermine the value of internal supportive reviews for the maintenance of academic quality. Similarly, Type III reviews risk being captured by the institutions they review, and more seriously, risk also becoming part of the evaluating/funding process of central government. The former would lead them to be rejected by government, the latter by the universities.

Conclusion: A broader conception of "quality" for academic reviews

The model of academic reviews presented in this paper is intended to allow us to see how different kinds of academic reviews contribute to one another, and how in combination they may contribute to the "quality" of academic work in a university, college or department. And this brings us to a different view of what constitutes "quality" in higher education. We surely want to judge not only an academic unit's capacity to conduct research or teaching, but also its capacity to govern itself, to define its own character and mission, and to act effectively in fulfilling that mission. A key question in a broader conception of academic quality might be "What is the capacity of this unit -- department or college or university -- to respond to change -- to changes in the map of knowledge; to varying student demand for different subjects; changes in funding by central government; changes in the institution's own mission; changes in the institution's relations with the private economy, and in its capacity to serve outside constituents. The appropriate response to change externally is not always a parallel change inside the institution: the universities in Hitler's Germany surely taught us that. And yet institutions must be able to change appropriately to changes in their political, demographic, financial and intellectual environments. How an institution responds to change points to deep seated qualities of the unit which must also show up in its research and teaching, activities that are more difficult for outsiders to understand and properly assess.

Essentially, we want to assess the quality of an institution's intellectual life -- the process of education -- rather than its outcomes, which are long delayed, difficult to recognize, difficult if not impossible to measure, and

mixed up with many other forces and factors outside the institution.²² It may be argued that the process of education is as difficult to measure as are its outcomes. That may be, but yet the process is going on before our eyes; we can see and hear intellectual liveliness and involvement and dedication, and even competence. And inside our institutions administrators and colleagues from other departments and units can engage with academic units in conversations about that life -- what they are doing, what they think they are doing, what they want to be doing. The allocation of resources to academic units will and should be affected by these conversations, but not mechanically; it is a mindless policy that assumes that academic weakness should always be punished by the withdrawal of support.

There is considerable evidence that the most successful universities and colleges have had the power to allocate their resources internally according to their own lights. This freedom and autonomy in no way limits the power of public authorities to determine the overall level of support they are prepared to commit to higher education, nor their broad authority to determine the size and shape of publicly supported systems of higher education. Governments properly make broad decisions regarding the division of academic functions among university systems and segments. But within these broad parameters, wise governments give their support to higher education through block grants to institutions, while encouraging them to make their internal allocations with discrimination and on the basis of deep knowledge about the workings of their component units. And for that, tough and rigorous internal reviews are essential, reviews of the broad life of academic units and not just their "outcomes."

I have argued that the culture of excellence, and the internal quality control procedures that reflect it, are strongest in the great research universities. Moreover, I have suggested that the growth of the evaluation and assessment industry reflects the growth of mass higher education, and of institutions in which a culture of excellence embodied in traditional quality

²² "On the matter of assessing educational outcomes, no matter how attractive the notion is, ... there is very little evidence that outcomes can be measured in any rigorous way on a broad scale. General achievement tests are too crude to assess vast differences in the preparation students bring to college, the differences in colleges themselves, or the differences in programs within colleges and universities. Students attend our institutions for a variety of reasons and with a variety of expectations, have a variety of experiences while in school, and leave with a broad range of outcomes. The value of an education is not measured at a single point in time. Indeed, as many of our alumni attest, appreciation for education often increases many years after graduation. And ... let me assure you that "the public" whose interest the would-be regulators would like to protect has not arrived at any agreement whatsoever as to what precisely it wants it universities to provide." Casper, *op. cit.*, p.16.

control procedures and values is not firmly institutionalized -- leading governments to feel the need to impose a structure of evaluative reviews from outside the institutions to provide incentives for raising academic quality where internal quality control is absent or weak.

Even under those conditions, I believe, external evaluative reviews are poor instruments for raising academic standards and the quality of teaching and research, for all the reasons sketched above. Ultimately it is the academic staff who will determine the quality of their own work; the evidence of lack of trust and indeed, of the deprofessionalization of academic work inherent in close management from outside through evaluations and assessments cannot improve the work of teaching and scholarship. Efforts, from outside or from inside, to raise the quality of academic work must be aimed at creating a culture of excellence in an institution, and at the inner motivations and values of its academic staff. The nature of academic work, and the quality of that work, will differ in institutions of mass higher education, and should not be measured against the same yardsticks used by the great research universities. The criteria used to assess the quality of work in colleges and universities need to be closely linked to the varying missions of those institutions, missions which become ever more diverse as mass higher education develops. The point here is that a culture of excellence in a primarily teaching college or university need not be keyed to the same criteria of quality used to assess work in leading research universities, and it may be supported by different procedures and mechanisms. But it will be marked by strong motivations to excel in whatever is done, and by a critical self-consciousness, a readiness on the part of staff and leaders alike to learn and innovate, to respond creatively to the rapid social and intellectual changes that mark our time. Outside governmental or quasi-governmental agencies have a role to play in monitoring and encouraging the emergence of this culture in institutions of mass higher education, but not through "evaluations" based on uniform criteria and linked to funding.

The current concern about the "quality" of higher education, arising out of the growing diversity of institutions and students alongside declining public support, can be a positive development if it leads universities and colleges of all kinds to become more self-conscious, more aware of their own activities and of variations in the quality and effectiveness of their departments and academic staff, more sensitive to ways of strengthening themselves and more motivated to act towards the improvement of all their functions. But universities must have freedom if they are to become more effective in these ways. Institutional autonomy is a necessary if not a sufficient condition for the development of a culture of excellence, one that embodies a wide range of quality control mechanisms, including internal reviews, through which academic excellence is achieved and sustained.

U.C. BERKELEY LIBRARIES



C114006026
