Final Report of the Assessment Task Force on Assessment of Student Learning at the Program Level

Carnegie Mellon University
May 2010

Susan Ambrose, Provost's Office and History (Co-Chair)
David Boevers, Drama
Steven Brookes, Computer Science
John Carson, Art
Milton Cofield, Tepper
George Duncan (Heinz, ret.)
Mark Fichman, Tepper
Brian Junker, Statistics (Co-Chair)
Gordon Rule, Biological Sciences
Paul Steif, Mechanical Engineering
Karen Stump, Chemistry
Klaus Sutner, Computer Science
Lowell Taylor, Heinz
Lynn Walker, Chemical Engineering

with the assistance of

Tarini Bedi, Eberly Center
Anne Fay, Eberly Center
Marsha Lovett, Eberly Center

Table of Contents

I. INTRODUCTION

Context

In the 2006-2007 academic year, Carnegie Mellon performed a self-study of its assessment practices as part of the preparation for the Middle States Commission on Higher Education's upcoming visit in 2008 to renew the university's accreditation. As a result of this initial self-study, we realized that, while Carnegie Mellon has always been data-driven, we are not as systematic or transparent as we could be in the assessment methods and processes we use. We also lack a shared language for discussing our assessment activities at the program, college, and institutional levels. These shortcomings are a result of the decentralized nature of the institution where local practices are rarely shared beyond departments and schools as well as the bottom-up approach with which we have always done assessment and which allows faculty, departments, and schools to use methods specific to the their discipline and local culture.

In this context, the Assessment Task Force (ATF) was created for a three-year term, beginning in 2007 and ending in 2009, with two goals. The first goal was to document more fully all of the forms of assessment that currently exists on campus. The second goal was to point the way toward future evolution of a sustained and deliberate culture and practice of assessment on campus.[1] The ATF focused predominantly on program assessment[2], which uses results from course-level assessments, surveys of graduates and their employers, faculty's judgments of needs in the field, and other sources to provide coherent, up-to-date data for use in decision-making. In this final report of the ATF, we review our activities with particular focus on the three central questions about assessment at Carnegie Mellon that we began with three years ago. These questions include:

1. What is the current state of assessment of learning outcomes across the entire campus at the program, department, and school levels?

2. What facilitates or hampers engaging in assessment activities at the department and school levels?

3. What should be the future of assessment practices and culture on campus?

Each of these questions will be considered in turn in the subsequent sections of this report. First, however, we present our general conclusions.

---

[1] The dean of each college appointed two faculty members from that college to serve on the ATF; membership changed minimally over the three years of the ATF's existence.

[2] Program assessment is a systematic, ongoing process of gathering, analyzing, and using information from various sources to improve programs. In educational institutions, program assessment measures program outcomes in order to improve student learning—more specifically, program assessment involves articulating what the program's graduates should know and be able to do as a result of their participation in the program (Hutchings & Marchese, 1990; Palomba & Banta, 1999; Huba & Freed, 2000).

General Conclusions

The ATF offers seven general conclusions about the assessment practices and culture at Carnegie Mellon:

1. *Assessment of student learning at both the course and program levels does occur at Carnegie Mellon.* Student learning is assessed using methods that include paper and pencil work, authentic activities including course projects and performances[3], and interviews with employers and other "consumers" of Carnegie Mellon graduates. However, there is little uniformity across campus in how this assessment activity is documented or integrated into formal program reviews and improvements. This is partially because departments are at different stages of development and have different attitudes toward assessment.

2. *A wide variety of beliefs about the purposes and roles of assessment in academic programs exists at Carnegie Mellon.* As a result, different faculty, administrators, departments, and schools have different motivations for participating in assessment— for example, general program redesign, documentation for an external accreditation board, evidence for instructor mentoring, and raw materials for marketing. Consequently, the commitment to assessment varies greatly across departments and schools at Carnegie Mellon.

3. *Faculty's view of assessment is more positive now than prior to the Middle States self-study.* From anecdotal evidence, the discussions that emerged as a function of the self-study process, as well as the continuing work of the ATF, helped to dispel negative myths about assessment and to increase faculty's awareness of and openness to the potential value of assessment in their teaching practices and programs.

4. *The ATF provided important access to attitudes toward and materials for assessment across departments and schools at Carnegie Mellon, but the ATF itself could not and cannot provide the leadership that is required for the short- and long-term future of assessment at Carnegie Mellon.*

5. *The need for resources and support to enhance assessment practice is expressed throughout the entire Carnegie Mellon community.* The departments with access to resources and support are more positive about assessment and more likely to integrate assessment into their practices, at both the course and departmental levels.

6. *Carnegie Mellon remains at the forefront of peer institutions despite the needs that the ATF's work has revealed.* Our review of selected programs at other colleges and universities did not find any other institution that was able to capture assessment practices as part of a continuous internal program improvement process. Rather, assessment at the institutions we reviewed tends to be treated as largely summative

---

[3] Assessment and other classroom activities are considered "authentic" if they closely model the kinds of activities that students will perform outside the classroom as employees, graduate students, citizens, and so on.

for students (in determining course grades) or programs (in creating documentation for external reports).

7. *For Carnegie Mellon to reach its full potential in effective assessment practices, assessment must be championed by a network of leaders at the university, college and program levels who are committed to the process and model the value of assessment by using data in their own decision-making.*

II. CURRENT STATE OF ASSESSMENT OF PROGRAM LEARNING OUTCOMES

Middle States Findings

As part of the self-study that occurred prior to the Middle States accreditation visit, Carnegie Mellon convened a working group on assessment that, together with satellite working groups in each college and school within the university, carried out an audit of all departments and academic units, focusing on: (1) a general description of the unit's undergraduate program; (2) program-level learning goals for undergraduates in that unit; and (3) curricular modules and milestones.

In addition, specific documents were collected from each program, forming an evidentiary basis for the responses in these three areas of the assessment audit. The following four general conclusions and recommendations were reached as a result of this audit and self-study process (Carnegie Mellon University, 2007, pp. 114-117) and were confirmed by the evidence gathered by the ATF:

1. *There is a high degree of alignment between university and program learning goals.* All departments have well-defined sets of core courses that, combined with college-defined general education requirements, provide graduating students with the depth and breadth of knowledge and skills that define a Carnegie Mellon graduate. There is, however, variability in how clearly and publicly articulated the goals are and in how systematically learning outcomes are reviewed and used to guide curriculum revision.

2. *All students are assessed on a regular and rigorous basis to ensure that they have acquired the knowledge, skills, and competencies expected of a Carnegie Mellon graduate.* These assessment activities include a wide range of formative and summative direct measures, including feedback to students in time for self-monitoring and help-seeking as needed. These measures are often supplemented with data from multiple evaluators, perspectives, and other indirect measures on each student. Many programs also use a capstone course, project, thesis, recital, or other performance in the senior year to assess students' overall proficiency in the major.

3. *In addition to formal survey data provided by Carnegie Mellon's Office of Institutional Research, colleges, schools, and departments use various combinations of other data to inform educational practices.* These other data include students' professional publications

and performances to placement data, interviews and surveys of various stakeholders (e.g., graduating students, alumni, employers), student focus groups, and student advisory councils. Colleges and departments also undergo periodic self- and external evaluation of their instructional programs based on the judgments of their own faculty and on recommendations from external experts, including advisory and accreditation boards. A few faculty-led education research projects have examined student learning outcomes in more depth and have had implications for teaching and learning practices both within and beyond the institution.

4. *Practices diverge widely across Carnegie Mellon in the collection and use of assessment data, especially for program monitoring, evaluation, and improvement.* Some colleges have well-articulated, transparent, and systematic processes for collecting and using assessment data. Other colleges' educational programs have recently undergone extensive review and revision, and these units are in the process of implementing systematic and transparent assessment and program review processes. Some departments systematically collect a broad range of data on student learning but are not using it as effectively as they could, while others collect a narrower sets of data but use it more effectively. Broadly speaking, we lack a shared language for discussing our activities at the program, college, and institutional levels. Many departments would benefit from knowing how other departments have proceeded, thus avoiding common pitfalls and progressing more efficiently and effectively.

ATF Activities and Findings

In the 2007-2008 academic year, the ATF invited 51 department heads and other faculty members to participate in basic data collection in order to broaden and deepen the collection of assessment practices begun for the Middle States accreditation process. We also conducted individual, in-depth interviews on assessment approaches, activities, and attitudes across the campus with 33 department heads and other faculty members; in these interviews, six colleges and 17 departments were represented. Throughout these discussions, we were also able to collect 29 assessment artifacts and examples of assessment practices at various stages of development across the campus for inclusion on the assessment website (http://www.cmu.edu/teaching/ assessment [Assessment Task Force, 2008a]). We also created a template that can be used to document a wide variety of assessment tools.

In the next three subsections, we discuss our findings regarding assessment culture, assessment practices, and the assessment website.

Assessment Culture. In our interviews with department heads and faculty, we were able to identify three distinct groups on campus that differ according to their attitudes toward and knowledge of assessment and, as a result, how they engage in assessment practices.

Established and Knowledgeable Users. The first group, which comprised about 10% of our contacts, consists of a small number of department heads and other faculty who view assessment as a valuable pedagogical and research activity. Most are already incorporating assessment into their pedagogical efforts and are actively seeking best

practices in assessment in order to be part of the national conversation on assessment. Members of this group would like access to experts to help them develop and/or rework their assessment approach, plans, methods, and/or tools.

Emerging Novices. The second group, which consists of the large majority of department heads and other faculty we interviewed, is just beginning to see the value of assessment practices beyond summative mid-term and final exams and, indeed, beyond the classroom. Members of this group acknowledge that they need guidance in directing their efforts, but most do not know what resources are available. They are not resistant to altering their approaches to incorporate clearly articulated outcomes and assessment of learning outcomes into their curricula or programs. Some members of this group have developed rudimentary assessment tools but are unsure about how to increase their effectiveness, put them into practice, and/or use the data they collect. Others are unsure about how to start the assessment process or how to convince their colleagues of the value of such a process. In general, this group is looking for resources to help them and was the group most likely to mention the need for a centralized repository, such as a website, that could be accessed on an ongoing basis and provide examples of assessment practices used across campus.

Resisters. The third group consists of department heads and other faculty—less than 10% of our interviewees—who resist the idea of assessment. Much of this resistance derives from the language of assessment, its association with external accreditation, and the belief that assessment is overly time-consuming. Related to these reasons for resistance is the concern that formative assessment (i.e., for improvement) is too easily conflated with, or passes as, summative assessment (i.e., for judgment for promotion and tenure).[4] We also encountered some concern that formal assessment inherently captures what is easy to assess rather than what we value as an institution. However, some members of this group indicated that centralized resources, dedicated to helping both individual faculty and program-planning committees, could reduce this resistance tremendously.

Assessment Practices. Many assessment practices across the university—beyond traditional course-level tests and quizzes—have emerged from very practical needs to address particular problems at the course and program levels.

Although ATF focused primarily on undergraduate program assessment, the issue of classroom assessment naturally arose in many of our discussions. There are many thoughtful motivations for classroom assessment beyond summative, grade-generating exams and quizzes. These include explicating grading criteria, clarifying expectations in process-oriented and project-oriented courses, encouraging students to apply basic concepts to problem-solving, and determining students' prior knowledge and assumptions about the field, the course, and even the nature of learning, among many other motivations. Assessment tools that address these issues include rubrics and rating scales completed by instructors, supervisors, or the students

---

[4] This is why the Eberly Center's consultations with instructors are kept in confidence with the instructor and are not shared with the instructor's department or program. This allows a purely formative interpretation of the Eberly Center''s evaluations to be maintained.

themselves; short quizzes at the beginning or end of class quizzes, as well as "instant" mid-class quizzes mediated by classroom "clicker systems," which identify general or specific areas of confusion; journals, process books, and detailed writing prompts, which allow students to reflect on their work; and fully or partially controlled experiments evaluating curriculum variations using both traditional assessments (e.g., pencil and paper tests) and technology-enabled assessments (e.g., online tutoring systems). Many of these classroom-level assessments are documented on the assessment website.

In considering program-level assessment, both in internal discussions and in conversations with departments and other academic units at Carnegie Mellon, the ATF identified four motivations for program-level assessment that are distinct from those described for course-level assessment. First, program-level assessment is an appropriate response to a variety of external demands, including providing information to advisory boards, making arguments about resources to central administration, and meeting the requirements of accreditation agencies. Second, program-level assessment maintains and improves a program's national ranking. Third, program-level assessment aligns content (i.e., courses taught) with effectiveness (i.e., graduates' preparation for careers and further education), which facilitates recruiting for a marketing a program as well as analyzing graduates' placement records in graduate school and professional employment. Fourth, program-level assessment regularizes a program's curriculum by using immediate post-tests and downstream pre-tests to set curriculum expectations and provide a quality check on curriculum delivery.

In some departments and programs, these goals are part of a broader, but often informal, culture of internal program evaluation. Faculty members in these units care deeply about their teaching and its contribution to the overall success of the program. Indeed, it can be seen from the above examples that assessing student learning outcomes (i.e., setting learning goals and assessing how well students meet these goals) follows naturally from assessment that is intended to increase program effectiveness. Some common tools that are used for program-level assessment include alumni surveys, employer surveys, exit interviews with graduating seniors, focus groups with current students, focus groups with recruiters, projects in capstone courses, senior theses, observations of students' final or senior performances, and portfolios compiled over the course of students' undergraduate study. Examples of some of these tools are included in the assessment website (Assessment Task Force, 2008a).

These program-level assessment tools usually are developed in each department or program as needed. Little sharing of materials or perspectives takes place across campus, whether programs share superficial similarities or have similar assessment needs (e.g., performance assessment in drama and project assessment in statistics or computer science). In addition, very few departments and programs explicitly use the data they collect to make curricular revisions, largely because there is a lack of knowledge about how to do so. The exploration into the culture and practices of assessment on campus also helped to inform the development of the assessment website.

Assessment website. Using our exploration of assessment culture and practices on campus as a partial guide, the ATF directed the creation of an assessment website (Assessment Task Force, 2008a), which includes examples of the assessment practices described above. The ATF created

the assessment website to capture the richness of practices and examples of both course and program assessment, to facilitate the sharing of materials and perspectives on assessment, and to provide some guidance in choosing and using assessments according to purpose, audience, and other needs.

The assessment website (Assessment Task Force, 2008a) includes descriptions of both course-level assessments, which can help instructors monitor and improve student learning and inform their teaching practice, as well as program-level assessments, which can help ensure that our programs are preparing our students to meet the demands of an evolving workplace. The website also provides information on why and how to design and implement assessment activities, analyze and interpret results, and apply findings to improve educational practices.

The Eberly Center for Teaching Excellence has assumed maintenance and expansion of the assessment website from the ATF for two reasons. First, the Eberly Center's work with individual faculty always emphasizes the importance of aligning course assessments with course objectives and instructional activities. Second, the Eberly Center also works with departments and programs interested in reviewing and revising their curricula, a process that always includes program assessment.

However, we have two concerns about the website's long-term viability. First, although the website was created with funds allocated to the ATF, there are no funds allocated for maintaining the site; this includes both technical maintenance and intellectual management. Second, a resource site like this remains useful only if it is regularly infused with new ideas and examples. Thus, it is critically important that there are key people within the university community (e.g., a designated assessment liaison in each program) who contribute new and/or revised examples and practices to the Eberly Center for this website. This need is discussed in more detail later in the report.

We cannot stress enough the importance of the website in helping to create a shared language and shared set of understandings about assessment across campus, consistent with the bottom-up approach to assessment that we use at Carnegie Mellon.

III. FACILITATORS IN AND IMPEDIMENTS TO ASSESSMENT PRACTICES

Throughout the first and into the second year of the ATF, we identified a number of impediments to and facilitators in the establishment of richer assessment practices on campus that might provide better feedback to students as well as better data on which to base curriculum evaluation and contribute to an internal program review, among other results. This information played a large role in the ATF's discussions of what would constitute a sustainable and deliberate culture and practice of assessment on campus. In the subsections below, we discuss these impediments and facilitators, respectively.

Impediments

We identified five key impediments to establishing an effective culture and practice of assessment on campus:

1. *Lack of resources and supporting infrastructure.* The most significant impediment is the lack of resources and knowledge available to help design assessment approaches, plans, methods, and/or tools.

2. *Faculty view of assessment as busy work.* Faculty who do not see the immediate benefit of assessment activities view assessment as busy work. This may be, in part, a result of faculty's previous negative experiences with assessment activities, such as poorly constructed assessments that did not provide robust information matching faculty's needs, or collecting data that was never analyzed or used.

3. *Lack of constituent engagement.* Even when provided with easy access to opportunities to provide feedback about a program, individuals (e.g., current students, alumni, employers, faculty) may not participate.

4. *Alienation and lack of a shared process.* Most faculty and department heads whom we interviewed feel that they have to develop assessment activities on their own. This is a process that is perceived as overly daunting and time-consuming. As a result, assessment practices are often in disarray, or directed by an "assigned"—or, at best, volunteer— faculty member. In many cases, the individual directing assessment practices has little input from colleagues, minimal institutional support, and limited knowledge. As a result, the assessment tools produced tend not to be well aligned with the real or perceived needs of other faculty and may receive limited attention. Such assessments seldom find their way into routine use, and the entire process often reinforces negative attitudes toward assessment and causes faculty to disengage.

5. *Lack of sharing of assessment practices or assessment knowledge.* Most faculty members reported very little sharing of assessment efforts even within the same department, let alone across departments and colleges.

Facilitators

We identified three key facilitators in establishing an effective culture and practice of assessment on campus:

1. *Leadership from department heads.* Department heads who support and encourage assessment practices seem to be key in creating an effective culture of assessment. Departments in which heads routinely talk about assessment of learning also seem to be departments in which the largest number of faculty members already use assessment tools.

2. *Access to resources.* Assessment practices are more likely to be used when faculty or departments have easy access to examples that can be used as-is or be easily adapted or to collaborative efforts to develop assessment tools within the department. In departments

that have developed online sharing of assessment practices, faculty are more likely to incorporate assessment practices into their teaching.

3. *Support from experts.* Individual faculty or departments who have sought help from the Eberly Center are more likely to have assessment plans and practices in place. These faculty have partnered with the Eberly Center staff to collaboratively develop assessment tools and engage in ongoing consultations to support their emerging assessment practices.

IV. THE FUTURE OF ASSESSMENT AT CARNEGIE MELLON

Statement on Assessment

As a major part of its second-year activities the ATF created a position statement on assessment (Assessment Task Force, 2008b), published on the assessment website with the approval of the Provost. An excerpt of the statement follows:

> At Carnegie Mellon, we live in an environment where assessment is an integral part of everyday activity. It takes place continually and on many levels—course, program, department and school. It takes many forms—formal and informal, summative and formative, qualitative and quantitative, standardized and customized. Our students are assessed by various constituents—instructors in a course, faculty across the department, peers, and audiences made up of the general public. We are data-driven at our core, and actively seek out authentic and meaningful ways to assess our students and our programs. Our choice of methods is guided by questions such as:

> *What will this process tell me about my students' knowledge, skills and growth?*

> *What will I learn about the strengths and weaknesses of our program?*

> *What information will this give me on how to improve my teaching or our program?*

> These kinds of questions help to ensure that our assessment practices align with our curricular goals, and provide useful and usable feedback.

The ATF identified three key features of the Carnegie Mellon approach to assessment by reviewing the Middle States self-study materials (Carnegie Mellon, 2007) and holding discussions within the ATF and between ATF and various academic units at Carnegie Mellon. These key features include:

1. *One size does not fit all.* Individual instructors, departments, and colleges are best positioned to determine how to most effectively assess students in a specific course and for a specific program's specific objectives and goals. For some departments or

programs, standardized competency exams developed by external agencies or boards may be appropriate for assessing student competency in certain concepts and skills. Others may find that internally created, highly customized assessments, which can be more easily and frequently revised to reflect a rapidly changing knowledge base and workplace, are more appropriate to assess their students' accomplishments. In most cases, especially when the consequences of decisions based on assessments are greater, using more than one type of assessment will be most appropriate, but the tradeoffs of different assessment tools vary greatly across campus.

2. *We assess what we value, not what is easy.* Since what we assess drives what students try to achieve, it is critical that we focus assessment on what we value. There is ample evidence in the behavioral economics literature to support the notion that rewarding one thing while valuing another (Kerr, 1975)—or giving equal rewards to easy but less valued and harder but more valued behaviors (Homstrom & Milgrom, 1991)—will produce disappointing educational results. Carnegie Mellon values graduates who are resourceful and creative in fluid environments, who are leaders, and who can cooperate in team efforts. It is critically important that, when these are the traits we wish students to develop, we use assessments that are valid for these traits, even if, at first, they are difficult to implement or lack established reliability and validity. Experiences with research and creative expression, as well as team projects, are nearly universal features of Carnegie Mellon, in part because this is where we can see resourcefulness and creativity in fluid environments, whether they are artistic, scientific, or professional. The development of clearly articulated performance criteria through repeated reflection and revision, represented in rubrics or other shareable formats, can move us toward more reliable assessment.

3. *Learning outcomes assessment should be situated within a broader educational context.* Assessment data can be used not only to support, guide, and evaluate student learning, but also to support, guide, and evaluate instructional practices and program design. Faculty continuously collect data on their teaching and their courses—much of it informal as they recognize an assignment that didn't work or explanations that weren't sufficient, as well more formal early course evaluations and focus groups. Departments also collect various kinds of data to help inform their decision-making and guide changes to their programs and practices. However, departments and programs on campus vary in the frequency in which they engage in, talk about, or use the results of assessment activities for student feedback and evaluation, internal monitoring and updating of educational programs, or other purposes.

A Selection of Assessment Practices and Infrastructure at Other Institutions

To better understand assessment needs and practices at other institutions around the country, the ATF reviewed assessment programs at a few selected schools identified by the Eberly Center's then-Director of Assessment. The institutions selected did not include our typical peers. Instead, they included the University of Virginia, Alverno College, the University of North Carolina, Texas Christian University, Clemson University, the University of Pittsburgh, the University of Georgia, Swarthmore College, University of Massachusetts (Amherst), the University of Washington, and Oberlin College. Within this group, we sought to distinguish programs intended

to support faculty-led assessment from either institutional research units or units whose main mission was development of faculty teaching practices. Four institutions—the University of Pittsburgh, Texas Christian University, the University of Georgia, and the University of Washington—seemed particularly relevant to Carnegie Mellon. We discuss the practices of each of these four institutions in more detail below.

University of Pittsburgh. The University of Pittsburgh is launching a new assessment program that should be operational during the 2009-2010 academic year. The program's goal is to assess every degree and certificate program, including any sub-programs that have distinct goals. The program describes the assessment process in five phases: (1) articulation of goals, (2) identification of three to five outcomes, (3) identification of assessment methods, (4) development of standards of comparison, and (5) development of process for faculty and/or administrators to review assessment data and close the loop between assessment and program improvement.

In this multi-phase process, faculty are responsible for developing and administering assessments, department chairs coordinate the assessment process, schools and regional campuses document assessments, and deans report the results to the provost annually. An "assessment matrix"—i.e., a worksheet for identifying key issues in the process above--is provided for faculty use, and an "assessment process page" combines brief descriptions of a rich array of assessment opportunities and methods with links to other assessment materials around the country.

Texas Christian University. Texas Christian University has an established Office for Assessment and Quality Enhancement. The vision and organization for assessment is clearly stated on the office's webpage and in an assessment manual. Assessment is divided into three levels—institutional, department/unit/program, and classroom—with the stated requirement of an action plan for each academic unit.

The assessment manual is a guide for departments and academic units and, to a lesser extent, individual faculty who conduct assessment to improve programs and services for students, to inform administration and external groups about a program's contributions within and beyond a university as well as to demonstrate what a program or unit is accomplishing in a way that can be persuasive to students, faculty, staff, and the larger community. The manual defines an effective assessment process as one that can help answer three related questions: (1) what are we trying to do?, (2) how well are we doing it?, and (3) how are we using what we discover to improve what we will do in the future? These questions are further elaborated in a set of nine "Principles of Good Practice for Assessing Student Learning" from the American Association for Higher Education Assessment Forum. The manual also elaborates on an assessment and learning cycle that involves gathering evidence about an existing program, interpreting that evidence in terms of the program's stated goals, and making appropriate decisions about how to maintain or improve the program.

The TCU manual also includes a rubric for evaluating units' assessment plans that is broadly similar to the assessment matrix that the University of Pittsburgh provides to individual faculty.

University of Georgia. At the University of Georgia, support for faculty-developed assessment is organized under the Faculty Development in Georgia Program. Several relevant internal publications are available to faculty, including *Outcomes and Assessment, Assessment Bibliographies*, and *Assessing Teacher Effectiveness*. The university also gives faculty a *Field-Tested Learning Assessment Guide*, which offers self-contained, modular classroom assessment techniques (CATs) and discipline-specific tools for science, math, engineering, and technology (SMET) instructors interested in new approaches to evaluating student learning, attitudes, and performance. Each technique and tool has been developed, tested, and refined in real college and university classrooms. The guide also contains an assessment primer to help faculty select the most appropriate assessment technique(s) for their course goals. The CATs are organized into categories (e.g., attitude survey, concept test), and each technique is described in detail in terms of teaching goals, description of activity, limitations, step-by-step implementation instructions, advantages and disadvantages, and theoretical support. Other resources, such as information about the National Teaching and Learning Forum, also appear in the guide.

Other campus units participate in supporting or coordinating assessment across campus, including the Office of Student Affairs Assessment, the Global Learning Center, and the Office of Institutional Effectiveness. Each of these units contributes an important component of developing an assessment culture, but these efforts do not seem to be well-coordinated across the university.

University of Washington. The University of Washington established an Office of Educational Assessment (OEA) in 1989 as a response to 1987 legislation mandating the monitoring and assessment of public education in the state. The OEA develops assessment tools and provides assessment services of various kinds across campus, including research services for individual departments and programs, course evaluation, optical scanning and scoring, and management of test administration of various kinds. The staff includes specialists in quantitative and qualitative assessment methods, and they collaborate with faculty and staff to create effective assessment strategies, particularly in the areas of assessment of college outcomes, program evaluations, and survey research. In reviewing the OEA's online materials, there is a thread of concern about measurement quality, comparability and reliability, and whether the office should focus on course evaluations (e.g., how reliable are they and how many are required to evaluate a course or a faculty member) or other modalities for assessment.

The OEA also makes available numerous examples of learning goals and curriculum and course mapping. These are provided as a resource rather than a service. The OEA also monitors the extent of assessment by method and learning goals across majors and departments. A recent report shows that most of the institution's departments articulated learning goals that were explicitly related to content (91%), writing (82%), critical thinking (75%), and methodology and research (74%) in addition to other learning goals such as ethics, team skills, experiential learning, and lifelong learning.

In summary, the OEA is a fairly large, sophisticated service organization primarily designed to facilitate and standardize summative assessment with a secondary emphasis on formative assessment.

Strategy for a Sustainable, Deliberate Culture of Assessment at Carnegie Mellon

For Carnegie Mellon to move forward in developing both a culture of and effective practices for assessment, we believe that there are three key areas in which progress must be made. First, those involved with developing, implementing, and maintaining academic programs—whether individual faculty, department heads, or other program personnel—need to articulate *shared values* about assessment. Second, *leadership* must be in place to encourage sustained faculty-driven assessment, of which the methods and the results should be shared across the university. Third, sufficient *infrastructure* in the form of people and resources must be available to support these activities. We elaborate on each of these areas below.

Strategic area #1: Shared values about assessment. In addition to providing instructors with feedback on teaching or students with evaluation of learning, there is a variety of ways that assessment helps to evaluate and improve academic programs. A shared understanding of the multiple purposes of program assessment can achieve will help create a common commitment to carrying it out. The six most prominent reasons for valuing assessment that were identified by members of our community in discussions in and with the ATF include:

1. *Diagnosis and improvement.* Assessment of both students and programs can help to diagnose many underlying causes for problems that have arisen, e.g., low enrollments, decrease in rankings, dissatisfaction with student placement after graduation, dissatisfied alumni, or low-performing students. The underlying causes can then be addressed.

2. *Formalizing the informal.* Assessment can be a tool in moving programs away from "anecdotal evidence" to a more systematic and ongoing process of data collection, analyses, and decision-making.

3. *Confirmation.* Assessment can help to assure that programs are meeting their self-defined goals (i.e., it can help determine whether you are accomplishing what you believe and say you are accomplishing).

4. *Historical documentation.* Assessment and its results can be part of the documentation for efforts, findings, and actions in developing and improving academic programs. It also can provide evidence justifying decisions, policies, and procedures for incoming deans, department heads, and faculty.

5. *Informational.* Assessment can provide information to professional external accreditation agencies (e.g., ABET, AASCB, NAAB, Middle States).

6. *Public relations.* Collecting data from alumni and employers not only provides good information for "diagnosis and improvement," but also the act of seeking their views can build good will with alumni who often complain that the institution only contacts them when asking for a donation. Alumni and employer views can also provide information for marketing our programs.

Strategic area #2: Leadership. Because of the decentralized nature of our university, leaders at all levels—president, provost, deans, and department and program heads alike—must be committed to supporting a culture of assessment. This begins with a clear understanding on the leaders' part of the value of program assessment so that they can lead by example (analogous to the leadership displayed by President Cohon in heading the University's Diversity Action Committee) and by establishing guidelines that keep assessment alive in the minds of faculty (for example, Dean Lehoczky's addition to H&SS faculty's annual reports of space for examples of course assessment tied to department objectives). As the value of assessment is established through these and other activities, it can foster a belief in and commitment to assessment among faculty and other program personnel.

Strategic area #3: Infrastructure. As buy-in from faculty increases, Carnegie Mellon will need to establish an infrastructure of personnel and resources to firmly establish an assessment culture across the university. This will be needed both to coordinate the sharing of materials and ideas across departments and programs and to support the adoption or development of new assessment methods to address needs as they arise within each program.

The ATF recommends a two-level hierarchy of personnel, consisting of a *Central Assessment Coordinator* at the university level and *Program Assessment Coordinators* within each academic department or program. The functions of these individuals are defined in the "Recommendations" section of this report.

V. MOST RECENT DEVELOPMENTS

Carnegie Mellon committed itself to the following three deliverables in its self-study report for the 2008 Middle States accreditation review.

1. The institution should have clearly articulated learning outcomes for all courses within 10 years.

2. Departments and schools should clearly articulate their program outcomes and make them readily available to current and potential students and other stakeholders.

3. Colleges should ensure that student outcomes data are systematically collected, synthesized, and used and that learning outcomes and subsequent curricular changes are documented.

The Periodic Review Report, due to Middle States in June 2013, requires that we show progress toward these commitments. To this end, Carnegie Mellon has set the following goals for these deliverables:

By December of 2012, 50% of courses will have learning objectives, and all programs will have *preliminary* outcomes with documentation of curricular changes made and the data used to inform them.

By December of 2016, in anticipation of the next Middles States accreditation visit, all courses will have learning objectives, and all programs will have a final set of program outcomes that are publicly available (e.g., on websites, in course catalog) and a process that systematically collects and uses data to inform curricula decision-making.

While many programs are involved in continually assessing and, if needed, changing their programs, most do not document such changes in a systematic way. To facilitate and support this process, a Program-Level Outcomes Assessment Chart (see Appendix A) was created to be used as a template for identifying measures relevant to important program-level goals, documenting findings based on those measures, and recording actions taken in response to those findings. A template for Departmental Reports on Learning Objectives in Course Syllabi (see Appendix B) was also created, which can be used to directly document Carnegie Mellon's progress toward its commitments on learning objectives in 2013 and 2016.

In their December 2009 meetings with the provost, deans, and department heads, heads agreed to designate a coordinator in each department and program to work with Susan Ambrose, Michael Murphy, and Beth Whiteman toward meeting the 2013 and 2016 goals. This hierarchical structure is similar to the one that the ATF recommends (see "Strategy" subsection above and "Recommendations" section below) to create a sustained commitment to effective assessment practices at Carnegie Mellon.

Initially, Ambrose and Whiteman are providing support to programs that seek advice on providing this documentation. As departments become aware of their obligations, however, we expect demand to grow to a level that exceeds the availability of these two individuals, so additional structure and people will be required. Separately from this, the Eberly Center's staff will provide assistance to faculty developing or revising course-level objectives, a role that they already play.

VI. RECOMMENDATIONS

The ATF believes that Carnegie Mellon can further capitalize on the momentum gained over the past three years. Success depends on maintaining the decentralized nature of current assessment practices while continuing the trajectory of supporting, sharing, and building upon efforts across the university. Coupled with the aforementioned need for commitment at all levels of our institution, the ATF recommends the following structure and resources.

Recommendation #1. The university should designate a member of the central academic administration, who has program assessment as her/his explicit role and responsibility, to drive the ongoing process of making assessment a more visible, formally recognized activity on campus, with broadly shared tools, experiences and practice, as outlined in Recommendations #2 through #5 below.

Recommendation #2. The university should hire a Coordinator for Program Assessment who works for the person designated in Recommendation #1 and who supports the work of the program coordinators as outlined in Recommendation #4. This individual would:

- keep track of the approximately 150 academic programs (B.A./B.S., M.A./M.S., and Ph.D.) to assure that each one has a documented set of outcomes and an assessment process that collects and uses data to measure success and continually improve programs (not only for Middle States but also for advisory boards, disciplinary accreditation agencies, etc.);

- assure that support is provided to the program coordinators as they document outcomes and their assessment process;

- interact with appropriate units on campus (e.g., Institutional Research and Analysis, Student Affairs, Office of International Education, Intercultural Communication Center) to gather, share, integrate, and/or analyze data toward providing comprehensive documentation of the university's efforts to enhance student learning; and

- develop resource materials and provide presentations to assist faculty and department heads with assessment and curricular review initiatives.

Recommendation #3. The university should provide additional capability, as needed, to work with program coordinators to help programs:

- create or broaden their approach to assessment;
- enhance their current assessment practices;
- identify or expand their methodologies for collecting, analyzing, and reporting learning outcomes; and
- use these data more systematically to inform future course and curricular changes by working with program coordinators, described in the next recommendation.

These individuals should be embedded within already-established university structures with synergistic goals and activities. The hiring should be triggered as demand for these services increases across explicitly defined thresholds.

Recommendation #4. Program chairs should designate an assessment coordinator within their respective units, who would receive release time in recognition of the importance of and time necessary to do this work. This person would be responsible for enhancing and documenting assessment activities and would serve as the liaison to the central Coordinator for Program Assessment. More specifically, the program assessment coordinator would help their respective program to:

- better articulate program outcomes;
- extend or expand current assessment practices, if needed;
- systematically document assessment practices and results used to inform curricular changes; and
- share their approaches, methodologies, and examples with the campus community, for example, through inclusion on the assessment website.

Recommendation #5. The high value that Carnegie Mellon places on successful assessment achievements should be made regularly and prominently visible on campus. For example, an effective program assessment that resulted in either affirmation of successfully meeting program outcomes or program changes directly tied to the collection and analysis of assessment data could be recognized in several ways. For example:

- an award each year for the program that develops and implements exemplary assessment activities (which would need to be defined); or
- visibility on the assessment website, if the capability for maintenance and updating of it manifests.

VII. APPENDICES

Appendix A: Program-Level Outcomes Assessment Chart

This form is intended to facilitate reporting program outcomes assessment to accrediting agencies, advisory boards and other internal or external audiences. For the purpose of following through on 2008 Self-Study recommendations, this information will be collected annually.

Date: _____ **Program:** _____   **Name of Person Completing Form:** _____

Program Outcomes[1]	Direct Performance Measures[2]	Indirect Performance Measures[3]	Major Finding(s)[4]	What Actions Resulted from Finding(s)?[5]

Suggestion: It may be easier to work the chart from right to left, beginning with documenting recent changes to the program curriculum in the "actions" column.

[1]Program outcomes identify knowledge, skills, attributes and/or capabilities students will demonstrate upon completion of the program. The outcomes need to be specific and measurable.

[2]Programs should gather data to measure each stated outcome through direct measures (i.e., students demonstrate their knowledge, skills, etc.).

[3]Indirect measures, where students, employers or others report their perceptions or observations of student/employee knowledge, skills, etc., can be provided but cannot stand alone as a sole measure of student performance.

[4]Programs should identify the major findings after analyzing data collected.

[5]Programs should provide evidence that the results have been applied to further the development and improvement of the program (i.e., actions that were taken as a result of data collection and analysis).

Appendix B: Departmental Reports on Learning Objectives in Course Syllabi

This form is intended to facilitate reporting on the quality of learning objectives in syllabi to accrediting agencies, advisory boards, and other internal or external audiences. For the purpose of following through on 2008 Self-Study recommendations, this information will be collected by departments biannually.

Department: _____

Department Head: _____

Syllabi reviewed for (check one): ☐ Spring 2010 OR ☐ Fall 2010

Please indicate the number of course syllabi in each category (category descriptions below):

Acceptable: _____

Need Improvement: _____

Unacceptable: _____

Acceptable: Learning objectives are present and
- *student-centered* (they focus on the knowledge, skills, abilities and/or attributes students should be able demonstrate by the end of the course);
- *measurable or observable* (they use action verbs -- e.g., apply, interpret, create, design, explain -- to make concrete and explicit the actions and behaviors students should be able to demonstrate by the end of the course); and
- *appropriately specific* (according to the level of students and course context).

Need Improvement: Learning objectives are present but
- *instructor-centered* (they state what students will be taught as opposed to what students should be able to do or know as a result of the course); and/or
- *not measurable or observable* (they do not indicate how students would demonstrate the level of their learning); and /or
- *too broad* (they don't guide the students toward the end goal) or *too specific* (they don't capture the breadth of understanding).

Unacceptable: Learning objectives are
- *missing.*

VIII. REFERENCES

Assessment Task Force, Carnegie Mellon University. (2008a). *Assessment examples & tools.* Retrieved December 2009 from http://www.cmu.edu/teaching/assessment.

Assessment Task Force, Carnegie Mellon University. (2008b). *Position on assessment.* Retrieved December 2009 from http://www.cmu.edu/teaching/assessment/ATF.html.

Carnegie Mellon University. (2007). *2008 self-study overview.* Retrieved December 2009 from http://www.cmu.edu/middlestates/2008-self-study. (Requires Carnegie Mellon password.)

Homstrom, B., & Milgrom, P. (1991). Multitask principal--agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization, 7,* 24-52.

Huba, M.E., & Freed, J.E. (2000). Applying principles of good practice in learner-centered assessment. In M.E. Huba & J.E. Freed (eds.), *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning* (pp. 65-90). Needham Heights, MA: Allyn and Bacon.

Hutchings, P., & Marchese, T. (1990). Watching assessment: Questions, stories, prospects. *Change: The Magazine of Higher Learning, 22*(5), 12-38.

Kerr, Steven. (1975). On the folly of rewarding A, while hoping for B. *Academy of Management Journal, 18*(4), 769-84.

Palomba, C.A., & Banta, T.W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education.* San Franciso: Jossey-Bass.