

Benchmarking as a Driver for Change³

Benchmarking has been promoted in the HE sector by a range of groups for at least 20 years. It attracted attention in the early 1990s when economic recessions increased competition for public revenues and reduced funding for HE. The recessions also heightened interest in cost control and efficiency measures as ways of increasing productivity in public agencies generally and universities specifically. Universities looked to the private sector for successful examples of cost containment and quality improvement, one of which was benchmarking (Astele 1995:2–4). The financial conditions of recent years have again stimulated interest in benchmarking as a management tool. Some national agencies like the UK Higher Education Funding Council (2012) see benchmarking as “a valuable tool to identify efficiencies and control costs” and to help colleges and universities “make better use of scarce resources.” The UK Higher Education Statistics Agency (2010) takes a similar view—benchmarking is a way to “improve efficiency.”

The interest in benchmarking is not confined to the search for efficiency measures. Benchmarking is also a response to “increasing competition and demands for accountability (which) are changing the ways in which higher education institutions operate” (Weeks 2000:59.) Competitiveness is most readily observed in market-based systems of HE where institutions vie for students, faculty, and resources. Some of these market-based systems have competition between public and private universities and others have performance-based competition between public universities. But even in less competitive environments, central funding agencies look for efficiency and productivity measures, such as benchmarks, to guide resource allocation decisions (e.g., between sectors such

³ Prepared by Alan Ruby, Senior Fellow, Graduate School of Education, University of Pennsylvania.

as health and education, or between types of education, such as vocational and general education).

One response to these changes in the external environment has been to look to industries with a track record of success in quality improvement for ideas and strategies that will improve productivity. The most frequently cited example of successful benchmarking is that of Xerox, which responded to increased competition and loss of market share by integrating benchmarking into its organizational strategy to successfully reduce costs and improve productivity (Epper 1999; Astele 1995; Achtemeier and Simpson 2005).

Financial constraints and competitiveness have produced some notable examples of benchmarking in HE. The American Productivity and Quality Center (APQC) started HE studies in 1996 and has continued to work with various national and regional groups. Cross-nationally, the Association of Commonwealth Universities (ACU) began promoting benchmarking as a “self-improvement” tool for organizations on selected themes like strategic planning and risk management in the same year (PA Consulting 2011:14). Some within-country groups of universities have also adopted benchmarking as a way of improving management information. The Group of Eight Australia (2012) produces an “executive dashboard” to enable its members to benchmark performance on key variables like student numbers and research funding and output. And within the U.S., the National Association of Business Officers of Colleges and Universities (NACUBO) promotes benchmarking as both an efficiency measure and a process for self-improvement (NACUBO 2012).

While these groups differ in scale (e.g., from the eight Australian research-intensive universities to the over 2,500 NACUBO members) and in mission, they all see benchmarking as an active process focused on institutional improvement. This underscores some of the key characteristics of benchmarking.

Increased attention to institutional accountability in HE has also generated greater interest in benchmarking. Innovations in public administration and the increased autonomy of HEIs have encouraged ministries and HE coordinating agencies to look for ways to monitor and analyze institutional performance without intruding into the detailed working of universities. This approach to performance management is compatible with a “corporate governance” approach to managing public systems of HE in which universities are given greater autonomy in return for enhanced accountability. In this environment, institutions are encouraged to increase productivity and, in some cases, compete for funding on the basis of performance against system-wide or institutionally specific benchmarks and to participate in QA programs (Harman 2011).

Combined with financial constraints, increased competition, and the need for greater accountability constituted a significant shift in the operating environment of HE. Institutions and national agencies hence looked for tools to enhance

efficiency, control costs, and improve performance. Given its notable successes in the corporate world, benchmarking was a logical, and common, response.

Key Concepts in Benchmarking

There are some generally accepted features of benchmarking and a measure of consensus about its benefits and its shortcomings. The most distinctive feature of benchmarking is that it is an active process that focuses on improving performance. It engages people in the workplace in a process of learning about what they do now, studying how others do what they do, and comparing the relative merits of the different approaches with the aim of making improvements. This is well captured by Epper (1999:26): “benchmarking involves first examining and understanding your own internal work procedures, then searching for ‘best practices’ in other organizations...and finally adapting those practices...to improve performance.” While benchmarking was once a term used by carpenters and surveyors to refer to a standard that was known to be true and reliable, it now refers to a “process of measurement using an external standard to measure internal and external tasks” (Weeks 2000:60) and to “systematically making comparisons to... make improvements and to accomplish change”(Achtemeier and Simpson 2005:117).

Benchmarking shares with comparative education fundamental design questions: the choice of comparators, who, how many, and from what domains? Responses usually fall into two groups: within-field and across-field benchmarking. Within-field benchmarking concentrates on comparing like institutions: a process of “peer to peer” comparison or a within-class or domain comparison; e.g., comparing research-intensive institutions or those dedicated to the health sciences. This can increase the relevance of comparisons and make it easier to transfer practices and policies because the context is largely the same. It can also limit the range of options and alternatives investigated because of the similarities of the institutions being compared.

A more wide-ranging approach is to make comparisons with institutions that are “best” at the process or practice under scrutiny. This is sometimes called “generic benchmarking” and can include comparisons with organizations that are outside the industry. The comparisons are not limited to HEIs but across fields, looking at the same process in other industries. For example, to compare efficiency in the distribution of text books and learning materials, it might be instructive to include Amazon or another online retailer in the comparator group rather than just looking at other universities. One argument in favor of this approach is that studying the best will be more informative than studying a similar institution. Another is that it focuses attention on a specific business process rather than trying to understand how to improve an institution overall.

It concentrates on the locus of change or the point of intervention and makes the change process more manageable and probably more achievable.

From the perspective of HEIs, benchmarking has six main benefits embedded in its comprehensive approach:

- It develops an organizational culture committed to quality improvement by involving many parts of the university and a cross-section of personnel (faculty, administrators, trustees, students, and researchers) in the task of studying ways of improving performance. It can also involve the wider community including parents, alumni, employers, and other social partners.
- It uses a systematic approach to appraising potential competitors or exemplars and looks at their component parts individually rather than in a summative fashion, like a research productivity index.
- It helps with strategic planning and forecasting by looking at processes and policies that might be adopted in the future and examines how they have or have not worked elsewhere.
- It acts a source of new ideas and points to some possible goals. In particular, it identifies “real innovation” and “demonstrated best practices” rather than simply the way in which universities with the best reputations do things (Epper 1999:30).
- By emphasizing data collection, analysis, and systematic inquiry, it adopts an approach to problem solving that is compatible with the overall mission of universities (see Astele (1995: 3-11) for a discussion of some of these benefits).
- It focuses on creating a model of action by getting a “sense of exactly *how* other organizations have improved their performance” (Epper 1999:31).

Benchmarking has its criticisms. It is relatively expensive. It takes money and time, especially as it involves a period of self-study, and it is comprehensive, looking at various aspects of the university in depth. It is also costly to independently identify, collect, and verify the data needed to assess processes. One way to contain costs is to use a consortia approach, where members of, say, a trade association share data and information freely and sometimes anonymously (PA Consulting 2011:30–31). This collaborative approach is used by the two early initiators of benchmarking, APQC and ACU.

Another criticism is that benchmarking’s roots in the corporate world, which values profit, client satisfaction, and tight control, make it inappropriate for HE, which values collegiality, shared governance, and academic expertise. This view overlooks the institutional benefits that can be gained from balancing the “external demands for accountability and efficiency...with internal concerns for improvement and effectiveness” (Achte-meier and Simpson 2005:126).

Other critics see benchmarking as instrumentalist or conservative, fostering change at the margins rather than looking for substantive or fundamental change.

By looking primarily within an industry or field, benchmarking narrows the scope of the search for improvements to things that are already being done, at the expense of inventiveness. These critics see benchmarking not as a source of innovation but as a process of adaption or movement towards the industry “norm,” promoting mediocrity not excellence (Astele 1995: 33–34).

Most of these criticisms about the scope of change are based on a narrow approach to benchmarking, when the comparison is limited to like institutions. They do not apply as readily to cross-field or generic comparisons.

Distinguishing Benchmarking from Rankings

The increased interest in rankings since the early 1980s comes from some of the same factors that stimulated interest in benchmarking. Notably, they are both influenced by a desire to increase productivity, but they differ in how they propose to achieve this end.

The popularity of rankings is due to their simplicity. Rankings make it relatively easy to compare complex institutions by reducing many variables to a single value to produce a rank order. To determine which university is superior, rankings often aggregate scores for: reputation; research commitment and productivity; revenue raised, held, and spent; and students attracted and selected.

This simplicity is also the weakness of rankings. Turner argues that university league tables are “excessively simplistic” and do the mathematically “indefensible ...adding indicators which have completely different scales and ... variations ...which are not comparable. This error is compounded by aggregating measures from institutions and “systems where diverse and competing goals” exist (Turner 2005:371). It is like comparing a small sushi bar serving only the chef’s selection to a school cafeteria feeding nutritionally balanced lunches to a thousand students.

Despite these and other limitations, rankings have value. For example, they are useful for those seeking to make decisions about where to apply to study. Rankings simplify the task of evaluating the competing claims of many institutions. They can also serve as an aid to decision making when more detailed information about various universities is not available, accessible, or affordable. (See Ruby 2011) on the utilitarian value of rankings.)

Both forms of benchmarking, within- and across-field comparisons, stress the importance of looking for means or paths to improvement. This distinguishes benchmarking from rankings. University rankings are fundamentally about competition. They are attempts to assess which university is “best” or which is “better” than some others. Initial attempts to formalize these assessments were based on notions of measuring institutional effectiveness, asking which

university was the most productive. This basic formulation persists to some degree in most ranking systems.. They tend to look at inputs, outputs, and outcomes. But many look at only one or two of these dimensions and rely heavily on the reputation of the university or program rather than on observable productivity of the institution, or how well it uses its inputs. (See Shin and Toutkoushian (2011) for an overview of the history of rankings and the different models behind various ranking schemes.)

The stated aims of the more widely known ranking schemes vary. The Times Higher Education rankings refer to improving academic decision making as “helping university leaders...make strategic decisions” (Baty 2012). The QS rankings offer a “multi-faceted view of the relative strengths” of universities (QS 2012) and Shanghai Jai Tong rankings focus on research performance (ARWU 2012). None of these three offers guidance about what might be done at the institutional level to improve quality or lift productivity. The dominant purpose is competition: which university is the best? If there is a theory of change behind rankings, it is that the desire to improve its ranking will motivate an institution’s members to perform to a higher standard or more efficiently. Shame or pride in an institution’s place on a ladder or “league table” will encourage its members to look for a better way of doing things or to change behavior in some desirable way.

Conversely, the theory of change behind benchmarking is more elaborate and sophisticated. Alstete (1995) ties it to the continuous improvement cycle of “plan, do, check and act” and to human learning theory. Weeks (2000) uses a five-step linear model of problem specification, analysis, planning, action, and reflection. PA Consulting (2011) uses a “strategy contingent” approach based on four questions: “Where are we now, what do we need to know, what information is available, and what can we learn?” There are other logic models or theories of change in the benchmarking literature, but all adopt a process similar to these three. All use a problem statement that includes an assessment of the current state, followed by research and data gathering on the way other institutions do, or have, worked. This is followed by analyzing those practices and adapting them to suit the institutional context or redesigning an existing process to integrate improvements. This new approach is then tried and evaluated. A version of this theory of change is embedded in the processes and protocols associated with the UGSC.

The University Governance Screening Card

The UGSC was conceived as a tool to examine complex institutions and to examine one key variable: governance. It captures the various elements that shape governance in universities. The elements and the way they are defined are discussed elsewhere in this report and in various other World Bank reports.

The UGSC does not produce an index of good governance. Nor does it provide an aggregate score that would allow universities to be ranked on a scale like research productivity. Rather, it produces a chart that shows how an institution functions on five dimensions of governance and compares that with how a leader of the institution perceives its operations. This reveals the degree of alignment between the university's self-perception of its governance practices and a quantitative measurement of them.

Concentrating on alignment helps the self-reflection or self-study process that is most commonly used in the accreditation of established HEIs in the U.S. (Alstete 2004:62) and which is an integral part of most QA processes in the European Higher Education Area. The prominence of self-study in QA and accreditation comes from the widespread belief that it is likely to lead to institutional improvement.

The design and protocols for use of the UGSC acknowledge that institutions vary; even within the same field, academic tradition, and region they are different. These differences limit the value of summing the various scores on the scorecard to rank institutions. Similarly, the design and protocols of the UGSC do not assume or identify a specific model or form of governance; rather, they identify dimensions where institutional performance and perceptions of performance can be analyzed systematically.

The UGSC is a useful tool for self-reflection by members of institutions (as discussed in Part 6). But what are the benefits of the UGSC for governments, for national or cross-national QA groups, or for groups of like institutions? More specifically, how can the UGSC benefit cross-institutional groups?

Cross-Institutional Uses of the UGSC

The most obvious benefit is seen in the formation of the “communities of practice,”⁴ as occurred in the initial rounds of the UGSC's application in MENA. Leaders of regional institutions worked together to deepen their understanding of their own institutions and of others. The UGSC provided a framework for dialogue and a sharing of practices, and gave leaders a common set of data and concepts with which to talk about the strengths and weaknesses of their institutions, and to subsequently identify strategies for improvement or change. In a sense, the UGSC provided them with a “language” to discuss the practice of university governance.

In general, communities of practice tend to be self-regulating in terms of membership and program of work. Their growth and development depend on

⁴ I.e., groups where people “share understandings concerning what they are doing and what that means” (Lave and Wenger 1991:98).

the value members derive from the activities and exchanges. As professional communities, they create, validate, and share good practices. Sometimes they codify these into standards which they promote and celebrate. They encourage practitioners to take responsibility for the growth and development of their profession and institutions and to use the standards to determine membership and recognize or accredit institutions.

For ministries and QA agencies, communities of practice are cost-effective forums for communication and improvement. They offer the benefit of the effective transfer of good practices between institutions without the administrative burden of central collection and verification of data. Another benefit is that benchmarking of universities pursued on a collaborative basis is cheaper in terms of time and money than acting independently. Analyzing a recent benchmarking study for the UK Higher Education Statistics Agency, PA Consulting (2011) identified the ready access to verified and reliable data and information about practices of like institutions and of leaders in the particular area as the biggest source of savings.

But ministries or QA agencies can have a proactive role. For example, a ministry can use the UGSC process to foster improvement by sponsoring and supporting institutional participation in national or cross-national studies because it will stimulate reflection, comparison, and improvement.

A ministry or QA agency can also look at the alignment between the desired shape of governance embedded in a nation's policies about HE and its current reality. If a nation favors and promotes broad participation in institutional oversight bodies, how is this reflected in practice?

A ministry or QA agency can suggest that it favors a participatory model of governance—through community engagement, the involvement of social partners, and/or faculty involvement. It can place a value on student “voice.” By looking at the ministry's ideal, the rector's perception, and the assessment from the data assembled through the UGSC process, all parties can triangulate their assessments of a particular process or domain of governance. In some cases, this will point to areas in need of national and institutional attention. At the national level, it might suggest the need for laws and regulations to codify and promote greater participation or for changes to laws to limit the dominance of particular groups or agencies in governance structures. At the institutional level, it may point to the need for the inclusion of students on academic councils or the direct involvement of faculty, employers, and trade unions on oversight boards.

Ministries and QA agencies can work together across national boundaries just as institutions do. The benefits of the shared communities of practice can also be realized by regional groupings or consortia of ministries or QA agencies. An example comes from work under the Bologna process led by a network of European QA agencies to “develop an agreed set of standards, procedures and guidelines” and ensure that there is an “adequate peer review system” for QA

agencies. This culminated in a set of standards and guidelines for the European Higher Education Area (ENQA 2005) and provided the basis for closer cooperation between HE agencies in Europe, including the establishment of a European Register of Quality Assurance Agencies in 2009. The creation of the Register was seen as an important step in “modernizing” HE in the pursuit of three goals: “enhancing employability” of graduates; “strengthening mobility”; and “improving data collection and transparency.” (ENQA nd.) The latest step in pursuit of these goals and an illustration of the practical benefits of close cooperation is a study commenced in October 2012 of different national practices in the publication of QA reports. Following an exploration of stakeholder needs for transparent and comparable data, the project team will examine the feasibility of creating a “European template for quality assurance reports” to increase transparency (ENQA 2012).

In the MENA region, the promise of this form and level of cooperation is illustrated by the work supported by the British Council to underpin joint work on fostering excellence in HE. Jackson (2009:87) concluded that “the logic of cooperation is compelling,” arguing that smaller states do not have sufficient opportunity or capacity to foster and support “effective peer-review systems and need the expertise offered by other countries to help establish common standards and good practice.” Regardless of scale and wealth, nations gain from sharing expertise, good practices, and materials.

There are also wider benefits, especially in regions where skilled labor moves freely across national borders. Employers gain by having greater confidence in the qualifications of people from other nations and have access to a wider pool of skilled professionals. And individuals gain by having their credentials recognized and more widely accepted, giving them access to a wider pool of job opportunities.

These economic and social benefits are increased when there is a shared market for skill and when there are significant numbers of people seeking work outside the country where they were educated. Similarly, the benefits to individuals and nations are increased when there are skill shortages that can be filled due to labor mobility.

Conclusions

The UGSC draws on the lessons learned from benchmarking in HE over the last 20 years by focusing on areas and processes within institutions as the most likely domains for improvement. This distinguishes the UGSC from university rankings, as does its concern with promoting institutional and cross-institutional dialogue about change and improvement. The uses of the UGSC within communities of

practice and consortia offer real benefits to HEIs and to agencies concerned with national policy for HE and QA. The UGSC and the processes associated with its use are a significant development in the use of benchmarking to promote change and improvement in HE.