Defining the Role of Academics in Accountability

Elaine El-Khawas, George Washington University, Washington, DC USA

A selection of papers for the IMHE 2008 General Conference will be published in the OECD’s *Higher Education Management and Policy Journal*. The opinions expressed and arguments employed herein are those of the author and do not necessarily reflect the official views of the Organisation or of the governments of its member countries. © OECD 2008

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to: rights@oecd.org.
Defining the role of academics in accountability

While the policy debate on accountability in higher education has been vigorous in many countries around the world, debate has focused primarily on broad objectives or approaches. Much has been accomplished certainly, with more than 70 countries today having quality assurance agencies with responsibility for overseeing the quality of higher education provision (INQAAHE, 2008). Typically, governments have taken steps to establish an accrediting or quality assurance agency and empowered it to develop mechanisms for review of academic programs. Subsequently, higher education institutions have been required to submit to external reviews or to gather and report statistical information on a regular basis. A recent World-Bank funded project extends this approach further, offering a global initiative that encourages existing agencies to share information on good practice with other agencies just getting underway.

In recent years, there have been new initiatives to move beyond such broad policy approaches. Especially promising is the increasing focus on student performance and outcomes. This has become a major emphasis among national and international organizations that have responsibilities for quality assurance. The Quality Assurance Agency in the United Kingdom, for example, has offered leadership in assessing student learning through its initiatives to establish subject-matter benchmarks (Quality Assurance Agency, 2008). In several subject fields, they have sponsored the development of a set of formal statements about the expected coverage of subject matter that are specific, relatively comprehensive, and also flexible in their use. For master’s level study, for example, benchmark statements are available for five subjects: business and management; chemistry; engineering; pharmacy; and physics. For each of these subjects, narrative statements describe the content of course materials and what should be learned by students in master’s level study. Narratives describe the information to be learned as well as the experiences that students should have and the skills that should be developed. A clear underlying premise is that institutions have flexibility in how they structure their instructional programs to meet such benchmarks. Similarly, the statements acknowledge that a range of assessment tools are appropriate. In master’s-level physics, for example, thirteen different assessment methods are described as appropriate,
including examinations, individual and team project reports, peer assessments and direct observations of student actions to demonstrate learned skills.

In the US, there have been attempts to foster attention to student outcomes throughout the last few decades. Such arguments have been found especially among academics, marked by the continuing insights offered by Alexander Astin (e.g., 1993) and Peter Ewell (e.g., Ewell, 1983; Ewell, 1993; Ewell, 2003). In the early 1990s, the American Association for Higher Education, a national nonprofit organization, gave considerable visibility and support to the development of methodologies and pilot projects to develop student outcomes (cf. El-Khawas, 2005). About the same time, U.S accrediting agencies – including regionally organized accrediting organizations and other accrediting agencies that focus on professional specialties – became more active and many strengthened their standards to articulate the necessity of having study programs regularly assess whether student performance is meeting the standards of quality that should be expected (Kezar and. El-Khawas, 2003; Eaton, 2003)). In recent years, there has been another spur to a focus on student outcomes by colleges and universities as a result of recommendations made by a national commission sponsored by the US. Secretary of Education. Several of its final recommendations urged widespread development of improved approaches to assessing student outcomes, including the development of templates or standard formats for informing student applicants and the general public about student performance and outcomes at each university (The Secretary’s Commission,2007; Dickeson, 2006).

These and similar initiatives in other countries – notably including the Hong Kong Council on Academic Accreditation, the National Assessment and Accreditation Council in India, and the Australian University Quality Agency – have usefully shifted the policy debate toward developing measures of student performance and achievement in tertiary institutions. However, despite such progress, there has still been limited attention to the critical role of institutional implementation of such new mandates. More attention is needed to the many complex tasks that institutions of higher education must undertake in order to define and implement systems to assess learning outcomes At best, colleges and universities are called to action, but without any specification of who, what, where or how the recommendations are going to be implemented. So too, the role of academics is neglected in most policy discussions about quality assurance. Yet, any effort to define student outcomes and to assess a program’s success in helping students to reach those outcomes is, in its essence, an academic task, not an administrative one. Academics have the expertise in subject areas; they have the experience with instruction and with strategies for achieving instructional objectives as central components of their day-to-day work. Giving members of the professoriate a central role in accountability is also consistent with long-standing principles of academic autonomy. In some countries, academic freedom is buttressed by formal legislation or by the wording of university charters, but it also is reflected in widely accepted norms for universities in almost all countries.

Active academic involvement at the institutional level is a vital component of any quality assurance system based on student outcomes. Implementing accountability meaningfully requires decentralized implementation that is linked to the differing circumstances of study fields and levels (Barrie and Ginns, 2007). Unlike elementary and secondary levels of schooling, where a single outcome or limited set of outcomes can be identified (e.g., reading proficiency, or performance on international mathematics tests), differentiated outcomes are necessary at the tertiary level in order to evaluate the merits of each of program designed to train young people for a wide variety of careers. Training in the arts, in social welfare administration, or in chemistry, for example, call for developing and assessing quite different
skills and knowledge areas. Any single test of competence – even in broad areas such as effective writing – cannot do justice to what should be assessed to identify well-trained individuals in these or most other areas of tertiary-level study. Furthermore, complex tests of competencies are needed to assess the outcomes of study programs that involve master’s and doctoral study. Law, medicine, and other areas of advanced study call for testing outcomes in subspecialty topics as well as in specific skills and competencies unique to their fields.

This paper seeks to contribute to current policy discussion on quality assurance by offering perspective on ways that institutional actions might be taken to move forward with the important work of assessing student performance and outcomes. Following some general comments on the nature of policy implementation at the university level, the paper describes processes that need to be undertaken, primarily within institutions of higher education, in order to carry out national or system-wide initiatives. The paper then reviews some critical issues that must be addressed within each institution in determining how to engage academics in the development of outcome measures for student learning. While the primary focus is on institutional-level actions, some discussion is also devoted to the potential benefits of engaging national agencies or organizations in projects to conduct the sophisticated research and development needed to prepare valid, efficient, reliable and meaningful measures of student performance and outcomes.

The Difficulties of Policy Implementation

As complicated as it can be to reach a clear policy decision, whether in a legislative body or within a governmental ministry, scholars consistently warn that the effort to implement a new policy can be fraught with other difficulties. The research on policy implementation (e.g., Sabatier, 1986; Ewell, 1993; Gornitzka, Kyvik and Stensaker, 2002; Kogan 2005) offers several perspectives relevant to the challenges of adopting a student outcomes process in institutions of higher education. Given the decentralized structure of universities, with the core tasks of instruction found at the lowest organizational level of separate and sometimes diverse departments and programs, it can be expected that policy will be translated into practice unevenly, interpreted differently in various settings, or fail to take hold in some units. In the best of circumstances, a tension must be managed between the “top-down” approaches offered by administrative officials and the “bottom-up” perspectives of academics who have more day-to-day experience with what instructional approaches are most promising and what problems are faced in actual practice. And, as the late Maurice Kogan emphasized, implementation approaches have better prospects if organizers recognize that the implementation of new policies can have “deep implications for academic work,” often calling for adaptation of long held values and assumptions (Kogan, 2005, pp. 59, 64).

For a policy of introducing assessment focused on student performance at the tertiary level, at least two major stages of policy implementation can be identified:

- Development of assessment mechanisms/tools
- Implementation: the use of assessment tools

The development stage requires technical knowledge about methods of constructing assessment instruments and ensuring their validity. It also calls for subject matter expertise. These are areas in which academics have a central role, allowing them to apply their subject matter expertise to the development of assessment instruments. It need not be handled at the institutional level, however, particularly because it also requires the development of consensus around a host of design decisions.
Several approaches can be taken, and can include academic representatives in the process at appropriate points. For some countries, a ministry of education might organize an orderly process for accomplishing this task. It could appoint a special commission, or contract with a well-qualified organization to conduct the developmental work and also to coordinate opportunities for academic reviews and commentary on proposed approaches. Other precedents exist that rely on national (and international) professional societies or specialized accrediting agencies to organize a consensus-building process to identify and decide on student outcomes appropriate to their specialties. Indeed, professional organizations in such fields as business management, engineering and nursing already have experience in developing student outcome measures as part of their own leadership on accreditation matters (El-Khawas, 2001). Representatives of national and international bodies, including experts from UNESCO, OECD and the World Bank as well as representatives of national quality assurance agencies and their affiliate organizations (e.g., INQAAHE, ENQA), are needed for their ability to identify issues and to encourage wide dissemination.

The implementation stage, in which assessment tools are used, calls for very different types of involvement. Institutional-level action is the focus, although institutions may obtain guidance and assistance from external sources. A number of separate tasks must be addressed within institutions of higher education, including:

- developing an administrative plan and schedule for phasing in new systems
- developing assessment instruments (or modifying existing ones) to fit each of the institution’s programs
  - interpreting assessment results to public audiences (i.e., public accountability)
  - interpreting and acting on assessment results for program improvement (i.e., program accountability)

Each of these tasks involves academics, in some cases based on small working groups or in other circumstances requiring formal opportunities for wide consultation and commentary by academics in a range of subject fields. Some of these tasks also involve administrators, working jointly with academics. A range of institution-level actors thus need to be involved in developing measures of student learning that are to have system-wide impact. Roles and levels of involvement would differ according to the sequence of tasks needed for both development and implementation.

There is a danger of overly simplistic planning. Many important steps are required. It is necessary to allow time to assess the results of initial changes, for example, and then to make further modifications. Even when good results are achieved, these results need to be replicated in additional settings or with other groups of students, typically leading to still more modifications before wide-scale adoption of an approach can be recommended. Without care and discipline at each stage, chances of disappointing results are substantial.

**Institutional Actions: A Process Outline**

Figure 1 outlines in some detail the tasks that must be undertaken at the institutional level to put in place a system for assessing student learning with performance and outcome measures. As it suggests, multiple steps are called for, and need to be coordinated. For a specific institution, the developmental work may involve new steps while, for others, prior actions might offer a useful foundation that could speed remaining efforts. Similarly, some study fields may already have taken some steps toward an assessment system while others
would need to start from the beginning. The outline suggests, at the least, that such steps cannot be done well if they are hastily mandated. The integrity of the entire process would suffer unless good decisions are made with each of many judgment points.

Comments on critical issues that arise during the developmental and implementation process are noted below. Other resources also might be reviewed for further perspective on processes of instructional design, program evaluation, and assessment/accreditation approaches, all of which are relevant to the development of good institutional-level procedures.

**Figure 1:** Institutional-level Steps to Introduce Assessment Approaches

<table>
<thead>
<tr>
<th><strong>Institutional Planning:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Develop, refine and test assessment instruments suitable for all study programs.</td>
</tr>
<tr>
<td>2. Develop and refine assessment criteria, cut-offs and consequences.</td>
</tr>
<tr>
<td>3. Develop a program improvement strategy, based on assessment results.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Program-level Implementation:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Develop an administrative plan and schedule for carrying out assessments.</td>
</tr>
<tr>
<td>2. Develop and conduct reporting procedures for public accountability, so that assessment results are explained appropriately to governmental agencies, to students, and to the general public.</td>
</tr>
<tr>
<td>3. Implement program improvement, so that assessment results are discussed by program faculty and changes made in light of problems that are identified.</td>
</tr>
</tbody>
</table>

**Critical issues in the planning stage**

1. **Who will be in charge:** To carry out the planning and implementation steps that have been outlined, universities and other tertiary institutions generally will wish to designate some form of general coordinating body, whether a task force of academic council, made up of academics who can represent the range of study programs offered by the institution. The Assessment Council will have long-term responsibilities and thus might be structured formally or given clear links to other standing units. In light of the range of its duties for development and implementation, the Council also needs administrative support, including research and information technology resources. The Council’s members are to be academics, preferably persons with experience at the institution and who have gained the regard of colleagues for their fairness and sensitivity to student interests. Specific expertise in the design of instructional materials would be especially valuable. Academic expertise is a critical ingredient, especially because procedures must be constructed that are appropriate to different study fields.

2. **Selecting instruments:** Typically, universities can find resources that offer assistance with the initial developmental task of identifying assessment instruments that may be appropriate for use. Accrediting agencies, nonprofit educational institutions or partner institutions in other settings may offer resources. Some study programs at the institution may have relevant instruments that can be adapted to other programs. Because any instruments may need to be adapted, it generally will be better to find and adapt instruments that have already been used than to conduct the lengthy process of newly constructing an instrument. The adaptation process will accomplish needed refinement, while it is also giving faculty direct experience with the strengths and weaknesses of specific assessment instruments.

Preliminary selections, to identify a few assessment instruments, can be made by the Assessment Council. A subgroup of Council members may suggest modifications that will reflect the range of study programs the institution offers. A process is needed to allow faculty
in each study program to review assessment methods and determine which methods can best fit their program. In some cases, substantial modification may be needed but, most programs will find they can choose and/or modify something from a adequate range of existing instruments.

3. **Determining criteria and cutoffs.** While sophisticated methods exist for developing assessment instruments, methodological guidelines offer only a starting point for the decisions that institutions must make for a number of consequential decisions. Any assessment instrument needs a metric: what is a passing score, or a failing score? What performance will be considered “substandard” or “weak”? Will a single score be used to separate those students with acceptable performance from others, or will several gradations be used (e.g., very good; good; good with limited deficiencies; etc.)? Related to these decisions, it is necessary to identify what consequences will follow for students receiving various scores or assessment results. For failing or below-adequate scores, will students be allowed to retake assessments (or only the subparts of the assessment that they failed)? Will they be required to retake coursework?

The Assessment Council must take responsibility for developing detailed policies that spell out such decisions and procedures. Necessarily the Council will rely on substantial commentary and advice from individual academics and programs on how their students and programs might be affected. Also, the Assessment Council should make recommendations that receive final action or approval by some part of the university’s academic structure; at some institutions, the chief academic officer – a provost, dean, academic vice president or similar position – will make the final decisions, based on the recommendations sent forward by the Assessment Council.

4. **Designing a program improvement strategy.** Program improvement is at the heart of an assessment effort. Procedures must be designed that accomplish two purposes: assessment results must, first, be reported in such a way that they lend themselves to program change and, secondly, procedures must arrange for a time period in which the faculty in the affected program are to make modifications that respond to needed changes.

This is often a weak part of institutional models for assessing student learning, perhaps allowing programs such a limited time to make changes that they can only make superficial improvements, or promises. To make significant improvement, programs need to be given resources, including the assistance of experts, to help them develop good solutions to issues that arise. They also need time to discuss, consider and test alternatives.

For most institutions, a single, coordinated model for program improvement can be developed that is applicable to all of the institution’s study programs. A regular schedule of program review/improvement might be designed in which, on a rotating basis, each program undertake a formal review every 2, 3 or 4 years.

**Critical issues in the implementation stage**

1. **Developing a schedule for implementation.** Several legitimate issues are likely to emerge as assessment procedures are being developed. Most institutions encounter pressures to “go slow,” for example, as academics become immersed in working out unexpected problems and as they try to establish new assessment procedures. Nevertheless, in terms of setting a schedule, the overall institutional perspective should prevail; senior academic leaders should be sensitive to the problems that inevitably arise but must also establish a definite
schedule for initial implementation of the new assessment system. The schedule should take into account the concerns raised by academics but it must balance these internal concerns with the expectations of external audiences for disclosure of meaningful assessment information.

In many settings, a phased schedule may be workable. Under one version, certain study programs would be able to have reportable assessment results relatively quickly; these might include programs in engineering, business and management, and other fields that have already been subject to assessment processes. Other programs, newly encountering assessment methods or facing special issues in obtaining meaningful results for their students, might issue their first assessment results at a later date. For some institutions, results might be made available at an earlier date for certain levels of study and, for other programs (e.g., for advanced degree programs) at a later date.

2. **When decisions will count.** A major decision during initial implementation involves whether results will “count” for students. In some settings, administrators have announced results and explained that they will be used only for purposes of program improvement; students would continue to be evaluated on the basis of traditional grading procedures and the individual results of assessments would not be used for decisions on individual progress or completion. In other settings, administrators determine that assessment results will not “count” against individual students during the first few years that an assessment program is introduced, but results are publicly announced. In such a situation, poor results may get considerable publicity and actually motivate students to work harder in anticipation of the time when results will count against them.

3. **Interpreting results for public audiences.** Some difficult decisions are needed to establish what information will be made available to students, their parents, and the general public, stemming from the assessment process. It is tempting for administrators to release results in such a way that little guidance is available on what the results mean. In the United States, as various states issued “report cards” on the performance of higher education institutions in the last few decades, some states issued web-based reports filled with quantitative grids of statistics and obscure definitions. Sometimes, multiple measures were reported, giving the reader little guidance on how to interpret the measures. Another approach, also of limited value, is to issue narrative statements that focus on mainly on the strong points that have emerged in assessment results, giving little attention to weaknesses that were found. While guidance may be obtained from the experience of accrediting agencies, most institutions will need to allow time for trying out different approaches, perhaps working with small groups of students, parents, local citizens or other constituency groups to arrive at useful reporting styles.

4. **Interpreting results for program improvement.** Perhaps the most critical result of an assessment system is the ability to provide each program with practical feedback on ways to improve instruction and student learning. Needed, first of all, are measures that are in some way aligned with instruction. A useful approach, which might be established early in planning, is to align the assessments with key elements of the instructional program. For programs that include a writing component, for example, assessment instruments might be used that directly test student writing ability. For other programs that might include an internship or external field work, the assessment might include several components that are tied directly to such parts of the instructional program. Problems of interpretation can still arise. A survey response that indicates that students “did not learn enough practical skills” can have different interpretations: were opportunities offered but in an ineffective manner? Or are different opportunities needed?
Often, further inquiry is needed, or follow-up questions need to be added to assessment instruments.

A balance is needed between allowing the program’s faculty to interpret, possibly affected by their own emotional involvement with the program, and the interpretations that others might offer. One useful approach may be to have an administrative officer, or a subcommittee of Assessment Council members who have special experience with interpreting results and revising programs, meet with a program’s faculty and discuss results and possible implications.

**Conclusion**

This discussion has offered an outline of processes that are needed within institutions of higher education to adopt a student outcomes approach to accountability. It has discussed some of the important steps needed but acknowledges that much greater detail could be given on each step. Indeed, many resources and prototypes can be found in the work of quality assurance and accrediting agencies that have developed assessments in specific subject areas (cf. El-Khawas, 2001). Throughout, the discussion has emphasized the need to foster the engagement of academics in considering how to use assessment instruments. The larger goal is to encourage their use of these new approaches to assess student outcomes in individual subject areas.


