

Applying Learning Outcomes Concepts to Higher Education: An Overview

Prepared for the Hong Kong University Grants Committee

Peter Ewell

National Center for Higher Education Management Systems (NCHEMS)

This paper provides an overview of concepts associated with student learning outcomes and their principal applications in higher education settings. Its primary intent is to provide a background for discussing how these concepts might relate to university education in Hong Kong based on experience gained in other settings. While the majority of this experience has occurred in the U.S., applications of student learning outcomes concepts can now be found in many other parts of the world including the United Kingdom, Australia, New Zealand, and various countries in Western Europe.

The paper first provides some background on the learning outcomes tradition then goes on to a) define key terms and concepts, b) describe the principal strengths and weaknesses of using these concepts in higher education settings, c) examine the most important applications of these concepts in institutional and national settings and, d) review the most important ways in which institutions gather evidence that student learning outcomes are being attained. A brief concluding section outlines some potential policy questions that the University Grants Commission might consider in reflecting how these concepts might be applied to the Hong Kong context. An Annex provides and briefly describes a range Web-based references that can be examined to deepen understanding of these core concepts, and similar references are included throughout the body of the paper as appropriate.

Background and History. The notion of “student learning outcomes” has always been at the heart of university teaching and learning. Faculties in university classrooms have from the outset had an implicit notion of what they wanted students to learn in them and they have incorporated these objectives with various degrees of consciousness into what they teach, how they teach, and the ways they assess student performance. But it is only recently that extensive and visible attention has been paid to identifying in *operational* terms what students at various stages of their educational careers should know and be able to do.

Although progenitors were visible in various parts of the world, the notion of establishing explicit “learning objectives” began most prominently early in the last century in the U.S. in elementary and secondary education, guided by the work of John Dewey and others of the “American Pragmatist” school. Early experiments in extending these concepts to postsecondary settings were signaled by the establishment several mastery based university programs in the 1920s and 1930s in the U.S., most notably by Benjamin Bloom through the University Examiner’s Office at the University of Chicago and Alexander Meiklejohn’s Experimental College at the University of Wisconsin. With the massification of the U.S. higher education system in the 1960s and early 1970s these

pioneers were joined by a diverse but extensive range of “competency-based” college programs featuring narrative transcripts in place of grades or marks, and periodic demonstrations of student performance judged against established criteria associated with each level of each ability being sought.¹ This movement was provided additional momentum by the development of procedures to assess learning resulting from previous employment or practical experience so that it might count toward earning a university credential—a process that was frequently pursued by assembling a “portfolio” of demonstrated achievements organized in terms of an established learning outcomes framework.² Further stimulation was provided in all nations by the development of professional standards and associated licensing and competency examinations and assessments in many occupational/vocational and professional fields. All of these examples rested upon formal application of a learning outcomes framework to define, assess, and credential individual learning in a university setting.

A somewhat different set of applications emerged in the same time period in the U.S., using the attainment of defined learning outcomes as a measure of institutional effectiveness. For example, as early as 1928 a pioneering study tested a comprehensive array of learning outcomes for all undergraduate students enrolled in colleges in the state of Pennsylvania to determine patterns of overall attainment and effectiveness (Learned and Wood 1938). But again, it was not until a good deal later that such applications became widespread in the U.S., largely in the form of the so-called “assessment movement” that began in the mid-1980s with government calls to examine the effectiveness of the funds invested in public institutions of higher education by looking at how much graduates had learned by the point of graduation. These aggregate applications were supported by a growing literature and body of technique that had been previously applied solely to educational research (for example, Astin 1977, Pace 1979, Feldman and Newcomb 1969), but which were now deployed in attempts to systematically improve the teaching and learning process. And by the 1990s, these same approaches had found applications to institutional and programmatic quality assurance in the form of accreditation.³

Defining the Territory. While familiar elements of the academic landscape like “teaching,” “courses,” “modules,” and “degrees” have evolved some reasonably common meanings through continuing use (at least within a particular national context), distinctions among such concepts as “outcomes,” “learning,” “assessment,” and “effectiveness” remain relatively underdeveloped. Many authors and university systems have defined them from their own perspectives and have naturally concentrated on those elements of the topic that are closest to their own interests. But discussing learning outcomes sensibly requires an approach that can appropriately distinguish a) different levels of analysis, b) different kinds of “results” of an academic experience, and c) different perspectives or viewpoints.

¹ Prominent examples included Alverno College, Hampshire College, The School for New Learning at DePaul University, and Governors State University.

² For example, the now well-documented procedures to “assess prior learning” developed by the Council on Adult and Experiential Education (Keeton 1981, Whitaker 1989).

³ Subsequent sections of this paper will treat these applications in more detail.

One way to begin to make sense of this topic conceptually, therefore, is to think systematically about each component of the core concern “student learning outcomes.” Doing so first requires discussion of what is meant by an “outcome” and how this is different from other dimensions of performance. Second, it demands distinctions among units of analysis—at minimum, individual students and aggregates of students grouped by characteristic, academic program, or institution. Third, it requires one to distinguish “learning” from other kinds of “good effects” that students may experience as a result of participating in a postsecondary experience. And finally, it necessitates specific consideration of how we know whether (and to what degree) any of these results has occurred, and to what causes we can attribute them. A brief tour of the terminology associated with this tradition is provided below (key terms noted in *italics*).⁴

- “Outcomes” vs. “Outputs.” While an *outcome* in current academic usage is clearly the result of institutional and student activities and investments, there is a fair degree of conceptual consensus that not all “results” are properly considered outcomes. Numbers of graduates, numbers of teaching hours generated by a faculty, or types of service or research products are clearly results of what an institution of higher education does. But they are more commonly defined as *outputs* of higher education. Other dimensions of institutional or program *performance* like efficiency or productivity are equally the results of what an institution does, and assessing them may be important for some evaluative purpose. But they are not the same thing as outputs. This latter kind of performance constitutes the central conceptual foundation of what has come to be called “institutional effectiveness” in quality assurance discussions in the U.S., which examines the extent to which an institution as a whole attains the performance goals that it establishes for itself. Although outputs and performance are predominantly institution-level concepts, moreover, *outcomes* are only visible at the institutional level by aggregating what happens to individual students. For purposes of this discussion, therefore, an “outcome” can be most broadly defined as something that happens to an individual student (hopefully for the better) as a result of her or his attendance at an institution of higher education and/or participation in a particular course of study.
- “Learning” as a Special Kind of Outcome. Similarly, relevant and valuable outcomes are not confined to learning because students can benefit from their engagement in postsecondary study in many other ways. Additional *behavioral outcomes* or *experiences* that may result include employment and increased career mobility, enhanced incomes and lifestyles, the opportunity to enroll for more advanced educational studies, or simply a more fulfilled and reflective life. Presumably these are related to learning in some way, and evidence that students

⁴ These “definitions,” of course, are merely the central tendencies of a large and diffuse literature that has evolved over many years. Readers should be aware that some of these terms are defined somewhat differently by different authors and some remain contested. But they are nevertheless reasonably consensual across a wide body of practice within the learning outcomes tradition. This particular treatment is adapted from Ewell 2001, 4-8.

have obtained such benefits is often used by institutions as a proxy for instructional effectiveness. But the learning outcomes literature emphasizes that such subsequent experiences should not be confused with actual mastery of what has been taught. Although equally an outcome and frequently examined by institutions, student *satisfaction* with the university experience should also not be confused with learning. Certainly, satisfaction is important—especially if it is related to *motivation* and *persistence* (and therefore continued opportunity to learn). Student learning outcomes, then, are properly defined by this tradition in terms of the particular *levels of knowledge, skills, and abilities* that a student has attained at the end (or as a result) of her or his engagement in a particular set of teaching/learning experiences.

- Learning as “Attainment.” Defined in terms of the levels of *attainment* achieved, however, requires learning outcomes to be described in very specific terms. While institutions, disciplines, and professions vary considerably in the ways (and the extent to which) learning outcomes are described, several broad categories are usually distinguished. *Knowledge* or *cognitive outcomes* generally refer to particular areas of disciplinary or professional content that students can recall, explain, relate, and appropriately deploy. *Skills outcomes* generally refer to the learned capacity to do something—for example, think critically, communicate effectively, collaborate productively with colleagues, or perform particular technical procedures—as either an end in itself or as a prerequisite for further development. *Attitudinal* or *affective outcomes*, in turn, usually involve changes in beliefs or the development of certain values—for example, empathy, ethical behavior, self-respect, or respect for others.⁵ Learned *abilities*, typically involve the integration of knowledge, skills, and attitudes in complex ways that require multiple elements of learning. Examples include leadership, teamwork, effective problem-solving, and reflective practice. All such taxonomies require institutions or programs to define *learning goals* or *learning objectives* from the outset as guides for instruction and as benchmarks for judging individual student attainment. Expressed in terms of *competencies* or *qualifications*, moreover, such goals describe not only what is to be learned but also the specific levels of performance that students are expected to demonstrate. *Certification* or *mastery*, finally, implies that these specific levels have actually been attained.
- “Learning” as Development. In many cases, institutions and programs describe student learning not just in terms of attainment, but in terms of *growth* or *enhancement*. While this construction emphasizes the unique contribution of the educational program to current levels of student attainment, it also requires some knowledge of what levels of attainment characterized a given student before enrollment. *Value added*, “*before-after*,” and *net effects* are terms that are frequently used to describe such longitudinal ways of looking at development. This perspective, of course, need not be confined to student learning. For example, many educational programs base their claims of effectiveness on things like enhanced income, changes in career, or even increased satisfaction. From the

⁵ One the most widely cited taxonomies is in Astin 1977, 8-9; see also Ewell 1984, 11-44.

standpoint of quality assurance, both attainment and development may be important. Certification of specific levels of knowledge, skill, or ability for a given program completer—for example in the form of a licensure examination—is thus intended to guarantee that the certified individual is able to perform competently under a variety of circumstances. Evidence of this kind is claimed as especially important for employers seeking to hire such individuals or the clients who seek their services. Evidence about value added or net effects, in contrast, will be especially important elements of “quality” for prospective students who are looking for institutions or programs that will benefit them the most, or for policymakers and the public who seek maximum payoff for the resources that they have invested. In either case, it is important to be clear about definitions: *student learning outcomes* in this tradition most typically refers to the attainment of the particular competencies reached by students on completion of an academic program: if *development* or *value added* is intended as well, this must be clearly signaled.

- Assessment and Outcomes. A final key concept in the learning outcomes tradition is *assessment*, which refers primarily to the methods that an institution or program employs to gather evidence of student learning and/or to certify attainment. But historically, the term has been employed in several ways. In quality assurance, the most common meaning refers to the collection and use of *aggregated* information about student abilities (either in absolute or value-added terms) assembled to examine the extent to which program or institution-level learning outcomes goals are being achieved. But the term *assessment* is also commonly used to describe the process used to certify *individual* students, or even in some cases to determine marks or grades. Both of these can be looked at from the point of view of attainment against established *standards (criterion-referenced assessment)* or from the standpoint of how the performance of an individual or group compares to others (*norm-referenced assessment*). The term *evaluation* also commonly refers to evidence-gathering processes that are designed to examine program or institution-level effectiveness. But the object of evaluation usually extends beyond learning outcomes to examine a much wider domain of institutional performance. Finally, all these applications can be undertaken from a *formative* standpoint (that is, to advise or improve performance) or from a *summative* standpoint (that is, to judge performance for a decision or the record).⁶
- Evidence and Outcomes. Differences in concept and terminology are also apparent when describing the informational results of assessment. Here, terms like *measurement* and *indicator* are frequently used, implying that legitimate

⁶ Indeed, many assessment taxonomies are based on a cross between formative vs. summative purposes and individual vs. institutional units of analysis. At the individual level, “formative” assessment is equivalent to student advisement while “summative” represents certification. At the institutional/program level, “formative” assessment represents internal evaluation processes designed to improve the program while “summative” review (as in accreditation) informs substantive decisions about whether it will be funded or continue to operate (Terenzini 1989).

assessment should yield only quantitative results. Measurements, however, are only a special kind of *evidence*, which has come to predominate as the descriptor for assessment results in quality assurance contexts. Evidence can embrace the results of both quantitative and qualitative approaches to gathering information, both of which can be useful in examining learning. At the same time, the term evidence suggests both the context of “making and supporting a case” and the need to engage in consistent investigations that use multiple sources of information in a mutually reinforcing fashion. But to count as evidence of student learning outcomes, the information collected and presented must go beyond self-reports provided by students and graduates through such means as surveys and interviews or employment placements to include the direct examination of student work or performance.

Chart 1 attempts to display some of the key terms in this tradition in a tabular format so that the relationships among them are apparent. Within each column, a variety of commonly used terms is listed, each of which is appropriate for certain purposes. Different combinations of these terms define typical applications in certification, instruction, or quality assurance discussed in the third section of this paper. For example, mastery or competency-based instructional designs take the student as unit of analysis and directly examine knowledge, skill, or ability outcomes from the perspective of attainment using one or more of the methods listed under evidence of achievement. Assessment approaches applied for quality assurance purposes (for example in U.S. accreditation or in the examination-based quality assurance systems of several U.S. states and Brazil) examine aggregate outcomes at the institutional or program level that may look at the same set of abilities documented by the same kinds of evidence. Both are significantly different from more typical approaches to program evaluation which also take the perspective of the institution or program, but which examine outputs along with many other aspects of performance, with outputs themselves including many things beyond student learning.

CHART 1
A Taxonomy of Terms Commonly Used in Connection with the
“Assessment” of Student Learning Outcomes

Units of Analysis	Ways of Looking at Performance	Ways of Looking at Outcomes	Ways to Review Performance
Institution	Efficiency	Behaviors	Evaluation
	Productivity	<ul style="list-style-type: none"> • Employment • Further Education 	
	Effectiveness	<ul style="list-style-type: none"> • Career Mobility • Income 	
Program	Output		Measurement
	Productivity	Satisfaction	Indicator
Student	Outcome	Learning	Assessment
		<ul style="list-style-type: none"> • Knowledge • Skill • Ability • Attitude/Disposition 	Evidence of Achievement
			<ul style="list-style-type: none"> • Examinations • Performances • Student Work
		Attainment	
		Development	

Advantages and Drawbacks. Proponents have over the years claimed many advantages of adopting a learning outcomes approach, regardless of the unit of analysis to which it is applied. Among the most general claimed benefits are the following:

- Clarity. Using the language of learning outcomes can help focus sharper attention on the objectives of the teaching-learning process. At the level of the institution or program, this can help foster communication and align curricular designs and instructional delivery across diverse teaching staffs. If articulated in the form of a regional or national qualifications framework, the same alignment can be sought across institutions offering similar credentials. At the individual student level, creating course or module syllabi that are structured around learning outcomes can help communicate expectations to students about what levels and kinds of performance are demanded, helping them focus their efforts more effectively. Finally, the clarity of a learning outcomes approach has considerable appeal to external stakeholders like policymakers and employers who by nature are inclined to judge the effectiveness of an enterprise in terms of its results.

- Flexibility. Learning outcomes specify the intended ends of instruction but leave open the means to attain these ends. This accords considerable flexibility for instructional provision. At the program level, very different instructional designs and learning environments can be configured to foster the same learning outcomes including self-paced approaches, approaches using distance delivery and other forms of technology, or modular designs that either break up coursework or alternate formal study with internships or other work experience. So long as it can be demonstrated that these varied experiences result in a comparable “educational product,” the formal means of instruction do not matter. Similarly, very different kinds of students can be accommodated through an outcomes-based approach. Different instructional paths can be devised to suit the individual needs of learners based on educational and experiential background, levels of knowledge and skills at entry, and personal learning style. As an auxiliary, a recognized outcomes framework can enable valuable past experience (usually obtained in the workplace) to be recognized through assessment and incorporated into a learning plan.
- Comparison. Credible learning outcomes can establish comparable standards through which to benchmark and evaluate the performances of institutions, programs, courses, or individual students. While traditional assessment and grading schemes already claim to do this, variations in local standards (as well as variations in the grading metrics used to signify individual levels of achievement across institutions and national contexts) render them far less useful for purposes of comparison. At the institutional or program level, such comparisons can be applied to support summative assessments of program performance for accountability purposes, or they can be used to chart progress or benchmark against peers as part of local improvement efforts. At the level of the individual student, comparisons of assessed outcomes with recognized standards or criteria can form the basis of certified attainment (as in licensed professions), or they can provide a sound basis for admission or placement either in comparison with other students (normative) or in terms of previously established criteria (summative).
- Portability. In a similar fashion, credible learning outcomes can form the basis for a system of credentialing student learning that can transcend established programmatic, institutional, and national boundaries.⁷ Diplomas or degrees representing the completion of particular courses of study can be mapped to appropriate arrays of competencies at various levels to establish comparability despite differences in nomenclature, program design, or length of study. More importantly in an age of growing student mobility and modularity of instructional provision, learning outcomes frameworks can be used to establish the relative comparability—and therefore transferability—of learning experiences *across* formal programs. The resulting portability of learning from one setting to another, if designed appropriately, can both increase the capacity and alignment

⁷ See for example Johnstone, Ewell, and Paulson 2002.

of a multi-institutional system of instructional provision and provide more accessible paths for different kinds of students to attain higher credentials.

Despite these many potential advantages to adopting an outcomes based approach, experience in many settings suggests a number of cautions. Among the most prominent of these drawbacks are the following.

- Definition. All of the advantages noted above are premised on the existence of meaningful, clear, credible, and assessable statements of learning outcomes. This requires learning outcomes statements that can succinctly and accurately describe the characteristic or ability in question at a sufficient level of generality to cross contexts, but with sufficient precision and consistency to enable a valid and reliable judgment to be rendered. This is not easy to do for many abilities, and it has proven nearly impossible for some.

Two considerations are important here, but they are not always considered explicitly in debates about the suitability of adopting an outcomes perspective. First, any agreement about definitions is always provisional—limited by the intended range of application for the planned framework, which should always be explicitly bounded. The acceptance of learning outcomes as credible markers of attainment will therefore always be appropriately confined to a particular *community of judgment* that consists of a specific faculty, groups of faculties within and across institutions, or (ideally) coalitions of faculties and stakeholders. Second, statements of learning outcomes only have meaning in the context of the assessment methods or bodies of evidence that render them *operational*. Apparent agreement on the language of a given outcome may mask important differences in the way various parties recognize that it is present—which may in turn reflect significant differences in the ways they actually construct the ability. This means that any consideration of definition for a proposed learning outcome cannot avoid the question of evidence—of how, operationally, attainment will be concretely recognizable.

- Legitimacy. Just as important as definitions of learning outcomes are accompanying *perceptions* of these definitions—especially on the part of members of an academic community. A first challenge here is philosophical, stemming from the healthy skepticism of many academics who believe that learning is “ineffable” and therefore not able to be meaningfully captured by simple learning outcomes statements, however they are constructed or assessed. This, of course, to some extent begs the question because academics assess student learning all the time from within the frame of their own disciplinary expertise. But philosophical objections of this kind have merit in limiting the often excessively universal claims of validity promoted by some outcomes schemes. And experience suggests that reductionism and reification will always constitute a prominent challenge in implementing an outcomes based approach.

A second related difficulty is language, because the terms and concepts underlying outcomes based approaches are fundamentally rooted in the contexts of business, education, and the social sciences. Business concepts (like those associated with Total Quality Management) provoke natural suspicion in much of the academy because they are associated with what many see as growing commercialization or “managerialism” in higher education.⁸ At the same time, education and the social sciences are not generally at the top of the disciplinary “pecking order” at most universities. Together, these perceptions mean that the initial legitimacy that any outcomes based approach will command will vary significantly and predictably by disciplines. For the professions, accustomed to external standards and frequently subject to licensing examinations governing entry, the approach will be largely familiar and should encounter little resistance. For other disciplines, care and time must be taken to allow the underlying concepts to be translated and internalized.

- Fractionation. By their very nature, outcomes schemes tend to break down holistic conceptions of learning. This tendency toward “fractionation” may have important, though unintended, consequences. From the operational perspective, the level at which learning is assessed may become too narrow, missing the essence of the integrated “ability” that is supposed to unite many discrete skill elements into expert practice.⁹ This has real implications for assessment because an assessment approach that concentrates on demonstrating student mastery of technique in the absence of contextual factors, and with all the necessary information about a problem supplied, may completely miss important elements of expert practice. At a deeper level, this tendency may subtly privilege an “additive” over a “developmental” view of the learning process—one that views learning largely as a process of incrementing a student’s current inventory of knowledge and skills with new elements one at a time, rather than one that emphasizes cognitive reorganization at a higher level. This may lead to instructional approaches that are excessively reductionist and assessments that are overly mechanical. Finally, the same tendency toward fractionation can reinforce already strong tendencies to break instructional programs down into smaller and smaller components. As noted, part of the attractiveness of outcomes based approaches is that the flexibility they provide may enable diverse educational experiences and learning from many settings to be knit together for individual learners. But that same tendency may exacerbate challenges of maintaining coherence and standards.
- Serendipity. Establishing a particular array of learning outcomes, no matter how well crafted, leaves out the unexpected. The approach therefore presumes that all of the valued and important ways that a learner can construct meaning in the

⁸ See for example Bruneau and Savage 2002.

⁹ This distinction is often framed in the claimed dichotomy between “education” and “training,” where the latter is believed to concentrate solely on transmitting discrete and applied skills while the former involves mastery of the deeper conceptual and contextual grounding of any ability. As should be apparent, this paper’s position is that this claimed distinction is not very useful.

context of a particular discipline or ability are known in advance. This presumption is likely to be less and less applicable at increasingly higher levels of attainment. Constrained serendipity may thus present less of a problem at more basic levels of achievement than for advanced study in any discipline. Similarly, this challenge may vary considerably by discipline. While all fields of study value autonomous scholarship at advanced levels, some—like the fine arts—will likely emphasize creativity and the development of an individual voice at quite early stages of an instructional program. As a result, the assessments developed to implement learning outcomes based approaches must be designed to accommodate unexpected demonstrations of the ability in question. More importantly, learning outcomes taxonomies themselves must be subject to periodic revision as more experience is gained about the actual dimensions of student performance.

Understanding the advantages and drawbacks associated with adopting a learning outcomes based approach is critical, but both are necessarily cast at a very high level of generality in the above discussion. The particular ways of making an outcomes approach operational in the form of application and assessment outlined in the following two sections help make these conceptual points more concrete.

Student Learning Outcomes in Practice. Learning outcomes approaches have been used at many levels, ranging from that of instructional design where the individual student is the object of interest, through institutions and programs where the prominent concerns are evaluation-based program improvement and quality assurance. Some of the most common examples of application under each of these headings are described in this section, together with associated strengths, drawbacks, and examples.¹⁰ Many of the applications described can be deployed at multiple levels, but they are discussed under the heading at which they most commonly occur.

National or State Level. There are many examples of applications of the learning outcomes concept at the national or sector level in different nations, as well as in individual states in federal polities like the U.S., Australia, or Germany. Most of these do not involve independent outcomes-based policies or initiatives, but instead include the learning outcomes approach (or some aspects of it) in broader funding, accountability, or quality assurance processes. Among the most prominent examples of such applications are the following:

- Institutional or State Performance Indicators. Indicators of student learning outcomes are most frequently deployed as part of the larger accountability framework based on statistical performance indicators that became prominent in U.S. states in the early 1990s and which spread to many other countries including Canada, Australia, New Zealand, and the United Kingdom and other countries in Western Europe.¹¹ Virtually all of these indicators in the realm of student

¹⁰ Many sources and websites associated with each of the examples given are listed in the Annotated Reference List which is included as an Annex.

¹¹ See Ruppert 1994 and Burke, Yang and Minassins (2002) for some discussions of the U.S. experience.

learning outcomes are derived from standardized examinations—either administered especially for the purpose of grounding overall judgments of quality, or derived from the many licensing examinations administered to govern individual entrance into professional practice. The most common of these schemes rate institutions on a comparative basis, while a few attempt to reflect overall achievement within a geographic or political unit. As an example of the latter, the National Center for Public Policy in Higher Education recently issued a report on a five-state demonstration project in the U.S. that created a “learning profile” for each state that included the assessed performances of samples of two-year and four-year undergraduates on authentic tests of advanced intellectual skills, together with comparative indices based on commonly-administered standardized professional licensure examinations and examinations used to admit students to graduate studies (Miller and Ewell 2005). Similarly, six individual states in the U.S. administer standardized examinations to students enrolled in public institutions as part of their wider accountability schemes (Ewell and Ries 2000). Meanwhile, sixteen states in the U.S. compile statistics on licensure pass rates for teacher education programs, as well as for vocationally-oriented two-year college programs, and include them in their performance reporting schemes (Burke 2005).

Outside the U.S., however, student testing to create performance measures purely for accountability purposes is relatively rare, but there are occasional instances. Brazil, for example, conducts program-level standardized examinations of completing students at public universities on a national basis to examine institutional quality with results reported by institution—an exercise that was recently converted to a sample-based “value added” measure by adding a pre-test of students administered early in their study career (Schwartzman 2004, de Moura Castro 2002).

Strengths or advantages of applying an outcomes-based approach to accountability through performance indicators include their credibility for external stakeholders and their ability to focus institutional and faculty attention on deliberately designing educational experiences to develop particular desired attributes of students and graduates. Challenges associated with this approach include the difficulty of finding evidence for all of the outcomes that are of interest that is valid, reliable, and acceptable to faculty across fields of study.¹²

- Resource Allocation and “Institutional Steering.” Information about student learning outcomes is only rarely used by states and nations to inform the process of providing resources to support institutions and programs. In the U.S., only the states of Tennessee and South Carolina have linked institutional resource allocation to collective student performance on standardized assessments (Burke 2005). In the former state’s long standing “performance funding” program,

¹² Finding credible end-point assessment instruments for every discipline offered at a major university, for example, was one of the most significant challenges associated with the establishment of the U.S. state of Tennessee’s “performance funding” scheme in the early 1980s (Banta 1986).

samples of students are tested in general education (in a before-after “value-added” design) and in the discipline corresponding to their major field of study (as a capstone). Higher scores mean more dollars allocated to the institution in the 5% set-aside pool associated with the program as a whole. In the latter, collective performance on licensing examinations (especially for two-year vocational institutions) helped determine the overall level of resources granted to institutions on the basis of a similar formula.¹³

But two indirect linkages between outcomes assessment and resource allocation are a good deal more common. The first of these makes institutional eligibility for the receipt of public funds contingent on institutions engaging in the assessment of student learning or on actual performance on assessments. For example, institutional accreditation in the U.S. requires institutions to engage in their own local assessment processes and achieving accredited status is a condition for the receipt of federal funds for U.S. colleges and universities. U.S. states like Oklahoma, Missouri, and South Dakota have similar conditions for the receipt of state support by public colleges and universities. The second indirect approach relies on consumer choice—informed by information about student outcomes and experiences—to influence the flow of resources to institutions through tuition and fees. For example, federal law in the U.S. requires all institutions receiving federal funds to make available information about degree completion rates to prospective students and their parents. In Australia, nationally-sponsored surveys of student outcomes and experiences are a required part of the accountability process but are seen to be equally important in informing student choice.¹⁴

Strengths or advantages of this approach again include its appeal to external stakeholders as a true incentive for institutional performance and its unmatched ability to get the attention of institutions in response. Associated drawbacks—especially if implemented in a formulaic “performance funding” context include the complexity of the process¹⁵ and the tendency for institutions to try to “game” the process by maximizing indicator values instead of attending to underlying instructional issues.¹⁶ As a result, the most effective and long-lasting examples of this linkage tend to be indirect through accreditation or the provision of consumer information.

- **Alignment of Standards.** Beginning in the early 1990s, learning outcomes frameworks have become increasingly prominent in national efforts to ensure that the degrees and other credentials granted by different institutions and programs are of comparable quality. As might be expected, the first of these efforts were in

¹³ The South Carolina performance funding program has since been suspended.

¹⁴ See Griffin, Coates, McInnis and James 2003; for a description of the survey process see www.itl.usyd.edu.au/Teval.

¹⁵ Together with the difficulty of finding suitable quantitative indicators for all outcomes that can be used in an allocation formula.

¹⁶ See Harvey and Newton 2004; Ewell forthcoming.

vocational areas of study—applications where external stakeholders have a substantial interest in ensuring uniformity.¹⁷ Resulting “qualifications frameworks” generalized to all fields of study have emerged in most of the English-speaking world outside the U.S. including England, Scotland, Ireland, Australia, South Africa, Namibia and New Zealand. In general, qualifications frameworks comprise a matrix where one axis consists of a set of generic abilities or traits that are expected as a result of postsecondary study while the other axis consists of a hierarchy of levels or standards at which the particular ability or trait is manifest. Individual cell entries in this matrix contain concrete descriptions of the ability in question that can be applied to any given institution or programmatic offering to ensure that it is in alignment. For example, the Scottish Credit and Qualifications Framework is intended to clarify the relationships between various awards and clarify entry/exit points for different levels of study for students, employers, and institutions. The framework embraces twelve levels extending to doctoral study, with abilities statements grouped under five basic domains: knowledge and understanding; practice (applied knowledge and understanding); generic cognitive skills; communications, computing and numeracy skills; and autonomy, accountability and working with others (SQF 2003).

Perhaps the most extensive effort of this kind was undertaken in the late 1990s in the UK, where benchmarks were established for 46 disciplines at the baccalaureate Honors level.¹⁸ Subject benchmarks are developed by subject specialists under the auspices of the Quality Assurance Agency and are intended to both align instructional provision at the subject level and to provide a basis for subject-level reviews of academic quality. Statements of intended learning outcomes are a component part of all subject level benchmarks, although these statements also contain additional standards related to instructional materials and pedagogy.

As a final example, a cross-national effort at standards alignment is currently under way in Europe under the auspices of the Bologna Process, which seeks to create a “common European higher education area” by 2010 through a uniform set of degree levels and standards at the bachelors and masters levels. While much of this effort is currently concentrated on program structure and the time needed to earn a degree, several initiatives are under way to ensure the alignment of outcomes (see Adam 2004). For example, the so-called “Tuning Project” has identified a set of core attributes of graduates at each level in a range of participating countries and has applied this framework successfully in nine disciplines at 130 institutions across a range of national contexts in Europe.¹⁹ A

¹⁷ For example, the National Vocational Qualifications Framework in the UK (see Atkins, Beattie, and Dockrell 1993).

¹⁸ See <http://www.qaa.ac.uk/academicinfrastructure/benchmark/honours/default.asp>.

¹⁹ See www.let.rug.nl/TuningProject/index.htm.

similar outcomes framework has been developed to apply to doctoral level study.²⁰

The primary advantage of developing qualifications frameworks is their ability to link diverse course and programmatic offerings under a single (though admittedly broad) set of standards that reflect agreed-upon equivalencies of achievement. This can both assure external stakeholders that a consistent educational “product” is being generated by all institutions and programs and can facilitate student mobility geographically and across different institutional and programmatic contexts. Disadvantages of this approach are the difficulty of casting the requisite statements of ability at the right level (broad enough to be applicable to all fields of study yet specific enough to provide guidance for assessment), and the tendency to homogenize standards across institutional and programmatic contexts which really are of varying levels and kinds by forcing them all to the “lowest common denominator” (Blackmur 2004).

- Accreditation. Accreditation has become a world-wide mechanism for certifying the basic acceptability of an institution or program based on self-study and peer review. The process began at the institutional level in the U.S. about a century ago and has since spread to other national contexts (principally Eastern Europe at this point) as a first step in aligning academic standards and providing a publicly credible system of minimum quality assurance. More recently in Western Europe, accreditation schemes are replacing or supplementing more complex and intrusive national quality review or audit processes (Westerheijden 2001). In parallel with institutional accreditation, a number of individual fields of study (most notably business and engineering) have developed programmatic accreditation standards which seek to align programs against common standards regardless of where the program is located.

The role of student learning outcomes became increasingly prominent in U.S. accreditation in the late 1990s at both the institutional and programmatic levels, partly through the stimulus provided by the federal government. At the institutional level, accreditation currently requires institutions to have adopted their own system of visibly assessing student learning outcomes against their own goals for learning. This requires them to a) set clear (“measurable”) goals for learning framed in outcomes terms for general education and for each program of study, b) establish a method for gathering and interpreting evidence of the achievement of these goals by students at various levels that goes beyond individual grades or self-reports and, c) visibly use the results of the assessment process to improve the teaching and learning process.²¹ In addition, institutional accreditation standards in the U.S. require institutions to include in their list of intended outcomes a particular set of student abilities (though these are cast at a

²⁰ The so-called “Dublin” Descriptors, see http://www.unibuc.ro/uploads_ro/35714/Dublin_Descriptors_2004Doctor.pdf.

²¹ The Council for Higher Education Accreditation (CHEA) has recently put forward a statement of principles for accrediting organizations and institutions agreeing to these points (CHEA 2003).

very general level), and the review process is in part directed at whether or not these valued outcomes are attended to (AACU 2004b).²² Specialized accreditation also address generic outcomes of this kind and frequently goes a step further by requiring particular kinds of learning results (especially if the field of study has a licensing requirement) or particular kinds of assessment features.²³ But despite this general trend, accreditation schemes in the U.S. have so far been quite varied in the degree of emphasis that they place on student learning outcomes in the review process and in the relative weight of outcomes-related factors in informing an accreditation decision in comparison to more traditional earmarks of program quality like resources, curricular design, and faculty credentials.

Advantages of addressing student learning outcomes within the framework of a broader accreditation approach include the ability to tailor both the outcomes of interest and associated assessment methods to important differences in institutional mission and context, and the reliance on multi-faceted peer judgment to determine if adequate and appropriate evidence of student learning is present (and being used). Prominent weaknesses include the challenge of finding and training peer reviewers with expertise in assessment and the tendency for traditional quality markers to supplant outcomes as indicators of “quality” when the two are mixed in the same process.

- Quality Reviews (Audit). A parallel method for assuring institutional or program quality evolved in the late 1980s and early 1990s in Europe and Australasia centered on quality process reviews. Most commonly called “audit” in the English-speaking world, these approaches are centered on examining an institution’s own internal evaluation and evidence-gathering approaches in depth, usually through application of a detailed review protocol which enables reviewers to “drill down” to the operational level through techniques like following audit trails to determine if internal evaluation processes are carried out consistently and as designed (Dill 2000).²⁴ For some of its proponents, the audit approach was intended as a conscious alternative to examining evidence of student learning outcomes (Massy 1997) because of the challenges of definition and measurement associated with the latter.

But with the advent of the Bologna Process in Europe, which emphasizes the need to create aligned academic standards centered on particular degree levels, this has

²² For example, CFR2.2 of the Western Association of Schools and Colleges (WASC) Senior Commission notes “college level written and oral communication; college level quantitative skills; information literacy; and the habit of critical analysis of data and argument; understanding of diversity; civic responsibility; the ability to work with others; and the capability to engage in lifelong learning” as intended outcomes that should be included (WASC 2001, p.20.)

²³ For example, the Accrediting Board for Engineering Technology has adopted eleven general outcomes statements that should be assessed in all programs: see <http://www.abet.org/Linked%20Documents-UPDATE/Criteria%20and%20PP/05-06-EAC%20Criteria.pdf>

²⁴ The current Hong Kong Teaching and Learning Quality Process Review (TLQPR) run by the UGC is, of course, a classic example of this approach.

begun to change. For example, pilot multi-national quality reviews were recently undertaken for bachelors programs in three disciplines in eleven European countries through the auspices of the European Association for Quality Assurance in Higher Education (ENQA).²⁵ This effort is now in a second phase and involves partnerships of national quality assurance agencies that undertake joint quality reviews under a common set of guidelines. Like many accreditation approaches, the ENQA joint reviews identify a set of generic outcomes that should be present in all bachelors degree programs and explicitly reference the Bologna-related Tuning Project competencies and the “Dublin Descriptors.” But the review process makes no attempt to directly assess these outcomes and instead concentrates on examining quality assurance processes and the adequacy of local assessment efforts.

An advantage of approaching learning outcomes within the setting of quality review and audit, like that of accreditation, is its ability to accommodate appropriate differences in institutional mission and context. In addition, audit-based quality reviews have the advantage of focusing evaluative attention on teaching and learning to an extent that is generally superior to more general accreditation approaches (Massy and French 2001). But a disadvantage of this approach is that its concentration on institutional processes and locally-defined standards to assure outcomes, rather than examining outcomes directly, may miss real differences in standards and student performance across institutions (Westerheijden forthcoming).

Institutional Level. The majority of institution-level applications of student learning outcomes concepts are currently located at U.S. institutions under the auspices of the “assessment movement.” As noted, this phenomenon began in the mid-1980s, directed primarily at gathering aggregate evidence of student learning and applying this information to the improvement of academic programs.²⁶ Similar efforts have emerged more recently in other English-speaking countries including Australia, growing out of widespread use of Outcomes-Based-Education (OBE) models in secondary schooling. More radical, but less frequent, are instructional designs implemented at the institutional or program level in which the entire academic program is organized around demonstrated mastery of defined outcomes as a substitute for the completion of time-based or content-based courses of study.

- Program Evaluation and Improvement. As noted, institutional accreditation in the U.S. requires all colleges and universities to define student learning outcomes, collect evidence of the achievement of these outcomes, and use the results to improve instruction. At most institutions, such “assessment programs” are organized along a familiar pattern in which faculty develop formal statements of student learning outcomes for each degree program and for general attributes and skills assumed to be common across all baccalaureate or associate degrees, then design their own methods for assembling evidence around these local definitions.

²⁵ See <http://www.enqa.net/projectitem.lasso?id=34837&cont=projDetail>.

²⁶ For an historical treatment, see Ewell 2002.

These procedures are generally embodied in a formal “assessment plan” with results collected and reviewed annually.²⁷ For institutions that have already established a regular and formal system of program review, assessment results are typically made a part of this process. In the most sophisticated examples, moreover, a professionally-staffed organizational unit at the institution such as a Center for Teaching and Learning or an Assessment Center coordinates these activities. Many of these initiatives began explicitly with standardized testing and commercial or published student surveys that could be readily deployed, though they were not tailored to the specific goals and contexts of the institution. Now, most institutions use locally devised methods including portfolios, samples of student work reviewed periodically using scoring guides or rubrics, or locally developed surveys of current students and recent graduates. Examples of institutional assessment initiatives aimed at program improvement in other nations tend to be based more on surveys and less on direct evidence of student academic achievement, but are otherwise similar in organization and intent.²⁸

Twenty years of experience with such initiatives has yielded mixed results in the U.S. When implemented sincerely and with the visible support of academic leaders and faculty, institutional assessment programs can have strong positive impacts on instructional organization and delivery. But if assessment programs are established only at the behest of external authorities and purely for accountability purposes, the evidence of student learning they assemble tends not to be used and the effort as a whole has relatively little impact (Ewell 2002, forthcoming).

- Competency-Based Instructional Designs. At the other end of the applications spectrum, some institutions and programs incorporate instructional designs that are all or partly based on the demonstrated achievement of specified student learning outcomes. Designs of this kind are usually termed “competency-based” or “mastery” programs and are most often encountered in applied fields of study where definitions of competency are clearer and the development of performance assessments more straightforward. Such programs differ from the more typical licensing approach, in which students must pass a certification examination to complete their programs but otherwise attend classes or modules configured in a traditional manner, because they incorporate sequences of performance assessments deployed according to a fully-developed outcomes framework throughout the students’ course of study. Class attendance, class participation, and graded exercises and assignments are incidental to achievement in that the only thing that counts is periodic mastery assessments. This means that students

²⁷ Good examples can be found at every level of U.S. higher education. For research universities, a particularly well-elaborated undergraduate assessment process can be found at North Carolina State University. For medium sized public institutions, good examples are Truman State University, James Madison University, and Southern Illinois University at Edwardsville (SIUE). Among two-year community colleges, good examples are Johnson County Community College and Sinclair Community College. Many examples of such programs are documented in Banta and Associates 1993 and Banta, Lund, Black, and Oblander 1996.

²⁸ Good examples are the University of Central England and Sheffield-Hallam University in the U.K.

can complete programs at their own pace and, in principle, need not attend classes at all. Examples here include Western Governors University (WGU) and Excelsior College which teach no students directly but certify the completion of each university requirement through examination or performance assessment.

Less radical competency-based instructional designs are offered by institutions in the U.S. like Alverno College, the School for New Learning at DePaul University, California State University Monterey Bay, and the University of Charleston where instruction in each discipline is governed by a multi-level outcomes scheme that resembles a single-institution application of a national qualifications framework. Mastery of each level of each outcome must be demonstrably assessed for students to graduate—a process managed by a combination of externally-administered performance assessments and common rubric-based faculty grading standards applied to key assignments within regular academic work. International institutions like the British Open University incorporate similar elements of this mastery-based design.

Advantages of competency-based instructional designs include the external credibility of degrees awarded on the basis of directly assessed abilities, the transparency of requirements to students (who can readily see what they have to achieve and how to get there), and their ability to accommodate asynchronous study that may involve quite different learning experiences taken in quite different sequences by different kinds of students. Drawbacks include the difficulty of establishing unambiguous competencies and valid and reliable assessments for all abilities of interest in a full undergraduate program and the acceptability by other institutions of credentials based entirely on assessed abilities absent recognizable grades and coursework.

- Managing Student Transitions. A somewhat related institutional application of learning outcomes is to govern the movement of students from one level of study to another. In some cases, movement into more advanced study requires a direct demonstration of a particular level of mastery through assessed performance. In the U.S., for example, all students in public institutions in the states of Florida and South Dakota must pass a battery of commonly-administered general skills examinations before moving on to their third year of collegiate study. More commonly, standardized examinations of this kind based on clearly-articulated outcomes standards are used to assess student readiness for particular areas of university study so that they can be placed at the proper level. In the U.S., standardized outcomes-based placement examinations of this kind are used in all public institutions in Texas, Florida, and New Jersey and several other states.²⁹

A more complex application of outcomes based approaches to student transitions addresses the growing phenomenon of student transfer from one institution to

²⁹ It is important to note that these are not “college admissions” tests such as the well-known SAT or ACT examinations in the U.S., which are used to select applicants for college admission based on aptitude; they are instead measures of actual performance in the general skills typically demanded by college-level work.

another: in the U.S., for example, almost two thirds of those who eventually receive a bachelors degree have attended more than one institution with almost a fifth attending three or more (Adelman 1999). In the U.S. and increasingly in Europe, the majority of these transitions are managed through transfer of credit arrangements.³⁰ But several U.S. states including Missouri and Utah, recognizing the burden of class by class articulation based on syllabus review, have recently explored common outcomes frameworks to manage student transfer. Under these arrangements, common general education outcomes are identified for transferable blocks of prior work that are certified by the sending institution using mutually agreed-upon standards certified by direct assessment or a periodic audit/review process. Similar institution-to-institution approaches to mutual recognition of transfer competencies have been established by a number of “feeder-destination” networks of institutions in the U.S. through projects like the Quality in Undergraduate Education (QUE) initiative (Henry 2006).

In the UK and Europe, meanwhile, rising levels of student mobility have raised similar concerns about how to recognize academic work across institutional boundaries. While the principal approach adopted through the European Credit Transfer and Accumulation System (ECTS) is fairly traditional, being based on a credit system that looks primarily at the amount of time spent in a course or program,³¹ more recent efforts at developing transferable modules are emphasizing the development of learning outcomes.³² Once again, however, these are not directly assessed using a common schema but rely on local faculties to assure that these agreed-upon standards have been achieved.

A final area where outcomes concepts are being more frequently applied is in the transition between secondary schooling and higher education. With the advent of the standards-based reform movement in elementary-secondary schooling in the U.S., for example, all states were directed to develop and implement high-stakes standardized exit tests for graduating students. In some cases, these are a condition of receiving a high school diploma and in all cases their results govern a set of sanctions that can be applied to low performing schools. But these examinations do not govern admission to postsecondary education, which remains a matter that is up to individual colleges and universities. This has led to considerable policy concern that the outcomes standards underlying secondary school exit tests and the abilities tested by colleges and universities be more seamlessly aligned (Conley 2003, Achieve 2004). Currently, most secondary school exit standards in U.S. states are lower than those established for “college-level work” by most colleges and universities. Many states are now attempting to remedy this situation through so-called “K-16” policy initiatives.³³ Others have

³⁰ Most U.S. institutions have credit-based class-by-class articulation agreements in place for transfer of credit with their most common “supplier” institutions and some states have mandated automatic transfer of credit among public institutions; see Schoenberg 2005.

³¹ See http://europa.eu.int/comm/education/programmes/socrates/ects/index_en.html.

³² For example, Moon 2002.

³³ Good examples are in Maryland (<http://mdk16.usmd.edu/>) and Georgia (<http://www.usg.edu/offices/p16.phtml>).

developed alternative university admissions approaches based on portfolios of student secondary school work reviewed by panels of faculty against an established outcomes framework.³⁴

Advantages of applying an outcomes framework to manage student transitions include the promise of more efficient movement from one setting to another because difficulties of equating quite different learning experiences with respect to duration and kind are avoided. Difficulties include the challenge of agreeing on commonly defined outcomes among institutions and of developing assessments or other demonstrations of adequate performance in the disciplines of interest that are acceptable to faculties drawn from quite different kinds of institutions.

Evidence and Assessment. Statements of learning outcomes mean very little without operational specifications of how they can be credibly recognized or demonstrated. But there are many ways in which this can be done, each with associated strengths and weaknesses. “Evidence” in this context refers to the body of information assembled and interpreted to sustain the contention that a given learning outcome has been attained. Most observers prefer this term to “measurement” because it allows non-standardized, qualitative information to enter such judgments and because it emphasizes that the information assembled may be drawn from multiple sources and is interpreted deliberately to make a case (WASC 2002). “Assessment” in this context refers to the specific approaches typically used to gather evidence of attainment, which may be applied at the individual student or the institution/program level.

Properties of Evidence of Student Learning Outcomes. Before proceeding to a brief review of typical sources of evidence of the attainment of learning outcomes, a number of basic concepts related to assessment-based evidence must be reviewed.³⁵

- Direct vs. Indirect. Evidence about learning outcomes is usually termed *direct* or *indirect* based on the distance from the cognitive construct of “learning.” The former uses assessment methods that attempt to examine learning itself such as examinations, assignments and tasks, or other forms of student work that demand observable deployment of the ability in question. The latter uses methods that do not examine learning itself but rather its consequences such as related behaviors (job placement, civic participation, etc.) or testimony about learning (e.g. self-reports about learning gain or related behaviors as reported through questionnaires or interviews). There are, however, several approaches to making this distinction operationally which typically differ in how they classify behaviors.³⁶ Direct

³⁴ The most fully developed example is the state of Oregon’s Proficiency-Based Admissions Standards System (PASS); see <http://pass.ous.edu/>.

³⁵ For good comprehensive treatments of these basic assessment issues in higher education settings, see Erwin 1991 and Palomba and Banta 1999.

³⁶ For example, some schemes consider employment (and reports on job performance supplied by employers) to be direct measures of learning outcomes.

evidence of learning outcomes is generally accorded greater credibility than indirect.

- Constructed vs. Naturally-Occurring. Evidence about learning outcomes can also be gathered using specially-designed and deployed methods that are applied outside the normal teaching/learning context (e.g. special examinations, surveys, behavioral inventories). Alternatively, such evidence can be inferred by analyzing existing student work from a more comprehensive or developmental perspective (e.g. portfolios, work samples, observations of clinical practice). The latter has the advantage of being produced automatically without requiring either special instruments or special attention to ensuring student motivation to perform. But the abilities such evidence reveals may not completely correspond to the learning outcomes of interest.³⁷
- Metric-Based vs. Expert-Based. Judgments of the adequacy or level of attainment on a given learning outcome are typically made in two quite different ways, which in some respects correspond to the nature of the evidence itself. Metric-based assessments consist of questions that automatically yield quantitative scores and sub-scores because the items that comprise them have clear right or wrong answers.³⁸ In this case, expert judgment only enters the process in the form of the construction of the assessment itself and of the scoring scheme built into its design. Expert-based assessments directly involve such judgments through human raters who apply their knowledge and experience to arrive at ratings of performance.³⁹ To ensure consistency among expert raters, scoring guides or “rubrics” are frequently employed that provide fine-grained descriptions of the ability being rated, together with examples of what each level of performance looks like. Rubrics can either be applied “holistically” to judge the complete performance or can be constructed separately to judge different aspects or sub-dimensions of the performance. Because of natural variability in human judgment, as well as the somewhat lesser precision associated with such evidence in the first place, expert-based judgments usually yield less reliable results than metric-based judgments. On the other hand, they may be more valid because they involve complex applications of trained expertise, and they are applied to bodies of evidence that are more authentic.
- Authenticity. This property of evidence about student learning outcomes refers to how close the process that generated the evidence is to actually practicing the ability in a “real world” setting. Performing a procedure or writing a passage of prose are thus viewed as more “authentic” forms of evidence than answering questions about either of these activities. Authentic forms of evidence are generally valued most by both academics and stakeholders because they are closer

³⁷ For example, observation of student performance in a clinical experience in a health-related field may be authentic and unobtrusive, but the clinical encounter itself may not generate an opportunity for the student to reveal an important area of knowledge and skill.

³⁸ The most common example is the standardized multiple-choice test.

³⁹ Traditional faculty grading or marking student work, of course, is an instance of judgment of this kind.

to the ideal type of the ability of interest. Evidence based on authentic methods also has “face validity”—that is, regardless of how valid it really is, it *looks* believable to lay observers. But evidence based on authentic methods is usually expensive to gather and is frequently difficult to interpret comparatively.

- Standardization. “Standardized” evidence of student learning is derived from methods whose measurement properties are known and specified so that results can be meaningfully compared.⁴⁰ Constructing standardized methods is time-consuming and expensive, so they are rarely authentic, are usually produced by testing or survey organizations outside the academy, and are generally administered to large numbers of test-takers to achieve economies of scale. Their chief advantage is the ability to credibly compare performances across individuals or test-taking populations, while their primary disadvantages are expense and potential lack of correspondence with the abilities of interest.
- Sampling. Sampling is often used when the performance of populations rather than individuals is the focus of interest. This can yield accurate estimates of population performance at a fraction of the cost of gathering similar information from all the population’s members. But samples must be of sufficient size and appropriate construction to generate accurate estimates—rules which are frequently violated in practice. Using samples also limits the conclusions that can be drawn because not enough cases may be available to investigate smaller groups within the sampled population. And because the information is generally based on only a small group of individuals, sample-based evidence may lack credibility for many stakeholders. Finally, the notion of sampling can be applied to the range of content addressed in any assessment as well as the population to be assessed. For example, a given examination, however it is constructed, can only ask students about a limited range of topics, which represents a selected fraction of the range of knowledge, skills, and abilities that a particular course or module addresses.

Sources of Evidence of Student Learning Outcomes. Evidence of the attainment of student learning outcomes can be drawn from many sources and only the most common can be briefly reviewed here.⁴¹ Most of these methods are deployed at all levels of analysis (institution, program, individual student) but a few of them are confined to populations of students (usually collected on a sample basis). And not all of these methods are mutually exclusive as some may be included as components of others.

- Examinations. Examinations are the approach most commonly associated with the assessment of student learning outcomes, but they can be of many kinds. Large-scale assessments generally rest on standardized forced-choice examinations designed for a specific purpose (e.g. individual placement or certification, assessing the comparative performance of institutions or programs,

⁴⁰ Standardized examinations based on “forced choice” questions are the most common examples, but authentic standardized methods such as performance tasks have also been constructed.

⁴¹ Again, the most comprehensive single treatments are Erwin 1991 and Palomba and Banta 1999.

etc.). But faculty-made examinations administered in intact classes or as capstone demonstrations of mastery at the completion of a program can be equally valuable as evidence of attainment if they are deliberately constructed to yield comparative information and clearly directed toward a defined set of learning outcomes.

Evidence based on examination is direct and easily compared. As a result, such evidence tends to be highly credible for external stakeholders. But examinations (especially if they are administered once in a high-stakes setting) are frequently limited in their ability to assess mastery or other deeper forms of learning.

- Tasks and Demonstrations. Tasks and demonstrations require students to directly deploy knowledge and skills in a particular setting to provide evidence of a desired level of performance. They may be deliberately constructed to generate such evidence (for example, a proof in mathematics or an engineering design) or they may occur naturally in the course of a learning experience when a particular situation arises (for example, an internship or clinical encounter). Evidence based on such situations is direct and authentic, but it is also difficult to interpret consistently across observers. As a result, effective use of these approaches as credible evidence of attainment requires developing and deploying carefully designed interpretive tools like scoring guides or rubrics (see below).
- Student Work. This broad category of evidence potentially includes all of the “work products” generated naturally by students in a particular course of study including written assignments, examinations, problems, laboratory reports, and field or clinical performances. Although such evidence has the advantage of having already been collected, a great deal of effort is generally involved in meaningfully assembling and interpreting it. As a result, learning outcomes assessment approaches based on existing student work typically involve sampling such work products deliberately or systematically to assemble a structured body of evidence, then applying scoring guides or rubrics to draw conclusions about attainment. Probably the most common method for assembling student work is the *portfolio*, in which examples of a given student’s work are chosen and examined together. Portfolios can be constructed analytically—for example choosing one or two pieces of work deliberately to serve as illustrations of a given outcome. They can also be constructed longitudinally by choosing pieces of work exemplifying the same outcome at different stages of a student’s academic career. In both cases, a decision must be made about whether to include best work or most typical work, but the former is usually the norm because of difficulties in determining “most typical.” As a result, individual portfolios are usually best at showing maximum levels of attainment.⁴² For more representative examinations of outcomes for a given population, cross-sectional samples of student work are often assembled on a random basis. Alternatively, teaching staff can be asked to sample two or three examples of median performances on a given assignment—

⁴² The outcomes-based accountability argument currently being advanced by many colleges and universities in the U.S. is based on this premise—that students show their best work through “capstone” demonstrations that occur naturally in their primary field of study and that university effectiveness can be best demonstrated by examining this body of work (AACU 2004a).

those that fall in the middle of the grading distribution—to provide a “typical” sample of student work. Finally, scoring criteria can be built directly into the grading process itself for selected assignments and the results aggregated for later interpretation.⁴³

- Behavioral Outcomes. Student behaviors before and after attendance can provide useful indirect information of the attainment of learning outcomes. For students in career preparation programs, for example, placement and advancement in their field of study is commonly used as a measure of program success. A similar measure of this kind is participation in further higher education—for example, graduate study or continuing professional development in field. Other later behaviors that are frequently claimed as success measures for both institutions and programs because they are related to intended learning outcomes may include civic behaviors such as voting and volunteerism, lifestyle behaviors such as health and consumption habits, or social and geographic mobility including income levels and place of residence. In most cases, such evidence is collected by surveys of graduates and former students. But in some cases, existing records maintained government agencies can be tapped to obtain such information so long as individual records can be identified and matched to student attendance/attainment records.
- Self-Reports. A final method of obtaining indirect evidence of the attainment of student learning outcomes is through the testimony of students themselves. Most commonly, students or graduates are simply asked to rate their own current levels of knowledge or skill across a set of named areas that correspond to the learning outcomes in question.⁴⁴ Other approaches involve asking students to report on their current education-related behaviors such as study habits, interactions with each other and teaching staff, the time they spend on their studies, and the like—factors that previous research have demonstrated are related to high levels of actual attainment. Most self-reports are collected by means of questionnaire surveys, but alternative methods include telephone or individual interviews and focus groups. There are ongoing debates about the validity and reliability of self-reports as proxy measures of actual student attainment, though considerable past work has established consistent positive relationships between actual and claimed attainment.⁴⁵ As a result, the credibility of such measures is somewhat limited. Self-reports are in general far more credible for student behaviors, though these are only indirectly related to learning outcomes. Self-reports are also the only method available for obtaining evidence on such non-cognitive outcomes as attitudes, beliefs, or dispositions. Finally, evidence based on self-reports is common in many outcomes approaches because of the ease and efficiency with

⁴³ One of the most common approaches to doing this is “Primary Trait” scoring (Walvoord and Anderson 1998).

⁴⁴ An alternative is to ask students how much they think they have grown or developed with respect to the ability in question, or how much their studies have contributed to their current ability—both of which provide a self-reported measure of “value-added.”

⁴⁵ For a short review and citations, see http://www.ed.psu.edu/cshe/abet/pdf/Self_Reported.pdf.

which such information can be collected if information about institutional and program outcomes is all that is required. Large scale sample surveys, for example, can be administered at a fraction of the cost of standardized examinations or more authentic forms of evidence that requires human grading.

Some Questions to Consider. This paper only begins to address the many potential applications of learning outcomes concepts to higher education and the issues associated with them. As more deliberate discussion of these concepts unfolds in Hong Kong, universities and the UGC should consider the following questions:

1. How are learning outcomes concepts and approaches *already* being used (implicitly or explicitly) in higher education in Hong Kong? Are there ways these applications might usefully be extended? What models or lessons can be drawn from these experiences that could be applied elsewhere?
2. How might learning outcomes concepts and approaches be helpful as Hong Kong universities move from a three-year to a four-year undergraduate instructional framework? What are the specific implications for developing and assessing the “generic” attributes that all baccalaureate graduates ought to possess?
3. What areas of university study might benefit particularly from applying learning outcomes concepts and approaches to the teaching and learning process?
4. How might greater use of the language of learning outcomes in program descriptions and course or module syllabi help undergraduate students in Hong Kong universities understand better what is expected of them? How might it have an impact on student learning?
5. How are learning outcomes and their assessment currently used by each university in its efforts to improve the quality of teaching and learning through staff and instructional development? Would greater emphasis on these topics be useful?
6. How might learning outcomes concepts and approaches be useful to Hong Kong universities as the curriculum reform in Hong Kong secondary education continues to unfold? What specific implications are there for the university admissions process, for early student assessment, and for the first year of university study?
7. What role do learning outcomes and their assessment currently play in each university’s internal quality assurance processes for undergraduate teaching and learning? Would a greater emphasis be helpful? How might this be accomplished and what pitfalls should be anticipated?

8. How might learning outcomes concepts and approaches help create greater alignment and integration of undergraduate teaching and learning across universities in Hong Kong? Would a formal framework for academic awards be helpful? What are the drawbacks of such an approach?
9. How might the use of learning outcomes concepts and approaches help universities in Hong Kong better communicate their missions and accomplishments to government and the public?

As a final note, it is important to emphasize that learning outcomes are not a panacea for reform. Adopting a learning outcomes perspective and emphasizing demonstrated student achievement has proven beneficial in many higher education settings across the world. But these gains have only been achieved through deliberate and balanced approaches that reflect stakeholder values and perceptions, and that keeps the ultimate goal of improving student learning clearly in mind. Conceptual reification and rigidity in implementation will always be the worst enemies of a thoughtful outcomes-based approach.

References

- Achieve, Inc. (2004). *Do Graduation Tests Measure Up? A Closer Look at State High School Exit Exams*. Washington, DC: The American Diploma Project, Achieve Inc.
- Adam, Stephen (2004). *Using Learning Outcomes: A Consideration of the Nature, Role, Application, and Implications for European Education of Employing "Learning Outcomes" at the Local, National, and International Levels*.
<http://www.scotland.gov.uk/library5/lifelong/tehea-00.asp>.
- Adelman, Clifford (1999). *Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.
- Association of American Colleges and Universities (2004a). *Our Students' Best Work: A Framework for Accountability Worthy of Our Mission*. Washington, DC: Association of American Colleges and Universities (AACU).
- Association of American Colleges and Universities (2004b). *Taking Responsibility for the Quality of the Baccalaureate Degree*. Washington, DC: Association of American Colleges and Universities (AACU).
- Atkins, M.J.; Beattie, J.; and Dockrell, W. B. (1993). *Assessment Issues in Higher Education*. London, England: Employment Department, Further and Higher Education Branch.
- Astin, Alexander W. (1977). *Four Critical Years*. San Francisco: Jossey-Bass.
- Banta, Trudy W. (1986). *Performance Funding in Higher Education: A Critical Analysis of Tennessee's Experience*. Boulder, CO: National Center for Higher Education Management Systems (NCHEMS).
- Banta, Trudy W. and Associates (1993). *Making a Difference: Outcomes of a Decade of Assessment in Higher Education*. San Francisco: Jossey-Bass
- Banta, Trudy W.; Lund, Jon P.; Black, Karen E.; and Oblander, Frances W. (1996). *Assessment in Practice: Putting Principles to Work on College Campuses*. San Francisco: Jossey-Bass.
- Blackmur, Douglas (2004). A Critique of the Concept of a National Qualifications Framework. *Quality in Higher Education*, (10)3, 267-284.
- Bruneau, William and Savage, Donald C. (2002). *Counting Out the Scholars: The Case Against Performance Indicators in Higher Education*. Toronto, Canada: James Lorimer and Company.

- Burke, Joseph C. (2005). *Reinventing Accountability: From Bureaucratic Rules to Performance Results*. J. C. Burke and Associates, *Achieving Accountability in Higher Education*. San Francisco: Jossey-Bass, 216-245.
- Burke Joseph C., Yang, P., and Minassins, Hendrik P. (2002). State Performance Reporting Indicators: What Do They Indicate? *Planning for Higher Education*, 2002, 31(1), 15-29.
- Conley, David (2003). *Mixed Messages: What State High School Tests Communicate About Student Readiness for College*. Eugene, OR: Center for Educational Policy Research, University of Oregon.
- Council for Higher Education Accreditation (2003). *Statement of Mutual Responsibilities for Student Learning Outcomes: Accreditation, Institutions, and Programs*. Washington, DC: Council for Higher Education Accreditation (CHEA).
- De Moura Castro, Claudio (2002). Success and Perils in Evaluating Brazilian Undergraduate Programs. *International Higher Education*, 35, www.bc.edu/bc_org/avp/soe/cihe/newsletter/News35/text010.htm.
- Dill, David D. (2000). Designing Academic Audit: Lessons Learned in Europe and Asia. *Quality in Higher Education*, 6(3), 187-208.
- Erwin, T. Dary (1991). *Assessing Student Learning and Development*. San Francisco: Jossey-Bass.
- Ewell, Peter T. (forthcoming). The "Quality Game:" External Review and Institutional Reaction Over Three Decades in the United States. *Dynamics and Effects of Quality Assurance in Higher Education*. Porto, Portugal: CIPES.
- Ewell, Peter T. (2002). An Emerging Scholarship: A Brief History of Assessment. T.W. Banta and Associates, *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass, 3-25.
- Ewell, Peter T. (2001). *Accreditation and Student Learning Outcomes: A Proposed Point of Departure*. Washington, DC: Council on Higher Education Accreditation (CHEA).
- Ewell, Peter T. (1984). *The Self-Regarding Institution: Information for Excellence*. Boulder, CO: National Center for Higher Education Management Systems (NCHEMS).
- Ewell, Peter T. and Ries, Paula (2000). *Assessing Student Learning: A Supplement to Measuring Up 2000*. San Jose, CA: National Center for Public Policy in Higher Education.

- Feldman, Kenneth A. and Newcomb, Theodore M. (1969). *The Impact of College on Students*. San Francisco: Jossey-Bass.
- Griffin, Patrick; Coates, Hamish; Mcinnis, Craig; and James, Richard (2003). The Development of an Extended Course Experience Questionnaire. *Quality in Higher Education* (9), 3, 259-266.
- Harvey, Lee and Newton, Jethro (2004). Transforming Quality Evaluation. *Quality in Higher Education* 10(20), 149-166.
- Henry, Ronald J. (Ed.) (2006). *Faculty Development for Student Achievement: The QUE Project*. Boston, MA: Anker Publishing.
- Johnstone, Sally M.; Ewell, Peter T.; and Paulson, Karen (2002). *Student Learning as Academic Currency*. Washington, DC: American Council on Education (ACE).
- Keeton, Morris (1981). Assessing and Credentialing Prior Experience, in A. W. Chickering and Associates, *The Modern American College*. San Francisco: Jossey-Bass, 631-641.
- Learned, W. S. and Wood, B. D. (1938). *The Student and His Knowledge: A Report to the Carnegie Foundation on the Results of the High School and College Examinations of 1928, 1930, and 1932*. New York: Carnegie Foundation for the Advancement of Teaching.
- Massy, William F. (1997). Teaching Learning Quality Process Review: The Hong Kong Programme. *Quality in Higher Education*, (3)3, 249-262.
- Massy, William F. and French, Nigel J. (2001). Teaching and Learning Quality Process Review: What the Programme has Achieved in Hong Kong. *Quality in Higher Education*, (7)1, 33-45.
- Miller, Margaret A. and Ewell, Peter T. (2005). *Measuring Up on College Level Learning*. San Jose, CA: National Center for Public Policy in Higher Education.
- Moon, Jenny (2002). *The Module and Programme Development Handbook*. London, England: Kogan Page.
- Pace, C. Robert (1979). *Measuring College Outcomes*. San Francisco: Jossey-Bass.
- Palomba, Catherine A. and Banta, Trudy W. (1999). *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Francisco: Jossey-Bass.
- Ruppert, Sandra S. (ed.) (1994). *Charting Higher Education Accountability*. Denver, CO: Education Commission of the States (ECS).

- Schwartzman, Simon (2004). *The National Assessment of Courses in Brazil*. Chapel Hill, NC: Public Policy for Academic Quality, The University of North Carolina at Chapel Hill.
- Schoenberg, Robert (2005). *General Education and Student Transfer: Fostering Intentionality and Coherence in State Systems*. Washington, DC: Association of American Colleges and Universities.
- Scottish Credit and Qualifications Framework (2003). *An Introduction to the Scottish Credit and Qualifications Framework* (www.scqf.org.uk).
- Terenzini, Patrick T. (1989). Outcomes Assessment with Open Eyes: Pitfalls in Studying Student Outcomes, *Journal of Higher Education*, 60, 644-664.
- Walvoord, Barbara E. and Anderson, Virginia J. (1998). *Effective Grading: A Tool for Learning and Assessment*. San Francisco: Jossey-Bass.
- Westerheijden, Don F. (forthcoming). States and Europe and Quality of Higher Education. *Dynamics and Effects of Quality Assurance in Higher Education*. Porto, Portugal: CIPES.
- Westerheijden, Don F. (2001). Ex Oriente Lux?: National and Multiple Accreditation in Europe After the Fall of the Wall and After Bologna. *Quality in Higher Education* (7)1, 65-76.
- Western Association of Schools and Colleges (2002). *Evidence Guide*. Alameda, CA: Western Association of Schools and Colleges (WASC), Senior College Commission.
- Western Association of Schools and Colleges (2001). *Handbook of Accreditation*. Alameda, CA: Western Association of Schools and Colleges (WASC), Senior College Commission.
- Whitaker, Urban (1989). *Assessing Learning: Standards, Principles, and Procedures*. Philadelphia, PA: Council for Adult and Experiential Learning (CAEL).

Annex

Annotated List of On-Line Reference Materials on Learning Outcomes Approaches

General Treatments of Outcomes and Assessment

James Cook University in Australia has constructed a useful general site on Outcomes Based Education. While primarily oriented toward applications in secondary schools, the site contains links to papers and resources on the topic of learning outcomes in general. <http://www.library.jcu.edu.au/Educ/obe.html>.

North Carolina State University has assembled one of the most comprehensive U.S. websites on assessing student learning outcomes. The site has links to basic resources on outcomes concepts and assessment, as well as links to individual institutional websites on outcomes assessment. <http://www2.acs.ncsu.edu/UPA/assmt/resource.htm>.

A similar comprehensive site has been prepared by Texas A&M University. <http://www.tamu.edu/marshome/assess/HTMLfiles/oabooks.html>.

The Association of American Colleges and Universities (AACU) has prepared a number of resources on the assessment of student learning directed especially at defining and examining “generic” learning outcomes associated with collegiate study. The Association has also just completed a guide for colleges and universities on this topic that is available through the website. <http://www.aacu.edu/issues/assessment/index.cfm>.

The Mathematics Association of America (MAA) has undertaken a major initiative on outcomes and learning assessment. The project web page provides an excellent annotated bibliography of sources including both general treatments of the topic and specific applications to mathematics. <http://www.maa.org/saum/>.

The Bologna Process in Europe has stimulated considerable discussion of potential applications of learning outcomes concepts to aligning national standards and developing a new system of bachelors and masters degrees. A working paper describing emerging approaches in Europe is available at <http://www.scotland.gov.uk/library5/lifelong/tehea-00.asp> and contains many links to individual articles and national efforts to apply learning outcomes concepts in Europe.

National/State Examinations for Accountability

In the U.S., several states use common learning outcomes assessment instruments as accountability tools for public universities. Illustrations are provided by South Dakota, which requires all students to pass a common examination as a condition for advancement (<http://www.sdbor.edu/administration/academics/RegentalProficiencyExaminationGuidelines.doc>) and Tennessee, which uses standard examination results as part of its performance funding program

(http://www.state.tn.us/thec/2004web/division_pages/ppr_pages/Policy/pprpolicyperformancefunding.htm).

Brazil uses a similar approach to assuring quality for both public and private institutions by administering examinations in each program of study in a “before-after” format. http://www.unc.edu/ppaq/Brazil_designed.html.

National Qualifications Frameworks

Qualifications Frameworks have been established in a number of nations over the past 20 years. Among the oldest are in New Zealand (<http://www.nzqa.govt.nz/qualifications/index.html>) and the United Kingdom (<http://www.qaa.ac.uk/academicinfrastructure/FHEQ/EWNI/default.asp>). Newer frameworks that are designed for more flexible deployment in consultation with universities are in Scotland (<http://www.scqf.org.uk/>) and Ireland (<http://www.nqai.ie/en/>).

Learning Outcomes in Accreditation

The Council for Higher Education Accreditation is taking the lead in the U.S. in promoting greater use of learning outcomes as an integral part of the accreditation process. A working “statement of principles” was a direct result of this effort. <http://www.chea.org/pdf/StmntStudentLearningOutcomes9-03.pdf>.

Linkages Between Secondary and Postsecondary Study

Stanford University’s “Bridge Project” is probably the most comprehensive research project in the U.S. on this topic. Its web site provides descriptions of the initiatives currently being undertaken by the project, publications, and links to other resources. <http://www.stanford.edu/group/bridgeproject/>.

The North Central Regional Educational Laboratory has produced a working paper on aligning outcomes standards between secondary and postsecondary study, which is accompanied by useful reference links. <http://www.ncrel.org/sdrs/pbriefs/94/94-1ovr3.htm>.

Many U.S. states have undertaken so-called “K-16” initiatives that are attempting to align standards between secondary study and higher education. Good examples are in Maryland (<http://mdk16.usmd.edu/>) and Georgia (<http://www.usg.edu/offices/p16.phtml>).

Outcomes/Competency Based Instructional Designs

A number of higher education institutions in the U.S. are organized entirely around an outcomes or abilities framework under which competencies are assessed to receive credentials. Two of the most illustrative are Western Governors University (http://www.wgu.edu/about_WGU/WGU_advantage.asp) and Excelsior College

https://www.excelsior.edu/portal/page?_pageid=57,45940&_dad=portal&_schema=PORTAL).

Other institutions teach regular classes, but the entire program of study is organized in terms of a set of common learning outcomes or abilities with modules and classes mapped to those abilities and student mastery assessed independently. Prominent examples include Alverno College

http://www.alverno.edu/about_alverno/ability_curriculum.html) and California State University at Monterey Bay (<http://csumb.edu/site/x7450.xml>).

An illustrative application of outcomes based concepts to teaching and learning outside the U.S. is the University of Western Australia

http://www.catl.uwa.edu.au/_data/page/77897/OBE_and_Assessment.pdf).