

Extending the Academic Sector Core Indicator Set For Academic Faculties and Departments

**Economics Department
October, 2010**

I. Context

The *Core Indicator Set* is an attempt to measure Concordia's institutional progress along the different elements of its mission. The attempt is significant for the reasons given in the document: self-deception is always a pitfall in the process of self-assessment. As well, there is the challenge of measuring ourselves against other institutions having more or less similar missions with similar resources. So having some numbers that measure progress across time and comparatively with a benchmark group is certainly useful. And the document is careful to underscore the conceptual difficulties in using indicators to measure some attributes that may have limited relevance for particular faculties, departments or individual faculty members.

We can all agree that the measurement process makes good sense and could be useful. None the less, in assessing the merits of the process, we can ask: to what ends are these measures useful? The methodological embarrassment that is the MacLean's survey is an object lesson in how to use comparisons or rankings derived from a set of measures for a dubious end. Somehow small differences among myriad factors are aggregated by the magazine into big differences in ranking. And since ranks are widely dispersed and memorable they become important pseudo statistics for fund raising and recruitment. The all too predictable consequence is that now most institutions in the country have some planning function whose primary task is to improve/massage/reconfigure inputs into the MacLean's process.

Of course, universities themselves trumpet the results of rankings in their marketing efforts. Business schools are notorious in this regard where, for example, top-100 stature in the *Financial Times* world rankings merits tens of thousands of dollars in self-promotional publicity. In this context, rankings have themselves become a growth industry.

So it is surely to be hoped that the point of the Concordia exercise is not simply to say at a certain point in time that we have pulled to within three (or so) points of the University of Victoria and that we remain comfortably six (or so) ahead of UQAM, all because we have increased our aggregate number four (or so) points since the last exercise. This is the stuff of marketing, of chauvinistic self promotion and hardly worthy of any serious effort. It undermines rather than contributes to our mission of teaching and preserving and advancing knowledge.

On the other hand, the measures could be useful as crude diagnostics. To take an example that will be discussed further below, the results of systematically poor teaching

evaluations could point to a problem. But this measure is more properly a matter for a Department Chair who would presumably consider the matter in the context of courses taught, course content, course level and so forth. How such a subjective assessment could be aggregated across Departments as part of a core indicator assessment of the institution's teaching competence is difficult to imagine.

One would prefer to reserve judgment on the merits of the exercise in the hope that the discipline of self-assessment would be advanced and will not degenerate into a facile ranking à la Maclean's. However, the document does appear to indicate otherwise. In reviewing the criteria that have been used to select the core indicators, the document points to *Realism* as an important criterion: "... indicators are intended to assist in the objective adjudication of competing requests for scarce resources."

Here is the rub. Indicators that are sufficiently high-level to be aggregated across Faculties and Departments and broadly construable to be compared with the other universities in our benchmark group will, it seems, be tuned finely enough to assist in the distribution of resources across University units. The indicators are not mere diagnostics that could point to an issue within a unit, but appear to have in themselves sufficient weight/significance to assist in the *objective adjudication* of requests for scarce resources. Given this intended use of these indicators, we should take a more careful look at their actual content.

II. The Indicators: an appraisal

a) Indicators that can be measured

There are eleven indicators that are presented as part of a core set, and three others that are proposed seemingly to be dismissed for issues in measurability (maybe; see below).

Seven of the eleven are precisely measurable and can be aggregated easily enough: (1) External Student Funding; (4) Granting Agency Faculty Funding; (5) Publications; (6) Undergraduate Graduation Rates; (7) Undergraduate Retention Rates; (8) Through Rates in Graduate Studies; (10) Employment Success for Graduates.

There are four indicators that are entirely subjective in nature: (2A & 2B) Level of Academic Challenge/Engaged Teaching Practices; (3) Teaching Quality; and (9) Student Satisfaction.

The three "supplemental" indicators are: (12) Faculty Teaching Load; (13) Added Education Value; (14) Community Academic Engagement.

From our point of view (a Departmental perspective), the challenge in assessing these indicators is to determine in what way they could reasonably be used to assist in the objective adjudication of claims for scarce resources. Concerning indicators in the measurable group:

1. Indicator (4): The university may see itself lagging overall in funding and decide to allocate scarce resources to Departments that have been successful in attracting external funding. This instinct is sensible in so far as good Departments should be given every opportunity to get even better. Suppose Dept. X gets more funding than Dept. Y. Do scarce resources go to Dept. X? If the reality is that Field X is heavily funded, it may turn out that Dept. X is in fact not doing well vis-à-vis other departments in the 'appropriate' comparison group (other Department X's at other universities) and that Dept. Y is actually doing quite well relative to their respective group. Similar variants of this argument run for publications, where a seemingly modest publication record in a field with a strict and time-consuming review process may in fact be outstanding. Of course in the case of publications, a quality adjustment rather than a straight addition of publications or pages is desirable, and while never easy to determine, is certainly more feasible at the discipline level than across departments. **Conclusion:** measures such as (1), (4), (5) should use intra-disciplinary comparisons to be relevant.
2. Once the principle that Departments are more properly evaluated across institutions is acknowledged, it becomes difficult to compare Universities at the aggregate unless one university dominates another across Departments. One could measure the proportion of units in a university that are among the top two in the benchmark group. This approach would make sense for academic units that are common to the benchmark universities. On the other hand, how does one include units unique to the university in a measure of comparative progress?
3. Indicators (6), (7), (8). Any measure of graduation or retention rates depends on context. A few years ago, the Laval Economics Department had an undergraduate first-year entry class of 20; we suspect retention and graduation rates were high in this group. In Economics, we have a large entry class having a significant component desirous of an eventual JMSB landing. If these students leave for one reason or other, will our rates be interpreted as a sign of poor management of our students? How does one go about measuring this? Moreover, if there is the JMSB issue perhaps we should actually obtain more resources earmarked for making the Economics undergraduate experience more attractive (as opposed to meriting a penalty for poor indicator performance). Indicators (6)-(8) could perversely point in the wrong direction if taken at face value. In keeping with the point made in 2 above, it seems unclear how to aggregate Department experience to determine a figure for the University as a whole, or to move from a gross university number to the individual unit.
4. There is very little data available currently that allows an evaluation of indicator (10): Employment success for graduates. Census data does not include information concerning the university attended. Some survey data on how the graduates from different business schools fare in the job market may exist, but how such surveys would help the university adjudicate the allocation of scarce resources across departments is unclear and is open to the type of manipulation that may also plague the next set of indicators.

b) Indicators in the subjective group

5. Indicator (2) Level of Academic Challenge/Engaged Teaching Practice. It appears that students will be asked how many assigned reading were part of the course; whether they made a class presentation, etc. – and these responses will be aggregated somehow into a number across all courses. There is little chance that anything meaningful could emerge from a questionnaire that has such a restricted sense of what goes on in a course. It seems obvious that student questionnaires on course content can only be constructed and evaluated within each Department, and even here should vary depending on the nature of the course or its role in the program. The results can hardly be then aggregated across disciplines.

In contrast to a few years ago, our Department's distribution of students across sections has changed considerably, and this change has importance in measuring our performance with reference to indicator 2. Currently, a plurality of classes is offered in groups of 60 or less, yet a plurality of students is registered in classes of 100 or more. This bifurcation arises as a result of the Department's recent decision to instruct most of our 201 and 203 students in classes of 300. The implications of this decision are obvious for these indicators (and also indicator 12): if student responses to surveys underlying 2A and 2B are to be aggregated by class, then the plurality of relatively small classes should induce responses more in keeping with the indicator bullet objectives than if answers are to be aggregated by individual student responses. Parenthetically, the same argument applies to indicator 3, bullet 4.

As well, some of the bullet elements in indicator 2 appear dated. For example, a department where professors post comprehensive lecture notes on a Moodle type facility, and thereby mitigate the need for students to purchase expensive texts, would be judged to be performing worse than one where less vigorous instructor practices prevail.

6. Indicator (9) Student Satisfaction. Universities are wont to tout how happy their students are. We will be in the position that a student who is not happy with a course, professor, the football team will answer negatively and without due consideration of the overall good of our University relative to, say, Simon Fraser. The point is that subjective assessments elicited by a non-neutral party will be manipulated if there is a good reason to do so.
7. Indicator (3) Teaching Quality. The document suggests that "psychometrically sound student evaluation of teaching performance tends to be reliable and valid." Some references on this far from accepted point would be useful. In any event, the Concordia experience in the Economics Department on teaching effectiveness based on web-based questionnaires would suggest the contrary in so far as sample

sizes have diminished considerably, a reality that undermines completely their statistical significance and relevance.

c) Indicators in the “supplemental” group

8. Indicator (12) Faculty Teaching Load. As the text in the document seems to suggest, this indicator would be based on a number such as the ratio of full-time faculty relative to the number of students or student FTEs. Faculty teaching load is not the concept at issue. If we taught 100 FTEs having a load of 2 courses per year, are we better or worse off than if we had each a six-course load. Of course, the answer from one student perspective depends on the number of part-time teaching contracts are needed to make up the difference or, from a different perspective, the quality of instruction. There may be interesting metrics here if we look harder: percentage of lower-level courses/students taught by full-time faculty. In any event, load is misleading.
9. Indicator (13) Added Education Value. This indicator is an indirect response to a reality that dooms Concordia in national ratings: the entry class is not as strong as say at McGill. So it is our value added that must be the focus of attention. Yet the University has done no real work in this area, at least any work that makes its way down to the Department level. Our Department did some work in the past on retention/ graduation rates/success in theory courses conditional on entry profile. Some pedagogical initiatives were undertaken as a consequence. In any event, it is difficult to see how we could get a measure of value-added that could be compared to that of other universities or even across Departments. We hope the Institutional Planning Office continues to work in this direction and to devote resources to Departments that facilitate the transition to university studies for students having lower academic collegial achievement. As presented, the indicator would suggest that Departments that graduated their students somewhat-may must be deserving of greater resources.
- 10 Indicator (14) Community Academic Engagement. The indicator is perhaps meritorious, certainly immeasurable.

d) What is Missing

There are several measurable concepts that are certainly relevant yet do not make their way into the proposed core indicator set. Class size is one example. The relation between the number of students taught by an academic unit and the resources given the unit to manage its teaching commitments is not investigated. But these considerations are matters for Departments and the various strategic options available to the unit, such as the choice of class size relative to its role in the Department’s program, a point discussed above. The *Core Indicator Set* prefers to measure globally, with the intent to work its way back down to the unit by pretending to have an objective handle on issues that are

more properly understood within the Department context. We have argued this approach is seriously flawed.

Finally, there is no mention of one readily-measurable statistic: administrative costs. Are the costs of administrating Concordia in line with the costs of those universities in our benchmark group?

III. Conclusion

It appears that one goal of this attempt at measurement is to show that the University is edging toward the top 5 in our comparison class. Fair enough. But we should not lose track of the reality that Concordia's version of Core Indicators is an alternative to the Maclean's rankings, with all the same shortcomings associated with subjective assessments. It reflects to some extent our own measure of our self worth. The issue is not that we should forgo such an exercise. Rather the point is to appreciate that the exercise will make little contribution to the hard issue of resource allocation. Our point is that the University seems to have lost track of the fact that it is made up of Departments and it is Departments to which students (especially graduate students) and faculty come. What able and ambitious graduate student or prospective faculty member with alternative possibilities will decide to come to Concordia because it is now ranked as the third best comprehensive university in Canada, while the department in which they will study (or work) is not ranked in the top 10 in Canada or the top 200 worldwide? Most of the measures discussed above would allow meaningless comparisons across departments within the university. To think that they could be used for serious ends such as the distribution of resources across departments is truly worrisome. We wish to emphasize that in their current form the indicators are useless as a device to compare departments across universities. It is the later measure that is relevant, that should be of concern and that should be the focus of administrative initiatives. The university is certainly more than the sum of its parts, but it is at least the sum of its parts. It is hard to imagine that if the parts are not strong, and demonstrably so, that the university can be considered strong.