DOCUMENT RESUME

ED 446 470 HE 033 224

AUTHOR	Smith, David C.
TITLE	"How Will I Know If There Is Quality?" Report on Quality Indicators and Quality Enhancement in Universities: Issues and Experiences.
INSTITUTION	Council of Ontario Universities, Toronto.
REPORT NO	COU-674
ISBN	ISBN-0-88799-348-6
PUB DATE	2000-03-00
NOTE	40p.; For companion report, "Will There Be Enough Excellent Profs?" by the same author, see HE 033 225.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Educational Planning; Educational Policy; *Educational Quality; Enrollment Trends; Evaluation Criteria; Evaluation Methods; Faculty Recruitment; Foreign Countries; Higher Education; *Institutional Evaluation; Program Effectiveness; Research; Standards; *Universities
IDENTIFIERS	Council of Ontario Universities

ABSTRACT

In this paper presented by the Council on Ontario Universities, quality indicators are considered in light of their relationship to the challenges Ontario's universities expect to face in the next decade. These challenges include a surge in enrollment; the need for new, first-rate faculty; and Ontario's potential role in leading innovation and research. The report makes five recommendations in the following five areas to policymakers and planners in Ontario's higher education arena regarding quality indicators and quality enhancement. First, Ontario's universities are at a critical juncture in the quality of their work, and in striving to enhance quality, they face major challenges. Second, quality refers to excellence in the work of universities. It is considered in terms of the institution, academic program or department, and the university sector. Third, quality indicators refer to evidence of characteristics associated with quality and overlap with performance indicators. Fourth, the paper looks at Ontario's specific experience with indicators at three levels. Lastly, international perspectives on quality are considered. One appendix is included. (HB)

"How will I know if there is quality?"

REPORT ON QUALITY INDICATORS AND QUALITY ENHANCEMENT IN UNIVERSITIES: ISSUES AND EXPERIENCES

by

David C. Smith

This report was prepared for the Council of Ontario Universities. The views expressed here are those of the author and do not necessarily represent those of the Council or any of its members.

It is a companion report to *"Will there be enough excellent profs?"* which examines prospective demand and supply conditions for faculty over the next decade. Both reports are complementary to the 1999 report by PricewaterhouseCoopers for COU entitled *"Will there be room for me?"* which examined the expected large increase in student demand for university education in Ontario.

March 2000

Council of Ontario Univerisites
Conseil des universités de l'Ontario
180 Dundas Street West, Suite 1100, Toronto, Ontario M5G 1Z8
Phone: (416) 979-2165 Web Site: http://www.cou.on.ca

COU № 674 ISBN 0-88799-348-6

TABLE OF CONTENTS

SUMMARY AND CONCLUSIONS

"Not everything that can be counted counts; and not everything that counts can be counted."

This remark, attributed to Albert Einstein in another context, summarizes succinctly a couple of basic points about quality indicators. They are not simply compilations of data that draw together anything that can be counted about universities. A careful sorting of data for relevance to quality is required. At the same time, not everything that is important for quality can be represented by objective indicators; qualitative judgements must play a role.

This report examines quality indicators but considers them from a broad perspective in the search for quality in universities. The following are major conclusions and points of summary.

1. Ontario's universities are at a critical juncture in the quality of their work. In striving to enhance quality, they face major challenges:

- To accommodate a large expected surge in enrolment in the coming decade;
- To find and fund first-rate faculty who will be needed to teach the larger number of students and to replace an unusually large wave of anticipated retirements; and
- To fulfil their basic functions of education and research and to assist in meeting rising expectations about the leadership role Ontario could play in innovation and research.

These challenges arise at a time when a prolonged period of public funding constraints has eaten away at the intellectual and physical infrastructure of the institutions, and much uncertainty about future funding makes planning difficult.

2. Quality refers to excellence in the work of universities. It is considered here at three levels: the level of the institution; the level of the academic program/department; and the level of the university sector.

3. Quality indicators refer to evidence of characteristics associated with quality. They are related to, and often overlap with, performance indicators, but the latter are usually directed more to management objectives and results than to portraying the meaning of quality.

4. Ontario's experience with indicators is reviewed at three levels:

At the institution level:

- The government's three mandatory indicators are a limited, incomplete approach to recording institutional quality.

- The *Maclean's* ratings have serious weaknesses but have stimulated efforts to inform students better about institutional characteristics.
- The initiative by institutions to develop indicators is progressing impressively in some respects but is still incomplete.
- The 1999 report of the Provincial Auditor has a focus on the mandate of the Ministry of Training, Colleges and Universities, but raises interesting questions for institutional quality and governance.

At the program level:

- Quality appraisals of graduate programs through the Ontario Council on Graduate Studies are highly respected domestically and internationally.
- With respect to undergraduate programs, audits of university processes for ensuring quality, undertaken by the Council of Ontario Universities, are at an early stage of development but are a promising approach.

At the university-sector level:

- Both the Ministry of Training, Colleges and Universities and the Council of Ontario Universities do important work on the functioning of the university sector in Ontario but, with the demise of the Ontario Council on University Affairs (OCUA) in 1996, there is no longer an organization devoted to this task. Unlike the U.S., there is not a range of private institutes and centres on postsecondary education that can serve this function.
- Recent studies for the Humanities and Social Sciences Federation of Canada indicate that differences among major discipline groups should be reflected in the development of indicators of performance and quality.

5. Ways of identifying, assessing and assuring quality are generating much interest in many countries:

- Reasons for this interest are not always the same. Thus, the most recent initiative in Australia is linked to providing assurances of quality in the export of higher education services.
- Three summary points are derived from the U.S. experience:
  - The development of performance-based budgeting that links state grants to a set of indicators has spread to less than half the states and tends to be limited to a small percentage of appropriations;
  - Private initiatives that rank U.S. institutions generate responses similar to those aroused by *Maclean's* in Canada, but there is an advantage to some competition in these ratings; and
  - U.S. success in combining mass higher education with exceptional quality in many institutions occurs not because of a "top-down" imposition of centrally administered conceptions of quality. Rather there is a highly competitive environment that provides a

stimulus to quality and that draws substantial private and public resources to the support of quality.

- Recent British experience with quality assessment has created great interest in Ontario. A few points taken from the report in the text are:
- Through a periodic research-assessment exercise, university departments are subjected to peer assessment of the quality of research of faculty members, and a rating (or type of quality indicator) is derived that affects the allocation of a substantial amount of research funds. The allocation is sharply skewed in line with these assessments of quality.
- With regard to teaching and learning, a separate quality review of all subjects in all universities is conducted. Extensive documentation is required, and site visits by teams of assessors are made. Grades are assigned for each subject review in each institution and are made public. Although the results do not directly affect funding, the grades are taken very seriously as reflecting on the prestige of a department.
- The British approach appears to have some beneficial effects on quality but at a substantial cost in terms of administrative burden and adverse effects on incentives. The approach is evolving, and critics are calling for reforms that would reduce the degree of intrusiveness of the present approach.

6. Issues and options in the design of an Ontario system of quality indicators at the three levels:

At the institution level: Quality indicators carry dangers, including the danger of excessive attempts to quantify at the expense of qualitative judgement; failure to recognize differentiation in missions of universities; undue costs in the development of indicators; and dangers of confusing symptoms and causes of quality. But there are benefits too, including salutary tests of claims to quality, improved information for student choice and better bases for satisfying demands for accountability. The following key features for quality indicators at the institutional level are proposed:

- A strong emphasis on mission-specific indicators approved by the governing boards as appropriate to the institutions;
- No broadening of a government-imposed set of indicators but an encouragement to institutions to develop among themselves some commonalties in the presentation of their indicators to facilitate sharing of information;
- A credible, non-interventionist mechanism for auditing data provided for indicators; and
- Accompanying statements about other quality assessment processes being used.

At the program/department level: The Connell report provides a helpful review and guide for further development of quality assessments of graduate programs in Ontario's universities. The undergraduate program review audit process, overseen by the Ontario Council of Academic Vice-Presidents (OCAV), is at an early development stage but provides a promising approach to check that processes for quality assurance are in place. It is not suggested that Ontario simply adopt the British model for subject reviews and research assessment at the program/department level.

At the university-sector level: Indicators and assessments of quality are developed in the Ministry of Training, Colleges and Universities and the Council of Ontario Universities. Also, on occasion, a task force, commission or panel has been appointed that briefly delves into quality measurement and enhancement issues in the university sector. The argument is made here for strengthening this type of work, whether through present institutions or through a special new organization.

7. ***A key conclusion*** is quality indicators should be considered as only a component of a broader, more powerful approach to quality enhancement. They reveal some symptoms of quality but they do not show the causes of quality nor the conjunction of resources and incentives that produce quality. This report concludes with an emphasis on fostering a competitive and collaborative environment for universities, within a university sector that is funded reasonably in line with competing university sectors.

8. Five general recommendations on quality indicators and quality enhancement are made:

General Recommendation 1:
Each university should release on its web site an annual quality assessment report with a list of quality indicators approved by its governing board as relevant to the institution's mission. Concerns about the reliability and comparability of data in the use of indicators should be reviewed regularly through CUPA or the new organization proposed below.

General Recommendation 2:
Assurance of quality in academic programs is best provided in Ontario through further development of the current graduate appraisal program and the undergraduate program review audit. At this time a superior model has not emerged elsewhere.

General Recommendation 3:
The annual quality assessment report should incorporate, in addition to a set of quality indicators, the general results of any graduate and undergraduate reviews that have been completed, and a statement of the various approaches to quality assessment practised in the institution and of any planned changes.

General Recommendation 4:
If great care is taken with its structure at the time of establishment, a new organization or institute would be helpful to monitor the quality assessment reports of universities, and conduct analyses of quality in the sector as a whole and in relation to university sectors in other jurisdictions. Moreover, it could advise on quality issues if new private or public universities or new types of academic credentials are proposed. Among the critical features of such an institute should be an emphasis on first-rate analytical and empirical work; small size; broad mandate; published reports; openness to new ideas; careful consultation with university stakeholders; a hands-off philosophy with respect to directing institutions; and periodic critical external reviews of its performance.

General Recommendation 5:
Primary emphasis should be placed on a strong competitive and collaborative environment with

resources comparable to those in competing university sectors. This environment will be the most powerful influence on quality.

Section One

CHALLENGE OF ENHANCING QUALITY

Ontario's universities are at a critical juncture in the development of quality. A prolonged period of public-funding constraints has eaten away at their human and physical infrastructure. Much greater pressure on resources is anticipated as a result of current projections of close to a 40% increase in enrolment pressures from the end of the '90s to 2010-2011. Over this period an unusually large planned retirement wave of faculty, along with an enrolment-generated need for more faculty, will mean a scramble for a very large number of new faculty, estimated, for example by PriceWaterhouseCoopers to total about 10,000, a number roughly equivalent to four-fifths of the current complement of faculty. The projections do not include replacements for faculty who leave their positions in Ontario universities for reasons other than retirement, and new faculty who are required to support the higher levels of research and innovation expected by governments and the private sector. Thus, the total projected figure is considerably higher. These projections are reinforced by a second set of projections undertaken by a committee of the Council on University Planning and Analysis (CUPA). The adjustments to these large changes will be taking place in an environment of greater international and domestic competition for the bright professorial minds universities seek, and students need.

Will such pressures pull quality down a steep decline? Or are there other forces that will produce great opportunities for quality enhancement? As important as such questions are, this report examines more limited, though related, questions about what is meant by quality and how it might be represented by indicators and peer-adjudicated assessments. These questions lead, of course, to more fundamental questions about whether such representations tell us anything about what causes quality in universities and what can be done about it.

Section Two

MEANING OF QUALITY AND PURPOSES OF INDICATORS

Quality is defined in *The Concise Oxford Dictionary* as "general excellence" or, in its adjectival form, as "possessing high degree of excellence." In the context of universities, the word might be applied to various levels of activity – a course, program, department, faculty, institution, or group of institutions. A particular course in economics might be described as low quality but the economics department or the university in which it is being taught as high quality. Quality refers to how well the desirable functions are being fulfilled, and the degree of fulfilment can be different in different parts of the university.

How does one determine the presence of quality? Most university presidents will be pleased to tell willing (and even unwilling) listeners about the superb quality of their institutions. Trustworthy as such people usually are, their assertions do not fully satisfy everyone, since quality might be simple to define but it is extraordinarily difficult, if not impossible, to measure objectively. One route to more convincing quality assessments is through various formal processes of developing subjective judgements. These processes could take the form, for example, of reviews by independent experts. Another route is to specify some representations or indicators of quality and use them to suggest an objective specification of quality. Though we may not know how to measure quality directly, quality indicators can be proposed as indirect representations of quality.

This latter route – the quality indicator route – to identifying the presence of quality has become highly popular in many countries in recent times. But it has not been a simple route that has easily or quickly converged on the goal. Reference has already been made above to the need to specify which parts within or across universities are being considered. Universities and their parts also have complex functions of education and research, and each function has distinctive quality characteristics. In addition, universities have different missions, and statements about quality should reflect these differences. Moreover, universities have various stakeholder groups, and their interests in what should constitute quality will differ. Students, faculty, parents, alumni, and public officials will often have different interests in quality and emphasize different aspects of it. Thus, it should not be surprising to find there is no ready-made set of quality indicators that are universally accepted and applicable.

A further complication is that indicators designed as indirect representations of quality may overlap indicators designed for other purposes. Particularly popular are performance indicators, which are often designed to help with the management of a university and, for that reason, carry at times a second label – management indicators. They show the extent to which performance approaches specified targets. They tend to be highly specific to an institution and weighted toward output measures. Nevertheless, in current usage it is sometimes difficult to maintain a clear distinction between quality and performance indicators.

A third shading of difference among types of indicators is sometimes drawn, arising again from the purpose to which indicators are put. It may be that the interest is centred on the accountability of the institution, or parts of it, to policy objectives set internally or externally. The extent to which there is a response to the specified objectives becomes important, and a set of indicators for responsiveness or accountability is emphasized. Here, too, there may be an overlap with indicators for quality or performance.

In this report, a purist's distinction among types of indicators will not be maintained. The focus here is on quality indicators and their relevance for quality enhancement. But, if there is some straying across the borders of indicator domains, the reader is asked not to be too upset, for such waywardness is the way of the world out there.

Section Three

ONTARIO'S EXPERIENCE WITH INDICATORS

Ontario has not rushed to embrace the use of indicators. But cautious steps have been taken by the provincial government and many of the 17 universities, and there is the controversial *Maclean's* magazine ranking of Canadian universities, based on a set of indicators and published annually since 1991.

GOVERNMENT'S MANDATORY INDICATORS

On February 13, 1998 the Ministry of Education and Training announced in a press release a requirement for three mandatory indicators:

> In order to assist students in making an informed program decision, postsecondary institutions that offer Ontario Student Loans will be required to make available to students for September 1998 information about graduation rates, placement rates and loan default rates for their programs. Institutions that do not currently collect information on graduation and placements must begin to do so for September 1999.

This step was not simply a requirement for three indicators to help students make informed program decisions. A memorandum from the Ministry, dated November 19, 1998, clarified its intention to use the indicator for default rates to reduce costs of defaults to the government's loan program, and the target for this performance indicator was not really universities. Thus, the memorandum stated:

> The 1998 default rates for each sector ranges from 12.3% for universities, to 25.4% for colleges, to 34.5% for private vocational schools and 12.1% for other private and publicly funded institutions. These rates are unacceptably high. The Ministry's business plan requires that the overall default rate be reduced to 10% within the next four years.

Grumbling from universities arose largely from difficulties of measuring graduation and placement rates, and from concerns about the value of such indicators in relation to the cost of assembling them. Issues of measurement were resolved by not leaving interpretations and collection of data in the hands of universities but by adopting the following definitions and methods of collection:

- Graduation rates released in May 1999 were compiled from "the percentage of Year 1 students in bachelors or first professional degree programs in 1990 who subsequently received a degree between 1991 and 1997. The data were compiled by the Ministry and verified by the universities."

- Graduate employment rates, also released in May 1999, were compiled from "the percentage of 1996 graduates of bachelors or first professional degree programs who were employed six months and two years after graduation. The rates were collected through a survey of university graduates conducted by the Ontario Universities' Application Centre."

The results for graduate employment rates were clearly inconsistent with popular notions of the unemployability of university graduates. The average employment rate of university graduates two years after graduation was 96.7%, far above the national average. The range of variation among 16 of the universities was between 96.2 % and 97.9 %, a range that does not convey to students much significance in the selection of a university, if employability is the consideration. The one university outside this range was at 92.6%, but it is in a northern community with a much above-average general unemployment rate.

The results for graduation rates showed much variation among universities and programs. Thus, this indicator ranged between 46.3% and 90.0% among universities. But it is a controversial indicator with serious conceptual and measurement issues surrounding it. It is influenced by such factors as admission policies and types of programs. The value for student use at this time is seriously limited.

BROADHURST TASK FORCE REPORT

In May 1993, a report was published by the Task Force on University Accountability, which had been appointed by Ontario's Minister of Education and Training, and which was chaired by William Broadhurst. Asked to make a comprehensive review of the accountability practices of Ontario universities, the Task Force emphasized in its report a central theme that the governing body of each university should be "the primary and most effective locus of institutional accountability." It went on to explain that:

> "... each institution should have a mission statement and accompanying academic and financial plans sufficiently specific to permit assessment of progress toward the fulfilment of its mission . . . the Task Force recommends that each governing body determine the measures to be applied to assess performance. It suggests that these include performance indicators."

An appendix to the report provided the first full-scale review of indicators in Ontario. It was prepared by a COU committee – CUPA Committee on Accountability, Performance Indicators, and Outcomes Assessment – chaired by Daniel Lang.

The CUPA committee sifted initially 80 potential indicators using, as the screening test, the characteristics of relevance, reliability, accessibility and clarity. Emphasis was placed on "flexibility of use from institution to institution and from program to program" (p.15). It also classified the indicators according to five purposes in their use – quality, performance, resources, mission, and responsiveness – and showed that a particular indicator can serve more than one purpose.

In addition to the call for the use of indicators as part of the exercise of institutional accountability by governing boards, the Task Force recommended the establishment of an external monitoring agency at arm's length from both the universities and government. It would monitor institutional accountability through a desk review of the university accountability reports, produced every two years, as well as an on-site review conducted on a seven-year cycle and based on peer review. The reports of the monitoring agency would be public, identify any weaknesses in procedures and encourage improvements where needed.

The work of the Task Force gave a strong impetus to thinking about indicators in universities, but the monitoring agency was not established. The recommendation was to establish it in association with the Ontario Council on University Affairs, but that intermediary body was soon to be disbanded.

INSTITUTIONAL INDICATORS DEVELOPED BY UNIVERSITIES

Universities in Ontario have long used indicators for internal purposes, but a survey conducted for this report reveals a marked acceleration in their development in the past four or five years. Of the 17 universities, 10 have developed a set of indicators, four are in the process of developing them, and the other three stated they preferred to emphasize other methods of assessing quality and performance. This widespread development in the use of indicators has been occurring largely since the mid-1990s.

The purposes for the indicators were not always the same. In some cases they were clearly expressed as performance indicators to assist management and the board. The publication, *Nipissing University Management Indicators,* is an excellent example of this approach. In other cases, the emphasis is more on indicators of quality to provide information for student choice. An example of this approach is the Queen's publication, *Measuring Excellence: 20 Indicators of Performance* (1998 Update).

Not all universities are enthusiastic about this route to performance and quality assessment. The three institutions that reported they were not currently engaged in the development of indicators emphasized their interest in, and practice with, other methods of quality evaluation. Some other institutions that do use indicators pointed out drawbacks and also noted the importance of other forms of quality evaluation.

A common theme in the responses was the need to recognize differentiation among universities – differentiation in missions and differentiation in observable characteristics. The usefulness of quality indicators depends crucially on recognizing this differentiation and on avoiding trying to shoehorn different institutions into a common mould.

MACLEAN'S APPROACH TO INSTITUTIONAL INDICATORS

Since 1991 *Maclean's* magazine has produced an annual survey and ranking of Canadian universities, using its selection of quality indicators and weights to be attached to them. Initially all institutions were ranked together under a single set of indicators, but now institutions are separated into three groups and ranked in each group by a weighted set of indicators that range between 20 and 22 in number. The three groups are:

Medical Doctoral: Universities with a broad range of PhD programs and research, as well as medical schools.

Comprehensive: Universities with a significant amount of research activity and a wide range of programs at the undergraduate and graduate levels, including professional degrees.

Primarily Undergraduate: Universities largely focused on undergraduate education, with relatively few graduate programs.

The rankings have been controversial, with criticisms arising – not surprisingly – more strongly from institutions that have not fared as well in the rankings. But there has been more than institutional pride involved. Three problems continue to cast doubt on the accuracy and relevance of the rankings:

- ***Integrity of the data.***
  Maclean's puts considerable effort into checking the accuracy and comparability of data used in constructing the indicators. But its resources for this task are not extensive, and considerable dissatisfaction is expressed by many university leaders about data that are used.

- ***Differentiation of mission and characteristics of universities.***
  The three-fold classification reported above provides only a very limited recognition of the different missions and basic characteristics of institutions. For example, York University and the University of Victoria are both classified in the Comprehensive group, but there are enormous differences in characteristics between them.

- ***Differences among students in their preferences.***
  Students will not have the same set of preferences with regard to desirable characteristics of universities. As a result, they weight indicators differently, and rankings will be affected. Class size, for instance, may be a heavily weighted indicator for some students but less heavily weighted for others.

Solutions to these problems have not been easy to find. It would probably help to have some competition in the survey and rankings in order to show that many different rankings can be compiled, depending on the indicators and weights selected. This variety is evident, as will be noted later, in the many rankings published in the United States.

A recent paper by Ian Clark on the evolving functions of the Council of Ontario Universities draws attention to the long tradition of "universities' vigorously maintained autonomy" and commitment to self-regulation in quality assurance of academic programs. He notes also that "one of the strengths of the Ontario university sector is the differentiation in mission, program mix, academic specialization and research focus among the 17 institutions." It is thus not surprising that Ontario universities have not warmly embraced rankings based on a restricted set of characteristics and weights.

QUALITY INDICATIONS OF PROGRAMS

A. Graduate Programs

Through COU and its Ontario Council on Graduate Studies (OCGS), a system for assessing the quality of graduate programs in Ontario has developed well. The oldest assessment program of its kind, it also maintains a high international reputation.

While it has evolved over time, the key features of the system have been reliant on peer adjudication by an expert, independent team; the submission of new proposed graduate programs for appraisal before they become eligible for public funding; and a review cycle that brings all graduate programs under periodic reappraisal.

A review of this quality-assessment approach was undertaken recently by Dr. George Connell, former president of the University of Toronto, and his report was released by COU in 1999. His recommendations are under review, but essentially his report reaffirmed the importance of this approach for quality assessment in Ontario.

B. Undergraduate Programs

Universities usually develop internal procedures for quality assessments of their undergraduate programs. Typically, reviews take place on a regular cycle, such as every five years. When the head of a program or department is up for renewal or replacement, a major review will often involve an external consultant.

In October 1996, COU took the initiative to develop regular audits of undergraduate program reviews in Ontario universities. Under this initiative, the processes that a university follows to ensure quality are examined by teams of highly respected, independent experts. Unlike at the graduate level, the undergraduate programs themselves are not subject to formal appraisal, which would be a huge task to do carefully. The Ontario Council of Academic Vice-Presidents oversees this initiative. A fuller description of the procedures is provided in the Appendix.

QUALITY INDICATIONS OF THE UNIVERSITY SECTOR

The unit of observation for assessing quality in the work of universities may change from a program, to an institution, to the university sector. As the level of aggregation rises, quality becomes less identified with specific teaching and research activity, and this separation is a source of concern. Thus, the appendix to the Broadhurst report, cited earlier, contained a warning:

> Despite various attempts to generalize about quality, in the end quality indicators operate best at the program level within individual institutions. The higher the level of their aggregation, the less useful quality indicators become. (p.3)

Yet, there is also a sense in which indications of quality need to be considered in a sector-wide context. The Ministry of Training, Colleges and Universities is expected to take this broad perspective. The Office of the Provincial Auditor took this perspective when it examined, in a recent report, the exercise of the mandate of the Ministry with respect to universities. The former intermediary body, the Ontario Council on University Affairs (OCUA), took this broader perspective. The Council of Ontario Universities brings the collective perspective of its individual member institutions to sector-wide issues, including quality issues. Interest has grown recently in how the quality of the domestic university sector can affect international and interregional competition. This interest is reflected in such reports as *A Road Map to Prosperity* (Ontario Jobs and Investment Board, 1999) and *Growing Ontario's Innovation System: The Strategic Role of University Research* (Heather Munroe-Blum, 1999).

Commissions, task forces, panels and other such bodies have also been appointed on occasion to assess the functioning of Ontario's university sector. But with the passing of OCUA, there is no body charged with an ongoing assessment. The significance of this gap is considered in a later section.

Section Four

SOME CHARACTERISTICS OF EXPERIENCES ELSEWHERE

PURSUIT OF THE NATURE OF QUALITY FROM "A" TO "U"

Ways of representing, assessing, and assuring quality in universities are the source of intense interest internationally. No country with an advanced postsecondary education sector has totally escaped, it seems, from much advice and some action on this matter. From "A" for Australia to "U" for the United States, the search for how to identify, assess and assure quality continues.

"A" for Australia

Australia's Minister for Education, Training and Youth Affairs, Dr. D. A. Kemp announced on December 10, 1999, yet another proposed model for that country's approach. Care was taken to make clear it was not simply a replica of what some other country is doing. As he stated:

> ... there is agreement that any new framework should take into account our federal structure, our universities and our history rather than slavishly adopting a British model (with its expense and intrusiveness), a U.S. model (which involves industry regulation within a very diverse system) or a European model (with close links between universities and the state.)

Indeed, there is an interesting distinctiveness to the model he proceeded to outline, a distinctiveness that is related to providing assurance of quality to assist the export of Australian higher education services in a competitive market. As he stated:

> Education is now one of Australia's major export industries in an intensely competitive market. While Australian universities compete with each other in this market, they also compete with the rest of the world. Our major competitors have external quality assurance mechanisms, and countries in our largest markets look to government verification of quality standards. To maintain market position, we need to be able to advertise that we have quality assurance mechanisms in place, that they are being applied, and that they are having a positive effect on outcomes.

A key feature of the new model is to be an Australian University Quality Agency that will conduct audits and prepare reports on claims made by institutions and thereby provide evidence "to assure students, the community and the outside world." The Minister states it will be a system that "recognizes the autonomy of higher education institutions," "places the responsibility for the quality of provision on individual universities," and "signals to the community and the rest of the world that the quality of the higher education system is assured through a rigorous, external audit of university quality assurance processes."

In theory, the newly announced Australian model initially sounds similar to the Ontario undergraduate program review audit that was discussed earlier in this report and is summarized in the Appendix. Two major differences are, first, that the Australian model is to "involve a whole of institution approach incorporating teaching and learning, research and management both onshore and offshore." Second, a new agency at arm's length from the government and universities is to be established that will audit institutional self-assessments on a five-year cycle and that has power to impose penalties. One must await implementation of the model to know how it is likely to develop, in practice.

Further Along the National Alphabet

Travelling along the alphabet, one should mark "B" for Britain for a visit because of the enormous interest and controversy surrounding its recent approach to quality. It is tempting to pause at "N" for the Netherlands and admire the quality assessment processes that country has in place, processes that share much international acclaim with Ontario's graduate appraisal processes. But one wishes for time also to understand steps taken in another "N" for New Zealand and in "S" for Sweden – and to check back on information that "H" for Hong Kong has some interesting goings-on regarding quality assessment. Skipping over experiences – tantalizing for study – in many countries, one comes to the big "U" for the United States.

"U" for the United States

One could pause for a long time to explore the case of the United States in view of its enormous diversity of postsecondary institutions and policies. But only three features of U.S. experience – judged to be of particular interest for the subject of this report – have been selected here and will be briefly highlighted.

One is the use of performance-based budgeting. In 1979 the State of Tennessee took the lead in implementing budgets that rewarded postsecondary institutions in relation to their performance against specified indicators. Many other states followed until, in 1999, 22 states had budgeting policies that tied at least some of the appropriations to public colleges and universities to their success in meeting certain measurable indicators. That number means, of course, that 28 states – over one-half – have not adopted this approach.

The difficulties that some states have encountered in implementing performance-based budgeting are well illustrated in an article in the *Chronicle of Higher Education* (July 2, 1999). The case of South Carolina is examined, which adopted a law in 1996 to implement performance funding. Sponsors of the law called for 100% of state funds to be tied to indicators. After three years and an enormous amount of effort, it is now expected that no more than 5% of any institution's budget will be so determined, and criticisms continue to cast some doubt on the future of the initiative. Not all assessments of the new system are negative, but criticisms include the amount of time and money spent on gathering and reporting data – some of which is viewed as irrelevant – and the danger of giving the legislature too much say in university policies. Also, university officials

complain that the expected improvement in state funding to accompany evidence of performance did not take place.

Is there a pattern for survival emerging from the many U.S. experiences with programs that use indicators to determine government funding? Joseph C. Burke, director of the higher-education program at the Rockefeller Institute of Government, studies performance financing and is quoted in the same edition of the *Chronicle* as stating: "The history of performance funding is that programs that fail to protect budget stability and campus diversity fail to last for long."

A second U.S. feature to highlight briefly is the variety of institutional rankings based on different sets of indicators purporting to reflect quality. The best known is the ranking by *U.S. News & World Report*, which began an annual ranking in 1983. Initially limited and based solely on a survey of reputations of colleges and universities, the ranking has become a larger, more complex exercise. In 1999 *U.S. News* ranked 1,323 institutions categorized by broad mission and region. Using a modification of the categories developed by the Carnegie Foundation for the Advancement of Teaching, the report classified institutions (with numbers surveyed in brackets) as: national universities (228); national liberal arts colleges (162); and regional universities (504); regional liberal arts colleges (429). Regional institutions are further subdivided into regions: North, South, Midwest and West. Data are gathered for up to 16 indicators judged to be relevant for academic quality, and indicators are assigned a weight based on judgements about their relative importance. The top institutions are ranked, and the remainder are grouped into tiers. Reporting on "How *U.S. News* Ranks Colleges" in an Internet article in August 1999, Amy E. Graham and Robert J. Morse summarized the indicators used:

> The indicators we use to capture the dimensions of academic quality fall into seven categories: academic reputation; retention of students; faculty resources; student selectivity; financial resources; alumni giving; and (for national universities and liberal arts colleges) "graduation rate performance," the difference between the proportion of students expected to graduate and the proportion that actually does. The indicators include input measures that reflect the quality of students, faculty, and other resources, and outcome measures that signal how well the institution does its job.

The choice of methodology affects the rankings. Thus, in 1999 the method of weighting standardized test scores, per-student funding and student-faculty ratios was changed. Rankings of universities with strong science and engineering programs were boosted, and the California Institute of Technology took first place. (Over the past decade it was consistently ranked in the top 10, but this is the first year it has been first.) It is small in size with only about 900 undergraduates. But in its favour for current ratings is its emphasis on science, and its spending per student on instruction, research, and education-related services. This spending is about $192,000 US, more than double that of any of the so-called "big three" - Harvard, Princeton, and Yale. The methodology this year allowed for the size of differences in indicators to be taken into account.

The use of weighted quality indicators to rank universities bears many similarities to that taken by *Maclean's* in Canada. And it raises similar sharp criticisms. Thus, an article in *U.S. News* quotes president Gerhard Casper of Stanford University as commenting: "... much about these rankings -

particularly their specious formulas and spurious precision – is utterly misleading." Such sentiments would flow easily from the mouths of many Canadian university presidents if

asked to comment on *Maclean's* rankings. But a distinctive and healthy feature of U.S. higher education is the broad range of guides and rankings prepared through other outlets, a reminder of the caution with which any particular ranking should be treated. Among these other outlets are:

- *Princeton Review* which categorized "The Best 331 Colleges" in August 1999 based on a survey of nearly 60,000 students nationwide. The survey emphasizes the qualitative and anecdotal rather than the quantitative, but provides candid assessments of campus life at many institutions.

- *Time* which publishes a College Guide and has joined the *Princeton Review* in jointly naming a college of the year.

- *Kiplinger's* magazine which replaced *Money* magazine this past year in selecting "best buys" by ranking institutions based on value, educational quality and student satisfaction.

- Other proposals for university guides have been made, including one from Stanford – so far unsuccessful – to develop an alternative to *U.S. News*.

A third particularly interesting feature of the U.S. postsecondary education scene from the perspective of the subject of this report is the evidence of quality that has emerged from the competitive behaviour of leading universities. Quality has not been imposed from above through governments requiring adherence to quality indicators. In the *U.S. News* ranking of national universities the top 19 are private (though they are heavily dependent on public research funds). Quality has developed through the competitive interaction of both private and public institutions, in some respects within regional clusters, and in other respects between regional clusters. Competition for faculty, students and funds has been a source of quality, but this competition has also been modified by collaboration in promoting key areas of higher education policy.

Alan Ryan, Warden of New College, Oxford, has written much on why he thinks the U.S. has been more successful than Britain in combining high quality and mass higher education. In an article last year in *Prospect,* he argued:

> To combine successfully mass higher education with elite excellence, it [the British university system] must follow the decentralised U.S. model. It must introduce differentiated tuition fees, allow a mixture of public and private funding, and give up the pretence of uniform standards....
>
> The U.S. combines mass higher education with elite excellence (in both its state and private sectors), but it only achieves this through allowing wide diversity in standards, salaries, tuition fees and so on. This is the direction in which Britain is inevitably heading; we can either get there in two years or 10....
>
> Princeton has a Nobel Prize-winning physics department because it is always looking over its shoulder at its rivals, not fidgeting about the research-assessment

exercise.

This decentralized, competitive environment has provided a stimulus to quality. It has drawn much financial support to the universities from the private sector. But whether through public research support or through state appropriations, it has also elicited strong financial backing from institutions of government.

REPORT ON THE BRITISH APPROACH

Limitations on time do not permit a broad exploration of experiences in other countries, and descriptions of international experiences can be found in a number of studies. For example, a recent useful one - released in 1999 - that emphasizes performance indicators was prepared for the Humanities and Social Sciences Federation of Canada by Janet Atkinson-Grosjean, Garnet Grosjean, Donald Fisher, and Kjell Rubenson of the Centre for Policy Studies in Higher Education and Training at the University of British Columbia. It was thus decided to concentrate attention in this section on recent British experience, which has attracted more interest in Ontario in recent years than experience in any other country.

The following brief examination is based on interviews conducted in England in early February 2000 and on a reading of some relevant documents and other material. The intention is to consider salient features of interest for discussions of higher education policy in Ontario, not to try to provide an authoritative detailed account of the British approach.

A. The Context

The current pervasive role of quality reviews and assessments in British universities is related to the historical evolution of British education in general, with its emphasis on elite values and reliance on inspectorial methods. But it has also been strongly influenced by fundamental reforms to higher education, particularly since the early 1990s and the passage of the Further and Higher Education Act of 1992.

The reforms included:

- A conversion of all polytechnic institutions into universities.
- A deliberate massive expansion of enrolment in higher education that has doubled the participation rate to about 33% between the mid-1980s and the mid-1990s. (In more recent years the rate has not risen, but the Prime Minister has set a target for a further rise to 50%.)
- New institutional mechanisms for the allocation of government funds to universities. The Higher Education Funding Council for England (HEFCE) was established, with counterparts in Scotland and Wales. (The allocation is through the Department of Education in Northern Ireland.)

Amid such changes, various factors have created especially strong pressures to develop new forms of quality assessment. Five factors are suggested here:

1. The polytechnics, at the time of their conversion to universities in 1992, had stronger forms of official quality reviews than universities. It was felt that adjustments in standardizing policies in the newly unified sector should not be in the direction of weakening the official quality review processes of polytechnics.

2. Britain's planned move to a mass higher-education system raised concerns about the extent to which quality – long viewed as a hallmark of British higher education – would be compromised. Interest increased in ways of assuring quality.

3. Boards of Governors (often referred to as Councils) do not play as clear and as strong a role in accountability processes as is the case in Ontario universities. As concerns about accountability for public funds grew in Britain – as they did in Ontario and elsewhere – less attention was directed to how boards might better serve accountability standards and greater attention was directed to mechanisms governments might develop directly or indirectly.

4. A feature of public attitudes in Britain in recent decades – again reflected in attitudes elsewhere – is a greater distrust of professionals. Universities have been caught up in this general shift in attitudes, a shift that has produced greater demands for more formal assurances of quality in professions.

5. The notion of students as consumers became more widespread. Tuition fees, which had long been insignificant, have moved up to about $2,500 Cdn (with a means test). Though still less than fees in Ontario and most parts of North America, their growth stirred greater interest in demonstrating the quality of educational services received by students.

As a result of such factors, British universities now have a wide variety of quality reviews, and they can be classified into six types. Three are long standing and similar to types in Ontario universities. One is the accreditation requirements for a variety of professional programs, such as medicine and engineering. A second is the range of internal evaluations of faculty through student assessments of courses and reviews of faculty linked to career advancement. A third is special subject reviews undertaken by universities, which usually enlist the advice of external academic consultants. With regard to the latter, British universities have also traditionally relied more than their counterparts in North America on external assessors of the appropriate coverage and standards for undergraduate examinations.

The recent and most substantial departures in British quality assessments lie in three other types:

- First, a periodic research-assessment exercise (RAE) has been introduced that assesses research in all subject areas in all universities. It is conducted by the HEFCE (and its Scottish and Welsh counterparts) and substantially affects the allocation of research funds to universities.

- Second, an ongoing quality review of teaching and learning in all subjects across all universities is conducted by the Quality Assurance Agency. It is not used to allocate funds directly, but its grades for academic departments are a source of intense concern.

- Third, there is an ongoing quality audit at the institutional level across all universities, conducted by the Quality Assurance Agency. In addition, at the institutional level, rankings of universities and subject areas are produced by leading British newspapers, using the results of the various reviews and adding and weighting other indicators.

The processes through which these latter three types of assessment have been implemented have not been static and are expected to change further. The following descriptions of them are thus only a sketch of current characteristics.

B. Research-Assessment Exercise (RAE)

The Higher Education Funding Council for England distributes public funds for research separately from public funds for teaching and learning, as do the funding bodies for Scotland, Wales and Northern Ireland. It currently distributes over $2 billion Cdn for research and between $6 billion and $7 billion Cdn for teaching and learning. The RAE, conducted periodically since the early 1990s, involves, in effect, the construction of a quality indicator at the subject level to guide the allocation of research funds by department or subject area.

The steps taken in a research-assessment exercise can be summarized as follows:

- A decision is made on the number of distinct subject areas to be assessed. Currently the number is 69; the classification of areas is not without criticism.

- Each university department in each subject area compiles extensive documentation of its research strengths. A university need not enter a particular department into the competition if it does not wish to, and it can withhold submitting the names of some faculty in a department, if it is felt they might draw down the chances for a higher grade. (Such exclusions, of course, reduce the base on which allocation decisions are calculated.) To simplify the amount of documentation for consideration by assessors, the top four publications are to be listed for all faculty whose names are submitted.

- Assessors, who have typically been of high professional standing in subject areas, review the documentation. Site visits are not made.

- The assessors rate departments in all the subject areas on a seven-point scale: 5-star, 5, 4, 3A, 3B, 2 and 1. These ratings have eight corresponding multiplier weights of 4, 3, 2, 1.5, 1, 0.5 and 0.

- The multiplier weights times the number of faculty listed in the submission determines the proportion of the research funds allocated to a university for a particular subject area. (The university is not required, of course, to spend these funds in that subject area.)

The results of these exercises have led to a highly skewed distribution of research funds. Now 75% of the funds are allocated to 20% of the universities. Most research support has been cut from the bottom third of institutions in the rankings. Research councils, which are not unlike Canadian granting councils, provide a second avenue for research support, but their competitions, it is said, are influenced by the RAE, as well as by government policy.

Three RAEs have been conducted since the early 1990s and a fourth is scheduled for the year 2001. Their future, after that date, is now under review, through working parties organized by HEFCE.

The general reputation of the RAE appears to be quite high in many quarters. There are favourable views of the quality of the assessors and the integrity of the processes. Particularly those who have been assessed favourably praise its approach for making tough decisions on the basis of quality.

Concerns arise about some of its incentive effects. Since research funds generated by a university's subject area will depend on its rating (converted to multiplier weights) times the number of faculty, a department must decide – as noted above – if holding back the names of some department members, judged to be less strong on research, is likely to raise the rating by an amount that more than offsets the effect of fewer faculty on the allocation.

Another area of this so-called game playing relates to the cut-off date for documentation. In a year of the RAE, March 31 has been the cut-off date for evidence of research strengths. It is said that appointments of faculty with impressive curriculum vitae are sometimes hurriedly made to meet this deadline, without sufficient consideration of the broader needs of a department. Publishers have been known to try to help by shifting priorities in publication schedules of books to meet a RAE deadline.

A more serious concern, especially in the early period of research assessments, was that the funding incentive strongly favoured research, and teaching suffered as a result. The elaborate processes and grading schemes that have now been developed for teaching and learning mean that this area can no longer be considered as neglected.

C. Reviews of Teaching and Learning (QAA Subject Reviews)

Under the Further and Higher Education Act of 1992, the HEFCE is required to include an assessment of the quality of education as part of the process through which it allocates public funds to universities. Initially, the HEFCE itself undertook quality assessments, but because of the difficulties of incorporating such a large task in its operations, a Quality Assurance Agency (QAA) has been established. This agency carries out reviews under contract to the HEFCE, which specifies the process to be followed in these reviews.

The purposes of the reviews have been to strengthen the accountability framework, to better inform students, and to enhance quality of teaching and learning. Unlike the case of the RAE, the reviews are not linked to the allocation of funds. They are more in the nature of an audit of the educational processes and are intended to provide an assurance of quality, regardless of the subject and institution a student might choose.

The QAA reviews are organized quite differently from the RAE assessments. They, too, are conducted by subject area, but the lists of subjects are not identical. The QAA list is slightly shorter at about 60 subjects and is expected to be substantially reduced in the next round through combining subjects. Assessment teams visit each department in a subject area, and thus the process is drawn out over a number of years, unlike the case of the RAE. Teams generally consist of a lead assessor chosen from outside the subject area and three to six assessors chosen from the subject area. A team generally visits a department for four days.

The material a department is required to prepare in advance of a team visit is specified in great detail by the QAA. The entire workings of a department are to be laid bare with minutes of meetings, follow-ups to motions, course outlines, examination papers and so forth, all organized for scrutiny.

The assessment team reviews the processes of the department in six areas or "cells" relevant to teaching and learning. One assessor usually takes primary responsibility for each cell. At the end of the four-day visit, the assessment team will usually agree on a grade for each cell, and 4 is the top grade. Thus, with six cells, the maximum over all grade is 24. If, in any cell, the grade is as low as 1, the result is deemed to be a failure. In this case the QAA will require the submission of an action plan for remedying the low grade and will re-survey the department in 12 months. If there should be a failure again on the re-survey, the HEFCE could withdraw funds but, so far, such an outcome has not occurred.

Despite the lack of direct impact on funding, the grades are taken very seriously by departments. The grades are reported on the QAA web site; the prestige of a department is viewed as being affected, as well as the future interest of prospective students and faculty members. Small differences in numerical grades mean a great deal. A grade of 20 out of 24 for a department that considers itself a leading department is the source of much anguish. A grade of 22 is a source of considerable satisfaction.

The numerical grades indeed do bunch closely. Thus, mathematics is currently under review, and the results to February 2000 of the review of 33 departments were obtained. For this group, numerical grades ranged only between 19 and 23, a very narrow span.

This enormous interest in numerical grades – and small differences in them – has developed recently. It appears that the earlier intention was to assign only vaguer terms such as "excellent," "satisfactory" and "not approved," and that, in the next round of subject reviews, more general terms will be given greater emphasis.

The QAA subject reviews are currently highly controversial. In their defence it is argued that they have had a positive influence on teaching as a result of the closer scrutiny of the organization and delivery of teaching services, and through the dissemination of innovative ideas in teaching. Heavy reliance on peer assessments has helped sustain respect for academic values. The processes of subject reviews have not been rigidly fixed, but are evolving in response to perceived difficulties. Thus, while there is an outcry about the excessively intrusive nature of the reviews, it is pointed out that the QAA is now signalling that the next round of subject reviews will have a "lighter touch."

The costs of subject reviews have not been definitively pinned down. But they are roughly estimated to be substantial, if imputed costs of faculty time are included. The enormous amount of work in compiling materials generally extends over a couple of years before the site visit and takes much of the time of at least one senior faculty member. In the case of quite a large department, it was estimated that about 150 vertical file boxes of materials were being collected and that the total cost of preparing the review, including imputed costs of faculty time, approached $400,000 Cdn. At a smaller university, the cost per subject review was roughly estimated to be lower at about $150,000 Cdn. Since a university can experience seven to eight subject reviews in a year and subject reviews are carried out across all universities, the total cost is very high. In addition, to such estimates of costs, however, are the costs of operations of the Quality Assurance Agency that are covered, in part, by a levy on universities. At a time when student-staff ratios have been rising, the costs to faculty members' time are an especially sensitive topic.

It was hoped that the subject reviews would help remove tiering in the university system, and it is the case that the subject review ratings of some new universities (former polytechnics) on some criteria are equal to, or ahead of, some older universities. But, especially when the subject reviews are combined with research-assessment exercises, a distinct tiering continues. Moreover, in mid-February, it was reported in the press that the government is now favouring a four-tier system of higher education in which 1) some universities concentrate on teaching, 2) some concentrate on research, 3) some seek an equal emphasis on teaching and research, and 4) some concentrate mainly on regional activities.

D. Institution Reviews, Ratings and Indicators

Less attention is generally drawn to a second activity of the Quality Assurance Agency. It is the quality review – a type of quality audit – of institutions.

Site visits to all universities are conducted by QAA teams, and an internal team is assembled by each university being inspected. The report that emerges from these reviews includes recommendations. But, a basic difference in relation to the QAA subject reviews is that there is no system of grading or ranking, and this difference makes the reviews less controversial.

Rankings or "league tables" of institutions are developed by British newspapers. Those published by *The Times Higher Education Supplement* are perhaps best known. The RAE research grades and the QAA subject review grades are combined as indicators of quality. Other indicators judged to represent quality – such as resources devoted to libraries – are added and weighted. *The Guardian* takes a different approach by focusing more on subject rankings among institutions. Thus, unlike Canada with its single ratings of institutions through *Maclean's*, the British have more rankings to consider, a helpful reminder that no one set of rankings should be viewed as sacrosanct.

The use of other indicators, in association with those derived from the complex research and teaching reviews, is an interesting development. In 1997 the National Committee of Enquiry into Higher Education (Dearing Report) recommended the development of performance indicators. The HEFCE followed up by establishing a Performance Indicators Study Group (PISG) and published in 1999 a report on performance indicators that also reflects, in part, the government's interest in social and economic aspects of university performance. An increase in the use of such indicators is expected. But they are unlikely to supplant current quality assessment methods. The idea of peer assessment with site visits is now viewed as deeply ingrained in the methods.

E. Some Summary Observations from an Ontario Perspective

This brief description of current British experience began with a recognition that distinctive conditions, institutions and issues will influence the direction taken in a particular jurisdiction. Also, it has not been the intention to provide a comprehensive, careful review of the British experience. The following five observations have been drawn from the British experience as being of particular interest and relevance to discussions of quality assessments of universities in Ontario.

1. Quality assessments have become a prominent and controversial feature of British higher education during the past decade. The research-assessment exercise has been somewhat easier to develop, perhaps because it is more in line with established practices in adjudicating research. The subject reviews of teaching and learning have stirred deeper concerns about their intrusiveness and have raised more worrisome questions about their value in relation to costs. Debate continues on the appropriate design of future quality assessments.

2. The British experience provides a warning about how easily administrative requirements can escalate and become oppressive, leading to a rise in costs, especially if costs of faculty members' time are imputed. In particular, the QAA subject reviews are viewed by many people as excessively costly and intrusive in relation to their benefits.

3. The British approach has been evolving and is expected to change significantly in the future. But some things are difficult to change or reverse. It is said that there is now an "industry" developing around quality assessment in universities, with the careers of many people linked to it. They become a force for continuation of activity consistent with their interests.

4. The effects on incentives can mean that the assessment processes develop "a life of their own" with unanticipated negative effects. Thus, in the British case, it is now recognized that there have been incentives for game playing. Universities and departments can affect results by excluding certain faculty or even departments from the reviews. Hiring may sometimes be geared more to impacts on ratings than to teaching and research needs. "Anti-poaching" rules have been introduced to curb incentives to raid other departments near the cut-off date for the RAE, but new regulations to correct other regulations often, in turn, create unanticipated side effects.

5. Assessment processes that end in ratings of departments and/or universities create competitive behaviour, some of which can be positive in spurring the search for greater quality. But any system of ratings has flaws, and there are dangers that institutions will be induced to "chase the ratings" and thereby distort their priorities in teaching and research.

In general, the recent British approach is fascinating and has important lessons, but is not an ideal model for mimicking in Ontario.

Section Five

ISSUES AND OPTIONS IN DESIGNING A THREE-LEVEL SYSTEM OF QUALITY INDICATORS

Assessments and assurances of quality in universities require, typically, a combination of objective indicators, which are representative of quality, and subjective judgements by assessors or appraisers, who are knowledgeable about the nature of quality in universities. This report has placed more emphasis on the portion of that combination related to indicators, but, as the British research-assessment exercise demonstrates, the two parts of the combination can be very closely intertwined.

The results of widespread international experience with quality indicators have not been greatly encouraging and, indeed, there is no experience one can point to as obviously an ideal model for Ontario to try to copy. Yet, in some form or other quality indicators will persist as long as people are concerned about the quality of universities. In the following, some dangers, benefits and options in the design of quality indicators will be briefly considered at the three levels of institution, program and university sector.

QUALITY INDICATORS AT THE INSTITUTION LEVEL

People want to talk about the excellence of their university or to know about strengths and weaknesses of institutions they are choosing among for themselves or their children. Because of the difficulties of measuring excellence or quality in absolute terms, indicators are used, and then there is the seemingly irresistible temptation to weight them and derive rankings of universities. *Maclean's* is the best-known Canadian example, but there are many examples in other countries.

A. Dangers

The principal dangers to try to avoid in developing institutional indicators can be summarized:

- First, there is a danger that institutional indicators will be standardized for comparative purposes, and this standardization creates a type of "one size fits all" or "boxing in" approach that suppresses differentiation of missions among universities. The result is that the indicators can mislead the user and can also create incentives that distort institutional development in the direction of undue homogeneity.

- Second, there is the danger of trying to quantify characteristics of quality beyond the point that it makes sense to quantify them. Because many relevant characteristics cannot be objectively measured, much scope should remain for qualitative judgement.

- Third, in the use of institutional indicators there may be a danger of confusing symptoms and causes of quality. The indicators should be accompanied by more analytical and empirical work on the causes of quality and the influence of various policy options for affecting quality.

- Fourth, rankings convey the mistaken impression that quality is in fixed supply. For example, it might be that all of the top 10 institutions in a ranking now are of higher quality than the top-ranked institution 20 years ago, but rankings do not allow for changing general levels of quality.

- Fifth, institutional indicators are generally expressed at a high level of aggregation and conceal strengths and weaknesses in the composition of activities at the level of program, classroom or laboratory. They provide only limited information to help choice of study and research in particular fields in a university.

- Finally, costs can become very high. Whether in England or in South Carolina, a particularly strong complaint about their experiences is the enormous amount of time and effort required to gather and assess the information for quality assessments. There is a danger that modest beginnings can turn into huge complex undertakings that incur burgeoning costs and increased use of political and bureaucratic methods to reach compromises on controversial areas of official assessments.

B. Benefits

Such dangers have tempered enthusiasm for the use of indicators, but there are also good reasons why interest in them has been persistent and widespread.

- Institutional incentives for quality can be strengthened by constantly being challenged to substantiate claims for quality with objective evidence, even if such evidence is limited.

- Quality indicators are part of a broader set of information required to satisfy public and private needs for accountability in the functioning of universities.

- Choice, especially student choice, can be assisted with better indications of sources of quality within and among institutions.

- Just as universities endeavour to assess as best they can the quality of academic work of their students and faculty, so universities must accept the need to assess the quality of the institution's work in teaching and research. To reject the latter is to invite greater external intervention to do it for them. As one Vice-President (Academic) wrote to me:

> To do anything else [than agree to participate in quality indicators] reinforces the view, held by many, that we do not wish to be held accountable. Anyway, use of quality indicators is part of the way we work in the academy, at least on the research side. There is no reason that it should not generalize to teaching or institutional performance.

C. Design of a System of Institutional Indicators

Certain features of a system of institutional indicators emerge from consideration of these dangers and benefits:

- Indicators should reflect the distinctive missions of universities and not inhibit differentiation. Thus, rather than facing the imposition of a specific set of indicators, institutions should be free to develop indicators consistent with their mission. The governing board of the university should approve which indicators best represent the mission of the institution.

- All institutions should be expected to release on their web site annually their chosen set of indicators up to a maximum number of, perhaps, 25. They would be part of the institution's regular report to stakeholders.

- Through a committee such as CUPA or through an organization such as that proposed in Recommendation 4, there would be a regular review and comment on the sources of data used in the indicators. The objective is not to create an intrusive or costly bureaucratic inspection of data but a general assurance of credibility of data used.

General Recommendation 1:
Each university should release on its web site an annual quality assessment report with a list of quality indicators approved by its governing board as relevant to the institution's mission. Concerns about the reliability and comparability of data in the use of indicators should be reviewed regularly through CUPA or the new organization proposed in Recommendation 4.

QUALITY INDICATORS AT THE PROGRAM/DEPARTMENT LEVEL

An executive head is accountable to the governing board of a university for the allocation of resources to programs and departments in line with the university's mission. The processes that are developed for this allocation would not be the same among institutions but would have strong similarities as institutions share information on best practices. This independence on decisions that affect quality is not absolute, and universities are increasingly required to provide forms of public assurance of quality of programs.

This report has not unearthed new forms of quality assurance of programs that Ontario should transplant directly from other jurisdictions. Rather to proceed along the current twin tracks of graduate and undergraduate quality assurance would seem to be the most appropriate direction.

In brief:

- At the graduate level, the appraisal process, which has been developed through the Ontario Council on Graduate Studies, should be the primary approach to ensuring a minimum standard of quality in Ontario graduate programs. The process, which relies on qualitative judgement of experts in the discipline, supported by a careful review of data, has basically served Ontario

well. It has recently been subject to an intensive review, and recommendations for its further development are under consideration. Some universities are developing more explicitly quality indicators for their graduate programs, but this is best left as an institutional responsibility.

- At the undergraduate level, the current initiative, referred to earlier and described in the Appendix, is in its early stages and is important to continue. Through the Ontario Council of Academic Vice-Presidents, a procedure is in place for the systematic auditing of undergraduate program reviews undertaken by Ontario universities.

General Recommendation 2:
Assurance of quality in academic programs is best provided in Ontario through further development of the current graduate appraisal program and undergraduate program review audits. At this time a superior model has not emerged elsewhere.

Universities typically have, in addition, many other initiatives relevant to quality assessment. They include course evaluations by students; annual faculty performance reviews; student satisfaction surveys; professional accreditation reviews; and undergraduate experience surveys. This extensive activity is not well known. It would be helpful for information to stakeholders if each institution prepared, for its annual report, a summary statement of the various approaches practised in the institution, along with information on the outcome of any graduate appraisals and undergraduate program audits.

General Recommendation 3:
The annual quality assessment report on the web site should incorporate, in addition to a set of quality indicators, the general results of any graduate and undergraduate reviews that have been completed, and a statement of the various approaches to quality assessment practised in the institution.

QUALITY INDICATORS AT THE UNIVERSITY-SECTOR LEVEL

Questions are becoming more pressing as to whether the university sector is performing to standards of quality required to meet competitive challenges in knowledge-intensive activities and, more broadly, to help students develop their intellectual potential to the fullest. Sources of answers to these questions are diverse and include work at the Ministry of Training, Colleges and Universities, at the Council of Ontario Universities, at faculty, student and staff organizations, and work undertaken in conjunction with occasional commissions and task forces. A diversity of sources is desirable, but it is suggested here that there is a need for some form of organization, charged with analyzing, on a regular basis, issues related to quality in the functioning of the sector. Moreover, questions about quality will arise if private universities or if new forms of academic credentials are proposed. This organization could be an important source of advice on such issues of quality.

The organization could take various forms. It could be a further development of a division of the COU or of the Ministry, though there would be distinct advantages to a more arm's-length relationship. But, essential to its success would be a reputation for first-rate work. It should be small. It should be free to roam widely over issues affecting the quality of higher education. It should analyze, comment, publish and distribute its work. As an analyst and commentator, it should not also be involved in the implementation of policy. To avoid going stale, it should have a significant flow-through of leading thinkers on higher education. To avoid becoming removed from the specifics of issues, it should be expected to consult widely with universities and their stakeholders. It, too, should be subject to regular critical review of the quality of its work, with reviews drawing on the perspectives of highly knowledgeable appraisers from outside the province and country.

General Recommendation 4:

If great care were taken with its structure at the time of establishment, a new organization would be helpful to monitor the quality assessment reports of universities and conduct analysis of quality in the sector as a whole and in relation to university sectors in other jurisdictions. Moreover, it could advise on quality issues if new private or public universities or new types of academic credentials are proposed. Among the critical features of this proposed institute should be an emphasis on: first-rate work; small size; broad mandate; published reports; openness to new ideas; careful consultation with university stakeholders; a hands-off philosophy with respect to directing institutions; and periodic, external reviews of its performance.

Section Six

SOME BROADER ISSUES IN QUALITY ENHANCEMENT

Quality indicators can play a role in efforts to enhance quality. They help us understand what quality means and what features characterize it. But they do not reveal the causes of quality nor do they reveal which policies will help produce quality. Issues of quality enhancement go well beyond quality indicators and quality assessments. The issues must be subjected to rigorous analytical and empirical work on determining quality in universities, and the new organization suggested in Recommendation 4 should be expected to stimulate more of this type of work.

The most powerful influences on quality enhancement can be classified into a) the framework of incentives to provide, within resource constraints, the best possible quality of education and research; and b) the severity of those resource constraints. The incentives become embodied in institutions and processes of higher education.

A sharpening of incentives for quality seems to require a reasonable degree of competition in the pursuit of goals identified with quality. Universities compete and should compete. They compete for the best faculty, the best students, the most funds, and the best learning environment. For all such attributes, "best" and "most" are defined by the nature of the institution and its mission, but there are varying degrees of overlap among institutions that strengthen competitive incentives. Quality requires that these incentives are sufficiently strong.

Quality also requires a reasonable degree of collaboration among universities. This collaboration is to be distinguished from collusion that results in monopolistic practices to the detriment of students and other users of the services of universities. This collaboration should be based on the more effective use of resources that can be achieved through the combination of teaching, research and administrative matters. As Dr. Ian Clark pointed out in the recent paper to an Oxford Roundtable, which was referred to earlier in this report, Ontario universities have long functioned in an environment of greater autonomy and lower regulatory burdens than is the case for universities in many other jurisdictions. Such conditions help foster competition. At the same time, Ontario universities have been highly collaborative in undertaking more common services and self-regulation than is the case for universities in many other jurisdictions.

An appropriate competitive and collaborative balance is a key factor in determining the degree of quality universities can achieve with given resource constraints. One of the great dangers in the development of quality indicators is a lack of appreciation of this factor and the supplanting of both competitive and collaborative incentives for quality with an onerous need for attention to oppressive regulations.

General Recommendation 5:
A strong competitive and collaborative environment with resources comparable to those in competing university sectors will be the most powerful influence on quality and should receive primary attention in building the quality of Ontario's university sector.

Postscript

This report and its companion report "*Will there be enough excellent profs?*" were completed prior to the announcement by the Minister of Training, Colleges and Universities on March 14, 2000 of the "2000-01 University Operating Grant Allocations and 2000-01 to 2004-05 Tuition Fee Policies." That announcement does not address major problems identified in the two reports and takes a position quite different from the reinvestment in universities recently announced in other parts of the country – notably in Alberta and Quebec – and in the United States. Also the decision to tie some funding to three indicators of graduation and employment rates is worrisome in light of the problems in the use of these indicators, discussed in this report.

APPENDIX

PROCEDURES FOR THE AUDIT OF PERIODIC QUALITY REVIEWS OF UNDERGRADUATE PROGRAMS IN ONTARIO UNIVERSITIES

Summarized by Roy Fischer

In October 1996, the Council of Ontario Universities (COU) approved the establishment of a procedure for the systematic auditing of the policies and processes in place at all Ontario universities for the conduct of periodic quality reviews of undergraduate programs. The procedure and guidelines specify that auditing of processes includes the examination of a representative sample of the quality reviews. Subsequently, in February 1997, the guidelines were amended to include the auditing of the mechanisms used by the universities for the implementation of new undergraduate programs. Authority for the organization and management of the audits is vested in the Ontario Council of Academic Vice-Presidents (OCAV). The detailed oversight of the audit procedure is devolved to a committee of OCAV, the Undergraduate Program Review Audit Committee (UPRAC) whose responsibilities are set out in Bylaw 1 of the OCAV Constitution. One of UPRAC's duties is to recommend to OCAV the schedule of audits, which started in 1997 with Brock University and the University of Ottawa, continued in 1998 with the University of Windsor and the University of Western Ontario, and includes, in 1999, Queen's University, Nipissing University, and Carleton University.

The audits are conducted at arm's length by at least three auditors who are appointed by UPRAC according to the criteria in the bylaw: "Auditors shall be chosen for their recognized strength in the development and operation of undergraduate programs. They shall not hold an administrative appointment in an Ontario university during their terms as auditors." The procedures to be followed by the auditors are spelled out in the UPRAC Audit Guidelines: Methodology for the Audit of Undergraduate Program Reviews, hereafter called the UPRAC Guidelines. They describe in some detail "the objective, structure and elements" that "any credible periodic undergraduate program review procedure undertaken by an institution must include." For convenience, these key review components of the guidelines will be referred to as the UPR Process.

The UPRAC Guidelines apply two tests: the conformity of institutional procedures and practices to policy, and the conformity of institutional policy, procedures, and practices to the UPR Process. At the beginning of the first audit cycle, most of the program reviews chosen for audit have been conducted under policies that were already in place before the guidelines were adopted. In consequence, in the early rounds, the test of conformity of practice with universities' own policies has proven especially important, since experience with institutional policy provides clues about particular strengths of the local assessment culture, or contrarily, of problems that need to be addressed as part of the more general effort of complying with the UPR Process. Even, however, after all universities have taken steps to bring their internal policies into line with the UPR Process, the auditors expect that the test of the conformity of practice to institutional policy will

remain important and should be subject to audit. A primary objective of the test is to identify good practices that should be emphasized and retained. The audit of such practices recognizes measures that go beyond the minimum requirements of the UPR Process, that promote programs of high quality, and that reflect the unique character and mission of each institution.

Regarding the conformity of institutional process as a whole to the UPR Process, the auditors have been sensitive to the appearance of unfairness that would result from the retroactive application of the standards of the UPR Process to institutional policies that were not, and could not possibly have been designed to satisfy them. At the same time, the auditors are bound to apply these standards, since it is now expected that all Ontario universities that have not already done so, will adopt program review policies that meet the minimum standards set by the UPR Process. The way the auditors have dealt with this potential appearance of unfairness, while at the same time discharging their obligation to apply the UPR standards, is to treat the components of the UPR Process as best practices, and to frame their suggestions and recommendations to assist institutions audited to comply most readily with the new standards adopted by OCAV.

In organizing their report and presenting their findings, the auditors have found it helpful to distinguish between recommendations and suggestions. Instances where the auditors considered the policies and procedures not to be in conformity with the UPR Process are cast as recommendations. Suggestions are offered in cases where, although the institution's measures are in conformity with the process, those measures could, in the opinion of the auditors, be improved.

The audit process carried out under the authority of OCAV was approved by COU in 1995-96. It was designed to satisfy the needs for accountability identified in the Broadhurst report and by the Ontario Council on University Affairs in its Advisory Memorandum OCUA 93-VI Academic Audit Review, while preserving the principles of university self-regulation and autonomy.

In 1996-97, the Ministry of Education and Training (MET) delegated to each university the collection and evaluation of information relating to the criteria used for the funding of new (non-core) programs.[1] The auditors noted during the conduct of the first audits that the UPRAC Guidelines did not deal with these MET requirements. Therefore, in the fall of 1997, the auditors recommended to OCAV that the audit should extend to the policies of universities pertaining to the establishment of new programs and, where applicable, should consider the integrity of the information that is collected to satisfy MET's criteria. OCAV agreed with this recommendation and modified its guidelines accordingly. All new program proposals should be backed by a thorough review. That principle has been endorsed by OCAV.

1.The non-core programs are those not specified in the Ministry's list of Group A programs in its memorandum of November 8, 1996.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").