

## CHAPTER ONE

# THE NEED FOR QUALITY BENCHMARKS IN UNDERGRADUATE PROGRAMS

The hour for recognizing the singular importance of undergraduate education is here at last. Department chairs, faculty at all ranks, deans and provosts, and even college and university presidents finally realize that educating undergraduates is arguably the most important mission of higher education today. For those of us who have watched the focus shift from graduate and post-baccalaureate professional education to the education of traditional and nontraditional four-year students, the change is both powerful and palpable. Teaching undergraduates well is now a dominant focus in higher education. We are not suggesting that educators and administrators should not attend to the important and pressing concerns of graduate education; rather, we believe that the best support graduate and professional schools can receive is to send them well-prepared students.

As a result of the new emphasis on undergraduate education, colleges and universities aspire to provide the best educational experience for student majors that their resources will permit. Their common goal is to assess what and how well students have learned during their time at the institution. The issues involved, similar to the work itself, are challenging. Where should institutional self-reflection about undergraduate learning begin? Within the administration? Among the collective faculty? Or should alumni be tapped for their feedback on how their time at

## 2 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

the institution shaped their futures? What about the opinions of other stakeholders, including the students themselves?

When institutions turn their focus to undergraduate education, we argue that the place to begin this important work is at the departmental or program level—the unit of analysis that has the most day-to-day as well as discipline-based impact on student learning. Assessment in undergraduate education is often aimed at general education; that is, the distribution requirements of liberal arts offerings that all students in a given college or university must complete. Evaluating student learning in general education courses is certainly important, but we believe that the breadth and depth of discipline-based knowledge acquired within department-based majors is the more appropriate forum to capture assessment activities that reflect the true accomplishments of the baccalaureate program. How well are chemistry majors learning foundational materials in basic and intermediate courses? Does this foundational material later help these student majors display the necessary critical thinking skills in the advanced courses in the major? As they near the end of their major courses, can chemistry students conceive and design experiments? Are they able to interpret and explain the results of their research using the discipline's vernacular? The same sorts of questions could be appropriately framed and examined for any other major area of study, whether it be architecture, urban planning, or Urdu.

One challenge is that there is little formal consensus about what constitutes program quality in undergraduate education in the arts and humanities, social sciences, and natural sciences. Ratings or rankings are either aimed at graduate programs (for example, the top twenty forestry programs in the United States, best graduate business schools on the West coast) or, as we will see, the placement of the institution as a whole on higher education's pantheon (for example, the list of best regional small universities in the southeast, the top one hundred national liberal arts colleges). Ratings and rankings have their place, but they tell a limited story regarding quality undergraduate education. We wrote this book to help educators at all sizes and types of four-year institutions, including teachers, researchers, scholars and administrators, to constructively evaluate and document the effectiveness of current programs aimed at educating undergraduates.

We realize that there is nothing new about conducting academic program reviews. As is well known, most colleges and universities have implemented a formal review cycle for individual departments or programs, typically every five years or so. The pursuit of “educational quality,” for example, has been a key campus concern for a couple of decades now (see Berquist & Armstrong, 1986; Bogue & Saunders, 1992; Boyer, 1987; Haworth & Conrad, 1997). Following the groundbreaking work of Boyer (1990), debates about the nature of scholarship have become standard procedure as campuses define the implications for what faculty should do in and out of the classroom. What is *new*, however, is the availability of helpful assessment tools for evaluating the strengths, challenges, and unrealized opportunities within departments or other programs. Such tools—including discipline-based learning guidelines, curriculum evaluation guides, standardized student surveys, rubrics for teaching and evaluating writing and speaking demonstrations, and instructor-designed measures tailored to evaluate whether the learning goals of assignments are met—constitute some other possibilities.

We advocate the use of a particular assessment tool: *quality benchmarks*; that is, reasonable, reason-based, and peer-sanctioned criteria that can be used to assess the performance of academic programs and departments. Benchmarks provide a guiding standard for comparing what *is* with what *could be* achieved with redirected effort, energy, attention, or resources (see, for example, Astin, 1993a; Banta, Lund, Black, & Oblander, 1996). Formative rather than summative, quality benchmarks allow constituents—department chairs, program directors, faculty, deans, and students—to review progress, identify problems, establish or revisit goals, and reflect on student learning outcomes or establish new ones.

The term “benchmarks” has a particular meaning that differs somewhat from terms such as “criteria” or “rubrics” and the like. Benchmarking is the process of assessing a program’s qualities by comparing them with—and trying to shape them to conform to—what are considered to be the best practices or ideal program features. In essence, benchmarking provides a snapshot of program performance that helps the relevant constituents understand how particular educational activities within key domains compare with acknowledged standards. The upshot of such

#### 4 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

benchmarking is that a program will undertake changes in order to improve teaching and learning. The aim of benchmarking is to improve some particular aspect of program performance (such as student research skills, faculty scholarly output), but it is part of a larger continuous process in which departments and the colleagues who teach and administrate within them continually seek to challenge their practices.

We argue that benchmarks provide a set of standard reference points for internal self-studies and external reviews. Of course, benchmarking processes also are routinely invoked when comparing one program to another or when sizing up whether a program has achieved the standards set forth by an accrediting agency. Regardless of whether the focus involves internal or external standards, the process of comparing achievement to an identified standard can serve multiple purposes: benchmarking can assess institutional reputation locally or nationally, verify goal achievements, and measure whether outcome targets have been realized, among other objectives.

The emphasis of the benchmarks we propose is internal. We suggest using these selected quality benchmarks to assist undergraduate programs in establishing quality objectives, monitoring progress toward their achievement, and ultimately achieving a level commensurate with department talent, energy, and resources. In the process, working with developmental standards can further define their program missions and document their effectiveness.

We also want to be clear that benchmarks are not national standards for accrediting departments or programs. We are not proposing that the quality benchmarks presented in this book should be used to compare a given institution with its peer or aspirant counterparts. Again, we propose that departments and other academic programs establish appropriate internal benchmarks for measuring their own progress on various dimensions. By doing so, we advocate that departments concentrate on their own goal setting and progress. Further, accreditation standards are generally all-or-nothing propositions, whereas quality benchmarks are designed to be developmental measures for assessing a program. We urge readers to keep these points in mind as they read this book and subsequently reflect on, evaluate, and improve their departments or programs.

## AN EXAMPLE: USING BENCHMARKS FOR PROGRAM ADVOCACY

Consider a familiar example on many campuses. A department chair wants to search for and hire a new faculty member to offer courses in a key topical area that currently receives little or no coverage in the program's curriculum. The dean or the vice president for academic affairs (VPAA) reviews the request and suggests that a self-study is in order before any additional faculty line can be considered. The dean also points out that several other departments are simultaneously claiming an equally pressing need for new hires but the projected budget for the next academic year cannot possibly accommodate all requests. To make things interesting, let's imagine that enrollment in the department is solid and steady, but some of the competing departments have seen moderate growth in the numbers of students they serve.

How can this department chair hope to convince the dean that the requested hire is truly needed and worthy of support? How can the chair convince herself that her request is justified, given the needs of her colleagues who are leading other departments? We'll present two possible scenarios.

*Scenario one.* The chair dutifully undertakes the writing of a self-study, the first in several years. Department colleagues are supportive but expect her to carry the burden of gathering information, drafting arguments for the position, and writing the actual proposal. The majority of department members agree that the hire is justified, but aside from drawing up a list of reasons for the hire from the department's own perspective and perceived immediate needs (for example, enrollment crunch, recent retirements, most recent new line was five years ago), little attention is paid to how the hire could fit into the larger constellation of department issues, including curricular changes, wider institutional service, and enhancing undergraduate education. In other words, the department operates in an academic "business as usual" model, coupled with the oft-heard claim of "too many students to serve, not enough resources."

Although the final proposal was compelling in many respects, the dean received equally cogent arguments from other

## 6 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

departments. In the end, the dean gave the hire to another department, one with fewer full-time faculty members than the others and, up to that point, larger class sizes. After spending considerable time and effort on the proposal, the department chair is somewhat demoralized; her colleagues are frustrated, disappointed, and somewhat aggrieved. What little collaborative spirit existed in the department disappears for quite some time.

*Scenario two.* Imagine that the perceived need to hire a new colleague spurs the interest and involvement of the entire department. They agree to an approach based on benchmarking. Instead of expecting the chairperson to do all the work, the faculty members divide up the tasks related to the potential hire. For example, some colleagues identify both new and established courses that the new person would teach (curriculum), others examine the hire's advising responsibilities (student development), and so on. With the help of colleagues, the chair drafts a proposal that documents the department's current mission and goals, highlighting particular areas that are *distinguished* (for example, a recently revised curriculum that meets national disciplinary standards, an outreach program to the local community) as well as those that are *effective* (such as undergraduate research presentations delivered at regional and national conferences), or still *developing* (for example, the department now requires that students give formal, oral presentations in upper-level courses), or that need attention because they are *undeveloped* (for example, faculty publishing has declined precipitously in recent years). (We define and explain the italicized terms in the next section of this chapter.)

The proposal specifically explains how a new hire with a desired specialty can contribute to the areas of strength while also supporting the developing and underdeveloped areas of the department. In the course of the self-study, members of the department streamlined a few procedural issues that ended up helping the department's budget. The comprehensive nature of the proposal clearly documented the need for the new hire. The department chair was gratified by the level of enthusiasm, camaraderie, and participation of her colleagues. The colleagues, in turn, realized that while there were some areas of concern, the department was clearly moving in the right direction. The dean concurred,

impressed by the amount of effort, goodwill, and careful, thoughtful planning. The proposal was much more focused and reasoned than the competing proposals from the other departments.

Scenario one is all too familiar on most campuses. Scenario two is less familiar. Let's imagine that the dean in scenario two is convinced by the evidence and grants the new faculty line. Everyone in the department is overjoyed, especially the chair. But consider this: even if such benchmarking does not always succeed—there are usually any number of legitimate reasons to postpone valid hires—the act of comparing the current state against quality performance benchmarks still provides department members, the chair, and the dean with a sense of which aspects of the program are working well (often very well) and where some improvements could be made. Most important, not all improvements require an outlay of capital or an expanded budget; some are procedural, others organizational, and some rely on a combination of good will and common sense. The point is that something beneficial grew from the program review that embraced the benchmarking approach.

What are some of the concrete advantages of using performance benchmarks? These include:

- Engagement of faculty and students in crafting and revising the mission statement of a department or program
- Formative evaluation of teaching and learning outcomes
- Curricular review, refinement, and revision
- Recruitment and retention of quality faculty and students
- Assessment of resource needs
- Long-range academic program planning
- Evaluation and demonstration of program quality

## BENCHMARKING AND PROGRAM ASSESSMENT FOR EDUCATIONAL RENEWAL

Benchmarks are used in higher education as a way to improve the climate for learning within departments (see, for example, Umbach & Porter, 2002; Umbach & Wawrynski, 2005). We recently proposed performance benchmarks to assist undergraduate psychology programs in defining their missions and goals as well

## 8 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

as in documenting their effectiveness (Dunn, McCarthy, Baker, Halonen, & Hill, 2007). The developmental framework we proposed garnered considerable interest among communities of teachers and administrators within the discipline of psychology. The interest we received led us to think more broadly about how quality benchmarks could be used to evaluate virtually any academic department or program. Indeed, our experiences as program reviewers, faculty members, and part- or full-time administrators informs us that departments, department heads, and deans all want reasonable, reliable, and professionally respectable methods for evaluating the performance of undergraduate programs.

We believe the developmental framework presented in this book will satisfy all parties because we emphasize formative assessment. We do not advocate the use of summative assessment—here defined as the use of benchmarks to reward (for example, add faculty lines, build new facilities) or punish (for example, eliminate faculty, trim the budget) a program for its standing. Furthermore, we designed our framework to help programs identify and tout what they already do well even in situations involving seriously constrained resources. Finally, the performance benchmarks used to identify areas of program strength can, in turn, be used to recruit and retain students, to seek funding via grants or alumni support, and to enhance the perceived rating of an institution. When benchmarks reveal that a program or areas within it are undeveloped or developing, faculty and administrators can then plan where subsequent efforts and resources are best placed to raise a program's standing.

Our framework explores the attributes of undergraduate programs by focusing on educationally related activities in eight domains: *program climate*; *assessment*; *accountability*; and *accreditation issues*; *student learning outcomes*; *student development*; *curriculum*; *faculty characteristics*; *program resources*; and *administrative support*. We conceptualize a continuum of performance for each attribute in each of the domains to characterize undeveloped, developing, effective, and distinguished achievement for undergraduate programs. We will discuss this continuum of performance in more detail shortly. Our goal is to encourage individual



departments at various types of institutions to evaluate what they currently do formatively while identifying areas for refinement or future growth. We believe that our recommended benchmarks can improve program quality, encourage more effective program reviews, and help optimally functioning programs compete more successfully for resources based on their distinguished achievements.

### CHARACTERIZING PROGRAM PERFORMANCE

Within any of the eight educational domains, we construed a program's performance attributes as characterized along a fourfold continuum from undeveloped to distinguished. Exhibit 1.1 lists and defines the continuum's characteristics. An undeveloped characteristic is one that is interfering with a department or program's ability to pursue its educational mission. The problem may be, for example, resource-based (such as insufficient lab space) or personnel-based (such as too many part-time adjunct faculty relative to full-time tenure-track colleagues), or it can be the result of political turmoil (such as loss of faculty lines following colleague retirements, personal conflicts between colleagues, rivalries between faculty camps) or the absence of leadership or organizational skills displayed by the chairperson or director. Whatever its source, an undeveloped quality (or qualities) adversely affects students' educational experiences.

#### EXHIBIT 1.1 CHARACTERIZING DEPARTMENT OR PROGRAM ATTRIBUTES: FOUR PROGRESSIVE, CHARACTERISTIC DESCRIPTIONS

---

*Undeveloped*—makes no contribution to undergraduate learning or is even counterproductive

*Developing*—makes a marginal, limited contribution to undergraduate learning

*Effective*—makes a suitable, satisfactory, and favorable contribution to undergraduate learning

*Distinguished*—makes an outstanding contribution to undergraduate learning in the department or program and discipline

---

## 10 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

When a characteristic is labeled as developing, the outlook is a mixed one. Consider a program's attempt to provide adequate curricular exposure to issues of diversity for students. Where an undeveloped curriculum would ignore or overlook the relevance of diversity matters, a developing one would create one opportunity, perhaps within a common course that all majors must complete. Ignoring the dearth of female physicists versus openly discussing why relatively few women pursue graduate education and active research in the discipline would be an example. On the one hand, recognition that sociocultural differences among people exist is in itself a favorable outcome; on the other, however, limiting the discussion to one part of one course obscures the importance of the issue and dilutes its impact on student learning. Hence we consider developing qualities to offer marginal benefits to students, so that the areas for improvement are relatively obvious.

An effective characteristic for benchmarking is one that offers a minimal or adequate, yet still positive, influence on undergraduate education and student learning in a program or department. An example of an effective characteristic is found in the place of service learning in a department's curriculum. Service learning involves teaching, learning, and thinking where discipline-based knowledge is used to tackle some problem or need faced by an entire community or a group (such as teenagers or seniors) within a community. An undeveloped curriculum would not present information on service learning, whereas a developing one would discuss common community-based contexts and review book-based examples. In contrast, an effective curriculum would actually offer students the opportunity to engage in learning about a local community's concrete and real needs through a disciplinary lens.

The fourth and highest level on our continuum is labeled distinguished in that such benchmarks appear to offer an extraordinary opportunity for students to become immersed in the challenges of a particular topic. If we continue with the service learning example, a distinguished departmental curriculum might be one in which students actively designed and executed a project (or, better still, projects) that helpfully addressed a community's

problem. The educational benefit is increased substantially because the outcome uses a department or program's expertise, improves town-and-gown relations, and benefits students who directly witness the constructive impact of discipline-based insights.

We developed this continuum to reflect the reality that departments and programs in humanities and the arts, the natural sciences, and the social sciences come in various forms. Some are large—numerous faculty members, large enrollments, significant resources—whereas others are small, housing few faculty, students, and resources. In some programs, scholarly activities (research, publication, performance, exhibitions) represent the hallmark, whereas teaching in all its forms, from lecture to discussion to online communities, is the *raison d'être* of other departments. Obviously, many departments fall between these extremes.

We do want to be clear about one thing: receiving a rating of undeveloped or developing on some quality within a domain is not meant to be pejorative or probationary; again, benchmarking is for developmental, not accrediting, purposes. Neither a department nor its members—nor its head or chair, for that matter—should be criticized for one or more of these two ratings among the eight academic domains. Again, we view our benchmarking approach as developmental, which means that forward progress is always a possibility—a department can move ahead, say, from developing to effective, on one quality within one domain with a bit of reflection and more helpful assessment strategies. Similarly, the presence of an undeveloped or developing rating should never be used for summative assessment, only for formative assessment aimed at improving the educational experience for undergraduate students—which, once change is implemented, will provide other benefits to the department or program and its members and constituents.

Finally, the benchmarking examples we offer throughout this book are just that—examples. There is no attempt to establish “one size fits all” departments or programs. We offer a variety of detailed examples that can be exported, revised, and rewritten by readers and their colleagues to fit local needs, traditions, and circumstances. Arguably, the conversation that involves such

## 12 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

retrofitting of the benchmarks is an equally important part of the process of program evaluation as is doing the actual benchmarking.

### CONSIDERING INSTITUTION TYPE

Institutions of higher education vary a great deal from one another. There are well over three thousand educational institutions of various types in the United States. Some are research-oriented universities that are known around the world, others are small liberal arts colleges that draw students from all over the country. Still others are regional colleges and universities whose student bodies tend to be filled by individuals from the local area. Selectivity ratings for admission differ; some schools admit few applicants, others admit virtually everyone who applies. Many schools are public; that is, they are supported by a state or local tax base. Other schools are private, dependent on fundraising through alumni donations, foundation or other grants, gifts, and income generated from their endowments.

As we wrote this book, we were conscious that we could not create an institutional profile or model for benchmarking that would fit every institution. Instead, we have tried to focus on the general sort of parameters for benchmarking that should be found at most institutions. Before you begin the process of benchmarking your department or program, be sure to consider whether the nature of the larger institution is at all relevant to the process. If your school has some unusual qualities that bear on your department—the type of faculty it attracts, the nature of the courses or the students who take them—then you may want to factor such issues into the benchmarking process.

Although we wrote primarily with four-year colleges and universities in mind, we believe that, with some slight modifications, our system is readily applicable to most two-year institutions. Indeed, we imagine that all readers will use this book with an eye to adjusting our evaluative dimensions, benchmarking descriptions, and other materials to fit their local customs and relatively unique circumstances. When a dimension or description does not “fit” or seem otherwise germane, just move on to the next one. Similarly, the benchmarking conversation may inspire other

dimensions that uniquely capture the program that have not been identified in the current proposal.

## (RECENT) PAST AS PROLOGUE TO THE EMERGING FOCUS ON UNDERGRADUATE EDUCATION

Why are colleges and universities suddenly so concerned with improving and evaluating what and how well undergraduates learn? Chiefly for three related reasons: accountability pressures, the assessment movement, and the desire to advance the recognition some programs or departments receive in the wider world. As we discuss the influence of these reasons within higher education, we consider how benchmarking can ameliorate their impact within departments or programs.

### ACCOUNTABILITY

As having an undergraduate degree is no longer perceived to be a luxury but, rather, seen as more or less a necessity for those who wish to pursue a professional career in many fields, the ivory tower is no longer a refuge from public scrutiny. As the cost of education has risen dramatically, the public and often their elected representatives want proof that there are true educational benefits associated with the cost of higher education. Public institutions have always faced accountability pressures. Private institutions are now increasingly under the public microscope as well. Various prestigious universities have announced tuition waivers and scholarships for low income and middle class families (see, for example, Crimson Staff, 2006; Damast, 2008). Some politicians have called for wealthy schools to spend more of their endowments instead of seeking to increase their worth still more through fundraising activities. Although no college or university community likes the idea of outside regulation, claiming academic freedom, educational expertise, or entitled tradition does not garner much support beyond the campus gates—colleges are being held accountable. To engage rather than combat such public scrutiny, institutions would do well to explain their strengths

## 14 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

and weaknesses while justifying the quality of the education they are able to provide to undergraduate students.

Concern over the power of public opinion, coupled with the realization that higher education's cost should be justified, has refocused campus energies toward demonstrating that educating undergraduate students is the primary mission of four-year institutions. The upshot is that faculty and administrators want to be able to demonstrate to various constituencies—families, trustees or regents, local or national government, and granting agencies, among other stakeholders—that their programs are effective, that there is a true synergy between teaching and learning. In the final analysis, the best measure or demonstration of accountability is the whole experience of students as they take their educational journey through an academic department. (We invite readers to pause for a moment and to think back to their own undergraduate experiences, paying special attention to the impact of their “home” departments on what they learned and, perhaps, who they have become in the intervening years. We suspect that those readers who had a positive experience in the department or program housing their undergraduate major will agree with us about its importance.)

No department or program is apt to be immune from accountability pressures; however, their specific nature may vary according to local customs, traditions, resources, the type of institution, and so on. Accountability pressures have led to the establishment and growth of assessment practices.

## ASSESSMENT

Simply defined, *assessment* refers to the measurement and evaluation of how well students are learning the material being taught to them (Mentkowski et al., 2000). Besides the aforementioned accountability pressures, most teachers and program directors harbor the authentic belief that they should be able to show that students are benefiting from their studies, justifying the bulk of educational claims. As a result, educators are shifting their focus from traditional outcome measures, like grades, to a broader focus on the processes involved in teaching and learning. Part of

the assessment movement has focused on finding, adapting, or creating assessment instruments that tap into how faculty members actually teach and how students really learn (see, for example, Maki, 2001; Zubizarreta, 2004).

We believe that when it comes to the importance of assessment, the focus should be a bit broader than on just what happens in the classroom. We are interested in this fundamental question: Is a department or program doing a good job of educating its undergraduate students? By the term “students,” we mean both majors and nonmajors, and typical or average learners as well as those who are considered gifted or talented. Can the program’s effectiveness be characterized and demonstrated? The proper focus is captured by Halpern (2004, p. 11), who notes that each department or program “has a list of outstanding alumni of whom they are rightfully proud, but few consider that their most outstanding alumni are hardly representative of the vast majority of current or former students.” Thus assessment is aimed at everyone who passes through a program, not simply high performers or other outliers. Much can be learned from focusing on and evaluating the experiences of most students.

What is the assessment philosophy and practice advocated in this book? Like many educators, we believe that there is no single or “best” assessment indicator. Thus one of the main reasons we rely on performance benchmarks is that they can characterize the various facets of a department or program. We also believe that the assessment of a program or department is best carried out by all its members. Ideally, the faculty and administrators within a unit must cooperate and do the work together; beyond organizational and record-keeping functions, subcommittees are usually suboptimal. Finally, there is no reason to assess the benchmarks of a department unless the faculty use the information wisely and constructively. Where collecting undergraduate student performance data is concerned, for example, the findings should never be used to evaluate faculty performance (that is, for purposes of retention, tenure, promotion, salary adjustment). Assessment, then, should not be used to mete out rewards or punishments but rather to continually improve the education offered within a program or department.

## 16 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

## ADVANCEMENT

After accountability and assessment, the third driver of the need to focus on undergraduate education is institutional advancement. Whether they are willing to play or not, colleges and universities are now very much a part of a nation-wide ratings game. Although a few have opted out (see, for example, Finder, 2007), most are willing to use whatever means available to get the word out about the strengths of what they have to offer. Attracting the best and the brightest students is now a matter of marketing—a term that is distasteful, if not upsetting, to the ears of many academics. Nonetheless, the advent of the ratings and ranking system of *U.S. News & World Report*, coupled with a small industry that routinely publishes “insider” guides to the “best” institutions in the country, has led to pressure to advance institutions publicly at all levels. As a result, departments and programs want to distinguish themselves from each other while making students, their parents, alumni, administrators, trustees, and often state legislators aware of what makes them especially distinguished.

Although we do not want to sanction, let alone take part in, the quasi-arms race of comparing competing institutions with one another with a winner-gets-more-and-better-students mentality, we think there is a point worth exploring here. Namely, that each and every department is apt to have some clear strengths or qualities that make it stand out from peer departments. These strengths might be related to the curriculum, characteristics of the faculty, program resources, or some other important dimension. Identifying such strengths by using performance benchmarks will allow a department to publicize these qualities both internally (to other departments or programs or even colleges within a university or university system) and externally (to other programs at peer colleges or universities).

Within the context of accountability, the other side of benchmarking is equally important: those areas within a program that need improvement can be recognized and then addressed. Benchmarking thus can be an impetus for planning for the future, a task with which many departments struggle. Our own experiences as program reviewers and department members remind us that resource constraints often preclude quick action



where areas for improvement are concerned. Nonetheless, we maintain that doing benchmarking and sharing the results with administrative units allows a department or program to justify specific needs in order to move to the next level of performance. If nothing else, advancing a department or program's agenda via benchmarking can lead to frank discussions among faculty, department or program administrators, and upper-level administration (such as deans, VPAs, provosts) about how best to use available resources to achieve shared ends. The value of the resulting collegiality should not be underestimated.

## LOOKING AHEAD: USING QUALITY BENCHMARKS FOR YOUR UNDERGRADUATE PROGRAM NEEDS

Readers who want to take a broad perspective on using quality benchmarks for evaluating the educational experiences of undergraduates are apt to read this book from cover to cover. Individuals who want to address a more specific issue, such as program climate, will want to turn to a particular chapter (in this case, Chapter Two). To accommodate both sorts of readers, we wrote the chapters in this book so that they could stand alone as much as possible. Thus although it might be desirable to complete the whole book before writing a self-study of your own academic department (see Chapter Twelve), you could nonetheless turn to Chapter Four now if you needed to immediately begin to think about student learning outcomes. To brief both types of readers, we provide a chapter-by-chapter overview and then offer concrete advice about how to most effectively use the guidelines offered in this book.

### CHAPTER-BY-CHAPTER OVERVIEW

The topic of Chapter Two is evaluation of program leadership. What kinds of behavior help department chairs and program directors realize the greatest success in working with faculty? Some programs connect well with the wider institution while others languish in isolation, which is often a direct result of skilled

## 18 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

leadership or its absence. Some programs thrive even when resources do not appear to be adequate because of the commitment of the faculty, whereas programs rich in resources may not take full advantage of their riches. We offer a developmental framework to capture a continuum of leadership quality across several dimensions of program climate.

Chapter Three establishes the importance of assessment, accountability, and accreditation to any undergraduate program. A brief history of the assessment movement in higher education is presented. In the course of discussing our assessment framework, we identify what kinds of information departments should routinely collect and consult in the course of program evaluation. We also consider the useful role of assessment data for program accountability and promotion. The chapter concludes with a discussion of the impact of accreditation on quality processes.

As already noted, student learning outcomes are the focus of Chapter Four. This chapter proposes how faculty and administrators can evaluate students' writing, speaking, research, collaboration, and information literacy and technology skills through the lens of a given discipline.

Chapter Five concerns curriculum evaluation. Relying on our benchmarks framework, this chapter presents the key factors that should be addressed by any undergraduate program's curricula. We also discuss how programs can tailor our standard framework to evaluate unique qualities present in their programs.

Student development is not only the concern of student services offices. Chapter Six examines student development through the lens of a sponsoring program or department. What happens to students in the course of their education within a program? How and to what degree are they changed and challenged? Do students feel a part of the program and the larger discipline? Programs often neglect concrete activities that can enhance the experiences of their undergraduate students. This chapter presents ways to determine and enhance current student involvement in the life of a program.

The assessment movement has not overlooked faculty evaluation, which can be a thorny issue for many institutions. Chapter Seven offers guidance on how to constructively evaluate faculty characteristics within a given department or program. Admittedly,

faculty members are wary of having their teaching or professional activities assessed, yet evaluation is very much a part of academic life. We reiterate the importance of formative rather than summative assessment in the constructive evaluation of faculty efforts in and outside of the classroom. Using our developmental framework, we discuss ways to assess teaching, scholarship, resource generation, service, professional involvement in the discipline, community participation, availability to students, and ethical conduct. We also discuss the importance of faculty development.

Higher education is increasingly expensive. Thus Chapter Eight will be of great interest to many readers, especially program directors, department chairs, and deans, because it deals with how to review program resources. Some programs are equipment intensive, others can seek external funding, and still others need relatively few resources save for a requisite number of faculty members. This chapter explains how our developmental framework can be used to characterize available educational resources. Special attention is given to using the results of the benchmarking exercise to seek additional resources, including faculty lines and equipment.

Aside from the availability of resources, one of the most practical concerns for any academic program is administrative support, the subject of Chapter Nine. The truth of the matter is that programs respond to support and encouragement from their academic administrations. Programs cannot excel if they do not receive quality support from administrations. This chapter presents a framework for characterizing administrative support for

- Carrying out a program's mission
- Following university bylaws and procedures
- Adhering to appropriate and fair evaluation systems
- Making teaching assignments
- Advancing scholarly work
- Recognizing a program's efforts

All programs can benefit from a candid assessment of program climate and leadership, student learning outcomes and the curriculum designed to meet those outcomes, resources both administrative and fiscal, and issues of student development. These

## 20 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

dimensions are sometimes specifically influenced by the nature of the discipline. Many books dealing with nascent issues in higher education adopt the tools, techniques, and approaches of the social sciences. Indeed, the authors of this book are all psychologists by training. Yet we are keenly aware that faculty members who teach in programs in the arts and the humanities often face special considerations and challenges. We wrote Chapter Ten to tackle some of these considerations, including the different perspectives on scholarship, such as performance or exhibition, different teaching approaches, and, of course, distinct resource needs. We also examine the special challenges of crafting meaningful assessments in specialized programs, such as interdisciplinary majors.

Following that, in Chapter Eleven we offer suggestions for evaluating support for the natural sciences. We address factors that are unique to development of student learning outcomes, assessment, and curriculum within the natural sciences.

We are mindful that readers will have different reasons for consulting this book. Some will want to prepare a benchmark-based self-study in anticipation of an internal or an external review; others may want to develop skills before serving as an invited external reviewer for a program or department. Chapter Twelve offers advice on how to conduct an undergraduate program review using quality benchmarks. The chapter ties together the frameworks presented in the earlier chapters by providing concrete guidance on effectively conducting an internal program review (self-study). To aid would-be external reviewers, we wrote Appendix A, a guide for how best to use benchmarking while serving as an outside reviewer of another institution. In Chapter Twelve and Appendix A, we provide recommended timelines and suggest what information should appear in evaluation reports.

Finally, we close the book in Chapter Thirteen by considering how we can better serve our students and our institutions by using formative tools like quality benchmarks. This brief concluding chapter discusses how quality benchmarks and the various developmental frameworks presented in the earlier chapters can benefit the experience of students and improve the quality of our colleges and universities. We sincerely believe our approach can refocus and invigorate the work being done within academic programs.

## HOW TO USE THIS BOOK

Now that you have an overview of the book, how should you plan and prepare to use it for benchmarking purposes once you are done reading? First, no serious program review, especially a self-study, can occur without the active, engaged, and interested participation of the faculty who teach within it. Second, these teacher-scholars need to agree with one another about the department's or program's mission. Thus, before undertaking any program review, the department's mission statement must be revisited and then renewed or revised. Once the mission is agreed upon and shared by the faculty, only then should a benchmarking review begin in earnest. When the review does begin, all colleagues within the program or department should have a role. Without a collective sense of purpose and commitment, the practice of benchmarking to identify strengths as well as areas requiring attention will not be taken seriously.

As noted earlier, some readers may want to focus on a particular issue, such as assessing the quality of a departmental curriculum. Others may want to assess the entire program. In the latter instance, a team approach may work best, one in which teams of two or three colleagues read a relevant chapter of the book and then focus their attention on using the quality benchmarks (and gathering appropriate data) linked with the domain of interest. Each chapter contains a matrix of relevant dimensions plotted along the four possible program attributes just described.

As we will emphasize throughout the book, no program should expect (or even aspire) to be distinguished in each and every area. Our framework is designed to help programs identify and tout what they already do well even in situations of seriously constrained resources. Finally, using performance benchmarks to identify areas of program strength can, in turn, be used to recruit and retain students, to seek funding via grants or alumni support, and to enhance the perceived rating of an institution. When benchmarks reveal that a program or areas within it are undeveloped or developing, faculty and administrators can then plan for where subsequent efforts and resources would best be placed in order to raise a program's standing.

## 22 ASSESSING AND DEVELOPING UNDERGRADUATE PROGRAMS

Each chapter ends with some Guiding Questions. These questions are designed to encourage readers to link a chapter's material to the current state of the relevant domain in their departments or programs. We believe that these questions—and others derived from them—are a good way for departments and programs to begin a self-study. Program chairs, administrators, and faculty member should join one another in trying to answer these questions as a starting point for a program review. Indeed, we want to reiterate the point that the most successful climate for promoting change and development within a program is achieved when all parties participate. Thus each chapter closes with a set of questions so that readers and their colleagues have a place to begin the process of benchmarking their department or program.

### GUIDING QUESTIONS

1. Before doing any benchmarking, how would you describe the state of undergraduate education in your department or program? At your institution more broadly?
2. What qualities, if any, make your program stand out from other programs in your division (humanities and arts, social sciences, natural sciences) at your institution? Would colleagues in other departments in your division agree with your assessment? Why or why not? Do you believe your dean or provost would share your opinion?
3. What qualities, if any, make your department or program distinguishable from others at comparable institutions (that is, those comprising your formal or informal peer group)? Would disciplinary colleagues at these other schools agree with your assessment? Why or why not?
4. What are your goals for department or program evaluation? What do you hope to accomplish?