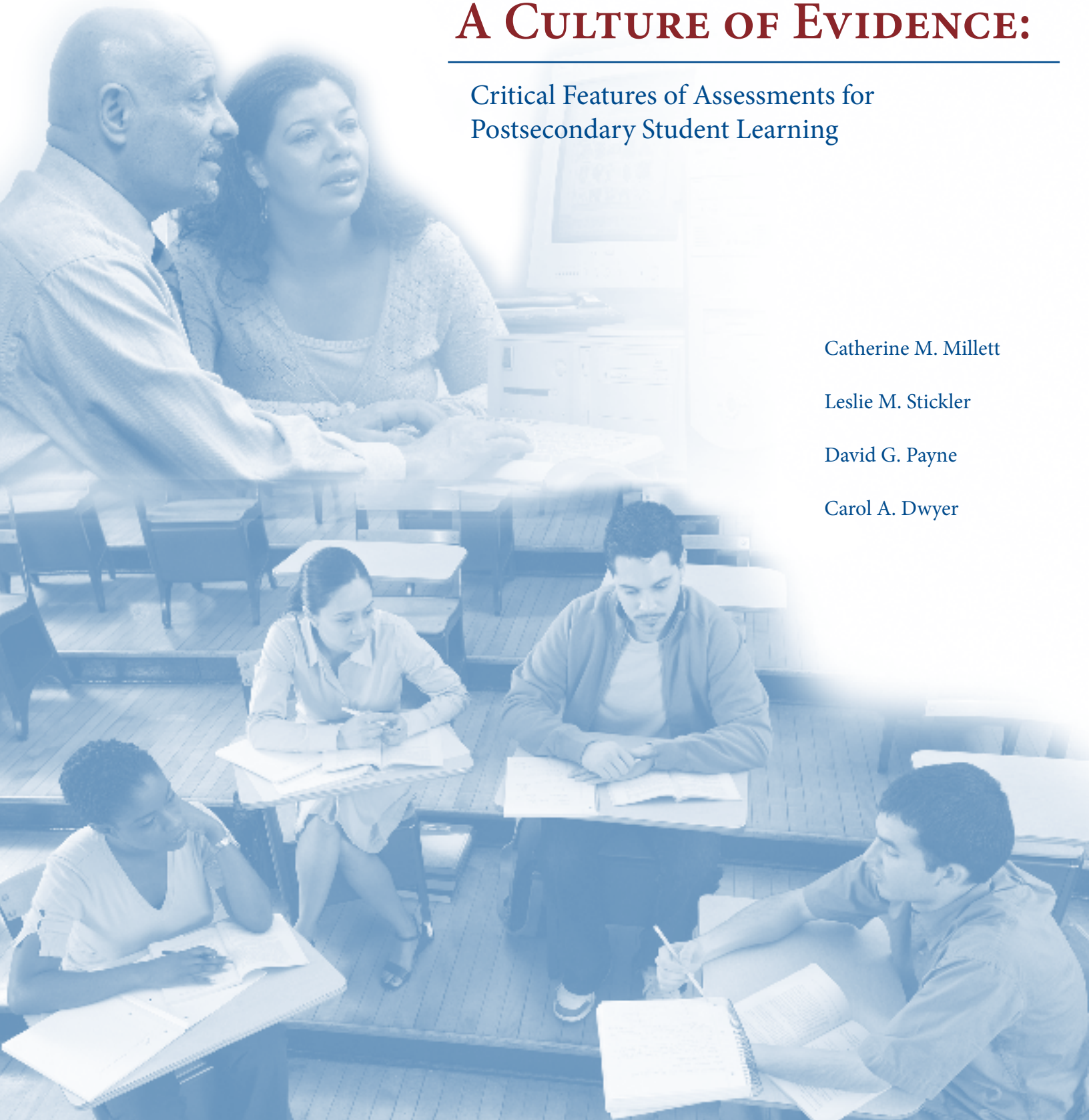
ETS®

Listening. Learning. Leading.

A Culture of Evidence:

Critical Features of Assessments for Postsecondary Student Learning



Catherine M. Millett

Leslie M. Stickler

David G. Payne

Carol A. Dwyer

Dear Colleague:

The current national dialogue regarding accountability in U.S. higher education has brought increased scrutiny to a number of facets of the functioning of our collective system of higher education. One area that has received considerable attention concerns student learning outcomes — how much do we know about the many aspects of student learning that take place during the time a student is in higher education, what inferences can be drawn from the available data, and how much data needs to be shared and with which stakeholders?

ETS is committed to helping the higher education community work through these issues. As a not-for-profit research, development, and assessment organization whose mission is to advance quality and equity in education, ETS is uniquely positioned to work with the higher education and assessment communities to provide objective information that can facilitate the national dialogue on improving learning in postsecondary education. As part of its social mission, ETS has made a commitment to bring to the higher education community a series of issue papers that address accountability issues in the student learning domain. In the first paper in this series, *A Culture of Evidence: Postsecondary Assessment and Learning Outcomes*, we reviewed accountability models and metrics that can be used in higher education.

In this paper, with the help of an advisory panel of national experts in assessment and higher education, we review the major tools in use today for assessing student learning and student engagement, an important aspect of the educational environment. The goal of this review is to provide a high-level overview of the major assessment tools so that higher education stakeholders can continue the national dialogue with even greater understanding of the current state of the art tools in assessing student learning in higher education. This paper provides an overview at the "30,000-foot level," which we believe will be useful to policymakers, national organizations, and two- and four-year college and university presidents and provosts.

We will soon present a third paper, which will provide a review of the critical conceptual issues in planning and executing a program to assess and improve student learning outcomes. We believe that this next review will be especially useful to those involved directly in establishing or improving systems to assess student learning.

We are pleased to be able to contribute to the conversations under way across the nation to find effective and practical ways to assess student learning.

Sincerely,

Kurt M. Landgraf

President and CEO

ETS

Acknowledgments

In 2006, ETS made a commitment to the higher education community to provide information that would help inform the national dialogue on accountability in higher education. The first product from that commitment was an ETS report, *A Culture of Evidence: Postsecondary Assessment and Learning Outcomes*. That report presented an overview of the assessment landscape and made a number of recommendations that should be undertaken as the nation moves in various ways toward addressing the dearth of empirical evidence regarding the learning that takes place in higher education.

The current report represents the second installment from ETS as it fulfills its commitment to bring together information that is needed to inform policy discussions, and institutional and system-wide planning involving the assessment of student outcomes. This second report, *A Culture of Evidence: Critical Features of Assessments for Postsecondary Student Learning* (hereafter referred to as *COE II*) was written in the winter-spring of 2006-2007. At that time, a number of important initiatives were underway in the higher education community regarding assessment and accountability. This report was clearly affected by that larger context.

One of the initiatives that had a major influence on the *COE II* report was the Voluntary System of Accountability (VSA) project led by the National Association of State Universities and Land-Grant Colleges (NASULGC) and the American Association of State Colleges and Universities (AASCU) and supported by a grant from the Lumina Foundation. The VSA project was highly active while the *COE* papers were being written, and the need to provide information to the task forces working on the VSA was a driving force in shaping the *COE II*. David Shulenburger, Vice President for Academic Affairs at NASULGC, and George Mehaffy, Vice President, Academic Leadership and Change at AASCU, played an especially important role in the *COE II* project. They shared important information on the workings and progress of the VSA project and also served as *COE II* advisory panel members. The *COE II* report has clearly benefited from the input of these two dedicated professionals and the membership organizations they lead.

This report has benefited immensely from the advice and sage counsel offered by the 13 members of our advisory panel. The advisory panel members are a distinguished group of research and educational policy leaders. This report has been improved in many ways through the active involvement of the panel, and we thank them for their efforts in helping us produce this report on a very aggressive schedule.

The advisory panel reviewed the criteria used for including assessments in this report, made suggestions for which assessments should be included in the report, and raised important methodological and policy issues that provided an overall context for *COE II*. The panel was also involved in reviewing early drafts of the report and identifying areas in which further clarity or detail would improve the overall utility of the report.

We are grateful to our colleagues Pamela Humphreys at the Assessment Resource Center at the University of Missouri-Columbia; David Chadima and M.J. Klemme at ACT, Inc.; Steve Klein at the Council for Aid to Education; Sharon Edwards at Project SAILS; Anthony Golden at PACAT, Inc. / NOMESys; Kay McClenney at the Community College Survey of Student Engagement at the University of Texas at Austin; and Jillian Kinzie at the National Survey of Student Engagement at the Indiana University Bloomington for providing thoughtful feedback and for checking the accuracy of the information presented about their assessments in Tables 2 through 13.

A number of ETS colleagues also made important contributions to *COE II*. Linda Tyler and Will Jarred from the Higher Education Division and Kelly Denson from the Government Relations & Public Affairs Division provided useful feedback on early drafts of the report and helped the authors to fine-tune a number of the data elements included in this report.

We acknowledge the assistance of these individuals and we are confident that the final report will have more of an impact as a result of their contributions. Any errors of fact or interpretation in this report are those of the authors.

Advisory Panel

The advisory panel for *Culture of Evidence II* and *III* played a critical role in defining many aspects of the documents and in providing input and guidance on the materials presented in the two documents. The panel members were asked to comment on any aspect of the project that they wished to. For example, one of the first tasks in writing *COE II* was to define the population of assessments that were to be included in the review; all panel members were invited to suggest assessments and to help the team establish the criteria for inclusion of assessments.

The panel members also served in more specific ways on this project. Given the limited time frame for completing these reviews, the panel members were assigned to one of three working groups.

These working groups were:

- **Culture of Evidence II.** This group was charged with working on the higher level and less technical version of the reports.
- **Culture of Evidence III.** This group worked on writing the final report produced in this series. This group obviously benefited from the work of the *COE II* group, but it also was charged with making the final report able to stand on its own and with integrating the work from the technical group.
- **"Technical" Group.** This group had responsibility for directing the development of technical aspects of the conceptual framework for assessing student learning outcomes that make up the bulk of the *Culture of Evidence III* report.

These three groups operated in a cascaded fashion. For example, after the larger panel had identified the assessments to be included in *COE II*, the *COE II* group began their work. The *COE III* group benefited from the accomplishments of the *COE II* group and attempted to extend this work to provide a framework and set of heuristics that institutions can use to structure their assessment efforts.

Each of the working groups reviewed initial drafts of both *COE II* and *COE III* as well as the final version of each paper. All advisory panel members were asked to review the final version of both papers.

The diversity of perspectives that each member brought to the project was one of the important contributions of the advisory panel. Some panel members are experts in assessment; others are involved in accrediting or have worked in the policy domain. The final products benefited significantly from the input of these experts and from their diverse backgrounds and experiences.

Advisory Panel Members and Affiliations

Name	Affiliation
Tom Bailey	Professor, Teachers College at Columbia University
Trudy Banta	Vice Chancellor for Planning and Institutional Improvement, Indiana University-Purdue University Indianapolis
Douglas F. Becker	Vice President of Development, Education Division, ACT, Inc.
Steven D. Crow	Executive Director, The Higher Learning Commission
Peter Ewell	Vice President, National Center for Higher Education Management Systems (NCHEMS)
Steve Klein	Director of Research and Development, Council for Aid to Education (CAE)
Paul Lingenfelter	President, State Higher Education Executive Officers (SHEEO)
George Mehaffy	Vice President for Academic Leadership and Change, American Association of State Colleges and Universities (AASCU)
Margaret Miller	Director of the Center for the Study of Higher Education, Curry School of Education at the University of Virginia
Liora Pedhazur Schmelkin	Senior Vice Provost for Academic Affairs and Dean of Graduate Studies, Hofstra University
Pamela E. Scott-Johnson	Professor and Chair of Psychology, Morgan State University
David Shulenburger	Vice President for Academic Affairs, National Association of State Universities and Land-Grant Colleges (NASULGC)
Belle S. Wheelan	President, Commission on Colleges of the Southern Association of Colleges and Schools

Table of Contents

Table of Tables

Introduction

Postsecondary education institutions should measure and report meaningful student learning outcomes

(U.S. Department of Education, 2006, p. 23)

Background

The Spellings Commission on the Future of Higher Education was charged with examining four key areas of higher education identified by Secretary of Education Margaret Spellings: access, affordability, quality, and accountability. The Commission's report, *A Test of Leadership: Charting the Future of Higher Education*, delivered six broad findings aimed at addressing these four areas (U.S. Department of Education, 2006). The six broad findings addressed access, cost and affordability, financial aid, learning, transparency and accountability, and innovation. In her September 2006 Action Plan, Secretary Spellings told the nation that her work would focus on accessibility, affordability and accountability (Spellings, 2006). The Spellings Commission and the hearings that were part of the Commission's work, the report the Commission produced, and the ensuing actions across the U.S. (e.g., the Department of Education initiatives regarding student learning outcomes, accreditation) have focused a national spotlight on the state of higher education. This report is intended to contribute to this ongoing analysis.

The Commission's attention to student learning, coupled with a steadfast interest in accountability, is aligned with ETS's institutional mission and priorities. In 2006, ETS made a commitment to the higher education community to help sharpen the national discussion on accountability—specifically accountability for student learning outcomes. ETS pledged to use its unique resources and its position as a not-for-profit organization dedicated to advancing quality and equity in education in order to bring information and data into the national debate on how to address the need for assessing student learning outcomes.

The first product to emerge from this commitment was a report titled, *A Culture of Evidence: Postsecondary Assessment and Learning Outcomes*. That report outlined a number of accountability models and metrics for the higher education arena. The report also provided an overview of the current national landscape on issues in postsecondary assessment and learning outcomes.

In *A Culture of Evidence (*hereafter referred to as *COE I)*, we asserted that postsecondary education's current state of knowledge about the effectiveness of a college education, as measured by knowledge of student learning in higher education, is limited. This lack of evidence of specific learning outcomes hampers informed decision making by institutions, by students and their families, and by the future employers of college graduates. In *COE I*, we proposed that a comprehensive national system for determining the nature and the extent of college learning could be developed, focusing on four salient dimensions of student learning:

- Workplace readiness and general skills
- Domain-specific knowledge and skills
- Soft skills, such as teamwork, communication, and creativity
- Student engagement with learning

We recommended that to understand the value that a college education adds to student learning, three measures must be addressed: Student input measures (What were student competencies before college?); student output measures (What were their competencies after college?); and a measure of change between the skills and knowledge students brought to higher education and the skills and knowledge they possessed at various points in their higher education careers (e.g., at the completion of the first year, the completion of general education requirements, or attainment of an associate or bachelor's degree).

The *COE I* report also addressed issues of fair and valid testing and outlined the characteristics that a comprehensive system of assessment of student learning might possess. The characteristics mentioned in *COE I* included:

- Regular (preferably annual) data collection with common instruments
- Sampling of students within an institution, rather than testing all students, with an option for institutions to test more
- Using instruments that can be used in pre- and post-test mode, and that have sufficient forms available for repeated use over time
- Using a variety of assessment formats, not limited to multiple-choice
- Identifying appropriate comparison or "peer groups" against which to measure institutional progress

In some ways, *COE I* represented a high-level view of what is needed to begin to build a national approach to assessing student learning outcomes. In the present report, we have attempted to bring some further precision to the overall picture. This is needed in order to make significant progress in addressing the regrettable dearth of widespread information on learning in higher education.

To foreshadow the present report and *COE III*, our aim is to make more widely available information regarding the "state of the art" in assessing student learning. As we move from discussions of student learning at the broad policy level and begin to address them at "ground level," i.e., on campuses across the country, we will need to have a fuller understanding of issues such as which assessment tools are currently available, and what conclusions can and cannot be drawn based on the results obtained with these tools.

Since the initial *COE* report was issued, a number of significant events have helped to further shape the national debate. First and foremost, the Spellings Commission released its much-awaited final report, *A Test of Leadership: Charting the Future of U.S. Higher Education* (U.S. Department of Education, 2006). Although there have been several other influential efforts to address key issues in U.S. Higher Education (e.g., the National Commission on Accountability in Higher Education led by State Higher Education Executive Officers [SHEEO], with support from the Ford Foundation), none have had quite the impact that the Spellings Commission has had.

Second, in response to the work of the Spellings Commission, several leading postsecondary organizations have taken steps to respond to the Commission's charge to "measure and report meaningful student learning outcomes" (U.S. Department of Education, 2006, p. 23). NASULGC and AASCU, with support of the Lumina Foundation for Education, are developing a Voluntary System of Accountability (VSA). Their goal is to develop a system of accountability that will facilitate comparisons of learning outcomes among institutions of higher education. Additionally, SHEEO is engaged with helping the states develop strategies for assessing student learning.

Overview of *COE II* and *III*

These two reports will in some ways address very similar issues; what differentiates them is the focus, or level, of critical review. In this section we will briefly outline the intended purpose and target audience for each of these reports.

Target Audience and Goals for *COE II*

The Spellings Commission identified a broad group of interested parties that need to be engaged in efforts to effect change. These parties included: colleges and universities, accrediting bodies and governing boards, state and federal policymakers, elementary and secondary schools, the business community, and students themselves. We acknowledge that, for the majority of individuals in these sectors of U.S. education, assessment is not the most salient part of their everyday work. Yet they recognize that assessment is at the heart of the educational enterprise and can serve as a tool to gauge what has been accomplished by institutions. Our goal in *COE II* is to provide user-friendly, high-level information on the current state of student learning assessments.

At the end of *COE I*, we provided a table that listed various assessments that have been cited in policy reports as promising measures of student learning outcomes, as well as assessments that are currently being administered by a significant number of colleges and universities. We recommended that the next step would be to convene an expert panel to populate the table and to determine whether each assessment accurately and appropriately measured workforce readiness and general education skills, domain-specific knowledge, soft skills, or student engagement.

In *COE II*, we are taking our work in this direction, with some additional refinements. Our aim is to provide to the nation's 2,500 four-year and 1,600 two-year postsecondary institutions, as well as to the many other stakeholders in U.S. higher education, information that will enable them to begin a critical review of the suitability of currently available assessments for their own purposes. In this report, we aim to address three issues:

1. What assessments are currently available for purchase in the four previously-identified domains of student learning?
2. What information is available to compare the different assessments on critical issues such as format, cost, etc.? How should these factors help guide selection of appropriate assessments?
3. What are the assessments that have served as proxies for student learning outcomes in the extant research that have not been validated for use as student learning outcome measures?

To summarize, then, the goals of the current report are to bring together, in a coherent and consistent framework, and employing consistent language, a guide to the most prevalent assessments of student learning. *COE II* is aimed at institutional leaders, legislators, Boards of Trustees, and other groups that, although interested in issues surrounding the assessment of student learning outcomes, will not be focused on the more technical aspects of the assessment tools.

Our hope is that this information will supplement existing efforts to measure and evaluate student learning outcomes through institutionally-based assessment. Such efforts by institutions and the six regional accrediting organizations (Middle States Association of Colleges and Schools, New England Association of Schools and Colleges, North Central Association of Colleges and Schools, Northwest Commission on Colleges and Universities, Southern Association of Colleges and Schools, and Western Association of Schools and Colleges) have been successful in documenting evidence of eligibility for accreditation and in tracking student progress over time within institutions.

Incorporating standardized assessments into the framework for evaluating student learning outcomes adds a valuable dimension of comparability to the picture, however. For example, although institutions articulating student learning objectives and outcomes may develop high-quality local assessments to evaluate their own programs and students, only standardized assessments can tell institutions how the success of their programs and students compares to that of similar institutions regionally or nationwide. This information can guide curriculum development and instructional planning as institutions look ahead to where they would like their programs and students to be relative to those of their peer institutions.

Target Audience and Goals for *COE III*

In *COE I*, we discussed the importance and characteristics of fair, useful, and valid assessments. A consideration of test fairness and validity is critical when selecting an assessment. The third report in the *A Culture of Evidence* series, to be published later this year, will take a closer look at test effectiveness and validity. *COE III* will expand on the present work to delve more deeply into important conceptual issues regarding the value of assessing student learning outcomes, again with the overarching goal of providing a common set of evaluation criteria that will serve the higher education community. The audience for *COE III* will be institutional assessment experts and those charged with implementing processes and strategies for assessing and improving student learning outcomes.

Validity and reliability are essential technical characteristics of an assessment that merit a brief introduction here. Validity refers to the degree to which evidence supports the interpretation of test scores. Reliability refers to the consistency with which an assessment measures the construct(s) that it purports to measure. A test that is reliable is not always valid — that is, a test that measures something consistently may not necessarily be measuring the construct of interest.

Validity is the most fundamental consideration in assuring the quality of any assessment. In their simplest form, the essential validity questions are: 1) How much of what you want to measure is actually being measured? 2) How much of what you did not intend to measure is actually being measured? and 3) What evidence do you have to support your answers to the previous two questions? (Dwyer, Millett, & Payne, 2006, p. 11). This view underscores the importance of examining tests within their total context. The validity of an assessment cannot be represented by a simple correlation coefficient; rather, determining validity requires a judgment about what inferences can be drawn from test data, including the intended or unintended consequences of using the test.

Regarding reliability, the assessments reviewed here have taken a very wide variety of approaches to reporting reliability. Approaches include measures of a test's internal consistency, measures based on a test/retest design, correlations between tasks, and correlations between scores given by different readers. Most of the data that we reviewed seemed strong; with a few exceptions, the values might seem impressive. But the more important consideration for potential users is whether the data provided are really addressing their concerns about an assessment's suitability for their student population and assessment needs. Issues of validity and reliability, as well as other important conceptual issues, will be treated in greater length in *COE III*. The third *COE* report will provide a conceptual framework and heuristics that can be used when implementing any program of assessing student learning outcomes. These heuristics are not offered as a simple checklist, but rather as a set of theoretical bases for evaluating the strength and efficacy of inferences about student learning outcomes.

Selection of Assessments for This Guide

As noted above, many assessment instruments have been discussed for possible use in postsecondary accountability applications. Not all of these instruments will ultimately be found to be appropriate for such uses, based on closer examination of their design and technical qualities.

The assessments selected for inclusion in this report were chosen after considerable analysis by the authors, and in close consultation with the *Culture of Evidence* advisory panel, as well as others involved in decision making about postsecondary education and assessment policy and practice. The assessments discussed here are by no means an exhaustive set, but rather include only those that were judged to be the most likely to be useful now or in the immediate future. This is thus a very selective review, including only those assessments judged to be most salient to current policy concerns.

Using the classifications that we proposed in *COE I*, we included assessments of general education skills and workplace readiness, domain-specific skills, and student engagement. For the purposes of this document, general education skills and workplace readiness encompass cognitive skills, such as critical thinking and mathematical reasoning, that are widely valued and applicable to most academic and professional endeavors. Domain-specific skills represent in-depth knowledge and skill in a particular subject area. Student engagement describes institutional practices and student behaviors that have been linked to student learning in the empirical research literature.

Although measures of these cognitive outcomes of higher education have been well-defined and extensively developed, the assessment of "soft skills" is still relatively new territory in educational testing and accountability. We were able to identify only one major assessment of soft skills that was currently commercially available for measuring student outcomes, ACT's WorkKeys® assessment in *Teamwork*. Several other soft skills assessments are under development, however: (a) Educational Testing Service's *ReadyEdge*™ assessments and learning tools, which includes self-report and situational judgment tests of community college students' behaviors and attitudes; (b) ACT's WorkKeys Personal Skills assessments, which measure job candidates' performance, talent, and fit for various positions; and (c) the Canadian Foreign Service Institute's (CFSI) Situational Judgment Test of Intercultural Effectiveness, a forthcoming component of the International Personnel Assessment (iPASS) tool. *ReadyEdge* assessments and learning tools, will be available in fall 2007; the WorkKeys Personal Skills assessments are scheduled to debut in spring 2007; and CFSI's Situational Judgment Test will continue piloting and development until winter 2007-2008.

These new tools will provide important new means of assessing an important aspect of the so-called noncognitive skills. However, because these measures are not currently available for purchase, they are not included in the present report. We have included only assessments for which the initial development is complete, that are published, and that are offered today for sale or use at either two-year or four-year colleges. In addition, we included only those assessments that have been explicitly and publicly proposed by their publishers for use in higher-education accountability applications and that have been used by a sufficient number of institutions to provide a sound basis for evaluating the technical characteristics of the test and for making appropriate score comparisons. This latter characteristic is important because one of the issues that has attracted considerable attention in the recent discussions of student learning outcomes is the importance of having a large enough number of users of the assessments to enable comparisons with other institutions.

The intended use of the assessment tool was another of our selection criteria. Some very well-known, technically sound, and widely-used instruments are not included in this guide because they were not designed for, or validated for, use as a measure of student learning outcomes. For example, neither the *Graduate Record Examinations*® (GRE®) General test nor the GRE Subject tests have been validated for the assessment of student learning outcomes, despite having occasionally been mentioned as possible tools for assessing student learning. As a consequence of the lack of validation data to support their use in assessing student learning outcomes, we do not include the GRE measures in this report, even though these measures do provide valuable information regarding students' knowledge, skills, and abilities. Similarly, insofar as we can determine, the publishers of the Law School Admissions Test (LSAT®), the Graduate Management Admissions Test® (GMAT®), the Medical College Admission Test® (MCAT®), and the SAT® tests have not conducted validation research that indicates that these admissions tests are appropriate for use as measures of student learning outcomes for postsecondary accountability purposes. The publishers of these tests also have not explicitly proposed their tests for such use.

Coverage of assessments for professional licensing or specific academic disciplines were also judged to be beyond the scope of this guide, although many institutions might benefit from a careful consideration of whether a set of specific discipline-based assessments might meet some of their institutional accountability needs. It is unfortunately the case, however, that many academic discipline associations have not yet articulated postsecondary student learning outcomes for their area of study; thus, this logical touchstone remains unavailable to postsecondary educators. There are also aspects of professional licensing assessments that currently limit their utility for higher education accountability purposes (e.g., the timing of assessments and the nature of institutional reports).

Information Provided About Each Assessment

Based on the criteria described above, we selected 12 assessments for this guide (see Table 1). Tables 2 through 13 provide a concise summary of key features of each of the assessments in a standard format. We have attempted, to the extent possible, to provide information directly from the assessments' publishers. All of the information in these tables was gathered from (a) the website of the assessment publisher, (b) personal communication with publishers' representatives, or (c) published reviews of the assessments (e.g., National Center for Education Statistics, 2000) that were subsequently evaluated by publishers' representatives. In all cases, the publishers' interpretations of their data prevailed.

Tables 2 through 13 include basic identifiers and descriptions of each assessment's purpose and uses, as well as the following information:

- *Intended population.* The type of institution that students attend (technical, community, or four-year college) and the academic level of students who take the assessments (freshman, sophomore, upperclassman, or graduating senior).
- *Items and forms.* What is the format of the questions or tasks? Are the questions or tasks in the form of Likert scales, multiple choice, essays, or other types of responses actively constructed by the student? Does the assessment provide more than one version, either as alternate intact forms, or through the use of an item bank? In what format is the test available (e.g., paper- or web-based)? Are there provisions for including locally-developed items, as well as those provided by the publisher?

- *Level of results.* Does the assessment yield information about individual students, cohorts or subgroups of students, or only about the institution (or program) as a whole?
- *Scores yielded.* Is the assessment conceived as a measure of student performance relative to a specific performance criterion (criterion-referenced); or is there information that would allow a user to compare students over time, between groups, or across institutions (norm-referenced)?
- *Comparative data availability.* Are there national comparative data available? Is there an option for the institution to select specific institutions for more targeted comparisons?
- *Cost.* There are many different models for how assessments are priced. We report pricing information as provided by the assessments' publishers.
- *Testing sample.* Does the assessment design require all students to take the assessment, or can students be sampled?
- *Pre-and post-testing.* If the assessment can appropriately be used for pre- and post-testing, what research designs can be used?
- *Time required.* How much time is needed or recommended for a student to complete the assessment at a pace sufficient to allow him or her to demonstrate learning in a valid manner?
- *Annual volume or institutional data pool of test takers.* Some publishers report the number of institutions that completed the assessment on an annual basis; others report the cumulative number of institutions that have utilized the assessment (in some cases over several years), and are therefore available for comparative purposes. These two measures provide information about how widely used the assessment is, and also about how many institutions are represented in the comparative pool. It is not necessarily the case that one method of using comparative data is intrinsically superior to the other. Data reported on an annual basis are by definition "fresher" than cumulative data, but cumulative data provide a larger base of comparison, which could be equally valid depending on the time frame covered by the data set, and the changes in student populations and institutional usage that may have occurred since data collection began.

To facilitate comparison across appropriate assessment tools, we have also summarized the major areas constituting each of the 12 assessments (see Table 14). The assessments are grouped by their four salient dimensions of student learning.

Conclusion

We hope that the information presented in this report will prove useful as the higher education community considers various approaches for measuring student learning in light of the accountability movement. Returning to the three questions posed to guide the development of this report:

1. We have identified 12 of the most prevalent assessments of postsecondary student learning that are currently commercially available.
2. We have provided information for each of the 12 assessments that delineates critical features, such as intended population, format, scoring paradigm, and comparative data. Our aim in providing this information is to enable postsecondary institutions to begin a critical review of which assessments possess those features that will best serve their unique accountability needs.
3. We also have touched on other tests that have served as proxies for student learning but have not been validated for use as higher education outcome measures, as well as a few assessments of noncognitive student outcomes that are currently in the late stages of test development.

We encourage readers to keep in mind as they consider the utility of the assessments reviewed here, as well as other assessment options, that the appropriate place to start a discussion about assessing student learning is — surprisingly to some — not with the selection of an assessment tool. Rather, the best place to start is by asking questions such these:

- What kinds of statements would we like to be able to make about students' learning?
- What evidence of student learning do we already have (e.g., portfolios), and what conclusions can be drawn from these data?
- What inferences about student learning can we draw from existing evidence, and how can we support and supplement these inferences with data from new assessments?

Although accountability to the higher education community and, ultimately, to the students and their families that comprise the "interested public" is an important consideration for every two- and four-year college and university, selecting the assessment best suited for accountability purposes must also be done in accordance with the specific student outcomes for which schools are being held accountable. For example, colleges and universities that focus on liberal arts education may be more interested in measures of general education skills and workforce readiness than in subject-area knowledge and skills; those that pride themselves on particularly intense discipline-specific preparation may emphasize measures of subject-area knowledge and skills in their accountability programs. Additionally, some institutions, such as community colleges, which respond to community needs for advanced educational opportunities for non-traditional students, may opt for assessments that will provide information about changes in student knowledge and skills over time rather than absolute achievement compared with other institutions.

Although our review has been limited to commercially-available measures of higher education outcomes for which comparative data have been compiled, we do not espouse the view that these assessments should be the sole measure of the value that an institution adds to its students. As institutions of higher education, colleges and universities must address academic outcomes in an accountability program; however, qualities such as satisfaction, ethical values, civic engagement and/or citizenship preparation, or many other areas not covered in this report, may also be outcomes that institutions seek to foster in their students. In such cases, institutions may wish to evaluate these outcomes as well, in order to develop a well-rounded picture of their students and graduates.

In this report, we have provided information at the "30,000-foot level." We hope that institutions that are evaluating their programs for assessing student learning outcomes, as well as institutions that are just beginning to put comprehensive systems in place, will find this report a useful starting point for their conversations. Any institution, policymaker, or legislator considering inaugurating a program of accountability should consider the critical features of the assessments described in Tables 2 through 13 in light of articulated student learning outcomes. Each college or university's unique student learning goals, institutional mission, and student composition, as well as external accountability requirements, will necessarily influence choices concerning the assessment features most critical to ensuring the success of their accountability efforts. The next report in the *COE* series delves more deeply into these and other issues as we construct a conceptual framework and set of heuristics that institutions can use to structure their assessments of student learning.

Table 1. Overview of Assessments

Assessment	Workforce Readiness/General Educational Skills	Domain-Specific Knowledge	Soft Skills	Student Engagement
Area Concentration Achievement Tests (ACAT®)		☑		
College Basic Academic Subjects Examination (*College BASE*)	☑			
Collegiate Assessment of Academic Proficiency™ (CAAP)	☑			
Collegiate Learning Assessment (CLA)	☑			
Community College Survey of Student Engagement (CCSSE)				☑
iSkills™ (formerly ICT Literacy)	☑			
Measure of Academic Proficiency and Progress™ (MAPP™)	☑			
Major Field Tests (MFTs)		☑		
National Survey of Student Engagement (NSSE)				☑
Standardized Assessment of Information Literacy Skills (SAILS)	☑			
WorkKeys®	☑			
WorkKeys® (*Teamwork*)			☑	

Table 2. College Basic Academic Subjects Examination *(College BASE)*

College BASE, developed by the Assessment Resource Center at the University of Missouri-Columbia in 1989, measures students' proficiency in core academic subject areas. *College BASE* is used to assess basic academic knowledge and reasoning skills, the success of academic programs, and student readiness for teacher preparation programs.

For more information, visit *College BASE's* website: http://arc.missouri.edu/collegebase/

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice (M/C)[1] ☒ Essay[2] ☐ Other constructed response ☐ Performance tasks ☒ Multiple versions[3] ☐ Item banks ☒ Paper-based format ☐ Web-based format ☐ Optional inclusion of locally-developed items	☒ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☒ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[4]	☒ National comparative data ☒ Self-selected reference groups available **Local Data** ☒ Scaled scores ☒ Percentile ranks	$7-$10.90/student one to four M/C tests $12/student for optional essay $2.50/student for materials **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	45 minutes per M/C test 40 minutes for essay			N = 135 institutions

[1] *College BASE* contains 180 M/C items in four subject areas: English (41 items), Mathematics (56 items), Science (41 items), and Social Studies (42 items).
[2] An optional essay contains one prompt.
[3] Available formats include the standard subjects option, the *College BASE* 1:1 (short form) option, and the basic skills option. The 1:1 option provides institutional summary data for all four subject areas, but assigns only 1-2 tests per student; the basic skills option assesses basic skills in English, writing, and mathematics.
[4] In addition to total scaled scores, *College BASE* reports scaled scores by subject and by skill cluster within each subject, as well as ratings of high, medium, or low for proficiency in each skill area.

Table 3. Collegiate Assessment of Academic Proficiency™ (CAAP)

CAAP, first developed by ACT, Inc., in 1990, measures students' general education outcomes in critical thinking, science, mathematics, reading, and writing. It was designed to evaluate college-level learning outcomes and institutions' general education programs as a way to inform educational interventions and curricular enhancement.

For more information, visit CAAP's website: http://www.act.org/caap

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice (M/C)[5] ☒ Essay ☐ Other constructed response ☐ Performance tasks ☒ Multiple versions ☐ Item banks ☒ Paper-based format ☐ Web-based format ☒ Optional inclusion of locally-developed items	☒ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[6]	☒ National comparative data by class level and type of institution ☒ Self-selected reference groups available **Local Data** ☒ Scaled scores ☒ Percentile ranks	$12/student for each M/C test $18.55/student two to five M/C tests $12/student for Writing Essay $360 annual institutional participation fee **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)[7]	40 minutes per module			N = 385 institutions

[5] CAAP consists of six test modules that can be used independently or in conjunction: Critical Thinking (32 M/C items), Mathematics (35 M/C items), Reading (36 M/C items), Science (45 M/C items), Writing Skills (72 M/C items), and a Writing Essay, which consists of two, 20-minute essays.
[6] In addition to total scaled scores, CAAP provides subscores in Writing Skills (2 subscores), Reading (2 subscores), and Mathematics (6 subscores), as well as performance evaluation in each content area.
[7] CAAP scores can be linked to students' ACT® or COMPASS scores to determine an institution's value-added to student learning.

Table 4. Collegiate Learning Assessment (CLA)

CLA, developed by the Council for Aid to Education with the RAND Corporation in 2004, requires students to analyze complex, ambiguous materials and to construct responses that demonstrate their abilities in critical thinking, analytical reasoning, and written communication. The CLA is designed to measure institutions' contribution, or value-added, to students' competency development and the effects of changes to curriculum and teaching methods.

For more information, visit CLA's website: http://www.cae.org/content/pro_collegiate.htm

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☐ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating senior	☐ Likert scales ☐ Multiple choice (M/C) ☒ Essay[8] ☐ Other constructed response ☒ Performance tasks ☐ Multiple versions ☒ Item banks ☐ Paper-based format ☒ Web-based format ☐ Optional inclusion of locally-developed items	☒ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[9]	☒ National comparative data ☒ Self-selected reference groups available **Local Data** ☒ Scaled scores ☐ Percentile ranks	$6,300 for cross-sectional sample of 100 freshmen and 100 senior students $28,000 for 4-year longitudinal study **Additional services available** ☒ Specialized score reports ☒ Research ☒ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☐ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)[7]	90 minutes for Performance Task 75 minutes for 2 Writing Prompts			N = 331 institutions

[8] The CLA includes two writing prompts: the make-an-argument prompt and the break-an-argument prompt.
[9] CLA scores can be expressed as value-added scores. These scores are the difference between the actual mean CLA score at a school and the mean the participating students are predicted to earn given their average SAT® or ACT® scores. The CLA program computes the difference in value-added scores between freshmen and seniors at a school to create a measure of that institution's contribution to student learning.

Table 5. *iSkills*™ (formerly ICT Literacy Assessment)

The *iSkills* assessment, developed by Educational Testing Service in 2006, measures cognitive and technical skills related to information and communication technology (ICT) proficiency. Designed to align with the Association of College & Research Libraries' (ACRL) Information Literacy Competency Standards, the *iSkills* assessment can be used to evaluate individual students' ICT proficiency and to inform curriculum planning and improvement.

For more information, visit *iSkill's* website: http://www.ets.org/iskills

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☐ Graduating senior	☐ Likert scales ☐ Multiple choice (M/C) ☐ Essay ☐ Other constructed response ☒ Performance tasks[10] ☒ Multiple versions[11] ☐ Item banks ☐ Paper-based format ☒ Web-based format ☒ Optional inclusion of locally-developed items[12]	☒ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☐ Criterion-referenced, proficiency scores[13] ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions	☒ National comparative data ☐ Self-selected reference groups available **Local Data** ☒ Scaled scores ☐ Percentile ranks	$22-$33 per student **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling
Testing Sample	**Pre- & Post-testing**	**Time Required**			**Institutional Data Pool**
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	75 minutes			N = 90 institutions

[10] The *iSkills* assessment evaluates seven information and communication technology proficiencies via 15 real-time, scenario-based activities: Define, Access, Manage, Integrate, Evaluate, Create, and Communicate information.
[11] The Core Academic Assessment is intended for students early in college, while the Advanced Assessment is appropriate for upper-level students.
[12] Institutions can add up to nine locally-developed student background questions.
[13] Students receive qualitative performance feedback by ICT proficiency and task, as well as total scaled scores and local scaled scores by proficiency component.

Table 6. *Measure of Academic Proficiency and Progress*™ (MAPP™)

The MAPP test, formerly Academic Profile, developed by Educational Testing Service in 1987 and revised in 2006, is an integrated, general education outcomes assessment. The MAPP test measures college-level skills in critical thinking, reading, writing, and mathematics and provides institutional and comparative data that can be used for planning, accreditation, and program review and improvement.

For more information, visit the MAPP test's website: http://www.ets.org/mapp

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating seniors	☐ Likert scales ☒ Multiple choice (M/C)[14] ☒ Essay ☐ Other constructed response ☐ Performance tasks ☒ Multiple versions ☐ Item banks ☒ Paper-based format ☒ Web-based format ☒ Optional inclusion of locally-developed items	☒ Individual student (Standard Forms only) ☒ Cohorts or subgroups (Standard or Abbreviated Forms) ☒ Institutional aggregate	☒ Criterion-referenced, proficiency scores[15] ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[16]	☒ National comparative data ☒ Self-selected reference groups option **Local Data** ☒ Scaled scores ☒ Percentile ranks	$15.50/student (Standard Form) $13.50/student (Abbreviated Form) $5/student for optional essay **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	2 hours (Standard Forms) 40 minutes (Abbreviated Forms)			N = 337 institutions

[14] The MAPP test is available in two formats: the Standard Form, which contains 108 M/C items, and the Abbreviated Form, which includes 36 items. Only the Standard Form is sufficiently reliable to yield individual student scores.
[15] The MAPP test yields proficiency classifications in writing, mathematics, and reading/critical thinking: Proficient, Marginally Proficient, or Not Proficient.
[16] Subscores: Humanities, Social Sciences, Natural Sciences, College-Level Reading, College-Level Writing, Critical Thinking, and Mathematics.

Table 7. Standardized Assessment of Information Literacy Skills (SAILS)

SAILS, developed by Project SAILS at Kent State University in 2005, assesses the degree to which college students' information literacy skills meet the Association of College & Research Libraries' (ACRL) Information Literacy Competency Standards. SAILS can be used to evaluate students' literacy skills and to identify areas for improvement.

For more information, visit SAILS's website: http://www.projectsails.org/

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice (M/C)[17] ☐ Essay ☐ Other constructed response ☐ Performance tasks ☒ Multiple versions ☒ Item banks ☒ Paper-based format ☒ Web-based format ☐ Optional inclusion of locally-developed items	☐ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[18]	☒ National comparative data by class level, major, and institution type ☒ Self-selected reference groups available **Local Data** ☒ Scaled scores ☐ Percentile ranks	$3/student Cap of $2,000 per administration $0.50/answer sheet for paper-based test **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	35 minutes			N = 71 institutions

[17] SAILS consists of 45 M/C items aligned to eight skill sets: developing a research strategy, selecting finding tools, searching, using finding tool features, retrieving sources, evaluating sources, documenting sources, and understanding economic, legal, and social issues.
[18] Scaled scores are reported for each information literacy skill set and for four of the five ACRL standards.

Table 8. WorkKeys®

WorkKeys, developed by ACT, Inc., in 1992, is a series of assessments used by educational institutions and employers to assess workplace-readiness and necessary job-related skills. The assessments reviewed here measure a student's proficiency in applying foundational skills to employment situations.

For more information, visit WorkKeys' website: http://www.act.org/workkeys

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice (M/C)[19] ☒ Essay ☒ Other constructed response[20] ☒ Performance tasks ☐ Multiple versions[21] ☐ Item banks ☒ Paper-based format ☒ Web-based format (some assessments) ☒ Optional inclusion of locally-developed items	☒ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☒ Criterion-referenced, proficiency scores[22] ☐ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions	☐ National comparative data ☐ Self-selected reference groups available **Local Data** ☒ Scaled scores ☐ Percentile ranks	Contact ACT **Additional services available** ☒ Specialized score reports ☐ Research ☐ Over-sampling
Testing Sample	**Pre- & Post-testing**	**Time Required**			**Institutional Data Pool**
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	30 minutes to 1 hour per test			N = over 3,000 educational institutions

[19] A WorkKeys assessment may include one or more of the following tests: Reading for Information (33 M/C items), Applied Mathematics (33 M/C items), Locating Information (38 M/C items), Observation (36 M/C items), Applied Technology (32 M/C items), Listening & Writing, Business Writing, and Teamwork (see Table 10).

[20] For the Listening and Writing test, examinees listen to six audio-taped messages and then write responses that convey the information to a potential coworker.

[21] Some WorkKeys assessments are available in Spanish.

[22] Scores on each test are matched to job profiles developed by content experts. These profiles reflect the skill levels required on one or more WorkKeys test to perform any of over 14,000 profiled jobs. If a student or job applicant meets or exceeds the Level Scores in the skills required for a job, he or she has demonstrated entry-level competency for that position in those skills.

Table 9. Area Concentration Achievement Tests (ACAT®)

ACAT, developed in 1983 and administered by PACAT, Inc./NOMESys, consists of subject-area outcome tests that provide curriculum-specific feedback. The ACATs can be used to evaluate how well students have mastered the content of their disciplinary majors and how well programs and curricula foster student achievement in these areas.

For more information, visit ACAT's website: http://www.collegeoutcomes.com/

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☐ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☐ Freshman ☐ Sophomore ☐ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice (M/C)[23] ☐ Essay ☐ Other constructed response ☐ Performance tasks ☒ Multiple versions ☒ Item banks ☒ Paper-based format ☒ Web-based format (some assessments) ☐ Optional inclusion of locally-developed items[24]	☒ Individual student ☒ Cohorts or subgroups[25] ☐ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions	☒ National comparative data ☐ Self-selected reference groups available **Local Data** ☒ Scaled scores ☒ Percentile ranks	$7.50-$17/student graduating seniors $5.90-$6.50/student two-year college students $6-$10.80/student for pretests **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☒ All students	☒ Different students[26] (cross-sectional study) ☒ Same students[27] (longitudinal study)	48 minutes to 2 hours (time varies by major area and curriculum/ content focus)			N = 550 departments (across 11 test areas)

[23] ACAT assessments are currently available in the following disciplines: Agriculture, Art, Biology, Criminal Justice, General Business, Geology, History, English Literature, Political Science, Psychology, and Social Work. Several forms that vary by length and content focus are available for most of these assessments.
[24] Items for ACAT assessments are written by faculty at participating institutions.
[25] Because these assessments are designed for graduating seniors in specific fields of study, only cohorts/subgroups of students take an ACAT in a particular area. That is, the cohort/subgroup score is the institutional aggregate for that test. ACATs in different areas are not equivalent and cannot be combined.
[26] As a measure of the success of academic programs and curricula, the ACAT provides longitudinal comparisons between each year's graduating cohorts.
[27] Academic departments and programs can use pre-tests available in several content areas to examine changes in individual students' subject-area knowledge.

Table 10. Major Field Tests (MFTs)

Major Field Tests were developed by Educational Testing Service in 1989. This series of tests is designed to measure academic achievement at the conclusion of major study. The MFTs assess mastery of concepts, principles, and knowledge, as well as abilities to analyze information, solve problems, and interpret materials. Results are used to assess and refine curricula and to compare student progress among similar programs across the campus or in other schools.

For more information, visit the MFTs' website: http://www.ets.org/mft

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☐ Technical schools ☐ Community colleges ☒ 4-year colleges and universities **Student Standing** ☐ Freshman ☐ Sophomore ☐ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice (M/C)[28] ☐ Essay ☐ Other constructed response ☐ Performance tasks ☐ Multiple versions ☐ Item banks ☒ Paper-based format ☒ Web-based format ☒ Optional inclusion of locally-developed items	☒ Individual student ☒ Cohorts or subgroups[29] ☐ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions	☒ National comparative data ☒ Self-selected reference groups available **Local Data** ☒ Scaled scores ☒ Percentile ranks	$24-$26/student for Bachelor's program graduates $25-$30/student for MBA program graduates **Additional services available** ☒ Specialized score reports ☒ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Institutional Data Pool
☒ Sample of students ☒ All students	☒ Different students[30] (cross-sectional study) ☐ Same students (longitudinal study)	2 hours for undergraduate tests 3 hours for MBA test			N = 2,277 departments (across 16 test titles)

[28] Major Field Tests are currently available in the following disciplines: Biology, Business, Chemistry, Computer Science, Criminal Justice, Economics, Education, History, English Literature, Mathematics, Music, Physics, Political Science, Psychology, Sociology and MBA.

[29] Because these assessments are designed for graduating seniors in specific fields of study, only cohorts/subgroups of students take a MFT in a particular area. That is, the cohort/subgroup score is the institutional aggregate for that test. MFTs for different fields are not equivalent and cannot be combined.

[30] MFTs provide longitudinal comparisons between each year's graduating cohorts to allow academic departments to gauge program effectiveness over time.

Table 11. WorkKeys® — *Teamwork*

WorkKeys' *Teamwork* test, developed by ACT, Inc., in 1995, requires students and job candidates to select the most appropriate responses to scenarios that depict teams in workplace settings. The *Teamwork* test involves choosing behaviors that simultaneously support relationships within the team and facilitate the accomplishment of work tasks.

For more information, visit WorkKeys' website: http://www.act.org/workkeys

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☒ Upperclassman ☒ Graduating senior	☐ Likert scales ☒ Multiple choice[31] ☐ Essay ☐ Other constructed response ☐ Performance tasks ☐ Multiple versions ☐ Item banks ☒ Paper-based format ☐ Web-based format ☒ Optional inclusion of locally-developed items	☒ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☒ Criterion-referenced, proficiency scores[32] ☐ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions	☐ National comparative data ☐ Self-selected reference groups available **Local Data** ☒ Scaled scores ☐ Local percentile ranks available	Contact ACT **Additional services available** ☒ Specialized score reports ☐ Research ☐ Over-sampling

Testing Sample	Pre- & Post-testing	Time Required			Annual Volume
☒ Sample of students ☒ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	64 minutes			N = over 3,000 educational institutions

[31] WorkKeys *Teamwork* includes 36 M/C items presented in a video format.

[32] Scores on the *Teamwork* assessment are matched to job profiles developed by content experts. These profiles reflect the skill levels required for over 14,000 profiled jobs. If a student or job applicant meets or exceeds the Teamwork Level Score required for a job, he or she has demonstrated entry-level competency in teamwork for that position.

Table 12. Community College Survey of Student Engagement (CCSSE)

CCSSE, developed by the Community College Leadership Program at the University of Texas at Austin in 2001, measures community and technical college students' engagement in their educational experiences via items assessing institutional practices and student behaviors. CCSSE can be used to expand educational experiences that have been associated with institutional effectiveness and student learning, persistence, and attainment.

For more information, visit CCSSE's website: http://www.ccsse.org/

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☒ Technical schools ☒ Community colleges ☐ 4-year colleges and universities **Student Standing** ☒ Freshman ☒ Sophomore ☐ Upperclassman ☐ Graduating senior	☒ Likert scales ☐ Multiple choice (M/C) ☐ Essay ☐ Other constructed response ☐ Performance tasks ☐ Multiple versions ☐ Item banks ☒ Paper-based format ☐ Web-based format ☒ Optional inclusion of locally-developed items	☐ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[33]	☒ National comparative data ☒ Self-selected reference groups available **Local Data** ☐ Scaled scores ☒ Percentile ranks	$1,500 to $11,500 sample size based on enrollment [34] **Additional services available** ☒ Specialized score reports ☒ Research ☒ Over-sampling
Testing Sample	**Pre- & Post-testing**	**Time Required**			**Institutional Data Pool**
☒ Sample of students[35] ☐ All students	☒ Different students (cross-sectional study) ☐ Same students (longitudinal study)	About 30 minutes			N = 550 institutions

[33] As CCSSE measures perceptions and self-reported behaviors rather than knowledge or skills, it is not scored in a traditional sense. Instead, frequencies, means, and standard deviations are reported for items and benchmarks. Benchmarks consist of sets of items addressing student behaviors and institutional practices that empirical research has identified as contributors to student learning, persistence, and personal development; the five benchmarks of CCSSE are Academic Challenge, Active and Collaborative Learning, Student-Faculty Interaction, Support for Learners, and Student Effort.

[34] CCSSE uses an institution's enrollment to determine how many students must be surveyed to yield reliable benchmark and subgroup reports.

[35] CCSSE uses samples of entire classes rather than individual students.

Table 13. National Survey of Student Engagement (NSSE)

NSSE, developed in 1999 and administered by the Indiana University Bloomington, measures four-year college students' engagement in their educational experiences via their participation in activities related to learning and personal development. NSSE can be used to evaluate and enhance educational practices that are associated with institutional effectiveness and student academic attainment.

For more information, visit NSSE's website: http://www.nsse.iub.edu/index.cfm

Intended Population	Items and Forms	Level of Results	Scores Yielded	Comparative Data Availability	Cost
Institution Type ☐ Technical schools ☐ Community colleges ☒ 4-year colleges and universities **Student Standing** ☒ Freshman ☐ Sophomore ☐ Upperclassman ☒ Graduating senior	☒ Likert scales ☐ Multiple choice (M/C) ☐ Essay ☐ Other constructed response ☐ Performance tasks ☐ Multiple versions ☐ Item banks ☒ Paper-based format ☒ Web-based format (some assessments) ☒ Optional inclusion of locally-developed items	☐ Individual student ☒ Cohorts or subgroups ☒ Institutional aggregate	☐ Criterion-referenced, proficiency scores ☒ Norm-referenced, scaled scores that compare students over time, between groups, or across institutions[36]	☒ National comparative data ☒ Self-selected reference groups available **Local Data** ☐ Scaled scores ☒ Percentile ranks	$3,375 to $7,500 sample size based on undergraduate enrollment [37] $300 annual institutional participation fee **Additional services available** ☒ Specialized score reports ☒ Research ☒ Over-sampling
Testing Sample	**Pre- & Post-testing**	**Time Required**			**Institutional Data Pool**
☒ Sample of students ☐ All students	☒ Different students (cross-sectional study) ☒ Same students (longitudinal study)	About 15 minutes			N = 600 institutions

[36] As NSSE measures perceptions and self-reported behaviors rather than knowledge or skills, it is not scored in a traditional sense. Instead, frequencies, means, and standard deviations are reported for items and benchmarks. Benchmarks are types of student behaviors and institutional practices that have been identified by the empirical research literature as contributors to student learning and personal development; NSSE's five benchmarks are level of Academic Challenge, Active and Collaborative Learning, Student Interactions with Faculty Members, Enriching Educational Experiences, and Supportive Campus Environment.

[37] NSSE uses institutional enrollment to determine how many freshman and senior students must be surveyed to yield reliable benchmark and subgroup reports.

Table 14. Summary of Major Content Areas Measured by Featured Assessments

General Education and Workforce Readiness

Assessment	Content Areas
College BASE	English, Mathematics, Science, Social Studies, and Writing
Collegiate Assessment of Academic Proficiency (CAAP)	Critical Thinking, Mathematics, Reading, Science, and Writing
Collegiate Learning Assessment (CLA)	Critical Thinking, Analytical Reasoning, and Writing
iSkills	Ability to Define, Access, Manage, Integrate, Evaluate, Create, and Communicate information
Measure of Academic Proficiency and Progress (MAPP)	Critical Thinking, Mathematics, Reading, and Writing
Standardized Assessment of Information Literacy Skills (SAILS)	Developing a Research Strategy, Selecting Finding Tools, Searching, Using Finding Tool Features, Retrieving Sources, Evaluating Sources, Documenting Sources, and Understanding Economic, Legal, and Social Issues
WorkKeys	Reading for Information, Applied Mathematics, Locating Information, Observation, Applied Technology, Listening & Writing, Business Writing, and Teamwork

Subject-Area Knowledge and Skills

Assessment	Content Areas
Area Concentration Achievement Tests (ACAT)	Agriculture, Art, Biology, Criminal Justice, General Business, Geology, History, English Literature, Political Science, Psychology, and Social Work
Major Field Tests (MFTs)	Biology, Business, Chemistry, Computer Science, Criminal Justice, Economics, Education, History, English Literature, Mathematics, Music, Physics, Political Science, Psychology, Sociology, and MBA

Soft Skills

Assessment	Content Areas
WorkKeys (*Teamwork*)	Ability to work collaboratively to accomplish tasks and to foster positive professional relationships

Student Engagement

Assessment	Content Areas
Community College Survey of Student Engagement (CCSSE)	Academic Challenge, Active and Collaborative Learning, Student-Faculty Interaction, Support for Learners, and Student Effort
National Survey of Student Engagement (NSSE)	Academic Challenge, Active and Collaborative Learning, Student Interactions with Faculty Members, Supportive Campus Environment, and Enriching Educational Experiences

References

Dwyer, C. A., Millett, C. M., & Payne, D. G. (2006). *A culture of evidence: Postsecondary assessment and learning outcomes*. Princeton, NJ: ETS.

National Center for Education Statistics. (2000). *The NPEC sourcebook on assessment, volume 1: Definitions and assessment methods for critical thinking, problem solving, and writing*. NCES 2000-172, prepared by T. Dary Erwin for the Council of the National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development. Washington, DC: U.S. Government Printing Office.

Spellings, M. (2006). *Action plan for higher education: Improving accessibility, affordability, and accountability.* Washington, DC: U.S. Department of Education.

U.S. Department of Education. (2006). *A test of leadership: Charting the future of higher education*. Washington, DC.

ETS®
Rosedale Road
MS 30-N
Princeton, NJ 08541

PRESORTED
FIRST CLASS MAIL
US POSTAGE
PAID
EDUCATIONAL
TESTING SERVICE