

Measuring the Quality of Higher Education

John R. Cameron *

Although quality is a genuine concern of educators and its importance a constant convocation theme, the problem of measuring the quality of higher education has not been well treated. The greatest single barrier to adequate consideration of the subject is the problem of attempting to quantify what are essentially subjective factors. Further complicating the problem are the inevitable institutional loyalties encountered by all researchers in higher education. Few academics or other graduates are very willing to concede that their degrees are inferior to those offered by another institution. Additionally, the kind of programme offered by any particular institution may be particularly well-suited to one student and not at all appropriate for another.

Yet there is a high demand for some reasonably accurate way to measure quality. Comparisons over time for a university or university system are becoming necessary if we are to evaluate how the massive influx of students in the past decade, and the accompanying increases in public funds going into the Canadian university system, have affected the quality of education offered. Interprovincial comparisons are useful for determining whether the practices found in one jurisdiction improve quality, making these practices worthy of adoption by other provinces. For example, does the highly-touted semester system result in a lower level of educational quality? Students also have an interest in determining the relative quality of the institutions they propose to attend.

Probably the most important reason for finding some way to measure quality arises in the context of the rapidly increasing costs of post-secondary education. The vast allocations of public funds now used for the support of universities are no longer regarded as beyond criticism. Provincial governments and the tax payers are increasingly

* *Mr. John R. CAMERON is Research Associate, Royal Commission on Education, Public Services and Provincial-Municipal Relations, Halifax. The opinions expressed in this paper are the author's own and not necessarily those of the Commission.*

interested in discovering what they are getting for the money they spend. We have gone off the gold standard in education finance.

It is frequently argued and generally held that a portion of the increased costs per student is attributable to improvements in the quality of the education being dispensed. Part of the increase in total cost is due to the vastly increased numbers of students pursuing higher education. Another portion of the increase is due to inflation. But some of it probably reflects improvements in quality. However, there is not a simple relationship between the amount of money spent on education and the quality of instruction.

Measuring quality is a murky field in which firm pronouncements one way or another cannot be made. Indeed, it is often said that quality is essentially unmeasurable. In this paper I will attempt to point out a few of the problems involved in the use of some of the more common quality indicators, and to suggest some areas in which better data might make more accurate measures of quality possible at least for comparative purposes. I am not confident, however, that a sound way of measuring the total educational quality of an institution can be found.

Several rather general and probably obvious points come immediately to mind. For example, even if one university system is demonstrably better than another, there may be institutions in the poorer system which are the equal, or the better, of the best universities in the preferred system. When comparing universities, even if its overall quality is lower one may be better than another in one or more specific fields. What is true of the past, for example, in comparisons based on 1965 data, need not be true of the present, and may be still less true of the future three or four years hence when the present entering class graduates.

A recent paper which considers a number of methods for comparing the quality of provincial university systems is Weldon's "Inter-Provincial Comparisons of Cost and Quality of Higher Education in Canada".¹ In this study Weldon uses such measures as the faculty-student ratio, qualifications of faculty as measured by the proportion of doctorates on the faculty, the size of the graduate school, and retention rates for high school graduates going on to university and for those with bachelor's degrees going on to graduate schools.

The employment of simple ratios of full-time staff to full-time students ignores a host of relevant factors such as differing proportions of part-time instructors or part-time students. To meet this point, Weldon contends that even if part-time students and instructors were included, the relative magnitudes of the ratios would be the same, and he presents evidence to show that there have been quite similar ratios of full-time to full-time equivalent students by region across Canada and over time (.84, range .81 to .87, in 1968-69).² Yet in 1968-69 the ratio of full-time to part-time winter session students in Manitoba (2.50) was considerably lower than in Nova Scotia (5.19).³ Further, the

Universities Commission in Alberta has calculated some interesting student/staff ratios. In that province in 1967-68 the ratio of full-time students to full-time faculty was 13.5 : 1, compared to 9.1 : 1 for full-time students to full-time equivalent faculty and 9.4 : 1 for full-time equivalent students to full-time equivalent staff. This is a sharp difference. Another interesting observation from the Commission's calculations was that ratios which included the full-time equivalents of part-time students and faculty tended to change slightly more over time than did the simple full-time student/staff ratios.⁴

Another defect of an analysis which simply includes faculty/student ratios is that it ignores variations in teaching load and factors relating to the range of academic programmes offered. When, as is commonly the case, a large number of courses are offered, some will have small enrolments as a result of students' choices while others will be very large. Very large first-year classes can mask quite small classes in later years. Graduate courses will generally have fewer students, and will also require more staff for supervising theses. Another relevant factor is the degree of interaction between students and faculty, and this need not depend at all on class size.

In elementary and secondary education the range of programmes offered by a school is often one indication of its quality. The possibility of a wide choice of subjects also makes a college or university more desirable, especially for students who are uncertain of their future direction. The ability to transfer to a more interesting course as an alternative to dropping out can reduce human wastage, presuming that university education is preferable to or more useful than other experiences for the individuals involved. Yet a college or university with a wide variety of offerings can teach them all poorly while one which concentrates in a few fields of excellence may offer a better education to the smaller number of students it suits.

The presence or absence of a graduate school and its size are not good measures of quality. Some of the best undergraduate instruction can be found in institutions where there is no graduate school to distract professors. On the other hand, a graduate school can attract well-qualified professors who are sometimes available to teach undergraduate courses, and can also provide graduate students who are available for tutorial work and occasionally to teach full courses. The presence of a large graduate school does not mean that the university is better than most in all aspects of its graduate programmes, although its over-all quality might be higher. How much effect a graduate school has on undergraduate education is problematic. The mere presence of graduate students may have a salutary effect in itself. Yet there is the case of one university with probably the best programme in Jacobite history in North America which offered no graduate courses in history. This was partly the result of a provincial policy against proliferating graduate schools, and partly because the university itself chose not to establish a master's degree in one discipline where there was little chance of offering either a number of specialities in the discipline or master's programmes in more than two or three disciplines.

The qualifications of faculty, generally the proportion of the faculty with doctorates, are sometimes used to approximate a measure of the quality of higher education. Staff qualifications do reflect part of the quality of teaching if possession of a doctorate is really indicative of experience in teaching and research and a greater degree of mastery of the subject taught. But although in many cases his research is improved, a professor's teaching is not necessarily better by reason of the fact that he has a doctorate.⁵

Ofttimes a professor can be specialized to the point at which he can no longer properly communicate the generalities of his subject. I have been given the (apocryphal) example of the university teacher who had spent five years writing his dissertation on the subject of the exact hour of Wolfe's landing at Anse-au-Foulon⁶, and who consequently spent half of his time in his introductory course on Canadian History on the events immediately preceding and following the Battle of the Plains of Abraham. There are also tales of mathematics professors existing in lofty clouds and jumping from conclusion to conclusion far ahead of the class. One of the more frequent complaints about universities is that professors, for all their knowledge, often need to be taught how to teach. The use of a Ph.D. to measure staff quality also ignores those who have gained their expertise by other paths than the academic, in business, government or through self-teaching.

Another set of quality measures can be established by investigating the pattern of higher education spending. Although the amounts of money spent are not necessarily indicative of quality, the way in which money is spent may be. For example, the level of spending on libraries may be a useful measure, as a well-stocked library is necessary for high quality instruction in some programmes, especially at the graduate level and in the humanities and social sciences. However, if the programmes offered by an institution emphasize other areas of instruction than those emphasized in the library collection, if relatively little research is done in the institution, or if there are good library facilities nearby, use of library expenditures as a measure of quality could be misleading.

Access to the library facilities of other institutions can replace part of an institution's spending on its own library. For example, in Halifax there are six degree-granting universities and colleges, a number of specialized collections, the city and legislative libraries and the provincial archives, all mutually interdependent at least to some degree. In Ottawa the two universities have the resources of the National Library, the Public Archives of Canada, the Parliamentary Library, the city library, collections of federal departments and agencies and each other on which to draw. To repeat an earlier point, a library can be poor in many areas but exceptionally good in one specialized field. Such a speciality will attract the students and scholars necessary to provide a very high quality programme in that field.

Any attempt to measure quality by comparing expenditure levels presumes that money spent is an adequate criterion of quality. But the existence of different costs for different programmes, and the different programmes requiring different mixes of instruc-

tion, research and library expenditures precludes the use of a single type of expenditure, or even the expenditure mix, as an indication of the quality of the programmes offered. Neither can the relative proportions of expenditures on instruction and research be used as a measure of quality unless a prior judgment is made on how much research is necessary to provide good teaching. The proportion of fixed costs to costs of instruction is a very good indicator of size, and possibly of administrative efficiency, but not of quality. For example, in Canada in 1968-69 administrative costs of small institutions (expenditures of less than \$1.5 million) were 14.9% of total operating costs, compared with 11.1% in medium universities (expenditures of \$1.5 to \$5 million) and only 5.7% for large institutions (expenditures of over \$5 million). Plant maintenance expenditures ranged from 13.1% for small institutions to 11.0% for medium and 9.8% for large institutions.⁷ The development of new programmes and efforts to catch up from a low base which differ from institution to institution make increases in total expenditures or in particular items of expenditures inadequate measures of attempts to maintain or increase quality.

Students' preferences might be regarded as a measure of at least what quality is thought to be, although patterns of migration are not necessarily related to the quality of education expected. There are probably institutions of comparable quality in most regions of Canada. Apart from this general criticism the available evidence is contradictory. The Macdonald Report contained a short migration analysis for graduates in the sciences and engineering which calculated retention rates for high school students moving to undergraduate programmes and for students with bachelor's degrees moving to graduate study. Because the respondents represent all ages, data on their migration from high school to university to graduate school present a picture cumulatively true of the past but not of the present or of any definable point in the past. For example, the retention of high school graduates in Newfoundland moving to universities was between 21% and 25%.⁸ In 1969-70, 85.6% of Newfoundland residents studying in Canada studied in Newfoundland.⁹ Indications of interprovincial flows of students show that there are a large net inflows into the Atlantic Provinces; but if the Macdonald Report's data are used as quality criteria the universities in the region are not regarded highly. However, the Macdonald Report does not attempt to use these flows as quality indicators. It regards population size and employment and study opportunities as the explanatory variables, not quality of education offered.¹⁰

Rather than reflecting students' judgments about qualitative differences, retention rates and migration patterns more likely reflect such diverse factors as the availability of desired programmes, students' sophistication or cultural isolation, even students' family incomes.¹¹ More sophisticated students and richer students are more likely to leave the province. Even the size of the province can have some bearing. In the smaller provinces, students may have to move out of the province to move the same distance as a person in a larger province could move without leaving the province. In the Atlantic Provinces a full range of programmes is offered only on a regional basis.

A number of studies outside Canada have used other measures of quality for which Canadian data is not available. To take one interesting example, Morgan and Sirageldin grouped colleges according to their selectivity, based on the percentage of applicants accepted by the college, the average CEEB test scores of recent freshmen classes, the ranking of recent freshmen in their high school classes, and related data on the students' potential. The use of the selectivity index as a measure of quality is based on the premise that a university can never be much better than its students and is not likely to be much worse.¹² If one accepts a selectivity index as a proper measure of university quality, it would appear that basically all universities are the same except for the student body. It is then the presence of larger proportions of exceptionally able students, undoubtedly an advantage, which differentiates between institutions of higher and lower quality. If this relationship holds true, it would appear that to rise from low to high quality all a university would have to do would be to implement a highly selective admissions policy. Given a sufficiently capable public relations officer there seems little doubt that such an approach would be effective. The general higher level of ability in the student body might raise the level of education for those less well endowed. Better students might also attract better professors.

On the other hand, such a measure may be an effect, rather than a cause. So many students seek admittance to a high quality institution that a very selective admissions policy which allows in only the best has to be adopted. Given the relatively slow rate of information dispersal about higher education, this measure may not reflect the present degree of quality, but rather one pertaining to some date or series of dates in the past. The measure is to some extent self-sustaining. Better students attract better students. And because they are abler and better motivated they do as well as or better than other graduates, possibly in spite of their university education.

Another interesting American example was an attempt by the American Council on Education to evaluate quality in graduate education.¹³ Faculty in thirty disciplines in over one hundred of the largest universities were asked to rate the quality of graduate faculty in their fields in each institution and to judge the effectiveness of doctoral programmes. Responses were broken down by rank of faculty (department chairmen, distinguished senior scholars, and knowledgeable junior scholars), region (east, mid-west, south and west), and the judgment of a small select panel of experts.

I am aware of the limitations of such subjective evaluations — that they are based only partly on first-hand knowledge, are influenced by hearsay, incorporate halo effects, are subject to time lags, etc. As one of our respondents quoted Dr. Johnson, "A compendium of gossip is still gossip." Despite some very real limitations, however, I believe such evaluations more clearly reflect reputation and accomplishment than do such "objective" measures as number of books in the library, Nobel laureates on faculty, Woodrow Wilson fellows enrolled, or dollars of research grants from government or foundations.¹⁴

To develop a proper measure of the quality of higher education, we must first define what we intend a university to do. If we adopt the "old notion that the functions

of a university are to preserve, disseminate and extend knowledge",¹⁵ possibly adding Schultz' suggestion that universities act as a discoverer of talent¹⁶, we can attempt to measure differences in the quality of achievement of these goals. On the other hand, assuming that the main function of the university is to educate, we have the unenviable task of measuring educative success in such areas as knowledge acquired, enhanced ability to reason, greater enjoyment of life, aroused social conscience and similar results even harder to evaluate.

The university's role as a discoverer of talent is probably unmeasurable, because there is no universal definition of what constitutes a talent, and further, even if there were, it would be next to impossible for the researcher to find out what proportion of talent a specific university was leaving undiscovered. It would be heroic to assume that there is an equal distribution of talent in all entering classes of all universities.

The university's role as a preserver of knowledge would appear to be a function of the system as a whole. Basically it involves compiling libraries, encouraging publication of worthwhile research (at last a good word for publish or perish) and employing knowledgeable people to teach students how to discover (and use) the knowledge preserved.

The extension of knowledge presents serious measurement problems. Any measure of the volume of research produced is inadequate. One discovery by a part-time professor at an obscure school could be of more importance than the entire research output of our largest university. When measuring research, quality is an essential part of evaluation, even though massive amounts of detailed work are also necessary. Yet quality in research is almost impossible to evaluate; the effects of some discoveries made thirty and forty years ago are still not fully assessed.

The third major objective of the university, and the most important, is education, or more restrictively, the dissemination of knowledge. A test to evaluate the comparative quality or effectiveness of universities in transmitting knowledge could conceivably be designed. Students entering university for the first time could be given a battery of tests such as SACU, the College Boards or even the GRE designed to evaluate their current knowledge of a wide range of subjects. Aptitude tests, rank in graduating class, high school marks, evaluations of potential success made by former teachers and principals, and other data relevant to motivation and ability would be gathered and used to standardize the students. Re-testing for knowledge would be undertaken, say, annually, until graduation, using the same tests. Improvements in test scores would be attributable in large measure to the information disseminated by the universities. Large divergences in standardized test score increments would indicate some divergence in the ability of different universities to impart knowledge.

Although ability will presumably remain constant through the three or four years measured, students' motivations will probably change. To the extent that changes in motivation are caused by the university or its immediate environs, it would appear

logical to assign their effects to the quality of the university's instruction in the widest sense of an atmosphere conducive to learning. Housing problems, personality conflicts with professors, perhaps even some emotional problems might be attributed to university policies. Other motivational changes would be external, such as a divorce or bankruptcy in the family. These external factors would distort any findings, but over a sufficiently large sample they might average out.

A study such as the one proposed could yield a good deal of worthwhile information. It could evaluate the comparative effectiveness of a three or four year programme, university entrance from junior or senior matriculation, a post senior matriculation year before university entrance (as offered by Quebec's CEGEPs), small town, rural or urban settings for universities, the trimester system, the importance of the presence of a graduate school or of faculty research to undergraduate programmes, the importance of the relative size of a university, and similar questions. It is unfortunate that such tests would have only a limited usefulness, for they can evaluate only one aspect of the university's educational role. Measurement of the quality of higher education must remain subjective, and we must be careful not to use objective measures of quality, which may not be valid in themselves, simply because they bear out our own subjective evaluations. In the measurement of quality, a wrong evaluation is worse than none at all.

Notes

1. K. Laurence Weldon, "Inter-Provincial Comparisons of Cost and Quality of Higher Education", (Toronto: Committee of Presidents of Universities of Ontario, 1970), mimeograph, 54 pp. Weldon's comments on quality are on pages 47-51.
2. *Ibid.*, pp. 8 and 24.
3. Based on Manitoba Universities Grants Commission, *Annual Report for the Year Ending March 31, 1969*, (Winnipeg n.d.), pp. 10-11, and information from the Nova Scotia University Grants Committee.
4. Alberta Universities Commission, *Annual Report, 1967-68*, (Edmonton: Queen's Printer, 1968). Details of the computations were given in Alberta Universities Commission, *Annual Report, 1966-67*, (Edmonton: Queen's Printer, 1968). (Neither report is paginated). Part-time faculty and students were calculated to be weighted one-third full-time faculty

and students. The Commission's 1969-1970 *Annual Report*, (Edmonton: Queen's Printer, 1971), gives 1969-70 ratios of full-time students to full-time faculty of 13.5 : 1, compared to 9.7 : 1 for full-time students to full-time equivalent faculty and 10.0 : 1 for full-time equivalent students to full-time equivalent faculty.

5. Cf. C.F. Poole, "The Ph.D. as a Qualification for University Teachers", *CAUT Bulletin*, 14, 2 (December, 1965), 61-65.
6. Three a.m. September 13, 1759, contrary to the four a.m. espoused by Stanley in G.F.G. Stanley, *New France: The Last Phase, 1744-1760*, (Toronto: McClelland and Stewart, 1968), p. 230, or the two a.m. suggested by Lanctot in Gustave Lanctot, *A History of Canada*, volume 3, *From the Treaty of Utrecht to the Treaty of Paris, 1713-1763*, Margaret M. Cameron, translator, (Toronto: Clarke, Irwin, 1965), p. 170.
7. Unpublished D.B.S. data. But see any edition of D.B.S. *Canadian Universities, Income and Expenditure* (81-212) for a similar picture.
8. John B. Macdonald *et al.*, *The Role of the Federal Government in Support of Research in Canadian Universities*, (Ottawa: Queen's Printer, 1969), pp. 311-28 at p. 327.
9. D.B.S., *Survey of Higher Education, Part 1: Fall Enrolment in Universities and Colleges, 1969-70* (81-204).
10. Macdonald *et al.*, *op. cit.*, pp. 314-316.
11. There is some evidence from unpublished tabulations of the D.B.S. Post-Secondary Student Population Survey to indicate that student mobility is correlated with parental income among other variables. See John F. Graham and John R. Cameron, *Federal Participation in the Financing of Post-Secondary Education in Canada (Excluding the Financing of Research)*, Study No. 3 of Financing Post-Secondary Education in Canada (Toronto: Council of Ministers of Education, 1971), mimeo., pp. 180-81.
12. Morgan, James, and Ismael Sirageldin, "A Note on the Quality Dimension in Education", *Journal of Political Economy*, 76 : 1069-77 (September-October, 1968), p. 1075. Cf. James Morgan, Martin David, Wilbur Cohen and Harvey Brazer, *Income and Welfare in the United States*, (New York: McGraw-Hill, 1962), p. xii, and James Cass and Max Birnbaum, *Comparative Guide to American Colleges* (New York: Harper and Row, 1964), where the index was calculated. It is interesting to note that Morgan and Sirageldin's study indicates that students who went to "better" colleges had higher annual earnings, a result to be expected given the method chosen to rank institutions by quality and the fact that no allowance was made for differences in student ability.
13. Allan M. Cartter, "Economics of the University", *American Economic Review, Papers and Proceedings*, LV, 2 (May, 1965), 481-94, especially pp. 485-89, although the very interesting discussion of quality continues to page 494.
14. *Ibid.*, p. 487. For a good review of American efforts to measure quality, with emphasis on the impact of universities on students, see Robert H. Berls, "An Exploration of the Determinants of Effectiveness in Higher Education", in Joint Economic Committee, United States Congress (91st Congress, 1st Session), *The Economics and Financing of Higher Education in the United States*, (Washington: U.S. Government Printing Office, 1969).
15. *Towards 2000: The future of post-secondary education in Ontario*, report prepared for the Committee of Presidents of Universities of Ontario by its Subcommittee on Research and Planning, presented as a brief to the Commission on Post-Secondary Education in Ontario, (Toronto: McClelland and Stewart, 1971), p. 25.
16. Cf. Theodore W. Schultz, "Resources for Higher Education. An Economist's View", *Journal of Political Economy*, 76 : 327-47, (May/June, 1968). The discovery of talent is "a process which provides students with opportunities to discover whether they have the particular capabilities that are required for the type and level of education at which they are working." *Ibid.*, p. 331.