



Stockholms
universitet

18/08/2009

Johan Falk
Professor of Spanish
Bologna Coordinator
johan@isp.su.se

Indicators we live by? *On the quantity-quality dilemma*

Abstract

In modern times the use of performance and quality indicators has become more and more pervasive in HE. The models for strategic planning and follow-up by means of indicators have increasingly copied other processing activities, in use in industry, business and administration. The aim of this contribution is to analyze some of the more problematic aspects of indicators, such as the selection, purpose, representativity, validity and reliability of indicators used in HE. Arguably, due to the fact that indicators have to be operational in order to be reliable, quality assurance systems in HE tend to overuse quantitative parameters. The computerized culture we live in feeds the idea that quantification is the solution. Awareness of the in-built conflict between quality and quantity is important when selecting indicators and interpreting outcomes. It seems indicator systems firmly rooted in quantification are nowadays giving way for a more qualitative approach. This tendency is noticeable at different levels of HE in Sweden, even in the models used for allocating resources to education and research. The Strategic Planning System at Stockholm University is goal- and indexed-based. General goals – specific goal – indicators – target values – measures form a chain where the indicator has a key position in the monitoring process. The qualitative rooting of the indicators is one of the central issues of the planning system.

1. Introduction

The use of indicators is spreading like a prairie fire in the realm of Higher Education and quality assurance. This is of course not a new phenomenon but as it has become a more and more pervasive instrument for allocating resources and establishing rankings of Higher Education institutions as well as departments at a lower level. That is why indicators deserve to be scrutinized and reflected upon. My intention is twofold: in the first place I would like to share with you some reflections on the use of indicators and for what purposes, in the second place, I will give a glimpse of the system of indicators used in the goal-oriented strategic planning that has been launched at Stockholm University some years ago. My focus will be

on the use of indicators within this system. My thesis is simple: the use of indicators in HE is nowadays incorporating more openly qualitative aspects, and I will exemplify this tendency from HE in Sweden and especially Stockholm University.

What is an indicator and what does it measure? This is a seemingly trivial question because performance indicators are of common use in all sorts of processes, in industry, in business and in education. As a matter of fact, we are constantly guided by indicators in our daily life and, even more so, our judgment of others is built up on indicators. If you see a person giving money to some charity organization, you will conclude that this is a generous and unselfish person. The act of giving money is an indicator of something else, something that is broader than the act I have witnessed.

2. Essential aspects of indicators

An indicator of some sort of performance – in research, teaching and learning – is something observable that represents something that is more general than itself. It is taken for granted that checking a complex activity can only be done by means of symptoms that stands something broader: for example, some aspects of the finish of a car stand for the quality of the whole car, the number of refereed scientific articles stand for degrees of scientific quality at an institution. This is in short the validity problem: Do indicators really measure what they are meant to measure? This question leads to several questions concerning crucial aspects of indicators:

- 1. To what purpose(s) and from which perspective are indicators chosen?*
- 2. How are indicators selected and by which processes?*
- 3. How should we judge the representativity and validity of indicators?*
- 4. Are indicators reliable in the sense that the same value shows the same state of things?*
- 5. Which are the causes of an increase or decrease in the outcome of the indicators? To what extent is it due to measures taken earlier?*

Each of these five questions could be the subject of a lengthy philosophical discussion that would not lead much further. Anyway, the questions point to important and problematic issues. For example, if the number of published doctoral thesis is chosen as an indicator of the success of research studies at a department and is as a criterion for assigning resources, at least three things should be kept in mind:

- A.** if this is a good measure both of performance and quality;
- B.** that all side-effects and bias that may arise from this specific indicator must be analyzed;
- C.** that the way the departments will react to this situation must be addressed.

What will an adaptive or evasive behavior on the part of the departments lead to in the long run, will the university benefit from institutions that try more get round the indicator than see what lies behind. All this might seem to be a very critical stand on indicators, but at the same time it must be stressed that indicators are indispensable because nothing can be assessed or monitored without them.

3. Success and fallacies of indicators

I will now shortly comment on some of the issues referred to and I will mention some of the problems and fallacies that indicators may give rise to. Exemplification will be drawn from the current debate on Swedish Higher Education.

3.1 The central issue is that of *representativity and validity*. To what extent does the indicator measure what it is meant to measure? Is citation index – to take one of the most common indicators used in the international rankings – a good criterion for the “larger” issue that is performance and quality of research-work. It is important to bear in mind that the domain where the indicator is applied is not homogeneous and that actors have very different opportunities. Take internationalization for example, for sure a complex matter. Is the number of incoming and outgoing exchange students a good measure of internationalization? I doubt it, but this is not to say that it is not interesting to know whether the number rises or not. It is one aspect, important yes, but it is not clear how representative it is. Things are very often confounded, because it is human – even for an audit group – to generalize from the token to the type.

It is often said that you measure what can be measured, even if it is not a very interesting aspect. Here is another example that illustrates the problem of validity. Is the number of students passing successfully a course or an exam a good measure of educational quality? Yes, to some extent, and these figures are needed to get an idea of the throughput. But in reality it says little about quality. We have here one of the main problems concerning indicator validity: that is, the opposition between quantity and quality. Getting back to the same example, I think that the rate of students achieving a paper for Bachelor Degree (the ratio between these and those who started the course) is a better indicator, but still we know nothing about the quality of the papers. My conviction is that a reliable perception of the quality in Higher Education – the output so to say – cannot be attained unless you go to the bottom line. Of course, all sorts of statistics are needed but it is difficult to draw conclusions about quality from quantity. The endeavor must then be to find operational cues that subsume various qualitative aspects of the case in point.

Things are moving in this direction in Sweden. In the next round of quality audits, the Swedish National Agency of Higher Education is planning to include a revision of degree papers. This is for sure a new step in order to measure more directly the quality of the learning outcomes, and there is a consensus on the fact that degree papers, combining quantitative and qualitative outcome, is a powerful indicator of performance and quality in HE. But, admittedly, this procedure is loaded with practical complications and that is what now is discussed at the Agency. Do the benefits gained by an evaluation along these lines compensate for the costs involved in undertaking an objective revision of all this material? Would it be sufficient to make a random control?

Summarizing, a basic claim on an indicator is that it reflects something essential about processes and outcomes that HE involves. Quantitative operators may be fallacious in a

quality assurance system if they are not clearly rooted in qualitative aspects of the outcome. As quantification is more operable – and operability must be taken into account - there will be a risk that quantity overshadows quality to the detriment of validity.

3.2 Indicators are selected for different *purposes*, but a claim must be that they reflect crucial aspects of the strategic planning, the outcomes and the demands that society makes on the HEI. There are more chances that reality moves in the right direction if the strategic planning and the indicators adherent to it are reproduced at different levels. There must be coherence between what the top management strive to improve and what departments do. Indicators at macro and micro level must be united in order to have effect.

The validity of indicators is intertwined with the purpose to which the outcome is meant to serve. If the purpose is to get a picture of the quality of learning outcomes, it would not be appropriate to not measure the number of examinees that year. If the purpose of the indicator is to climb up the ranking ladder, of course the relevant aspects of the ranking system should be carefully followed up, for example citation index, student exchange, gender and social balance, number and spread of master's programmes. A discussion of the purpose of the indicators will be an essential element in the selection process, and dependent on the goals that are set up.

The relation between indicator, purpose and the assessed phenomenon is often evident and unproblematic, but not always. Since the big computerized data systems came into use in the eighties, HE is haunted by quantification and data collecting, as elsewhere in Western societies. Quantitative terms like 'production of credits', 'percentages of BA graduates', 'average time for master degree', and so on. Certainly, it is important to know about the performance, and we know much more than before. Society has justly demands on HEI. But quantification has become more or less pervasive, with a risk of reducing quality to quantity. When gross figures of throughput (number of credits) are used as the principal instrument for planning and budget dimensioning, the relation to quality becomes problematic. In Sweden, and elsewhere, there are two indicators that control the distribution of resources and budget processes at different levels going from the Ministry down to the department level. That is number of enrolled students and number of credits taken. This, of course, concerns only education, not research.

This is not a bad system in itself, it is just quantitative by nature. Of course, it is "corrected" throughout the process: rectors prioritize and indicate new directions, faculties may reallocate funding, start up new courses and programmes. Still, quantification tends to favour what is big, according to normal market principles, and it would be difficult to touch upon an academic subject or a department that fulfill the quota that has been set up. Quality assurance policy is not absent, far from that, it has been developed and integrated during the last 20 years, particularly due to the Bologna process. But, still, it is treated as a second track. My point is that when quantitative criteria prevail in the steering processes, there is a risk of a

disassociation between sheer performance and quality of outcome. The links must be secured in a more organic way.

In Sweden a new system for evaluation of performance and quality is being discussed for the purpose of assigning resources. This debate originates from a report published in 2007 (*Resources for quality*) where the need for a qualitative-based funding system was put forward. The proposal I will comment on bears both upon the graduate cycles and the postgraduate.

The idea behind the recommendations in the report is to let one part of the appropriations to HEI depend on quality indicators and be exposed to competition between the HEI, and the other part be determined by the current norm that is based on production. As to the third cycle (postgraduate studies and research), the resources would be partially based on citations drawn from *ISI Web of Science*. The indicator brought up is the ratio between number of normal-cited articles (according to the world average) and the resources disposed of by the HEI for a specific area of research. A normal index has to be calculated for different areas. Statistical data from economic funding during a previous period of three years and data from citation index would make up this indicator.

In the discussion about this system several objections have been raised. In the first place, *ISI Web of Science* almost exclusively covers natural sciences, whereas humanities and social sciences have a tiny share. In the second place, it is clearly biased towards the Anglo-Saxon world. These shortcomings are evident for everyone and have led to the decision to build up a national base – SwePub – that should cover all scientific publications and constitute reliable basis.

As to graduate level a similar system has been brought up. Annual funding would be stripped and assigned as the result of different and quality-based criteria. For example, it has been proposed that a certain amount of money will be allotted to strengthen the research-supporting elements in undergraduate courses and programmes, another part will directly depend on the audits (which could include evaluation of exam papers) and a subsequent classification of departments and subjects. To sum up about the purpose of indicators and quantity-quality dilemma, I think that there is a growing tendency to get at indicators that show qualitative performance and that these should be used in the management of resources.

3.3 *The reliability of indicators* addresses the issue whether the measurement is secure and that repeated polls are trustworthy. No doubt, the more quantitative and well defined an indicator is, higher the reliability. The number of credits taken, the number of articles published in a defined set of scientific journals are reliable indicators, numbers do not cheat. When it comes to macro processes – like distribution of resources at high level – it is important to rely on indicators that are easily handled, and these are normally quantitative.

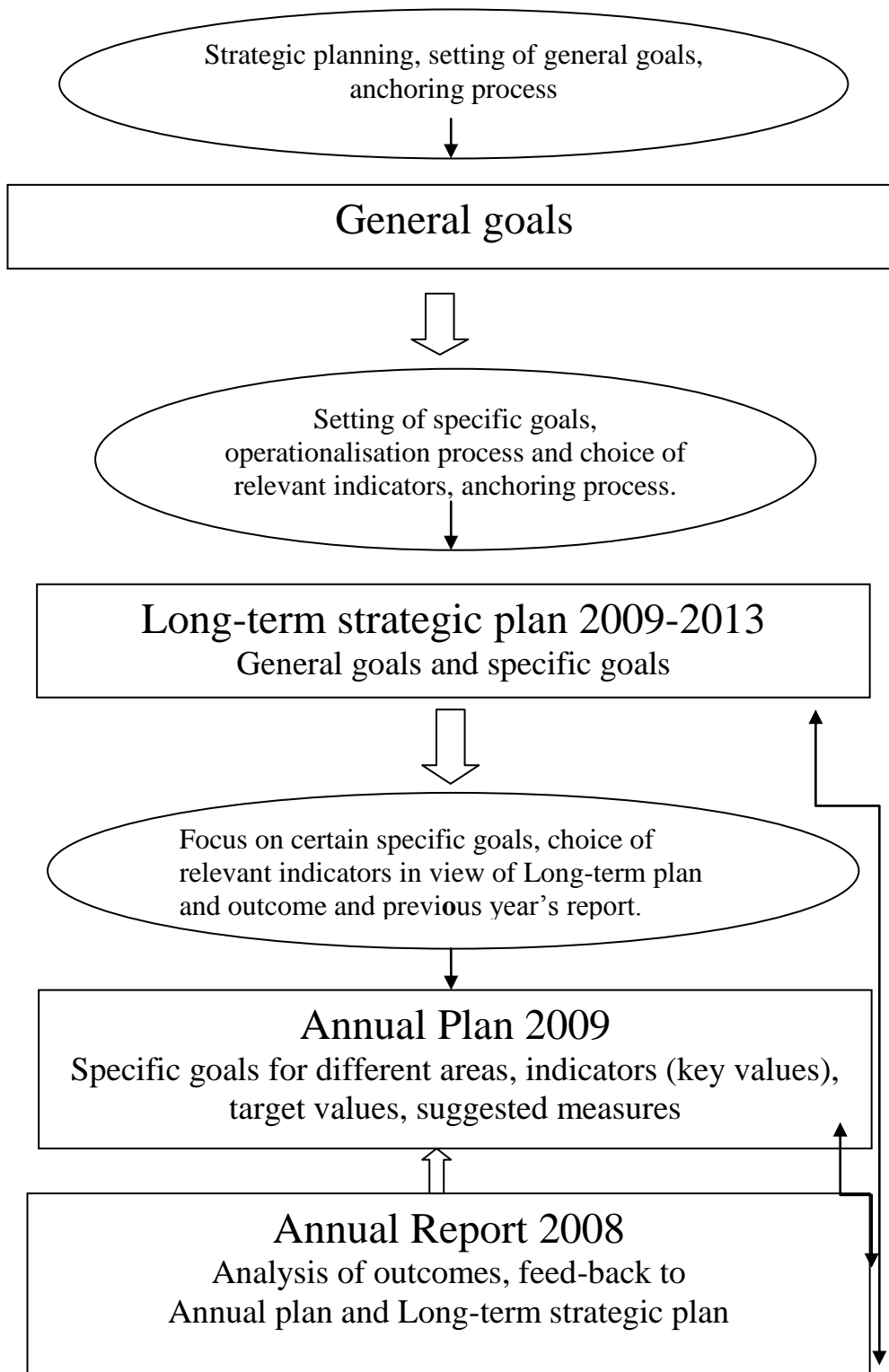
There seems to be an in-born conflict between reliability and validity. What is secure - indicators with high reliability - is often of scarce interest when you deal with complex activities. Counting unities may have a low validity for evaluation of teaching and research. On the other hand, criteria supposed to have a high validity, for example an indicator based on reports and classification of exam papers, may be very difficult to formalize and are, by this fact, less reliable. But the case stated here is not one of Scylla and Charybdis, there are surely ways to find reliable indicators that tell something important of the activities, and ways to make highly valid indicators manageable and useful in strategic planning. The challenge is to find them.

My impression is – valid for Sweden and perhaps for other European countries - that assessing and measuring the quality of outcome will be more important in the future and that management, dimensioning and allocation of resources will depend more on quality indicators in a strict sense.

4. Strategic planning – the case of Stockholm University

Stockholm University, with nearly 50,000 students, more than 2,000 different courses and programmes, is a typical capital university. The university has as clearly decentralized organization, transferring huge responsibilities concerning economy, course development, exchange and personnel to the faculties and departments. That is one important prerequisite for the strategic planning that the University is carrying out:

Strategic planning and indicators of performance



The core part of the Strategic Planning System at Stockholm University is the Annual Plan. In this plan a number of specific goals are picked out, others that have been fulfilled are removed. For each specific goal an indicator is stated as well as its target value for a specified period, for example 2011. Examples of measures and steps to be taken complete the picture of the actions within a specific area and of the expected outcome. The system is pushed by the previous report and pulled by the goals set up for the three-year-period to come.

In the 2009 Plan there are five focus areas, and nine general goals have been singled out, for instance: *quality in education, quality in research and postgraduate studies, profiling, competence supply and competence rising*, funding. Each of these goals is decomposed and made operable, but should be complemented by other kinds of goal-oriented actions. Just to give a glimpse of how the system works, for research and PhD-studies two specific goals are in focus in 2009:

Specific goals

- Increase of number/ratio of publications in leading scientific journals.
- Improvement of social and economic conditions of the PhD-students.

Indicators

- The number/ratio of publications in leading scientific journals shall increase.
- The ratio of students that finish their PhD-studies in four years shall increase.

Target values

- Increased values 2011 compared to 2000.

As such, this system for monitoring strategic development is not original. But it is built up in a consequent way and it has the necessary flexibility to alter focus according as environment changes or goals are fulfilled, without losing the long-term track. The discussions preceding the selection of indicators have been arduous because one important ambition has been to find operable but quality-driven indicators. It is easy to subscribe to what Scott L Elton once said: "Quality assessment must move from its stress on accountability for past performance to concerns about future performance and that way must be found to influence quality enhancement strategically." This challenge is also the challenge of how combine quantity with quality.

TOPICS FOR DISCUSSION

1. Discuss the choice of indicators for assessing the quality of a course? The following aspects should be taken into consideration: manageability, purpose, validity and reliability.
2. Discuss the eventual (negative) side-effects of the following three indicators: a) assigned resources are based on the number of credits taken (previous years) b) quality of teaching/education is based on ratio of teachers having a PhD c) assigned resources for research are based on the number of refereed papers.
3. Discuss the good and the bad aspects of a management system based on an abundant use of indicators. Does it promote a quality culture or is it just a new superstructure for its own sake?

References

Barkeman, Eva. 2009. "Innan citering styr fördelning bör verktyget testas mer [Before citation is used to govern funding the instrument should be tested more carefully]". *Forska* 2:2009. Swedish Council of Sciences. 8-9.

Elton, J. 1994. "Partnership, quality and standards in Higher Education". In D'Andrea, V. & David Gosling. *Quality in Higher Education*. Vol. 2. Issue 2. 95-104.

Harvey, Lee. 2009. "Demotrating quality". *Trends in Quality Assurance. A Selection of papers from the 3rd European Assurance Forum*. EUA. 5-9.

Hazelkorn, Ellen. 2009. "The emperor has no clothes? Rankings and the shift from quality assurance to world-class excellence. *Trends in Quality Assurance. A Selection of papers from the 3rd European Assurance Forum*. EUA. 10-18.

Kettunen, Juha. 2008. "A conceptual framework to help evaluate the quality of institutional performance". *Quality Assurance in Education*. Vol. 16:4 322-326.

"Selecting Performance Indicators". *Performance Monitoring and Evaluation TIPS*. USAID Center for Development, Information and Evaluation. 1996:6. (5/8/2009 http://pdf.usaid.gov/pdf_docs/pnaby214.pdf)

Stensaker, Bjørn. 2008. "Outcomes of Quality Management Assurance: A Discussion of Knowledge, Methodology and Validity". *Quality in Higher Education*. Vol 14, Issue 1. 3-13.

Technical Committee (TEC). "Monitoring and evaluation indicators for IWRM strategies and plans" (8/8/2009 http://www.gwpforum.org/gwp/library/Tec_brief_3_Monitoring.pdf)