How Academic Program Review Can Foster Intellectual Conformity and Stifle Diversity of Thought and Method

Michael L. Skolnik

How Academic Program Review Can Foster Intellectual Conformity and Stifle Diversity of Thought and Method

Among institutions in Western society, the university has been one of the most persistent and vociferous in its espousal of noble ideals. Though these ideals — which date back to the Middle Ages and are proclaimed in university mottoes, mission statements, and promotional literature — vary slightly from one institution or jurisdiction to another, they typically include such values as intellectual excellence, honesty and courage, moral purity, disinterested pursuit of knowledge, and unfettered search for truth.

Like other agents that espouse commitment to noble ideals, the university has encountered practical difficulties in pursuing them: the problems arising from what Isaiah Berlin [10] refers to as the inevitable clash of different values; the fallibility of human character; and the conflict among differing interpretations of what is right, good, or true, combined with intolerance of one for another's interpretations.

For example, excellence has sometimes been thought to conflict with freedom of inquiry. Deficiencies of moral character have been revealed with alarming frequency in reports of fudged experimental results. A recent book by an eminent historian [26] describes the average American university historian as "manipulative, mean-minded, provincial, and envious." In reviewing a book which asserts that the most important social function of the university should be to provide moral leadership in society, Clark Kerr observes that there is no evidence that morality is higher on campus than off, and he suggests that

---

Michael L. Skolnik is professor in the Higher Education Group at the Ontario Institute for Studies in Education and the Graduate Department of Education of the University of Toronto.

Copyright © 1989 by the Ohio State University Press
the university get its own house in order before preaching to others [34, pp. 596–97].

As for the third category of difficulties in pursuing the ideal, Winchester [61, p. 272] observes that “intolerance for new and different ideas, although it has no place in the university, is often found there.” In a study of the gatekeeping role of scholarly journals, Lindsey [39, p. 120] concludes that “the critical, the controversial, and the imaginative are not allowed to see the light of day.” Cambridge economist Robert Kuttner [37] argues that appointment, tenure, publishing, and funding in economics are controlled in such a way as to restrict opportunity for any who don’t subscribe to conventional thinking in the discipline. Ideas, he maintains, are selected not on the basis of their validity, but by how well they conform to economic orthodoxy. That intolerance for nonconformist thought intrudes not only upon professors’ research, but also upon teaching, is suggested by Cude’s study of the Ph.D. process: “The values most inculcated are endurance, obedience and caution, and the professional skills most exercised are memory, diplomacy, and conformity to the intellectual status quo” [19, p. 52].

Concerns about intellectual conformism and suppression of unorthodox thought and method are as old as organized intellectual activity itself, evoking such names as Bruno, Copernicus, and Galileo. The most trenchant of modern criticisms along these lines are found in Veblen’s, The Higher Learning in America [60], and more recent critiques seem restatements, elaborations, or confirmations of his prescient observations. Veblen portrayed the American university as “encouraging publications largely for the sake of institutional prestige, rewarding mediocrity as often as merit, and exerting enormous pressure on dissident faculty to conform” [31, p. 129]. Since Veblen, concerns, from a variety of perspectives, about academic dogmatism, intolerance, and conformity have mushroomed, until, as one commentator notes, “the cynicism and sense of self-deprecation within academe has become an open secret” [54, p. 10].

For example, the literature on diversity, one of the most cherished values in American higher education, has revealed what many ob-

---

1William Cude is a former Ph.D. candidate in English Literature at the University of Alberta who was denied his degree due to what he calls “departmental politics,” even though nine of the eleven chapters of his thesis were published as articles in four scholarly journals. While his own experience appears to provide the motivation for the study, the book [19] provides a penetrating and rigorous analysis of the doctoral process in the North American university.
servers regard as an alarming degree of uniformity and conformity with respect to method of inquiry, conceptualization of professionalism, and paradigm of knowledge [11, 32, 42, 49]. Following Mills, numerous commentators on academic freedom have warned that the chief threats are not from outside the university, but from within: "The real restraints are not so much external prohibitions as control of the insurgent by agreements of academic gentlemen" [44, p. 297; see also 20], giving rise to what might be called the Pogo View of Academic Freedom ("we have met the enemy and it is us"). Schrecker's study of the university during McCarthyism concludes that "the academy did not fight McCarthyism, it contributed to it" [51, p. 340]. Echoing Schrecker, and in the spirit of Veblen, Michael Taves, Book Review Editor of Telos, a journal of post-critical thought, observes: "Universities are deadly to serious intellectual work. The university ethos fosters mediocrity, boredom, and gutlessness. It has become a haven for conformist intellectuals who value patronage and status over intellectual quality and challenge" [59, p. 5].

While most who have written in this vein have focused upon the ways in which the pressures toward conformity have demoralized professors and students and have inhibited innovation and creativity within the university, some have emphasized the resultant impoverishment of society. Margaret Fulton, president of Mount St. Vincent University in Nova Scotia, observes that the "bigotry and pedantry" once believed to be due to religious domination of the university has merely recrystallized, now attributable to the intellectual arrogance of secular scholars [21, p. 236]. She laments that the most exciting and promising scholarship addressing critical social and environmental issues is now done by a new breed of scholars outside the university because such work is scorned within the university. In a book which surveys the critiques of academe since Veblen, Russell Jacoby describes how academic culture and bureaucracy have produced a disconnection between the university and public intellectual life. The result is the disappearance of the public intellectual; hence "public culture relies on a dwindling band of older intellectuals [for example, Howe, Galbraith, Mumford] who command the vernacular that is slipping out of reach of their successors" [31, p. x].

Jacoby notes also that the corrupting influence of academic conformity transcends political ideology. Radical academics have flourished, but in their professional norms and behaviour, their methodologies, and their obscurantist writing, they are indistinguishable from establishment academics: "Scads of Marxist economists, but no Sweezy or
Braverman”[p. 234]. On the same theme, Luke comments: "Radical, Marxist, feminist, or socialist assumptions, social scientists, economists, or historians don't really rock their disciplines' boats. They still grind the same data through accepted or alternative methodological machinery, but very few challenging or new findings are coming out" [40, p. 8].

Through the quotations in the preceding paragraphs, I have attempted to give a brief sampling of the concerns expressed by an increasing number of critics of the contemporary state of academe in relation to its ideals of free inquiry and unfettered search for truth, wisdom, and knowledge. Ultimately, it is near impossible to judge how accurately these portrayals describe the reality of an enterprise that is as diffuse as higher education. Based upon the frequency and vehemence of such commentaries, however, I am inclined to think that there is at least an element of validity in them. A question that arises then, is what are the factors and mechanisms through which an institution that is committed to the search for new knowledge exerts such pressure on faculty to conform to the status quo?

Factors Which Contribute to the Suppression of Non-Conformity

As stirring as their prose may be, the critical commentators have produced little analysis of the factors or processes which contribute to the intellectually stifling effects which they report. For some, the root of the problem would seem to lie in selection processes, that is, persons of particular character weaknesses are drawn to academe. This theme is prominent, for example, in C. Wright Mills’ critiques of academe. No prominent (and prolific) academic has described the average professor in such unflattering terms as Mills.

Men of brilliance, energy, and imagination are not often attracted to college teaching. The Arts and Sciences graduate schools, as the president of Harvard has indicated, do not receive "their fair share of the best brains and well-developed, forceful personalities." . . . The graduate school is often organized as a 'feudal' system: the student trades his loyalty to one professor for protection against other professors. The personable young man, willing to learn quickly the thought-ways of others, may succeed as readily or even more readily than the truly original mind in intensive contact with the world of learning [45, p. 130].

Theories of institutional behaviour which depend upon psychological traits are quite difficult to prove [8], and those who express similar sentiments as Mills are sometimes dismissed by colleagues as suffering
from career frustrations, mid-life blues, or other species of alienation. A more serious weakness in the selection hypothesis is that it emphasizes only one side of the fit between individual and institutional characteristics. Mills acknowledges this when he notes that “after he is established in a college, it is unlikely that the professor’s milieu and resources are the kind that will facilitate, much less create, independence of mind. He is a member of a petty hierarchy. . . . In such a hierarchy, mediocrity makes its own rules and sets its own image of success” [45, p. 131]. We are thus led to consider institutional characteristics which give rise to the pressure to conform.

Some have suggested that intolerance of new or different ideas in the university is to be expected because universities have a bureaucratic structure, and bureaucracies are naturally conservative [61, p. 272]. However, organization theorists have questioned whether the bureaucratic model accurately describes the university [5, 48], and the question would still remain as to why this particular bureaucracy embraces its particular values.

Some insight into this question might be provided by Jencks and Riesman’s depiction of The Academic Revolution [32]. That revolution is said to have involved the transfer of authority in academic matters from the president to the faculty, a shift from teaching to research as the primary focus of the university, and the emergence of research-preoccupied disciplinary communities as the arbiters of institutional life. In the course of the revolution, a single set of norms of professionalism — those described by Merton [43] and Parsons [47] — came to dominate the culture and reward structure of universities. In his earlier work, Riesman [49] had described the academic procession in which all institutions tended to emulate the most prestigious research universities. The combined effect of the academic revolution and the academic procession was a national convergence “of aims, methods, and probably results” throughout American higher education [32, p. 26]. In this scenario, the chief forces contributing to conformity were the evolution of a unified national (or international) market for scholarship and the single-minded pursuit within that market, by all concerned, of a single scarce commodity, prestige.

This explanation, however, leaves many questions unanswered. The Jencks and Riesman depiction of the academic revolution is largely descriptive. It sheds little light on why all professors and institutions should embrace the same set of norms, why all should be so prestige-seeking, and why the segmentation which divided subsectors previously should collapse in the mid-twentieth century. More impor-
stantly, it fails to explain why those who are not inclined to accept majority norms should be pressured to conform. Focusing on post-World War II dynamics, it also fails to explain why the writings of Veblen should have such apparent currency for so many in the late twentieth century.

Potentially more fruitful insights into the pressure toward conformity can be obtained through a synthesis of recent literature on academic cultures, political models of decision making in the university, and the sociology of knowledge. In tracing the evolution of the organization of the university "from community to federation," Clark [15] describes the university as a federation of divergent disciplines with differing values and paradigms, which defy effective coordination. Yet, if it is to be viable as a corporate entity, let alone flourish, not only coordination is necessary, but some articulation (and enforcement) of a common purpose and vision within the university [1]. But given the conflicting interests and perspectives of divergent disciplinary groupings, decision-making can occur only through a political process [4]. In relating how he came to develop a political model of university organization, Baldrige highlights as typical the remark of one of his administrator interviewees: "this place is more like a political jungle, alive and screaming, than a rigid, quiet bureaucracy" [p. 21].

While there are many different groupings in this political jungle and many different views as to what the university should stand for, the distinction most often noted in the literature is that between the sciences and the humanities, with the social sciences somewhere ambivalently in between [13, pp. 359–71]. The humanist, Clark notes, is committed to personal interpretation, the scientist, to interpretations that can be verified by others [p. 132]. The gulf in world view is enormous, and as Halsey and Trow [25] note, a professor's larger world view has an extraordinarily large bearing upon his more narrow academic beliefs and attitudes, for example, upon how she would define proper academic behavior or quality of programs in the university. Many of the questions posed by the humanist would be dismissed by the scientist as metaphysical or mystical, thus of dubious legitimacy for the university. Scientists, on the other hand, are preoccupied with counting, and this explains, for example, their preoccupation with citation indices as measures of quality, reflecting what Lindsey describes as a logical positivist bias in evaluation.

In the internal struggle to shape the university, the sciences, Bloom [13] notes, sit at the pinnacle of prestige and influence, and other dis-
ciplines attain legitimacy by emulating their methods and paradigms.\footnote{Sometimes this emulation produces bizarre results. In a study [27] in which economists were asked to rank the quality of various journals, the investigators tested their hypothesis of prejudice in favor of quantitative and theoretical journals by inserting a fictitious one in the list, "The Journal of Economic and Statistical Theory." The non-existent journal was ranked in the top third, and 19 percent of respondents claimed to be familiar with it.} In fact, Winchester observes that "research" has come to be understood to mean "doing things rather like one would if the subject matter was physics or chemistry" [61, p. 282]. The reasons for the dominance of the sciences include: that the roots of the contemporary university lie in the logical positivist paradigm which was such a powerful force in late nineteenth and early twentieth-century academe; that the sciences have been spectacularly successful in the twentieth century; and that the sciences offer a definitive, quantifiable model for the university in contrast to the divergent, subjective ones offered by other disciplines.

However, the hold which the sciences and logical positivism have over the university has its downside with respect to tolerance of intellectual deviation. Science, as Kuhn [36] describes it, tends to operate largely within the framework of an accepted paradigm of knowledge, and to be suspicious, if not hostile, toward approaches and ideas which lie outside that paradigm. Barber notes that a major theme in biographies and autobiographies of scientists is the intolerance with which new ideas are met and not infrequently suppressed ("even men who are not Cambridge mathematicians deserve justice," complained Oliver Heaviside, whose important contributions to mathematical physics were ignored for twenty-five years [quoted in 8, p. 542]). Given the way that science "deals with its own," it would hardly be surprising to find a similar attitude of intolerance in the sciences toward alternative paradigms of scholarship that emanate from other disciplines — especially when one notes the evangelical ethos within the sciences which "aims to purge men's minds of anything which smacks of pseudo-science" [50, p. 32].

In spite of the frequent reference to it by scientists, what Barber calls "the resistance by scientists to scientific discovery" has not been a subject of major study in the history and sociology of science. My purpose in raising it here is not to offer it as a full-blown theoretical explanation for the university's inhospitality toward non-conformist scholarship, but merely to note the ethos of science as a factor which may be relevant to the question toward which this paper is addressed. It may well
be that I am unfair in attributing a portion of the responsibility for the pressures toward conformity to the ideology and ethos of science, and that there is in the man or woman of the intellect, of whatever disciplinary orientation, a natural urge to have his or her views of what constitutes legitimate norms of knowledge, inquiry, and professional behaviour prevail to the exclusion of other views. Orwell's observation that most who oppose a tyranny of ideas do so in order to make way for a different tyranny of ideas suggests that the latter may indeed be the case.

From this brief survey, it may be seen that attempts to explain the pressures toward intellectual conformity in the university have focused upon personal traits and attitudes of professors and administrators or systemic structural characteristics of the university—or the interaction between the two sets of factors. In either case, the problem would seem to be quite intractable, for neither of these sets of factors is readily amenable to change. However, this depiction of the problem may be unnecessarily pessimistic, as it ignores the role of institutional mechanisms and processes for planning and decision-making through which conformity is achieved—and which are far more amenable to change than human character or institutional structure. Insofar as latent pressure toward conformity inheres in the latter two sets of factors, it becomes manifest only through particular decision-making processes. These processes may serve to reinforce the latent human and institutional tendencies toward conformism, or to counteract those tendencies, at least to a degree. One of these decision-making processes which has come to have considerable impact upon the phenomena in question is program evaluation. In the remainder of the article I describe how one particular approach to program evaluation works to stifle academic diversity and innovation. In this discussion, I will draw upon the experience of one jurisdiction, Ontario, as this approach and some of its consequences are so well illustrated in the Ontario experience. However, the model of program evaluation used in Ontario is similar to that used in many other jurisdictions, and it is the consequences of this evaluation model to which I wish to draw attention.

This case will illustrate how several factors which have been noted in the contemporary North American university—concern about funding limitations, consequent preoccupation with accountability and quality, and dominance of the sciences—have come together within the framework of a naive model of program evaluation to restrict both academic development and service to students and the community.
Before relating this case, some comments on program review and academic quality are in order.

Program Review and Academic Quality

Although academic program reviews have been conducted in one form or another for more than fifty years, they have become considerably more prominent, organized, and influential during the past fifteen years [18]. While program review traditionally was the prerogative of institutions, all fifty states and ten Canadian provinces now have some process for state-level review [7, 16]. Retrenchment has been a major impetus to the growth of program review activity, and Barak reports that twenty-eight state higher education agencies have the authority to discontinue programs based upon their reviews [6, p. 54].

Though there may be numerous objects of interest in program review — for example, societal need, student demand, efficiency, client satisfaction, or articulation with other programs — the central focus in most program review is quality [18, p. 30]. Yet the literature on quality emphasizes that quality remains an elusive concept [53], and a recent study warns that the state of the art for assessing program quality in higher education is "primitive" and "tentative" [2, p. 34].

Contributors to the large and growing body of literature on program review have attempted to develop categories for classifying the various models and approaches used in program review, but the value of these classifications is limited by the tendency of many program review processes to incorporate elements of different models. Some writers emphasize the distinction between qualitative and quantitative approaches and others the distinction between criterion-referenced and norm-referenced methods. In the former, which is also referred to as "objective-specific," the evaluation starts with the objectives of the program as articulated by those responsible for its operation and thus allows for variation among programs with respect to their objectives or the relative weights given different objectives. Norm-referenced approaches focus on evaluation criteria and measures that are defined by those external to the program and are applied alike to all programs falling within the purview of a particular review. Norm-referenced approaches usually emphasize standardized performance indices and have been criticized for embracing a reductionist view of education and for exerting pressure toward homogenization of programs [17, 22, 38].

Conrad and Wilson move beyond such binary distinctions and offer
a quadrangular classification of evaluation models, the categories being: goal-based; responsive, decision-making; and connoisseurship. The goal-based approach corresponds to the criterion-referenced concept described above. Its chief advantage lies in enabling systematic attention to performance of a particular program in relation to the unique objectives and environmental context of the program. Its chief weakness is its inability to deal with significant outcomes that are unrelated to goal statements [23].

The responsive model [23, 58] emphasizes the political, as opposed to technical, nature of evaluation and involves efforts to identify and negotiate among “the claims, concerns, and issues put forth by members of a variety of stakeholding audiences” [24, p. 7]. Given the variety of conceptualizations of quality within the university community [3] and the conflicting interests of the various stakeholders for each program within and outside the university, negotiation is probably necessary for an evaluation to have legitimacy among all concerned. However, the explicitly political nature of this model is a drawback to many who prefer to maintain that quality is a technical matter.

The decision-making model is organized around the types of decisions that those who have such authority for a program feel are incumbent, and it emphasizes the collection of the corresponding information. Its strength is the linking of evaluation with decision-making. Its weaknesses are those associated with rational decision-making models (for example, being able to anticipate all decision alternatives in advance) and possible divergence of values between decision-makers and those responsible for the program, which could result in a denial of the credibility of the evaluation [18, p. 49].

In the connoisseurship model, the evaluator is supreme, and his or her values, concerns, and prescriptions of goals drives the evaluation. Given the potential variability of orientations among connoisseurs, it is impossible to generalize about the substantive criteria in this type of evaluation. A consequence of that variability is probably a very low “test-retest” reliability in evaluations of this type. Also, the supremacy of the connoisseur results in the legitimacy of this approach often being a very contentious issue.

Although the Conrad-Wilson typology is a static meta-model, Guba and Lincoln [24] have provided an interesting portrayal of the evolutionary dynamics of evaluation models, which has some overlap with it. They trace the evolution of ideas and practice in evaluation through four generations. They characterize first generation models as technical, with the evaluator’s role as that of technician, and heavy emphasis
on quantitative indices, as befitting the era of scientific management right after World War I. In reaction to the inherent emphasis on summative evaluation in the first generation, the second generation, 1940s and 1950s, emphasized description of patterns and strengths of programs in relation to stated objectives and was geared to formative evaluation. In the 1960s the pendulum swung back again toward summative evaluation, and the role of the evaluator changed from describer to judge. While the second-generation, descriptive emphasis corresponded to the goal-based model, the third-generation, judgmental emphasis corresponded to the connoisseurship model. Growing recognition of value pluralism and distrust of connoisseurs have led to a move away from the judgmental model and to the increasing use of the responsive model in which it is accepted that evaluation is a socio-political process, and that rather than being the judge, the evaluator performs the roles of collaborator, learner, teacher, reality shaper, and mediator.

From this brief sketch, it may be seen that some of the generalized models of evaluation would foster toleration for diversity and nonconformity in academic methods and orientations (goal-based and responsive models), whereas others could have the opposite effect (decision-making and connoisseurship models). Further, the literature on evaluation offers no conclusive demonstration of the superiority of one model over another, although that statement in itself might be taken by some as an argument for the responsive model, the only one to endorse value pluralism explicitly. The next section describes the program appraisal mechanism in the Ontario university system and attempts to place it within the conceptual framework outlined above.

Program Appraisal in Ontario Universities

Provincial-level program review in Ontario has been confined to the graduate sector, although recently requirements have been introduced for review of new undergraduate degree programs in specialized or professional fields. New graduate programs must pass a review in order to qualify for funding, and that review focuses upon such factors as societal need, student demand, and intrinsic merit. In addition, in order to continue to be eligible for funding, all ongoing graduate programs must pass a periodic quality appraisal, and it is this periodic appraisal process that I wish to address in this section.

The appraisals are conducted by the Appraisals Committee (hereafter referred to as the Committee) of the Ontario Council on Graduate
Studies (OCGS), which is the graduate studies arm of the Council of Ontario Universities (COU), the association of the universities. Though in principle the COU can only transmit advice on funding matters, the practice is for the body that advises the government on university matters, the Ontario Council on University Affairs (OCUA), to accept OCGS appraisal results and to recommend to the government termination of funding for programs that fail an OCGS appraisal, and for the government to accept such recommendations. Thus, the Committee has the effective power to terminate a program.

The appraisal system was developed by the COU during a period of system retrenchment in which the universities feared increased government intrusion in their affairs and wanted to demonstrate that they could make tough decisions ("If we don't do it, the government will do it for us"). The desire of senior university administrators to show the government that the system worked, together with the norms of gentlemanly (all the university presidents are males) conduct in the club of universities, have produced strong inhibitions against challenging negative appraisals. When heads of programs that are negatively appraised attempt to make a formal appeal, they are thus likely to be restrained by their own universities. Besides, formal appeals can be lodged only on procedural grounds, and then only to the Committee which made the decision in the first place. Given these constraints on appeals, it is difficult to accept the view in a recent report on the appraisals process [46, pp. 9–10; hereafter referred to as the Report] that the fact that no use has been made of the appeals mechanism is an indication of the perceived credibility of the process. As will be seen, the absence of formal appeals could as easily suggest the lack of credibility of the process and hence the futility of appealing.

The procedures and criteria for the appraisals are outlined in OCGS By-Law No. 2. The criteria are rather broad, but certain emphases are discernible in the By-Law. First, the appraisals clearly are not intended to be goal-based, as there is only one brief reference to program objectives. Second, in the categories of information required — which apply uniformly to all programs — the main emphasis is on resources and procedural requirements of programs.

Third, the area of strongest emphasis is the quality of faculty. Almost half the list of items to be included in submissions to the Committee pertain to faculty publications, honours, research grants, and so on. The longest subsection in the section entitled, "The Basis of Appraisal," is on quality of faculty, which is described as "undoubtedly the most important single factor governing the quality of a graduate pro-
gram.” The subsection goes on to assert that the principal evidence of strong intellectual leadership by faculty is the volume of peer adjudicated publications as well as other accomplishments which have been subjected to peer review: “It is essential that the intellectual vitality of the group [of faculty] be maintained through regular participation of the bulk of its members in suitable activities, the validity of which can be verified by peer review.”

Now, the By-Law does not say literally that the real basis of the appraisal is the volume of publications in refereed journals and academic presses and of research grants received from granting councils. But anyone familiar with the bureaucratise and euphemisms which flavour such documents could easily draw that inference. This interpretation is supported by comments in the Report which relate to the characteristics of programs receiving different ratings. Ratings range from “a” to “d,” with the former denoting very good quality, the latter normally leading to termination, and the “temporary” ratings of “b” and “c” indicating deficiencies which must be addressed prior to a reevaluation in which these ratings are changed to either “a” or “d.” The Report notes that two characteristics of all programs which received “a” ratings were that faculty regularly received grants from the national granting councils and had sustained outputs of publication in scholarly journals and books which are reviewed in such journals. The other characteristics are adequate library, computing facilities, and resources; curriculum and program requirements comparable in scope and difficulty to similar programs in other jurisdictions; and relatively high grade point requirements for admission. No mention is made of instruction or student outcomes, quality or relevance of work, or service to external clientele. It is also noted in the Report that examples of concerns which result in a “c” rating include “a low volume of recent faculty publications, low or deteriorating research grant levels, equipment deficiencies or deteriorating library resources.” Clearly, maintaining what the Committee considers to be an adequate volume of publications and grants is a necessary, and possibly sufficient, condition for a successful appraisal.3

In terms of the models described earlier, the appraisals fit under the connoisseurship category [62, p. 9]. The connoisseur is the Committee,

3Among the letters that I have received from faculty in Ontario universities concerning bad experiences with the appraisals, is one from a former professor in a clinical program in which major changes had been made — effectively, in the professor’s view — gutting the clinical experience and supervision for students. One of the reasons for the changes was alleged to be to provide more time for faculty to write. The consequence was a substantial increase in the number of faculty publications, and an “a”
which consists of twenty-one members of the professoriat selected by a
nominating committee of OCGS for three-year terms. The Committee
may appoint program consultants to advise it, but the decisions are
made by the Committee. In 1985, the last year of the four-year period
examined in the Report, the majority of the members of the Committee
were from the sciences [46, Appendix C]. In addition, there were three
from political science and one each from psychology, classics, philos-
ophy, economics, history, English, and French. None were from edu-
cation, though that is the field which offers professional expertise in
evaluation of educational programs. There were none from newer,
nontraditional, softer, professional (except medicine), or interdisci-
plinary departments, except for the historian who was cross-appointed
to women's studies.

The results of the appraisals from 1981–82, when the system was
implemented, through 1984–85 show marked variation by discipline
[Appendix B]. The range in percentage of programs receiving less than
an "a" rating is from 0 percent in the natural sciences to 30 percent in
the humanities. Within the humanities, those fields which share the
least common ground with the sciences fared the worst: half or more of
the programs in classics, religion, and education were given less than
an "a" rating, and the Province's only Ph.D. program in adult educa-
tion got a "d" and was terminated. The extraordinary success of the
natural science programs is remarkable in view of the widespread con-
cerns of the university community during this period about inadequate
funding and the expectation that this could result in serious quality
problems for those areas which require expensive equipment and re-
search support.

As striking as the discrepancy between the sciences and the humani-
ties is, there is no comment on it in the Report. I have been able to think
of only two possible explanations for this. One is that Ontario profes-
sors in the humanities are significantly less competent than their peers
in the sciences. However, I know of no evidence to support this hy-
pothesis — other than the appraisals.

The other explanation is that for this particular group of connois-
seurs, the implicit standard for what constitutes quality, and hence the
operational criteria for the appraisals, reflect the prevailing norms of
the sciences. Thus, what are purported to be measures of quality are, in
reality, measures of the degree of conformity to the model of graduate programs which obtains in the sciences.

The methodology employed by the authors of the Report precluded any consideration of the possibility of such systemic bias. They concentrated on a case by case analysis, watching for impropriety and inconsistency [p. 13]. Having found none, they gave the process high marks for integrity and credibility. Of course, the process could be high in integrity (and I believe that it is) and still be subject to the kinds of unconscious, systemic bias described earlier in this article. As for credibility, the Report’s conclusion is limited by the authors’ failure to elicit any input from those whose programs were appraised, that is, program heads and faculty. If credibility in a quasi-judicial process means anything, it means credibility to those who are being judged, not just credibility to the judges.

Discussion

In the connoisseurship model used in Ontario, the connoisseurs are selected by a self-perpetuating elite which represents the status quo and gives dominant weight to the academic values and perspectives of the sciences. The connoisseurs are not accountable to faculty or students, and it is the connoisseurs’ tastes in academic matters which determine the de facto criteria for the appraisals, leaving little room for the diverse objectives and concerns of the various program constituencies. Because these tastes are not made fully explicit before the fact, there is no opportunity of public debate within the university community on the appropriateness of the values and criteria which drive the appraisals.

As the appraisals carry the ultimate sanction of program termination, they exert a powerful influence on faculty to conform to the connoisseurs’ expectations. This influence is manifested primarily through occasional terminations, rather than through mass terminations. In considering the inhibiting and conformity-inducing effects of the process, three major issues arise.

The first involves the substantive character of research, including such issues as research topic, problem definition, ideology, and research method. In principle, the connoisseurs have adopted a stance of neutrality on these matters, and there is no evidence of a direct attempt to exert control in this area. However, the strong emphasis placed upon volume of publications in scholarly journals and grants from national granting councils may result in an indirect restraint on non-
conformist research. This will be the case insofar as the most prestigious journals and granting councils themselves shun nonconformist research. The literature on this subject referred to earlier suggests that many observers believe that such gatekeeping is quite prevalent. An appraisal system such as the one in Ontario creates pressure on faculty to select the safest research topics and methodologies.

It is relevant here to note also the considerable variation among disciplines in opportunities for scholarly publishing and research grants. Lindsey reports that because of differences in funding practices and patterns, there are more spaces for articles in the sciences than in other fields. For example, in the United States in the 1970s, he estimated that there were 3–4 journal spaces per working scientist in biochemistry, compared to 0.5 in the social sciences [39, pp. 122–23]. He reports a rejection rate of 29 percent in biochemistry, compared to rates of 70 percent to 80 percent in the social sciences. Further, Lindsey observes that there are differences in the time required for publication in different fields. Articles in the sciences tend to be shorter and literature reviews tend to concentrate upon recent publications, whereas social sciences and humanities articles are developed in a longer-term context and normally require more preparation time [pp. 123–24].

Preoccupation with quantity of publication, as opposed to quality, is no doubt responsible for the great amount of “academic gobbledygook” which many critics have reported [9]. Besides stimulating a torrent of pedestrian publishing, the practice of assessing quality by counting inhibits work which has a long gestation period and penalizes those whose inclination is to publish only when they have something highly significant and polished to offer. Overall, the emphasis upon quantity of publication can work to throttle certain types of research and cause systemic bias in appraisals across disciplines. In particular, imposing on the humanities the sciences’ norms for volume of publications may inhibit creative scholarship in the humanities.

Regarding research grants, it should be noted that the most well-endowed of the national granting councils in Canada are for medicine and natural sciences, and for many disciplines, especially in the (non-medical) professions, there is no corresponding granting council. When the grant applications of researchers in fields that are marginal to the central interests of the main granting councils are adjudicated,

\[4\] Irving Louis Horowitz has described the shift in responsibility for judging tenure from academic departments to editorial boards of journals [28, p. 164]. What we have here is a similar shift in responsibility for judging the quality of academic programs from educators to editorial boards.
these researchers are at a distinct disadvantage, because their research paradigms and methodologies often deviate from the norms to which the judges are accustomed. Further, the emphasis on volume of research grants as an index of quality may operate as a sanction against reflective research done in a library or professor's office for which small or no grants are required and discriminate against those disciplines where such activity is thought to be the most productive. It is arguable whether greater spending is a plausible indicator of higher quality.

The second issue relates to the balance of faculty effort among instruction, research and service. There are wide differences of opinion, within and across disciplines, on the appropriate balance among these functions. The Committee’s preference for evaluating programs almost exclusively on the basis of research outputs is more appropriate for the sciences—due to the strong complementarity there among these three activities — than for other realms of the university, where these functions are more competitive than complementary. For example, in the sciences, doctoral thesis research is often a constituent part of the supervising professor’s own research. This is not typically the case in the humanities or professions where research is more of an individual effort, and time spent in supervision is time which does not contribute to publication. Even within the sciences, publication rates in experimental fields tend to exceed the rates in clinical fields, where the instructional demands are so much greater [12]. The disdain with which teaching is frequently looked upon in the university reaches absurd proportions when what are purported to be, and funded as, instructional programs are evaluated almost solely on the basis of research outputs, with little or no attempt to assess student development outcomes.

As for the service function, there are within some applied disciplines strong traditions of and public expectations for faculty providing extensive service to various clientele outside the university. In these fields, not only does service compete for time with scholarly publication, but frequently the nature of the service activity is such as to not directly foster scholarly publication. For example, Judge [33] describes the “no-win” situation of faculties of education confronted si-

Gordon Inglis, member of one of the award committees of the Social Sciences and Humanities Research Council of Canada, has observed that new procedures of the Council are making grants decisions more centralized and elitist. He warns that the new procedures threaten to impose on the social sciences and humanities “a vision of the feudal system that holds sway in the ‘harder sciences’ and make our research far less independent and far more conservative” [30, p. 17].
multaneously with conflicting expectations of public educational agencies on the one hand, and scholars in the “pure” disciplines, on the other. And Lindsey [39, pp. 85–89] notes the characteristically low rates of scholarly publication in social work compared to other disciplines, which is likely related to the priority service is given in that field.

It would be misleading to suggest that the service work in professional fields consists of pedestrian activity that is divorced from scholarship. Much of the emphasis is not on service for its own sake but as a means to gaining a deeper understanding of the problems of professional practice [29]. In many professions, a new paradigm of the “reflective practitioner,” which emphasizes the development of theory through personal reflection on experiences in practice, is replacing the older linear view of the relationship between theory and practice, which was rooted in logical positivism [50]. In the traditional view, researchers and practitioners inhabit different worlds, with the former identifying problems and the latter producing solutions. In the newer view, research and practice are both approached most effectively when they form an integrated whole of activity. Operating in this paradigm, faculty may labour long in the trenches between scholarly publications, in contrast to producing the continual flow of articles and books assumed in the Ontario appraisals model. The development of new models of professional practice and corresponding approaches to research is but one area in the growing literature on alternative paradigms in the disciplines that is inhibited by the type of appraisal model used in Ontario [for others, see 14 and 52].

Three points related to the developments which Schön [50] and Hunt [29] describe are relevant to our discussion of the ways in which the present appraisals model mitigates against the advancement of knowledge. First, in academic fields which subscribe to this model, faculty will have less time for and less interest in proliferating articles in scholarly journals than their counterparts in fields where the dominant view is of a linear relation between research and practice, with the former specializing in writing and the latter in doing. The principal means of scholarly communication will be other than print, for example, colloquia and demonstration projects, means that are systematically discriminated against in the appraisals process.

Second, the reflective practitioner model calls into question a hallowed assumption underlying the appraisals: that scholarly publication is the best, or only, indicator of a professor’s competence to train graduate students. This assumption was already shaky even in the traditional paradigm of the university, because there has never been any
empirical support for it, and the phenomenon of professors who publish little but are regarded by their colleagues as outstanding trainers of graduate students is well known in most graduate schools.

Third, the assumption that graduate programs can be evaluated only by measuring volume of research output is both illogical and unfounded. Evidence on student developmental outcomes would seem to be at least as relevant as measures of faculty research outputs, but there has been no attempt to develop such measures in Ontario or elsewhere at the graduate level. Possibly the preference for assessing research accomplishments over those of teaching and service reflects a belief that research is easier to measure, as in the story of the drunkard who lost his keys in a dark alley but looked for them under the street lamp because the light was better there. However, the quality of research is as difficult to measure as the quality of teaching or service. It is only because we have accepted the convention of assessing research by counting publications and dollars of research grants that the illusion is created that research is easier to assess. In this connection, A. E. Malloch suggests that when professors cannot even imagine another way of evaluating performance than by publications and research grants, “a threat to academic freedom is beginning to take shape.” He adds that presently intellectual activity expresses itself at least as often in process as content and voices concern that scholarly activity other than that reflected in scholarly journals and grants is seen as “suspect or illegitimate” [41, p. 27].

The third issue pertains to the responsibility of scholars to communicate to the interested general public. Earlier we noted Jacoby’s lament for the disappearance of the public intellectual and the consequent impoverishment of public culture, which he believes is a result, above all, of the culture of the university. Though there may be attitudes within the university which mitigate against writing for the public, it is only through a mechanism such as the Ontario appraisals process that such attitudes are operationalized in a system of rewards and punishments that shape the corresponding faculty behaviour.

Ironically, when asked their opinion about the legitimacy or importance of communicating with the public, many academics tend to regard this as a responsibility of scholars. At a symposium on this issue, sponsored by the Social Sciences and Humanities Research Council of Canada, most of the participants rejected the view that popularizing knowledge is appropriate for a journalist but not a scholar for whom it is a burdensome distraction [57, pp. 17–18]. In principle, it might be possible to address this responsibility by writing simultaneously for
specialist and generalist audiences. For example, Steve Klees has argued that the major claim to expertise on the part of social scientists is not "our specific theoretical and empirical worldview but that we are trained to be articulate, thoughtful examiners of ideas and their implications. . . ." [35, p. 607]. However, that this is not a generally held view will be appreciated by anyone who has had a manuscript rejected by a journal with a comment that although it provides a good treatment of the issue, it is a journalistic rather than a scholarly piece. Irving Louis Horowitz, a journal editor of note, has summed up the situation well: "In the social sciences . . . most journals have such finely tuned methodological criteria for publication that the method, rather than the findings or theory, becomes central" in determining what gets published [28, p. 167]. Thus, insofar as communicating with the public is a responsibility of scholars or their chosen creative outlet, it is pursued only at great peril when an appraisals system such as that of Ontario looms over their heads.

The appraisals process has additional implications with respect to variation in views about teaching, but because these do not necessarily involve restrictions on inquiry, they will be noted only briefly. The value that only those with the highest undergraduate marks should be admitted to graduate school conflicts with the desire of some faculties to consider other factors in admission that may be more relevant to career success and may foster social change through helping minority and the socially disadvantaged to penetrate various professions. The very rigid model of program requirements assumed in the appraisals has retarded institutional differentiation and prevented even the most tentative experiment with nontraditional education in Ontario universities [55]. The demand for such education among Ontario students has resulted in a number of American universities offering such programs in Ontario. Presently, the off-campus programs of American universities in Ontario are outside the jurisdiction of the appraisal system, but there has recently been considerable controversy surrounding the recommendation of the Council of Universities that these programs be subjected to the appraisals process. [56, pp. 73–74].

Conclusion

In the foregoing I have attempted to identify some of the ways in which a system of program evaluation such as Ontario's works to suppress diversity, innovation, and nonconformist approaches in the search for knowledge. Admittedly, a system of program appraisal may
be only one of several factors that generate pressure toward conformity in academe. However, it may be one which is more amenable to change than the more deeply rooted factors, such as the incentive and reward system which confer promotion and status on individual professors, or the inherent cultural traits and tendencies discussed in the first part of this article. As well, a program appraisal system such as Ontario's can be a powerful vehicle for the dysfunctional effects of paradigmatic ethnocentrism through the practice of placing in the hands of a small number of people the authority to impose uniform definitions and standards of quality across all disciplines and programs. The weaknesses inherent in this practice are exacerbated in the Ontario case by the disproportionate representation of one branch of the university—the sciences—and by the absence of opportunity for faculty at large to participate in the determination of the methods and criteria for assessing quality.

The Ontario experience calls into question the appropriateness of a total system-wide application of the connoisseurship model; that is, having a single group of connoisseurs make quality judgments for all programs. This feature of the system for appraisal of graduate programs stands in sharp contrast to the accepted practices for accreditation of professional programs. In the latter case, the responsibility for program evaluation is vested with multiple groups of discipline-specific connoisseurs, drawn from the various disciplinary communities. Graduate programs in most disciplines are as specialized as the professional programs that are subject to accreditation, and frequently more so, for example, compare a doctoral degree to a first professional degree. Given the great diversity in perspectives, priorities, and methods among different graduate programs, it is difficult to see any justification for the approach which involves the single connoisseur looking over all disciplines. Indeed, the very idea of "connoisseurship" over the enormous range of programs that characterize the contemporary university is a contradiction in terms.

Some might worry that having different authorities for appraisal of each discipline would contribute to the weakening of the corporate coherence of the university, exacerbating what Allan Bloom calls the "decomposition of the university" [13, pp. 347–456]. However, Bloom argues that one of the major factors contributing to this trend has been the inappropriate pressure on the humanities and social sciences to emulate the methods of the sciences, thus sapping the latter branches of their vigour, if not their raison d'être. At any rate, the differences among the disciplines are a reality that no appraisal system can change;
an inappropriate appraisal system can only retard development of
some disciplines by casting them in a framework which is suitable for
others. Ultimately, what binds the disparate parts of the university to-
gether is their mutual interdependence — which in no small way is
derived from their diversity — and their common commitment to the
search for truth. Both would be enhanced considerably by a decentrali-
ation of the appraisals process in which the autonomy of multiple
centers of connoisseurship were recognized.

References


to the Academic Community*. Washington, D.C.: Association of American Col-

1985.

and Sons, Inc. 1971.


8. Barber, B. “Resistance by Scientists to Scientific Discovery.” In *The Sociology of
1962.

Canada’s Universities on the Road to Ruin*. Toronto: McClelland and Stewart,
1984.

11–18.

Bass, 1983.

12. Bleything, W. B. “On the Workload of Faculty: Defining Faculty Workload.”

1987.


15. Clark, B. R. “Faculty Culture.” in *ASHE Reader on Faculty and Faculty Issues in
Colleges and Universities*, edited by M. J. Finklestein, pp. 129–42. Lexington,


