Measuring Up on College-Level Learning

By
Margaret A. Miller
Peter T. Ewell

October 2005

The National Center for Public Policy and Higher Education

National Center Report #05-8

Contents

Foreword

Measuring Up 2000, the first state-by-state report card on higher education performance, gave all 50 states an Incomplete in the category of learning. Although *Measuring Up* evaluated, compared, and graded the states in other key categories of higher education performance (including preparation for college, participation, completion, and affordability), the report card found that "there is no information available to make state-by-state comparisons" of higher education's most important outcome, learning. The primary purpose of the Incomplete was to stimulate a more robust discussion and debate about what states should know about college-level learning.

Shortly after the release of *Measuring Up 2000*, an invitational forum of public policy, business, and education leaders was convened by James B. Hunt Jr., governor of North Carolina, and hosted by Roger Enrico, vice chairman of PepsiCo, at the PepsiCo corporate headquarters in Purchase, New York. The purpose of the forum was to advise the National Center on next steps to address the issue of student learning at the state level. The forum recommended that the National Center begin by using information already available on college outcomes as the building blocks of a model to collect comparative state-by-state information on learning. Forum participants urged the National Center to move ahead with a "demonstration project" to determine whether or not it was feasible to collect information on learning at the state level that would be useful to state policy leaders.

The National Center was fortunate to enlist the help of Margaret Miller, professor at the Curry School of Education, to lead the National Forum on College-Level Learning, a five-state demonstration project to develop a model of college-level learning for the states. Peter Ewell, vice president of the National Center for Higher Education Management Systems (NCHEMS), was the senior consultant to the project. The Pew Charitable Trusts supported the project through a grant to the Institute for Educational Leadership (IEL).

The most recent edition of the report card, *Measuring Up 2004*, included a brief summary of the results of the demonstration project. This report provides a more comprehensive account of the project, its findings, and conclusions, as well as information that will be useful to states that may wish to replicate

the model. The report concludes that providing comparative state-by-state information about learning outcomes is not only feasible, but also important and useful for policy.

The model described in this report enables states to gather information that addresses two critical questions:

1. What is the "educational capital," or the knowledge and skills of the population, that states have available to them for developing or sustaining a competitive economy and vital civic life?
2. How do all the colleges and universities in the state (that is, public, private, not-for-profit, and for-profit) contribute to the development of the state's educational capital?

This approach is different from asking or requiring individual colleges and universities to assess or evaluate student learning. Colleges and universities can and should be accountable for assessing student learning and reporting results, but the measures used by individual institutions may not add up to a comprehensive assessment of educational capital for the state as a whole. The statewide approach, as shown by the demonstration project, allows comparisons among states, providing information about a state's relative standing to the rest of the nation in developing the knowledge and skills of its population.

In a knowledge-based global economy, the fortunes of states depend on the knowledge and skills of their residents. The demonstration project has shown that states can assess their educational capital feasibly and effectively to provide useful information for policymakers and educators in identifying problems and stimulating and targeting improvement. State leaders are urged to participate in similar efforts to expand their state's understanding of the knowledge and skills of their residents in order to enhance the economic and civic vitality of their state.

Patrick M. Callan

President

National Center for Public Policy and Higher Education

Introduction

At the symposium of policy leaders marking the release of *Measuring Up 2000*,[1] which was the first 50-state report card on higher education, one of the most dramatic moments was the unveiling of a U.S. map representing each state's performance in learning—the sixth and final graded category in the report card. In contrast to the brightly colored patchworks portraying grades for each of the states in the other five categories, the learning map was a uniform gray (see figure 1). A large question mark superimposed upon it represented the Incomplete that all states had earned in that category. The conversation among those at the symposium ended without a satisfactory answer to the sharply posed question: "Why can't we grade the states on learning, if that is the most important result colleges and universities produce?"

Figure 1

In *Measuring Up 2000* and *2002*, all states received an Incomplete in learning.

At one level, institutions and states actually know a good deal about what their college students know and can do. Apart from the many times students' work is evaluated in class, every institution must determine its success in educating students in order to meet the requirements of regional accreditors.

---

[1]*Measuring Up 2000: The State-by-State Report Card for Higher Education* (San Jose: National Center for Public Policy and Higher Education, 2000). Subsequent editions of *Measuring Up* were published in 2002 and 2004, and the next edition is planned for 2006.

Participants
National Forum on College-Level Learning
November 27–28, 2001
Purchase, New York

The Honorable Garrey Carruthers
President and Chief Executive Officer
Cimarron HMO, Inc.

Gordon K. Davies
President
Kentucky Council on Postsecondary Education

Thomas Ehrlich
Senior Scholar
Carnegie Foundation for the Advancement of Teaching

Roger A. Enrico
Vice Chairman
PepsiCo., Inc.

The Honorable Jim Geringer
Governor of Wyoming

Milton Goldberg
Executive Vice President
National Alliance of Business

The Honorable James B. Hunt Jr.
Womble, Carlyle, Sandridge & Rice

Glenn R. Jones
President and Chief Executive Officer
Jones International, Ltd.

Ann Kirschner
President and Chief Executive Officer
Fathom

The Honorable John R. McKernan, Jr.
Vice Chairman
Education Management Corporation

Charles Miller
Chairman
Meridian National, Inc.

Moreover, most states have some kind of statewide assessment requirement in place to improve performance, to give state officials a sense of what their investment in higher education has yielded, or both. But unlike the information collected in the other categories of *Measuring Up*, there are no comprehensive national data on college-level learning that could be used to compare state performance in this area.

The information states do have on collegiate learning is incomplete for their own purposes as well. When every campus within a state assesses its students' learning differently, the state has no effective method for interpreting the resulting information because there are no external benchmarks against which to measure a given program's or institution's performance. Even those states that employ common measures statewide for public colleges and universities know virtually nothing about the learning results of their private institutions. Nor do they know how the learning of their college-educated residents or current college attendees compares to the learning of those in other states.

Subsequent to the release of *Measuring Up 2000*, the National Center's Board of Directors considered eliminating the learning category. The board concluded, however, that the category—and the idea behind it—was too important to abandon. Subsequently, The Pew Charitable Trusts decided to sponsor an investigation into how to generate grades in that category. As a result of that decision, the National Forum on College-Level Learning was born.

The National Forum, Phase One

The National Forum on College-Level Learning began with interviews of higher education and policy leaders around the country, during which three questions were posed:

1. Should the National Forum attempt to assess student learning in comparable ways at the state level?
2. If so, what questions should be answered by whatever information the National Forum collects?
3. How should the National Forum go about collecting the information?

In November 2001, a group of higher education, policy, and business leaders considered the same set of questions at a meeting in Purchase, New York (see sidebar for list of participants). Their answers echoed those of the leaders interviewed earlier:

- **Should the National Forum attempt to assess student learning in comparable ways at the state level?** The answer to this question was a resounding "yes." Meeting participants observed that national pressures to assess collegiate learning, dating back to before the congressional ratification of the National Education Goals in 1994, were not dissipating. In fact, they were increasing. Moreover, it was "outrageous," as one participant put it, not to know more about higher education's most important product. Finally, without information about learning results, *Measuring Up*—as a state-by-state report card on higher education—would always present an incomplete picture of the success of higher education policy in the states.
- **What questions should be answered by whatever information the National Forum collects?** Participants formulated two state policy questions that any information gathered about learning should answer:
    1. "What do the state's college-educated residents know and what can they do that contributes to the social good?" This question became known as the "educational capital" question, because it sought to measure the level of educational capital within each state.
    2. "How well do the state's public and private colleges and universities collectively increase the intellectual skills of their students? What do those whom they educate know, and what can they do?" This second set of questions was directed toward finding out how the higher education system in each state (including public and private institutions) was performing as a whole.
- **How should the National Forum go about collecting the information?** To answer this question, participants adopted a model proposed by the project's advisory committee, developed with the assistance of a panel of assessment

Participants
National Forum on College-Level Learning, con't.

Lillian Montoya-Rael
Executive Director
Regional Development Corporation

Michael Nettles
Professor of Education and Public Policy
University of Michigan

Steffen E. Palko
Vice Chair and President
XTO Energy, Inc.

The Honorable Paul E. Patton
Governor of Kentucky

Charles B. Reed
Chancellor
California State University

Sean C. Rush
General Manager
Global Education Industries
IBM Corporation

Edward B. Rust, Jr.
Chairman and Chief Executive Officer
State Farm Insurance Companies

Ted Sanders
President
Education Commission of the States

The Honorable Jack Scott
State Senator
California State Senate

Kala M. Stroup
Commissioner of Higher Education
State of Missouri

experts convened prior to the meeting (see sidebar for National Forum staff and advisory committee members). The model's key components included:

1. information drawn from existing licensure and graduate-admission tests that many students take when they graduate,
2. results from the National Adult Literacy Survey (NALS), and
3. results of specially administered tests of general intellectual skills.

The State of Kentucky, as it turned out, already had access to information on student learning that fit into the first two categories of the proposed model. That is, the state had assembled scores on some licensure and graduate-admission tests and was willing to collect more. Secondly, it had administered the Kentucky Adult Literacy Survey in 1996, a replica of the NALS. With the generous cooperation of the state, the model was applied to Kentucky as an illustration in *Measuring Up 2002*, using the partial information that was available. Results were encouraging enough for The Pew Charitable Trusts to fund the next phase of the National Forum's work, in which five states would undertake a demonstration project to implement the model in full.

The National Forum, Phase Two

The states that joined Kentucky in the demonstration project were Illinois, Oklahoma, Nevada, and South Carolina—several small and one large state from various regions of the country. Between 2002 and 2004 the project team assembled information on the NALS and on graduate-admission and licensure tests for each demonstration state. Meanwhile, the states administered general intellectual skills tests to a random sample of students at a representative sample of public and private two- and four-year institutions within their borders. The four-year institutions also attempted (unsuccessfully as it turned out) to collect information about their alumni's perceptions of their own intellectual skills. Also, both two- and four-year institutions in each state administered surveys aimed at gauging students' engagement with their collegiate experience, since research suggests that engagement is associated with learning. The engagement measures were subsequently dropped from the model, since they are not direct measures of learning.

The results of the demonstration project were published in *Measuring Up 2004*. All five participating states were awarded a "Plus" in the learning category in acknowledgment of their successful implementation of the model. They had demonstrated that college-level learning could be assessed in a way that makes interstate comparison possible, that these assessments were consistent with other information that *Measuring Up* had revealed about these states, and that the information could be useful to policymakers in each state.

Experience with the demonstration project suggests that it is feasible to extend this approach to other states and eventually to create a nationwide benchmark for learning. While the project encountered difficulties in the logistics of administering tests, institutional commitment and preparation, and student motivation to participate, these challenges are typical of a first effort of this kind. With increased preparation and resources, these barriers can be overcome. To facilitate this process, detailed explanations of the logistics and costs associated with implementing the National Forum's learning model are contained in the appendix. The next edition of the report card on higher education, *Measuring Up 2006*, will report results for additional states in this category.

Why Measure Learning at the State Level?

Even with generous support from The Pew Charitable Trusts, the implementation of the demonstration project was challenging, and it required serious commitment and leadership from the participating states. Contributing to the purposes of a nationwide report card on higher education would not have been sufficient motivation for these states to make an effort of this magnitude, without an accompanying belief that the project would be useful to them.

Fortunately, they did believe in its usefulness. In Kentucky and Oklahoma, the project supplemented or completed existing statewide accountability systems. In South Carolina, it dovetailed with work being done on an accountability project supported by the Fund for the Improvement of Postsecondary Education (FIPSE). Leaders in Illinois and Nevada believed that the project would produce information that could be used to improve their higher education systems.

But what does this approach to assessing college-level learning tell states

Kentucky's Experience with the Demonstration Project

Upon initiating a major reform of postsecondary education in 1997, Kentucky developed an accountability system focused on a public agenda and organized around five key questions:

1. Are more Kentuckians prepared for college?
2. Are more students enrolling in college?
3. Are more students advancing through the system?
4. Are college graduates prepared for life and work?
5. Are Kentucky's communities and economy benefiting?

For each of these questions, the Council on Postsecondary Education developed specific outcome measures called "key indicators" of progress. Valid measures allowing comparisons across states were available for the first three questions, and Kentucky has demonstrated progress on most of these measures. But the fourth and fifth questions were more challenging. Kentucky's participation in the demonstration project assisted the state in developing indicators to address question four. The state's results are also helping to answer questions frequently posed by stakeholders who are external to higher education about the quality of the education being provided to the dramatically increased number of students now enrolled in postsecondary institutions in the state.

that their existing assessment approaches do not?

First, it tells states how much educational capital they have—an asset that every state needs to advance its economic, civic, and social welfare. It is virtually a truism now that education and training beyond high school is necessary for individuals and states to be players in the global economy. In addition, the pressing, complex challenges of our political life and the sophistication of attempts to influence the electorate, so vividly demonstrated in the 2004 national elections, require critical thinking skills that are increasingly essential to the workings of a democracy. Finally, the decisions individuals must face in everyday life—ranging from how to ensure the best schooling for their children, to planning for retirement, to completing the myriad forms that control access to services—have become so challenging that education increasingly differentiates those who are able to negotiate them successfully from those who are not. Certificates and degrees are increasingly inadequate proxies for educational capital. It is the *skills* and *knowledge* behind the degrees that matter.

Secondly, this approach to assessing college-level learning tells a state the extent to which its institutions are collectively effective in contributing to its store of educational capital. Until now, when states have raised the question of learning, the unit of analysis has always been the institution. The model's focus on the state as a whole permits states to ask broader questions that are quite different from how well individual institutions are performing. Among these questions are:

- How well are we doing in serving the various regions of the state?
- Are there achievement gaps among population groups that we should be concerned about and address collectively?
- How well are our workforce-development efforts working?
- Are we producing enough well-trained professionals in areas that are critical to the state's welfare?
- What economic development options are available to our state—or are denied to us—because of the educational capital resources we have?

- Do we have the range of college preparation programs or graduate opportunities needed for the economy and lifestyles that our residents want?
- How does the mobility of the graduating college population—coming here to work and live, or leaving our institutions to go elsewhere—affect our responsibilities to our residents or our ability to create the community life and employment opportunities we want?

A collective examination also enables cost-benefit analyses to be performed concerning the learning that the state's system of higher education is producing in relation to the state's investment. Armed with answers to these kinds of questions, a state can undertake further analyses, target resources where they are most needed to address urgent state priorities, and promote collective solutions to collective problems.

Third, as is true for all the *Measuring Up* categories, a state can benchmark its performance against that of other states and against itself over time, to chart progress and identify good practice. Given sample sizes that are large and representative, institutions too can see how well they perform relative to their peers on a few key assessment measures. These external benchmarks can serve to anchor their more extensive campus-based assessment methods, which continue to be essential to improvement.

Finally, this model represents a way to address the growing national mandate for accountability without creating a federal program. The No Child Left Behind (NCLB) Act has demonstrated the urgency with which the public is demanding a commitment to standards and educational equity through evidence of learning—an urgency that is beginning to be felt in higher education as well as in K–12 schools. The implementation of NCLB has highlighted the dangers of adopting federal solutions to national, state, and local problems. Because much of the information used in the National Forum's model derives from existing databases—and because the tests that are administered are voluntary and sample-based, and are not high stakes—the National Forum's approach is cost effective, minimally intrusive, and nonpunitive for students and institutions.

The results of the demonstration project for Kentucky suggest that the state's two-year institutions (where most of the recent enrollment increase has occurred) are doing a comparatively good job in preparing graduates for life and work. Students at the universities are faring less well on the direct assessments administered through the project, and the state as a whole remains challenged by low literacy levels in its general population.

Kentucky plans to seek $600,000 in recurring state funding to expand the application of these measures of student learning, in order to further investigate these conclusions and to develop baseline data that will allow the state to set the same kinds of improvement goals for learning that it created to measure progress in other areas. Discussions to refine and develop the measures are already underway with postsecondary institutions. These discussions are focused on integrating the National Forum's efforts with a parallel initiative to develop a competency-based assessment of general education outcomes based on the Greater Expectations project administered by the Association of American Colleges and Universities. In the final analysis, efforts to increase participation in postsecondary education must be judged in terms of the extent to which these increases prepare graduates to be successful citizens and workers who contribute to the quality of life of their communities, the state, and the nation.

A Model for Measuring College-Level Learning

For the National Forum's demonstration project, the learning category was constructed much as the other five performance categories in *Measuring Up* had been created. Indicators that reflected various dimensions of state performance were grouped under several overall themes, or clusters, and each was weighted:

- **Literacy Levels of the State Population** (25%). This cluster of indicators reflects the proportion of residents who achieve high levels of literacy. It directly addresses the question, "What are the abilities of the state population?"

For the demonstration project, the data used were the same as those included in the benefits category of *Measuring Up* and were based on the 1992 National Adult Literacy Survey (NALS) for residents ages 25 to 64, updated using the 2000 census. The NALS assessment poses real-world tasks or problems that respondents must perform or solve in the following areas:

1. *Prose literacy*: reading and interpreting texts;
2. *Document literacy*: obtaining or acting on information contained in tabular or graphic displays; and
3. *Quantitative literacy*: understanding numbers or graphs and performing calculations.

With new data available from the National Assessment of Adult Literacy (NAAL) in 2004, it will be possible to sharpen this cluster of indicators to capture the literacy levels of the college-educated population rather than of the state population as a whole. Due to limitations in the statistical procedure used to update the 1992 NALS, however, this was not possible for this analysis.

- **Graduates Ready for Advanced Practice** (25%). The

Oklahoma's Experience with the Demonstration Project

Oklahoma welcomed the opportunity to participate in the National Forum's demonstration project because it dovetailed well with the existing assessment and accountability initiatives of the Oklahoma State System for Higher Education. These initiatives include:

- a mandated, system-wide, college student assessment policy that has been in place since 1991 and that includes assessment of general education and program-level outcomes;
- the Educational Planning and Assessment System (EPAS), which links 8th and 10th grade assessments to the ACT and other information about college preparation;
- the federally sponsored Gaining Early Awareness and Readiness for Undergraduate Programs (GEAR UP), which focuses on school interventions down to the 5th grade and features an information campaign targeted to families to encourage college attendance;
- the Report Card on Oklahoma Higher Education, which includes many of the same measures as *Measuring Up*; and
- a partnership with the Oklahoma Business and Education Coalition (OBEC), the Oklahoma State Department of Education, and Achieve, Inc., which is leading to a comprehensive standards and benchmarking study.

Participation in the National Forum benefited Oklahoma in several ways. First, the project provided institutions with an opportunity to experiment with state-of-the-art assessment measures like the Collegiate Learning Assessment (CLA) during tight budget times and to expand their use of

measures in this area reflect the contributions of colleges and universities to a state's stock of educational capital. This cluster of indicators examines the proportion of the state's college graduates (from both two- and four-year institutions) who are ready for advanced practice in the form of vocational/professional licensure or graduate study. It addresses the policy question, "To what extent do colleges and universities educate students to be capable of contributing to the workforce?"

For the demonstration project, the measures were based on the proportion of college graduates (that is, associate's or bachelor's degree holders) within each state who have demonstrated their readiness for advanced practice through:

1. *Licensure examinations*: taking and passing a national examination required to enter a licensed vocation/profession such as nursing or physical therapy;
2. *Competitive admissions exams*: taking a nationally recognized graduate-admission exam such as the Graduate Record Examination (GRE) or the Medical College Admissions Test (MCAT) and earning a nationally competitive score; or
3. *Teacher preparation measures:* taking and passing a teacher licensure exam in the state in which they graduated.

- **Performance of the College Educated** (50%). This cluster of indicators focuses on the quality of the state's higher education "product" by addressing the all-important question, "How effectively can students who are about to graduate from two- and four-year colleges and universities communicate and solve problems?"

For the demonstration project, the measures consisted of two general intellectual skills assessments:

the ACT WorkKeys, which had been piloted in Oklahoma already. The project also reinvigorated statewide conversations about: (1) using common assessments to help align courses and learning goals throughout the system, and (2) establishing common general education competencies. Institutions were also encouraged to use learning assessment data in a recently established performance-funding approach.

Findings from the demonstration project were shared with numerous groups, including the presidents of the 25 public institutions, the vice presidents for academic and student affairs, the faculty advisory councils, the chairmen of all governing boards, and business leaders. All results were also provided to campus assessment coordinators for further analysis or local use. The findings indicated a possible writing deficiency among Oklahomans that has since been confirmed in discussions with the academic officers. Recently, much emphasis has been placed on improving math preparation, followed by reading; as a result, writing may have been overlooked. Another finding that Oklahoma took note of was the relatively low number of students prepared for graduate school: few seek advanced education, and many of those who do so do not achieve competitive test scores.

A number of initiatives in Oklahoma are planned as a result of the state's participation. First, the state plans to build on the work of the National Forum in collecting licensure and graduate examination scores, which has been a difficult task in the past. The state also hopes to explore other ways to compare teacher certification information. Finally, by hosting a follow-up meeting to the National Forum project in Oklahoma next year, assessment coordinators plan to consider other national measures that were not included in the demonstration project and hope to expand collection of these measures beyond the five pilot states.

1. *At two-year institutions*: the WorkKeys assessments administered by the American College Testing (ACT) Service. These assessments examine what students can do with what they know. Items on reading comprehension and locating information, for instance, might require students to extract information from documents and instructions; questions in applied math might test their abilities in using mathematical concepts such as probability or estimation in real-world settings. The business writing assessment requires students to prepare an original essay in a business setting.
2. *At four-year institutions*: the Collegiate Learning Assessment (CLA). The CLA is an innovative examination that goes beyond multiple-choice testing by posing real-world tasks that a student is asked to understand and solve. For example, students could be asked to draw scientific conclusions from a body of evidence in biology or examine historical conclusions based on original documents. Or they might be asked to write a persuasive essay, and analyze and then refute a written argument with logic and evidence.

Measures included under the first two clusters above—"literacy levels of the state population" and "graduates ready for advanced practice"—are available nationally and can potentially be calculated for all 50 states, although the smaller size of the national samples and the reduced number of state over-samples between the NALS in 1992 and the NAAL in 2004 make this difficult to do for the smaller states. The National Forum therefore endorses the recommendation made in the report of the National Commission on Accountability in Higher Education, sponsored by the State Higher Education Executive Officers (SHEEO), that the size of the next adult literacy survey be increased (see http://www.sheeo.org/account/accountability.pdf). Measures included in the third cluster above—"performance of the college educated"—will require special data-collection efforts similar to those undertaken by the five demonstration project states in 2004.

As with any data used to determine a grade in *Measuring Up*, values for each of the indicators within each cluster must be compared with a common standard. For the calculations in the five other categories in the report card, this standard is set by the best-performing states. Because the demonstration project involved only five states, however, the standard chosen was the national average on each measure. For those cases in which national data were unavailable, the five-state average was used.

Results for Participating States

The set of measures collected according to the National Forum's model can be used to create a "learning profile" that communicates graphically each state's strengths and challenges with respect to collegiate learning. Each state's performance on this profile is reflected by how many percentage points *above* or *below* the national or state benchmark its own performance lies. Horizontal bars on each state profile correspond to each of the measures and portray an overall pattern of performance for each state in relation to other states (see figure 2, page 12). Bars to the left of the vertical line in the center of the display indicate how far below the national benchmark the state falls on a comparable scale for these measures. Bars to the right of the vertical line indicate how many percentage points above this benchmark the state performs. For example, 26% of Illinois residents achieved high scores on the National Adult Literacy Survey (NALS) in prose literacy, compared with 24% nationally who did so—a difference of 8.3% in favor of Illinois (26 – 24 = 2; 2 ÷ 24 = .083), as shown in the top bar of figure 2. Deviations of only a few percentage points on a given measure indicate that the state's performance is not markedly different from that of other states, while larger deviations (that is, about ten points or more) suggest that the state is above or below most others on this dimension of performance.

It is important to emphasize that the evaluation of learning results presented for each state should be confined to raising issues for discussion and making broad comparisons. Because relatively small numbers of students were tested on the direct measures of student learning, and because the extent to which this test-taker population is representative of all graduates of two- and four-year colleges in each state is unknown, results should be treated with caution. Readers should look primarily at the overall *pattern* of such results without making too much of the individual values for each measure.

Illinois

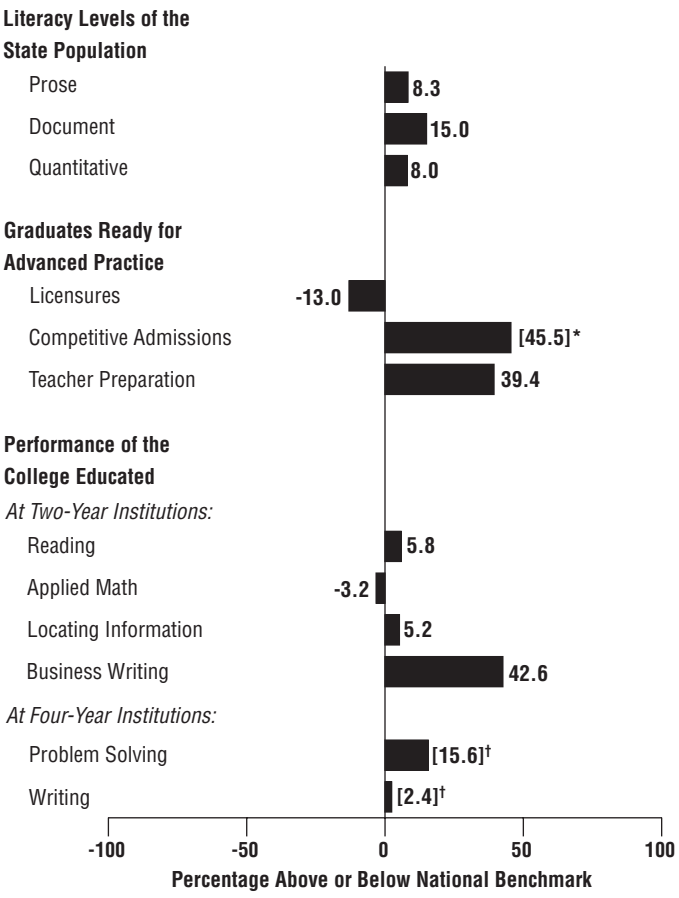
Illinois has historically had a strong and well-funded higher education system and enjoys a diverse economy and relatively high levels of educational attainment. These strengths are reflected in its above-average performance with respect to literacy in all three areas: prose, document, and quantitative literacy (see figure 2, page 12). For example, 26% of Illinois residents who took

the NALS earned the highest scores in prose literacy, compared with a national average of 24%. The other four demonstration states scored a high of 22% and a low of 18% on this measure. Results on the other two literacy measures show equivalent performance differences between Illinois and other states.

But higher education institutions in Illinois also tend to emphasize traditional over vocational fields of study. The state's community colleges, for instance, are all comprehensive community colleges rather than technical colleges. This emphasis is reflected in the cluster of indicators reflected in "graduates ready for advanced practice." For example, about ten percent fewer of the state's two- and four-year college graduates take and pass vocational or professional licensing examinations than is the case nationally. But four-year college graduates in Illinois take graduate-admission examinations at just above the national average, and they perform well on them. Fifty-two percent of Illinois graduates who take such exams achieve nationally competitive scores, compared with only 31% in other states.

Figure 2

Illinois Learning Measures



* GRE scores used as part of the calculation of "competitive admissions" for Illinois were based on the national average because of missing data for key institutions. All other test score data are specific to Illinois.

† These scores must be qualified because of the limited number of institutions participating.

Illinois students also perform at above-average levels on all direct measures of student learning with the exception of the applied math skills of students at two-year colleges. For instance, 26% of the state's test-takers at two-year colleges achieved top scores on the WorkKeys business writing test, compared with only 18% across the five states. And 61% of Illinois' test-takers at four-year institutions achieved top scores on the CLA problem-solving measure, while only 53% did so across all five states. While not as dramatic, score differences between Illinois and other states on the other assessments administered for the project within this cluster of indicators were similar. This favorable outcome probably reflects the strong high school backgrounds typical of Illinois high school graduates: Illinois received an A in preparation in *Measuring Up 2002* and a B+ in *Measuring Up 2004*.

But Illinois does face a challenge with respect to the performance of

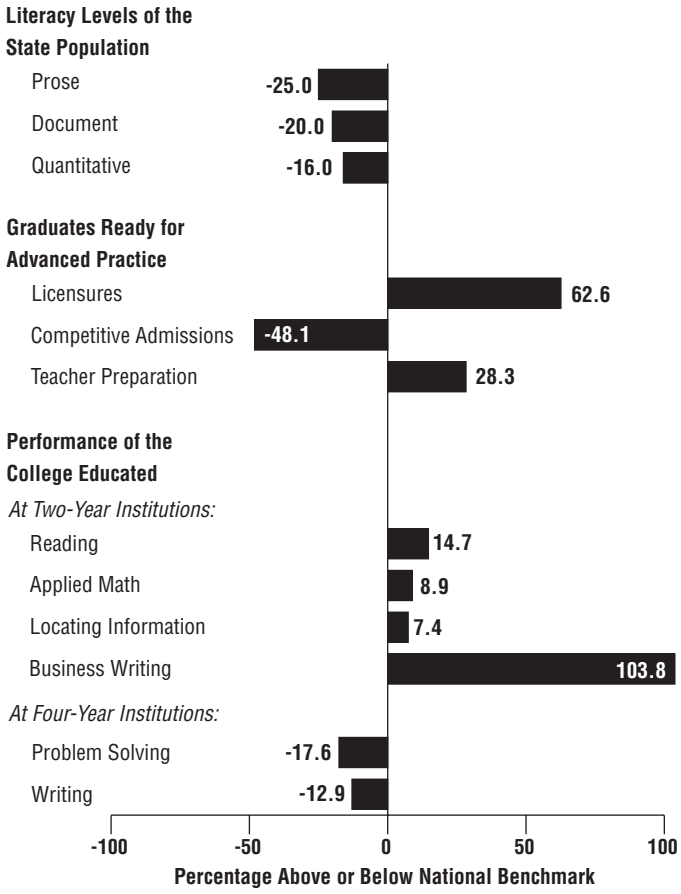
its minority students in higher education. African-American and Hispanic students in Illinois score not only below white students, but also at significantly lower levels than their counterparts in other states. Across all six assessments, white students in Illinois were more than twice as likely as their nonwhite counterparts to achieve high scores. For example, 55% of Illinois' white test-takers at two-year colleges achieved top scores on the WorkKeys reading for information test, while only 21% of the state's nonwhite counterparts achieved this level of performance.

Kentucky

Kentucky has recently made major investments in both K–12 and postsecondary education, in large part because it faces substantial challenges with respect to literacy and educational attainment. These challenges are reflected in literacy performances significantly lower than those of other states (see figure 3). For example, only 18% of Kentucky's residents scored in the top performance categories in prose literacy, while 24% did so nationwide. The state also lags the nation, but not as severely, in document and quantitative literacy.

Figure 3

Kentucky Learning Measures



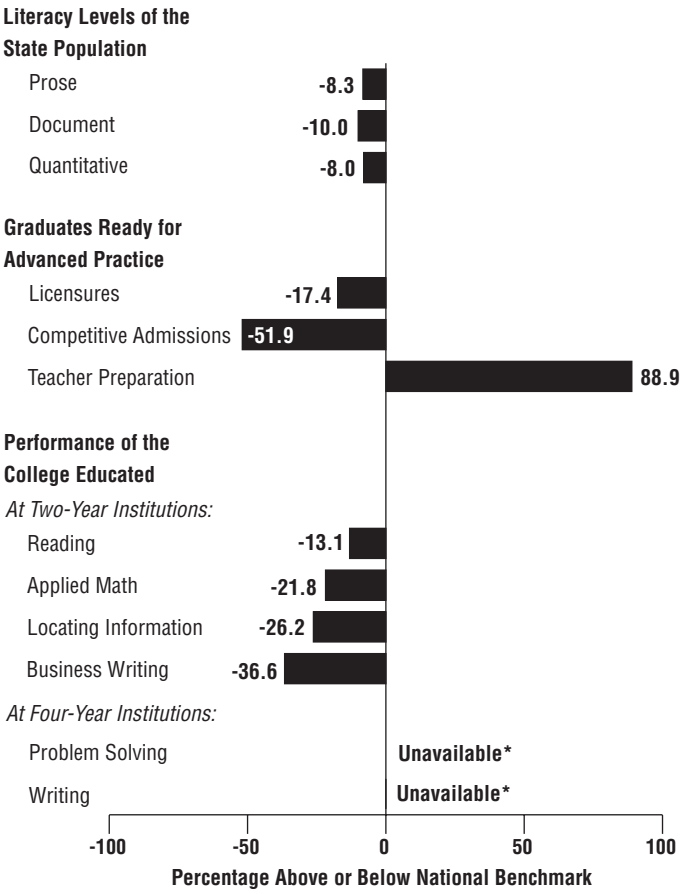
Kentucky has recently made a significant investment in its community and technical college system, allowing these institutions to play a much stronger role in workforce development. These investments appear to have paid off with higher-than-average proportions of graduates taking and passing licensing examinations in fields like nursing or physical therapy. For example, about half again as many Kentucky graduates of two-year colleges take licensing exams as do their counterparts in other states, and 86% of the Kentucky test-takers achieve passing scores, compared with 84% of those taking such examinations elsewhere. However, the state remains less competitive with respect to the proportion of four-year college graduates taking and performing well on examinations governing admission to graduate schools. Only about three-fourths as many Kentucky graduates of four-year institutions take exams like the GRE as do their

counterparts in other states. The proportion of test-takers achieving nationally competitive scores on such tests is 23% in Kentucky, compared with a national average of 31%.

Kentucky's two-year college students perform at high levels on the WorkKeys exams, especially business writing. More than a third (37%) of the state's test-takers achieved top scores on this measure, while only about 18% did so in other states. The performance of four-year college students in Kentucky is less competitive and constitutes a challenge for the state. For example, 44% of Kentucky's test-takers at four-year colleges and universities achieved top scores on the task-based CLA problem-solving assessment, compared with a five-state average of 53%. There are also notable performance gaps between white and African-American students on all these examinations in Kentucky, although the state's African-American students do perform better than their counterparts in other states.

Figure 4

Nevada Learning Measures



* These data were unavailable due to insufficient numbers of test-takers and logistical problems with test administration.

Nevada

Nevada has a unique economy and a small, nonselective, higher education system composed entirely of public institutions. The state has performed at the lower end of most *Measuring Up* scales with respect to preparation (receiving a D in *Measuring Up 2004*) and educational attainment. The state's below-average results on the literacy measures reflect this standing (see figure 4), as the state's residents consistently score about ten percentage points below the national average across all three indicators.

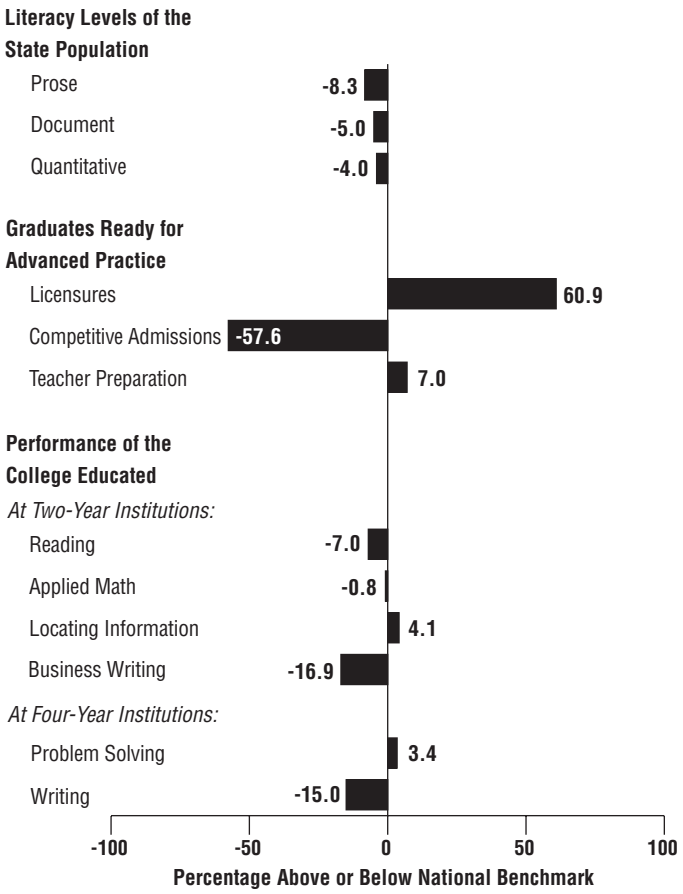
Meanwhile, Nevada faces an unprecedented teacher shortage as its K–12 system tries to keep up with an expanding population. The state appears to be meeting this challenge, as shown by an unusually high proportion of graduates taking and passing teacher licensure examinations: almost twice as many college graduates do so in Nevada as across the nation. But the state is far less competitive in the other two indicators of "graduates ready for advanced practice." About 20% fewer Nevada students take licensing

examinations as compared with other states, though the pass rates of Nevada students are competitive nationally. At the same time, graduates of four-year institutions in other states take graduate-admission exams at about one and a half times the rate of Nevada's graduates, and only 22% of Nevada graduates who take these exams earn competitive scores, while 31% of their counterparts elsewhere do so. These results probably reflect lower levels of student preparation upon entering college and the fact that fewer students graduate from college in Nevada; the state received an F in completion in *Measuring Up 2004*.

Because of problems encountered in the testing process beyond the state's control, direct evidence of the quality of outcomes is unavailable for four-year college students in Nevada. The performance of the state's two-year college students is below the five-state averages on all four of the skill areas tested. This was especially the case on the WorkKeys business writing exam, where only about 11% of Nevada test-takers scored at the highest level, while more than 18% did so across the five states.

Figure 5

Oklahoma Learning Measures



Oklahoma

Oklahoma has recently been active in attempting to improve the quality of its higher education system. The state faces substantial educational challenges with respect to baccalaureate attainment (it is in the bottom half of the 50 states) and the quality of its preparation of students in K–12 schools (the state earned a C– in preparation in *Measuring Up 2004*). The literacy levels of Oklahoma's residents reflect these challenges (see figure 5). The state lags behind national averages in the proportions of its residents achieving top scores by about five to ten percent across all three measures.

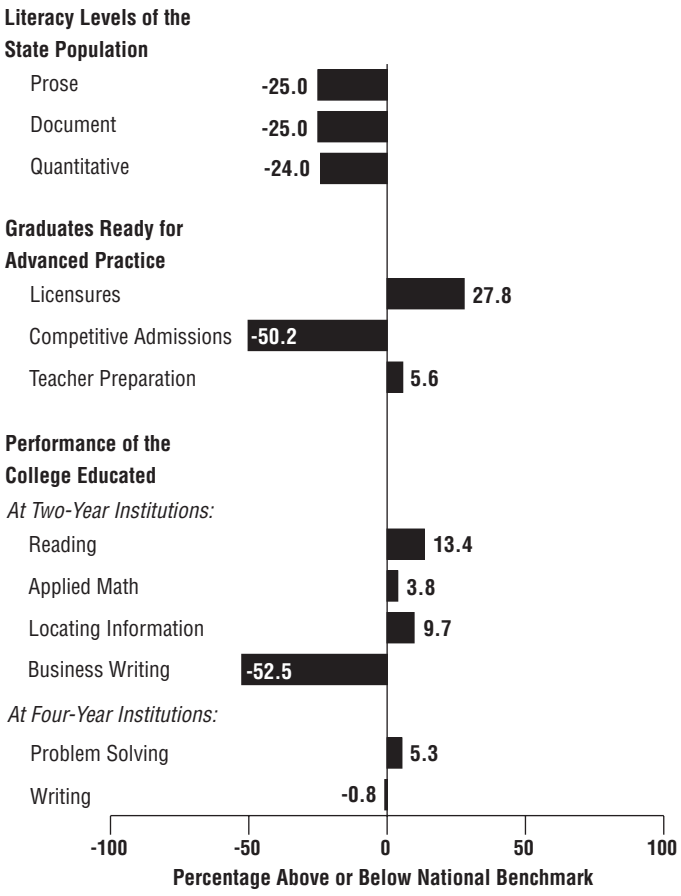
Oklahoma's higher education system is heavily and deliberately oriented toward workforce preparation, and this emphasis is reflected in its performances on measures of the readiness of graduates for advanced practice. The proportion of its graduates taking professional and vocational licensure examinations is well above the national average, with almost two-thirds more two-year

college students in Oklahoma taking such examinations than in other states. Furthermore, the pass rates of Oklahoma's two-year college students on such exams are also competitive, matching national averages on such tests. But about 60% fewer four-year college graduates take graduate-admission examinations in Oklahoma than in other states, and only 23% of these test-takers achieve top scores in Oklahoma, compared with 31% elsewhere.

Students at two- and four-year institutions perform at or just below national averages on direct measures of student learning. While Oklahoma's minority students perform at lower levels than white students in the state on all six measures, the performance levels of Oklahoma's nonwhite students are about the same as those in other states. Written communication skills, though, constitute a particular policy challenge for Oklahoma across all population groups. This is reflected in below-average performances on the WorkKeys business writing exam taken by the state's two-year college students (15% achieved top scores in Oklahoma, compared with more than 18% across the five states), on the CLA writing assessment taken by the state's four-year college students (just under 32% achieved top scores in Oklahoma, compared with more than 37% across the five states), and in prose literacy for the general population (22% of Oklahoma residents achieved top scores on the NALS, compared with more than 24% for the U.S. population as a whole).

Figure 6

South Carolina Learning Measures



South Carolina

South Carolina's "educational pipeline" loses many students early, with almost half of the state's ninth graders failing to graduate from high school within four years. However, those who do make it to college are comparatively well prepared, and the state's colleges and universities have very good rates of college completion (earning the state a B in completion in *Measuring Up 2004*).

This bifurcated pattern is also reflected in the learning measures assembled for South Carolina by the demonstration project (see figure 6). Literacy levels are well below national averages, with the proportions scoring in the top categories on the NALS lagging on all three measures. For example,

19% of South Carolina residents achieved top scores in quantitative literacy compared with 25% for the nation as a whole. But the performance of South Carolina students on direct measures of student learning is mostly above average. For example, on the WorkKeys reading for information test taken by the state's two-year college students, 65% achieved top scores while only about 57% did so across the five states. Similarly, over 56% of the state's test-takers at four-year institutions achieved top scores on the task-based CLA problem-solving assessment, while only 53% did so across the five states. This above-average performance, however, is not reflected in writing. At the two-year level, about half as many of the state's test-takers achieved top scores on the WorkKeys business writing assessment as did test-takers across the five states (9% vs. 18%). At the four-year level, South Carolina test-takers were about as likely as their counterparts elsewhere to achieve top scores in writing (both at about 37%).

Much of the policy challenge for South Carolina lies with the state's African-American student population, which constitutes more than a quarter of all students enrolled. South Carolina's African-American students not only perform at levels below those typical of the state's white students, but also frequently score lower than their counterparts in other states. To take an extreme case, 62% of South Carolina's white test-takers at two-year colleges achieved top scores on the WorkKeys applied math exam while only 13% of the state's nonwhite test-takers did so. Both of these statistics can be compared with 25% of nonwhite test-takers achieving top performances across the five states. While performance gaps on the other five measures between white and nonwhite students in South Carolina are not as large as those in applied math, the pattern of results is similar.

In the realm of graduates ready for advanced practice, the proportion of South Carolina graduates who take and pass licensing examinations governing entry to vocational and technical professions is above average. Two- and four-year college students take such examinations at a rate about 20% higher than is typical nationally, and their pass rate is 88% in South Carolina, compared with a national rate of 84%. This probably reflects South Carolina's historic commitment to two-year technical colleges and the many applied programs (especially in health-related fields) offered by the state's four-year colleges. On the other hand, only about two-thirds as many four-year college graduates take graduate-admission examinations in South Carolina as do so across the country. The proportion achieving nationally competitive scores on such examinations in South Carolina is 25%, compared with 31% nationally.

Two Challenges for Learning

The five-state demonstration project provided the National Forum on College-Level Learning with an opportunity to examine two important challenges the nation faces in the realm of collegiate learning. The first is a notable gap in the performance of white students and students of color on the direct measures of learning. The second is the uneven performance of states in preparing future teachers. Because both topics were incidental to the main purpose of the demonstration project—to generate and interpret comparable information on student learning across states—the results reported here represent only a beginning of this discussion. Nonetheless, the results do suggest the magnitude of the task the nation may be facing in these two important areas.

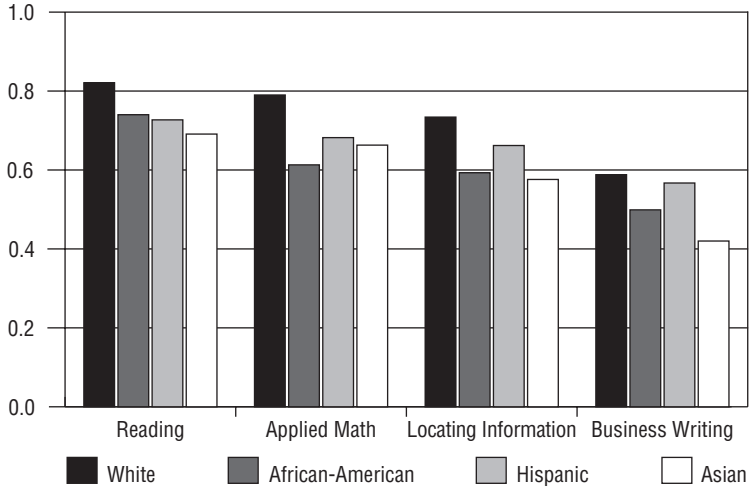
Performance Gaps by Race/Ethnicity

Measuring Up 2004, like its predecessors, reported several areas of performance in individual states where differences in outcomes or experiences for students of color had a substantial bearing on overall state performance. If such differences could be narrowed or eliminated, the report argued, the state in question would be substantially better off. Not surprisingly, the National Forum's demonstration project revealed parallel differences in college-level learning across racial/ethnic groups in all five participating states.

Figure 7

WorkKeys Examinations: Mean Scores

(Two-Year College Students)



Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test (for example, a score of 4 with a maximum of 7 yields a standardized score of 0.571).

Figures 7 and 8 present standardized mean scores[2] for all six of the examinations used: the four WorkKeys exams administered to two-year college students (see figure 7) and the two Collegiate Learning Assessment (CLA) tasks administered to four-year college students (see figure 8). The results are broken down by four major racial/ethnic groups across

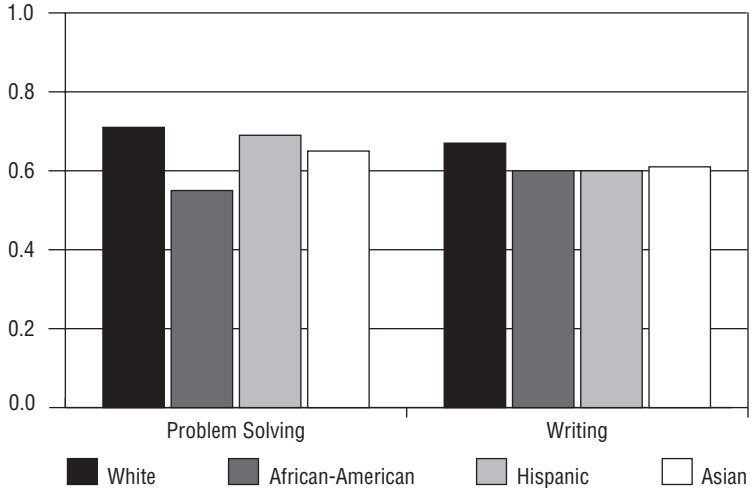
[2]See note to figure 7 for an explanation of how scores are standardized.

all five participating states. Although individual patterns vary by examination and the performance gaps are in some cases small, white students consistently do better than their counterparts from other racial/ethnic groups. These performance gaps are particularly wide for African-American students at two-year colleges in applied math and business writing, and for African-American students at four-year institutions in task-based problem solving. Equally notable is the fact that Asian students did not outperform white students in applied math in two-year colleges or in problem solving at four-year colleges and universities, despite the fact that other studies show that they tend to do so consistently on more conventional algorithm-based math exams. In contrast, the examinations employed in the demonstration project emphasized the use of mathematical tools and concepts in more complex problem-based settings where language skills are important.

Figure 8

Collegiate Learning Assessment: Mean Scores

(Four-Year College/University Students)

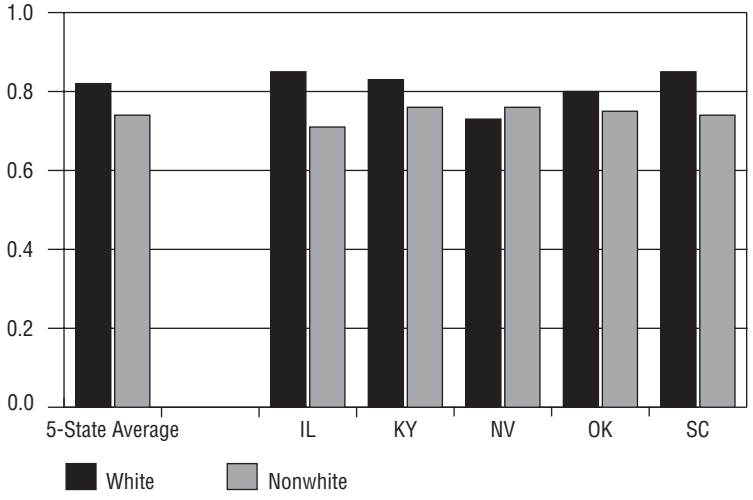


Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

The numbers of students tested in the demonstration project were insufficient to confidently explore patterns of performance across all racial/ethnic populations within individual states. But enough cases were available to examine performance gaps between white students and students from all other racial/ethnic groups on a state-by-state basis. Figures 9 to 14 present standardized mean scores for each of the six examinations for white and nonwhite test-takers in each of the five participating states, and in the nation as a whole.

Figure 9

WorkKeys, Reading: Mean Scores

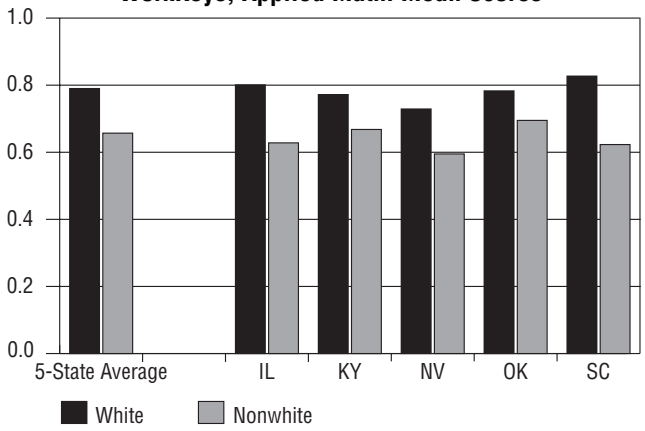


Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

As is apparent in these figures, a performance gap between white and nonwhite students is present for virtually every examination in every state, suggesting a widespread and

Figure 10

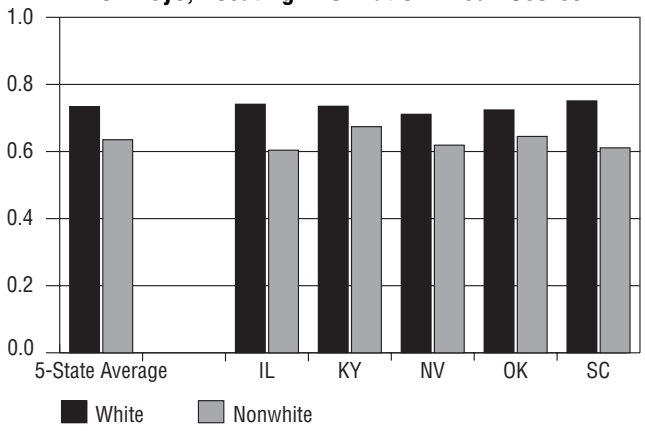
WorkKeys, Applied Math: Mean Scores



Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

Figure 11

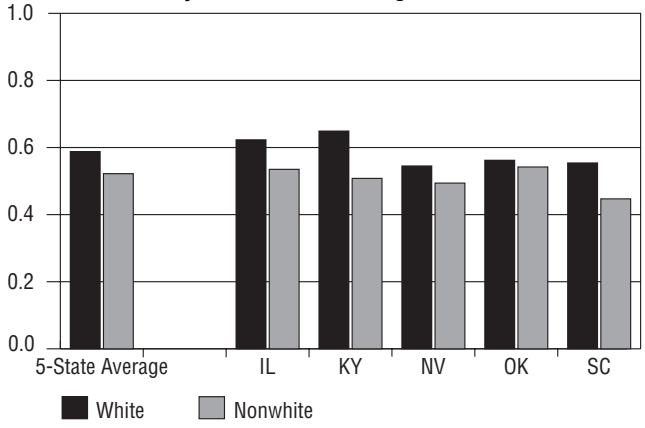
WorkKeys, Locating Information: Mean Scores



Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

Figure 12

WorkKeys, Business Writing: Mean Scores



Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

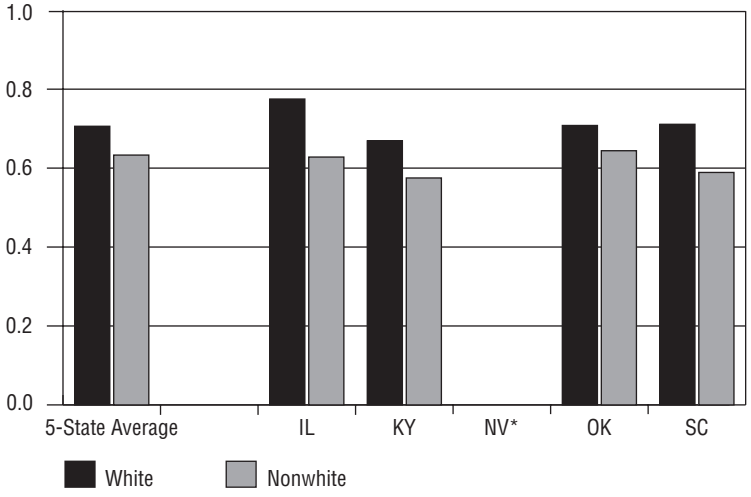
systematic pattern of adverse impact. Such differential patterns of performance for these two groups may significantly affect the overall results for individual states. In Illinois, for example, nonwhite students frequently perform at lower levels than their nonwhite counterparts in other states. Yet overall, Illinois is one of the strongest-performing states on these measures because the impact of nonwhite performance is masked by their limited numbers. The state's overall performance would have been even higher had these substantial performance gaps not been present. Similarly, nonwhite students in South Carolina also perform at relatively low levels compared with whites across all of these examinations. But the impact of this performance gap on overall state performance is far higher than in other states because nonwhite students constitute a substantial proportion of the South Carolina student population.

The demonstration project is certainly not the first learning assessment initiative to discover such performance gaps. Indeed, virtually every published report of a large-scale testing program in the United States at a national or state level shows similar gaps in performance. Calling particular attention to this issue in the context of *Measuring Up* and the National Forum, however, is compelling for at least two reasons. First, performance gaps based on race/ethnicity and income are already a persistent theme across many of the dimensions examined by *Measuring Up*, ranging from college preparation to collegiate access and persistence. For example, Oklahoma has made progress in narrowing the gaps in college completion between white and minority students over the last decade, but African-American students are still only three-quarters as likely to complete a degree as their white

counterparts. In Illinois over the same period, the proportion of Hispanic students receiving certificates and degrees has increased from 8 to 11 per 100 enrolled; nevertheless, Hispanic students in Illinois remain only about half as likely as white students to complete certificates and degrees. Parallel performance gaps in learning measures only serve to further confirm that this is a national problem worth significant policy attention.

Figure 13

CLA, Problem Solving: Mean Scores



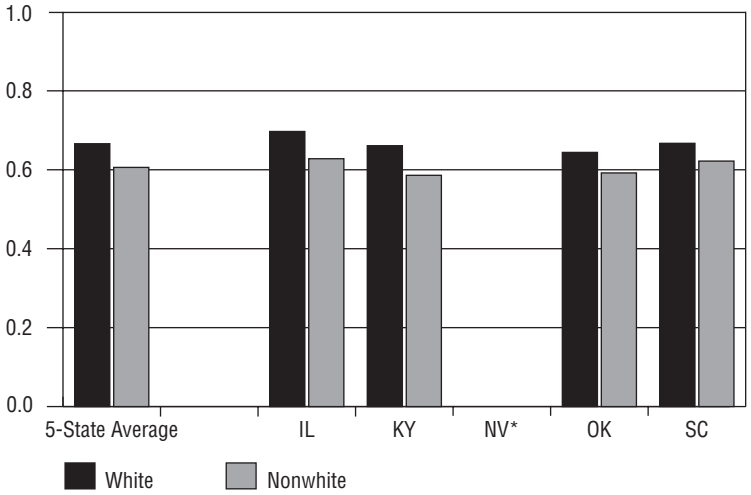
* Data for Nevada were unavailable due to insufficient numbers of test-takers and logistical problems with test administration.

Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

Second, the notion of educational capital that forms the conceptual foundation of the National Forum's work emphasizes the need to educate *everybody* in order to sustain economic and civic vitality. Performance gaps in learning, if they continue, will seriously erode state and national competitiveness—especially as the diversity of young adults in many states increases. This threat has already been documented for many states through data on inequities in educational attainment. A recent report by the National Center for Higher Education Management Systems (NCHEMS) concluded that total personal income (and associated state tax revenue) for Nevada, for example, would be $2.2 billion higher than its projected base of $43.9 billion by 2020 if Hispanics, African-Americans, and Native Americans achieved the same levels of education as whites.

Figure 14

CLA, Writing: Mean Scores



* Data for Nevada were unavailable due to insufficient numbers of test-takers and logistical problems with test administration.

Note: To allow comparisons, the test results have been standardized by converting them to a scale from 0 to 1, with 1 being the highest score possible on the test.

This is occurring at a time when other nations are rapidly overtaking the United States in the proportion of young residents earning a baccalaureate degree, and they are doing so largely *because of* such performance gaps with respect to income and race/ethnicity. According to data provided by the Organisation for Economic Co-operation and Development (OECD), for example, overall baccalaureate attainment rates among young American adults (ages 25 to 34) are now

lower than their counterparts in four other countries. Disaggregating these data reveals the fact that attainment rates for young white males in the United States approximately match the overall attainment rates achieved by residents in these other countries. Our growing underperformance internationally is a direct result of this nation's inability to increase the attainment levels of its nonwhite young adult population.

Educating Future Teachers

Since the publication of *A Nation at Risk* more than 20 years ago, states and the federal government have been engaged in substantial efforts to improve the quality of America's elementary and secondary schools. Higher education's critical responsibility in this effort is centered on preparing future teachers who are masters of their subject areas and are ready to take on the challenges of increasingly crowded and diverse classrooms. Yet our knowledge of how well the nation's colleges and universities are fulfilling this crucial responsibility remains limited. As reported in *Measuring Up 2004*, we know that most states have made progress in "teacher quality," as measured by the proportion of teachers in their K–12 classrooms who are teaching in the field in which they majored in college. But how much do prospective teachers actually know about the subjects they are preparing to teach? Title II of the Higher Education Act, whose reporting requirements went into effect in fall 2002, was designed to provide answers to that question by requiring publicly available reports on the pass rates of teacher candidates in each subject—with data available by institution and by state. Each state, however, can choose its own examinations. Even if states use the same exam, each can establish its own standards for licensure. As a result, published Title II reports do not provide consistent information to compare performance across states.

The demonstration project provided a limited opportunity to address this condition because three participating states—Kentucky, Nevada, and South Carolina—use many of the same examinations to certify teachers. These tests are provided by the Educational Testing Service (ETS) as part of the Praxis examination series and are used to assess both subject-area knowledge and basic skills in these three states. Teacher candidates first are tested in the basic skills of reading, writing, and mathematics using the Praxis I battery—either as an exit standard for certification or to enter many institutions' teacher education programs and become a candidate for certification. They then must pass a combination of other subject-area examinations in the Praxis II battery, depending on the particular certification they are seeking. A total of 66 different Praxis examinations were used by these three states from 2001

to 2003: 9 in basic skills and the remaining 57 in subject areas ranging from agriculture to teaching visually handicapped children. By obtaining actual test results from these three states instead of publicly reported pass rates, the states' performances could be directly compared. The two remaining demonstration project states—Illinois and Oklahoma—both employ their own, noncomparable teacher examinations and therefore could not be included in this analysis.

Following *Measuring Up's* established procedure of benchmarking each state's performance to that of the best-performing states, the highest standard for passing each Praxis examination in any of the 50 states was identified using published Title II reports. This score was then established as the standard and was applied to actual student scores in the three demonstration project states. The results show for each state: (1) the reported pass rate at the level actually used for teacher certification within the state and (2) what the state's pass rate would have been if the highest state standard in the nation had been applied (see table 1). If the highest national standards had been applied to each of these states, the percentage of students passing the teacher licensure exams would have been noticeably lower in all three states.

Table 1

Pass Rates for Praxis II: Subject Knowledge
Percentage of Students Passing

	Reported Pass Rate Based on State's Own Standards	*Recalculation of Pass Rate Based on Highest State Standard for Passing*
Kentucky	81.3%	57.4%
South Carolina	77.4%	59.1%
Nevada	89.9%	71.9%

The situation is more complex with respect to Praxis I, the basic skills examinations administered to future teachers, because of differences among the states in the way these tests are used. In Kentucky, for example, Praxis I exams are used by many (but not all) institutions to govern entrance to teacher education programs. As a result, that state's published Title II reports accurately show 100% pass rates on these exams, because all certified teachers would have had to attain the state's designated passing score in order to enter a teacher education program in the first place. In South Carolina and Nevada, on the other hand, many individual institutions establish additional entrance standards in basic skills. Teacher candidates in those two states must meet state-established basic skills standards in order to be certified—just as they must meet such standards with respect to subject-area knowledge. As a result of these differences, the actual pass rates within these two states are lower than in Kentucky (see table 2, page 24). As a final note, these examinations truly do test basic skills, reflecting levels of functioning in reading, writing, and mathematics that are typical of the ninth-grade level—which may be a relatively low standard for prospective teachers, regardless of where they are in the teacher-preparation pipeline.

Table 2

Praxis I: Basic Skills
Percentage of Students Passing

	Reported Pass Rate Based on State's Own Standards	*Recalculation of Pass Rate Based on Highest State Standard for Passing*
Kentucky	100.0%	31.1%
South Carolina	69.1%	37.8%
Nevada	82.8%	52.7%

As is evident from table 2, overall state performance is strikingly different for these three states when the nation's highest performance standards are applied. Only about a third of test-takers in Kentucky and South Carolina, and about half in Nevada, performed at the levels expected in the state with the highest standard for passing (Virginia). Supporting some of the test results reported earlier, these differences were particularly apparent in mathematics for Kentucky and in writing for Nevada and South Carolina. For example, South Carolina test-takers achieved a mean score of 174.1 on the Praxis I in writing—above the state's standard for passing of 173 but below the standard of 178 established in Virginia.

Are these differences important? States have long maintained that local conditions vary and that standards for licensing teachers should be set to match them. And it could be argued that current standards are adequate for practicing teachers and that states such as Virginia and Maryland may simply have set standards that are inappropriately high. In addition, many states are facing unprecedented teacher shortages (among them, most strikingly, Nevada), and it is questionable whether establishing higher standards for licensure would really help them to educate more children better under these conditions. Each state must make its own determination of what such differences, if they are detected, really mean. The key point is that current Title II data reporting, which is confined to local pass rates, does not enable states to detect such differences in the first place. Results of the National Forum's demonstration project show that many states can in fact benchmark their performances in teacher education against external standards if they choose to do so, which would in turn make such a policy conversation possible.

These two analyses—concerning performance gaps by race/ethnicity and educating future teachers—are examples of the kinds of issues that can be addressed when comparable statewide information about collegiate learning is available. With larger sample sizes, states would be able to perform other, more thorough analyses and comparisons with the information this model produces.

Conclusion
Modest Investment, Collective Benefit

Despite substantial challenges, the National Forum's five-state demonstration project achieved its principal objective of showing the feasibility of assembling indicators of collegiate learning on a comparable basis across multiple states. The resulting information about educational capital is consistent with what we already know about higher education in the five participating states, and the costs of obtaining this information are modest (see appendix for information about those costs). But in today's climate of constrained resources, why should a state support such an investment at all?

One answer is accountability. States spend millions, and sometimes billions, of dollars each year on higher education but have in the past been able to produce little information that can demonstrate to residents and taxpayers the effectiveness of these investments. Learning represents the inescapable bottom line for the nation's colleges and universities. On these grounds alone, not having information about learning presented in succinct and comparable form is increasingly hard to justify—just as participants in the National Forum's initial meeting at Purchase, New York, concluded more than three years ago.

But if accountability were the only reason to pursue this agenda, states might legitimately pause. Even more importantly, information about educational capital can complement other state-specific information on the strengths, challenges, and benefits of higher education to help forge a powerful public agenda for action. For example, a state's leadership can use information that has been broken down by geographic area or by population group to identify concrete problems and thereby begin to mobilize public action. These steps have already been taken in some states that have disaggregated data from *Measuring Up* by population or region in areas such as participation and affordability. Although the vignettes presented in the previous section of this report, "Two Challenges for Learning," remain preliminary, they illustrate how a useful disaggregation of information about learning can be accomplished.

Using consistent statewide information about learning outcomes in this way—to identify the specifics of a collective policy challenge that *all* institutions can help address—is far more attractive than the customary (and feared) use of performance information to reward and punish individual institutions. At the

same time, having comparable information across states can help policymakers identify best practices and track progress. These are essential conditions for improving performance, and the demonstration project shows they can be purchased at an affordable price.

The National Forum on College-Level Learning has put to rest the question of whether assessment can be done in a way that allows for meaningful state comparisons. It can. The implementation of the model described here requires leadership, hard work, and resources. But the principal recommendation from the National Forum's work is that states should adopt the model because the information it produces:

- is valid and useful for state policy;
- supplements existing accountability approaches and campus-based assessment efforts;
- leads to informed discussions about a public agenda for higher education; and
- can help provide motivation toward achieving that public agenda.

In its report, the National Commission on Accountability, convened by the State Higher Education Executive Officers (SHEEO), has called for all states to adopt the National Forum's model for assessing learning at the state level. In an era of heightened accountability, with not only states but also the federal government interested in determining what value they get for the investment they make in higher education, it is time for the higher education community to take the lead in determining how that value should be assessed. The model presented here is not perfect, and it was not perfectly implemented in the demonstration project. But it is as promising a place to begin as any yet seen in this country.

Appendix
How to Implement the Model for College-Level Learning

If state policymakers choose to adopt the model presented by the National Forum on College-Level Learning for assessing student learning, they need to address several issues, including:

- the roles of officials at the state and campus levels;
- the logistics of survey and test administration, and analysis of the resulting information; and
- the resources that will be required.

This appendix addresses each of these issues, based on the experiences of states that participated in the pilot project.

State- and Campus-level Leadership

Assessment at the state level cannot be accomplished without strong and consistent leadership at both the state and campus levels. Since governing or coordinating boards generally assume strategic leadership for state higher education systems, this generally is the logical place for state-level responsibility for such an initiative to be lodged. The unwavering and clear commitment of the board and its chief executive (generically known as the State Higher Education Executive Officer, or SHEEO) is a necessary condition for obtaining the campus cooperation that is crucial to implementing the model.

Active involvement of the SHEEO is particularly important, especially in the early stages of implementing such an effort. He or she is in the best position to generate the political impetus and board support needed to move the initiative forward. The SHEEO is also best able to create buy-in from the campuses by convincing them of the value of this kind of assessment for various stakeholders, reducing the threat of inter-institutional comparison that cross-campus assessment might seem to pose, encouraging and supporting campus leadership, and using the results to help create a public agenda for higher education and enrich the state's accountability system.

In the five-state demonstration project, a senior-level governing or coordinating board staff person was assigned to lead the effort in each state.

That state leader worked with the campuses to move the project forward by consulting with senior campus leadership, channeling resources to the institutions, keeping campus personnel informed through the various phases of implementation, and consulting with them about the strategies and protocols for testing.

Through periodic meetings and regular email communication, the state project directors provided campus leaders with crucial administrative and moral support. In a survey administered at the end of the pilot project, campus coordinators stressed the importance of such lines of communication between themselves and the state project leader. Especially important to them was timely information about the purposes and value of the project, the psychometric properties of the assessments to be used, and effective implementation strategies. The coordinators suggested in particular that the state director involve them early in the design of sampling and testing procedures, since they knew best how to recruit students for local assessment efforts.

Campus leadership is as important as state leadership. The president's role is to communicate the purposes of the project—as well as the value and uses of the information that it produces—to faculty, staff, and students. He or she also needs to ensure that the campus coordinators are provided with the resources of time and money required to do the job and that the results are disseminated and used to benchmark the results of campus-based assessments of student learning.

Campus coordinators are the people most directly responsible for the success or failure of the effort. In the demonstration project, they explained the sampling and testing protocols established by the project team to the personnel who recruited the students and administered the assessments, and they worked with them to develop incentive strategies. (Since the biggest difficulty the pilot ran into was enlisting students to take the tests, the budget for this work now includes a payment to students for participating.) Campus coordinators at four-year institutions also collected other needed information, including GRE scores (available from ETS at a nominal fee), information from the Integrated Postsecondary Educational Data System (IPEDS), and SAT scores for students taking the Collegiate Learning Assessment. Campus leaders need to ensure that the people administering the tests are consulted early in the project and have sufficient time for implementation, adequate staffing and resources, and an understanding of how the information will be used at both the campus and state levels.

Campus leaders also need to motivate students to show up (and do their

best) and to coordinate this project with other campus activities so that it becomes a useful supplement to campus assessment efforts, not simply a data collection "add-on." The logistical challenges of this approach also include those associated with administering online surveys and tests, including scheduling rooms and computers for testing, dealing with software problems, and accessing technical support.

Logistics

Overall, states should allow about a year and a half from the beginning of the project to the point at which they will have usable data. This allows sufficient time to assemble the measures needed in the first two clusters—"literacy levels of the state population" and "graduates ready for advanced practice"—which must be obtained directly from testing companies and other agencies. It also allows sufficient time for the detailed planning that is needed to administer the direct measures of student learning on selected college and university campuses.

Literacy Levels of the State Population: National Assessment of Adult Literacy (NAAL)

The National Assessment of Adult Literacy (NAAL) has replaced the National Adult Literacy Survey (NALS) as the nation's most direct measure of educational capital. Like the NALS, the NAAL is a household survey that assesses the prose, document, and quantitative literacy of a representative sample of the nation's adults, both the college-educated and those with a high-school diploma or less. In 1992, when the NALS was administered, a number of states over-sampled their residents in order to get results that were also representative at the state level, but only six did so when the NAAL was administered in 2003. Even without the over-sample, however, it is possible to approximate a representative state-level sample in the larger states by employing the methodology used for the adult literacy measure in the benefits category of *Measuring Up 2004*.[3]

The NALS was used in the demonstration project to determine the level of literacy of the state's residents, calculated as the proportion of its residents scoring at the highest levels (4 and 5) on all three tests. Given large enough sample sizes, it could also help states determine the value added by a college

[3] *Technical Guide Documenting Methodology, Indicators, and Data Sources for Measuring Up 2004* (San Jose: National Center for Public Policy and Higher Education, 2004).

education with respect to the literacy levels of the population. The original plan in the demonstration project was to use information from the newer NAAL, which was due for release in 2003. Because of a delay in that release, however, the National Forum instead created a placeholder by using simulated literacy data, created by applying a regression procedure to adjust 1992 NALS results on the basis of population characteristics drawn from the 2000 census. For at least the next five years, actual NAAL results will be available as a ready source for states to calculate this measure. The NAAL results will enable states to sharpen this cluster of indicators to capture the literacy levels of the college-educated population rather than of the state population as a whole. Due to limitations in the statistical procedure used to update the 1992 NALS, however, this was not possible for this analysis.

States are strongly encouraged to over-sample their residents in any future administrations of the adult literacy survey. Unfortunately, it has been administered no more frequently than once a decade or longer. (SHEEO's National Commission on Accountability recommends in its report that the federal government administer it more often and to larger numbers of people.) In the meantime, states with over-samples on the NAAL or states that are large enough for the statistical approximation can use the information produced by the NAAL to calculate the required performance indices.

The budget in this appendix presumes that the National Center for Higher Education Management Systems (NCHEMS) will analyze state data from the NAAL to produce state-level statistics, from which it will then calculate index scores for the literacy levels of the college-educated population. States that choose to do this on their own will have to obtain detailed breakdowns of the performance of college-educated individuals on each of the three assessments for their states and for the nation as a whole.

Graduates Ready for Advanced Practice: Licensure and Graduate-Admission Tests

The demonstration project used state-level results from those licensure and graduate-admission tests that satisfied three criteria:

1. national and state-level performance data are available;
2. the tests are required in order to practice a profession or enter graduate school; and
3. possession of a two- or four-year college degree is required to take the tests.

In addition, each state supplied data on the results of its teacher examinations.

Measures in the "graduates ready for advanced practice" cluster were computed for the demonstration project using the same methodology applied to Kentucky in *Measuring Up* 2002. This consisted of defining a particular level of performance on each test that could be used as a benchmark, above which a particular test-taker could be deemed "ready for advanced practice." In the case of licensure examinations with established national standards, this level of performance was passing the examination and being licensed. In the case of graduate-admission examinations, a criterion score was set at a level generally accepted as "competitive" with respect to gaining admission to a graduate program. The number of individuals achieving this level or higher was then counted. This number was divided by the total number of applicable degrees (baccalaureate or associate) associated with the credential and separately reported for nine licensure examinations and five graduate-admission tests. Fields included in the licensures list included nursing, clinical pathology, physical therapy, respiratory therapy, radiology, and physician's assistant. Admissions examinations included Graduate Record Examination (GRE), the Graduate Management Admissions Test (GMAT), the Medical College Admissions Test (MCAT), the Law School Admissions Test (LSAT), and the Pharmacy College Admissions Test (PCAT).

All test scores except GREs can be obtained directly from national sources. GRE scores were compiled by asking participating institutions in each state to request their scores from ETS, which can be done via a standard report for a small fee. This meant that the number of degrees used in the denominator of the calculation had to be adjusted to include only those institutions reporting GRE scores.

Comparing performances across states is problematic for teacher education because of differing standards in each state, as well as the use of different test batteries. The measure for teacher education used for the demonstration project was the number of individuals passing licensure examinations in the state (obtained from Title II reports) divided by the number of applicable degrees for individuals entering teaching obtained from the Integrated Postsecondary Educational Data System (IPEDS). "Applicable degrees" were defined as "education" plus all fields of study recognized in secondary education when counting teachers teaching "in field" for the teacher quality indicator included in the preparation category of *Measuring Up 2004*.

The budget presented later in this appendix assumes that a third-party

organization like NCHEMS (which played this role in the demonstration project) will contact the testing companies and assemble and analyze licensing and graduate-admission testing data. If states choose to undertake this task themselves, another responsibility that the state coordinator must assume is communication with the testing companies and collecting and analyzing the data they provide. For states that pursue this avenue, the following description of the procedures may be helpful.

Several months prior to the final assembly of information, the state-level coordinator should collect from those sites the pass rates for regular, first-time examinees on each examination at the state and national levels for the three most recent years for which data are available. State-level results for some tests are posted by state on the Web sites of the administering organization. This is the case, for example, for the MCAT, the GMAT, and the two nursing exams: the National Council Licensure Examination for Registered Nurses (NCLEX-RN) and the National Council Licensure Examination for Practical Nurses (NCLEX-PN). The state coordinator will need to directly contact the organizations that conduct the other examinations to solicit state-level results because they do not provide the information publicly (some may charge a small fee for this service). The GRE board has refused thus far to release state-level information for its test. But scores can be obtained and aggregated if the state coordinator asks each four-year institution to request from ETS the standard institutional report for the most recent three years.

Scores on all available professional licensure and graduate-admission examinations for all three years must be aggregated to create a single index score for each type of examination. The basic method for doing so involves determining the number of eligible students in the state who pass their licensure tests or achieve a competitive score (that is, one that will gain them admission to graduate school) on a graduate-admission test. The resulting number of "graduates ready for advanced practice" is then divided by the total number of applicable degrees associated with the credential, separately aggregated for licensure examinations, graduate-admission tests, and teacher licensure examinations.

Before using these data to construct index scores, a number of initial calculations are required to make them comparable:

- **Subscore Aggregation**. For tests with multiple subscores but no total score, subscores must be aggregated to create a single indicator of performance, weighting each subscore equally. The same procedure is used to average the number of individuals passing or scoring at or above a particular level where multiple subscores are present.

- **Standardizing Scores**. To adjust for differences in test-score scaling, summary test-score performance data should be indexed to a standardized value range of 0 to 1, depending upon the top score possible on a given test (for example, a GRE score of 450 with a maximum of 800 yields a standardized score of 0.5625).
- **Time Period Aggregation**. Up to three years of the most recent data should be used in these calculations to create an "average year." This approach allows more data to be used in cases where the number of test-takers in a given state is small. In cases where three years of data are available, data from all three should be aggregated and divided by three. In cases where two years are available, these two should be combined and divided by two.

After these initial adjustments, the resulting data consist of comparable summary performance statistics for each test, including number of test-takers, mean and median scores, standard deviation, and number passing or achieving at or above a designated score. From these data, the "graduates ready for advanced practice" indicator can be calculated. The following steps are used to create this indicator:

1. Determine the number of individuals ready for advanced practice. For licensure tests, this is the number of individuals passing the examination. For admissions examinations, it is the number of individuals achieving at or above a given nationally competitive score (GRE=600, GMAT=600, LSAT=155, MCAT=10, PCAT=215).
2. Determine the appropriate number of graduates associated with each potential test-taking population using IPEDS data. In most cases, these are baccalaureate degrees, but in some cases they are associate degrees and in others, both. For teacher examinations, the denominator used was the total number of baccalaureate degrees in education plus all other fields of study listed as providing a "qualified" teacher in the teacher quality measure used in the preparation category of *Measuring Up*. If multiple testing years were available, degree data were similarly aggregated by year to create an "average year."
3. Create a ratio between these two numbers. This is the fraction of educational capital within a state that is represented by this test.
4. Add the resulting fractional contributions to educational capital for each of the states under consideration and for the nation.

Pass rates and raw or composite scores of students taking the teacher

licensure tests can generally be obtained from the state's department of education. Pass rates and the numbers taking each test are also posted on the Title II Web site at www.title2.org. The measure for teacher education used in the demonstration project was calculated by taking the number of individuals passing licensure examinations in the state and dividing it by the number of applicable degrees as defined under item #2 above. If raw or composite scores can be obtained, however, the state can perform the kinds of comparative analyses illustrated in this report (see tables 1 and 2, pages 23–24) by applying the highest passing standard among the states on each test as reported on the Title II Web site.

Performance of the College Educated: General Intellectual Skills Tests

The demonstration project used two test batteries to assess the general intellectual skills of students. The two-year institutions administered four American College Testing Service (ACT) WorkKeys tests. The four-year institutions used the Collegiate Learning Assessment (CLA), which is administered by the Council on Aid to Education (CAE), a subsidiary of the RAND Corporation.

The ACT WorkKeys assessments principally examine what students can do with what they know. Items on reading comprehension and locating information, for instance, are focused on how well test-takers can extract information from complex documents and instructions, while items on applied mathematics test students' ability to use mathematical concepts like probability or estimation in real-world settings. The WorkKeys writing assessment also requires students to complete an extended essay. The WorkKeys battery used in *Measuring Up 2004* included four tests—reading for information, applied mathematics, locating information, and business writing—and the results of each test are reported separately. Additional information about the WorkKeys examinations is available at www.act.org/workkeys/.

The CLA goes beyond typical multiple-choice testing by posing multifaceted tasks—anchored in an academic discipline—that a student is asked to understand and solve. The CLA battery used in the demonstration project consisted of two types of assessments—a set of four authentic tasks and a set of two writing prompts drawn from the Graduate Record Examination (GRE). Because they are different kinds of assessments examining essentially different skills, performance on them was reported separately—problem solving for the tasks and writing for the GRE prompts. Additional information on the CLA assessment is available at www.cae.org/content/pdf/CLA-OpportunityToParticipate.pdf.

Administering the WorkKeys and CLA examinations constitutes the greatest challenge to implementing the National Forum's model. The subsections below describe: (1) the sampling procedures used in the demonstration project to select potential students to participate; (2) the administration of the tests; and (3) the analysis of results.

Sampling Procedures. The design for collecting testing data requires a total sample of some 1,200 test-takers for each of the two test batteries in a given state. This necessitates a cluster-sampling approach: first, a sample of institutions is drawn, and second, the sample of students to participate from each institution is selected. This sampling approach represents a compromise, based on the conflicting need to attain some degree of statewide representativeness and the desire to include enough test-takers at participating institutions to enable them to use the resulting data for local purposes. The basic sampling plan thus envisions about 75 to 100 test-takers at 12 to 15 four-year institutions and at an equivalent number of two-year institutions in a given state. However, many states may wish to select more institutions, or more students at each institution, to participate. Indeed, several states in the National Forum's demonstration project chose to do so. In Nevada, where there are only two four-year institutions and four two-year institutions, all were chosen, and the numbers of students targeted for testing at each was higher. In Kentucky and Oklahoma, all public institutions were invited to participate, with the institutional sampling frame used only to select private institutions.

In each case where a selection of institutions must be made, the universe of applicable institutions (four-year public, four-year private, and two-year) should be divided into groups of roughly comparable institutions. Variables used to construct these groups should at minimum include institutional size, type, disciplinary mix, selectivity, urban/rural location, full-time/part-time ratio, and racial/ethnic distribution. The resulting sampling groups can then be checked by running statistics for various combinations of potential selections within them to ensure that they produce samples that closely resemble known statewide distributions on such variables as full-time/part-time breakdown, gender, race/ethnicity, and disciplinary emphasis. The typical result for a state will be five to seven distinct groups of institutions within each category of institutions (public four-year, private four-year, and two-year). The first group in each cluster will consist of institutions that are required to participate because they are large, unusually selective, or otherwise distinctive. But given the need for flexibility in recruiting institutions, each state has the discretion to select a given number of institutions within the remaining sampling groups.

Once participating institutions are identified, the next step is to randomly select a group of students to be invited to participate in the testing. Accordingly, a set of sample-selection guidelines have been developed for use by participating institutions. The target population for sampling includes all students officially enrolled in the most recent fall term who are expected to complete a two-year or a four-year degree the following spring (identified by numbers of credits or courses completed). Institutions should be directed to randomly select an initial sample of students who meet these criteria, together with two backup samples to be used to replace members of the initial group who decline to participate. Institutions in the demonstration project were provided with several methods for conducting the random selection procedure and for employing the backup sample (see http://measuringup.highereducation.org/docs/technicalguide_2004.pdf, pp. 80–83).

Test Administration. The CLA and the WorkKeys batteries should be administered using protocols supplied by the vendors, customized for use in the demonstration project. The CLA assessments are typically completed in a Web-based format. Each CLA test-taker should be asked to complete either one task or two GRE prompts. Each CLA test-taker in the demonstration project also completed the National Survey of Student Engagement (NSSE), although results of this survey are not included in this report because they are not direct measures of student learning. The total testing time for the CLA battery administered in this way was just over two hours. Each WorkKeys test-taker should be asked to complete: (a) the applied mathematics and the reading for information examinations, or (b) the locating information and the business writing examinations. The tests were completed in a paper-and-pencil format in the demonstration project but will soon be available from ACT in a computer-based format. Each test-taker in the demonstration project also completed the Community College Survey of Student Engagement (CCSSE), although results of this survey are not included in this report. The total testing time for the WorkKeys battery administered in this way was about one and a half hours.

Additional testing materials should be supplied to each campus in case more students than expected show up for testing. Members of the initial sample should be invited to participate by means of a letter from the college president accompanied by recruitment materials (samples of both that can be adapted by each institution are available on the National Forum Web site at http://collegelevellearning.org). Members of this initial targeted group who decline participation or do not reply should be re-contacted in a week. If the response is still negative, the institution should recruit replacements from the pre-selected backup samples, selecting students with the same or similar majors and, ideally,

the same full-time/part-time attendance status.

Encouraging students to participate was the largest single challenge of the demonstration project and will likely be a challenge for any state trying to implement the National Forum's model. Campus coordinators reported that the length of the tests (up to two hours) was a problem, especially for working students and for seniors, who are the most difficult students to recruit. For that reason the budget below includes a $75 payment for each test-taker. Some states may prefer to give students academic credit instead, but this may require individual score reports for each of the students. With individual score reports, the WorkKeys can be used as a work certificate and the CLA can be noted on the student transcript. This would require students to take three WorkKeys tests or two tasks from the CLA, thus increasing the investment of student time and state money. Whichever incentive is chosen, it needs to be of significant value to the students.

Once the student samples are selected, campuses must schedule a number of dates near the end of the fall semester for testing. The exams need to be proctored, which means that rooms must be reserved if a paper-and-pencil format is used. If a Web- or computer-based format is used, campus coordinators must find a sufficient number of available computers. This was not an easy task in the demonstration project: increasingly, students are bringing their own computers to campus, and campus computer laboratories are in many cases getting smaller instead of larger (CAE is currently exploring the possibility of letting students use their own computers). Campus coordinators should begin room scheduling as soon as the institutions have agreed to participate. In the demonstration project, some four-year institutions determined that they needed to take the CLA to their Institutional Review Boards (IRBs), a process that also should be started early if it is required. The RAND Corporation (whose Council on Aid to Education oversees the CLA) has its own IRB, which has reviewed the exam. As a consequence of that review, the Council on Aid to Education requires students to fill out privacy and consent forms.

Analyzing and Reporting Results. The completed exams are sent to ACT and CAE respectively, where the results are analyzed and reports created for the institutions and the state. Campuses that have adequate numbers of test-takers and that can supply IPEDS data and SAT or ACT scores for the participating students can also request from CAE an analysis of how well their students test compared with what might be predicted based on the performance of similar test-takers who have taken the exam before.

Results described in this report for these examinations are based on the proportions of test-takers scoring above a given level on each of the tests given. For CLA, this level was based on adjusted scores of 26 and above, calculated separately for task-based problem solving and the GRE-based writing sample. For the WorkKeys tests, the levels differed because the scales for each of the four tests differ—high scores are six and above for reading and for applied mathematics, five and above for locating information, and four and above for business writing. Finally, results for the demonstration project were weighted as needed by race and ethnicity, gender, and institutional size to make them more representative. Such a procedure should be followed by any state undertaking its own analyses if it finds substantial and consequential differences in response rates between men and women or across demographic groups. Test-takers from larger institutions should also count more in computing the state's aggregate score than those from smaller institutions, in proportion to how much of the state's total undergraduate full-time equivalent (FTE) enrollment that each represents. Results should then be compared with the available national or multi-state norms on each examination.

Resource Requirements

Although the demonstration project has shown that the National Forum's approach to assessing college-level learning at the state level can be cost effective, it does require resources. The project team consulted the project budget and surveyed state and campus coordinators to determine both the time and money required to complete the tasks described above. Future implementations of this model will be done under different circumstances and will be affected by a host of variables that cannot be predicted, including increased charges for materials and services. Moreover, states may want to implement the most pared-down version of the model or a more robust one that generates more information but requires greater resources. Therefore, the cost estimations that follow are approximate.

Time

The first variable is personnel time. The project team did not try to translate these time estimates into salary dollars, given the many variables that affect compensation levels, such as the locality and the seniority of the people working on the project. The best source of information for the time spent by state leaders on the project came from the surveys of the state coordinators. Since they did not keep timesheets, they could only estimate the hours they had worked. Estimates for state leaders averaged about 180 hours spent on

the project over two years. Much of the model's groundwork has already been laid and will not have to be replicated. But if the states were to take on some of the responsibilities of the project team, such as negotiating with the testing companies or analyzing the data, then their time devoted to the project could increase to as much as 250 hours. Campus coordinators were also surveyed, and they reported an average of about 100 hours spent administering the project. These duties included selecting the sample, recruiting students, administering the tests, and providing the project team and testing companies with information.

Administrative Costs: $25,500

The survey also asked for information about the costs to the coordinating or governing board and to the campuses of administering the project, and that information was pooled with the project budget. It appears that these costs were minimal: roughly and on average $1,500 at the state level and $1,000 at the campus level (multiplied by approximately 24 campuses). At the state level this would include expenses such as the cost of meetings, and at the campus level, the postage and telephone charges associated with the recruitment of students.

Materials

In the pilot project, the costs of materials were covered by the grant from The Pew Charitable Trusts. In the future, these costs will be assumed by the state.

As table 3 (page 40) reveals, the range of potential costs to the states is broad. For instance, a state would incur non-personnel costs of *less than $87,000* if it chooses: to implement the model as part of a five-state consortium (which would enable economies of scale in collecting and analyzing data); not to pay students but give them academic credit instead; to administer tests to only 100 students per institution; and not to administer the National Survey of Student Engagement (NSSE), the Community College Survey of Student Engagement (CCSSE), or the College Results Survey. By contrast, a state that pays students, implements the model by itself, administers 200 tests per institution, and uses the CCSSE, the NSSE, and the College Results Survey could spend *up to about $370,000*, including personnel costs for data collection and analysis.

Table 3

Costs Per State

Item	*Description*	*Cost Estimation*	*Minimum Cost*	*Maximum Cost*
WorkKeys	$11 per student (the mean for two tests) x 1,200	$13,200	$13,200	$13,200
CLA for 200 students per institution — or — CLA for 100 students per institution	$6,500/institution x 12 — or — $4,500/institution x 12	$78,000 — or — $54,000	- - $54,000	$78,000 - -
GRE reports (3 years)	$225/institution x 12	$2,700	$2,700	$2,700
Licensing exam analysis by testing companies	$250/exam x 3	$750	$750	$750
Data Analysis*				
Cost to Set Up Data System		$25,000	- -	$25,000
Fixed Cost for Data Analysis (NCHEMS)		$50,000	$15,600 for a five-state consortium	$50,000
Marginal Cost for Data Analysis (NCHEMS)		$7,000/state		- -
Optional costs				
Student payment	$75 x 2,400	$180,000	- -	$180,000
NSSE/CCSSE	$1.50/student x 2,400	$3,600	- -	$3,600
	$300/campus x 24	$7,200	- -	$7,200
College Results Survey	Web site for 3 months	up to $4,400	- -	$4,400
	Survey, $2.00/graduate x 1,200	$2,400	- -	$2,400

* The cost to set up the data system is calculated based on the investment that the National Center for Higher Education Management Systems (NCHEMS) has already made to do so. If a state chose not to use NCHEMS for data analysis, it would need to replicate that work and would incur that cost. If the state did use NCHEMS, that investment would not need to be made again. The fixed and marginal costs for data analysis represent the costs associated with analytical work performed by NCHEMS. If only one state participated, that cost would be $50,000 (the fixed cost). If more than one state participated, the cost represented by each additional state would be $7,000 (the marginal cost). For example, if five states participated, the cost would total $78,000 ($50,000 + [$7,000 x 4]), and each state would pay $15,600.

Note: CLA stands for the Collegiate Learning Assessments. GRE stands for Graduate Record Examination. NSSE is the National Survey of Student Engagement. CCSSE is the Community College Survey of Student Engagement.

About The Authors

Margaret A. Miller, project director of the National Forum on College-Level Learning, is a professor of higher education policy at the Curry School of Education, University of Virginia, and editor of *Change* magazine. Having begun her academic career as an English professor at the University of Massachusetts, Dartmouth, she went on to serve as the chief academic officer for the State Council of Higher Education for Virginia and then as president of the American Association for Higher Education. She has been a member of the *Measuring Up* advisory board since its inception. She currently also serves as a TIAA/CREF institute fellow and vice chair of the National Center for Higher Education Management System's board of directors.

Peter T. Ewell is vice president of the National Center for Higher Education Management Systems (NCHEMS). His work focuses on assessing institutional effectiveness and the outcomes of college, and involves both research and direct consulting with institutions and state systems on collecting and using assessment information in planning, evaluation, and budgeting. He has directed many projects on this topic, including initiatives funded by the W. K. Kellogg Foundation, the National Institute for Education, the Consortium for the Advancement of Private Higher Education, the Lumina Foundation, and the Pew Charitable Trusts. In addition, he has consulted with over 375 colleges and universities and 24 state systems of higher education on topics related to the assessment of student learning. He has authored six books and numerous articles on the topic of improving undergraduate instruction through the assessment of student outcomes. In addition, he has prepared commissioned papers for many agencies, including the Education Commission of the States, the National Governors Association, the National Conference of State Legislators, and the National Center for Public Policy and Higher Education. In 1998, he led the design team for the National Survey of Student Engagement (NSSE). A graduate of Haverford College, he received his Ph.D. in political science from Yale University in 1976 and was on the faculty of the University of Chicago.

The National Center For Public Policy And Higher Education

The National Center for Public Policy and Higher Education promotes public policies that enhance Americans' opportunities to pursue and achieve high-quality education and training beyond high school. As an independent, nonprofit, nonpartisan organization, the National Center prepares action-oriented analyses of pressing policy issues facing the states and the nation regarding opportunity and achievement in higher education—including two- and four-year, public and private, for-profit and nonprofit institutions. The National Center communicates performance results and key findings to the public, to civic, business, and higher education leaders, and to state and federal leaders who are poised to improve higher education policy.

Established in 1998, the National Center is not affiliated with any institution of higher education, with any political party, or with any government agency; it receives continuing, core financial support from a consortium of national foundations that includes The Pew Charitable Trusts, The Atlantic Philanthropies, and The Ford Foundation.

152 North Third Street, Suite 705, San Jose, California 95112
Telephone: 408-271-2699 • FAX: 408-271-2697
www.highereducation.org

National Center Publications

The National Center publishes:

- Reports and analyses commissioned by the National Center,
- Reports and analyses written by National Center staff,
- National Center Policy Reports that are approved by the National Center's Board of Directors, and
- *National CrossTalk*, a quarterly publication.

The following National Center publications—as well as a host of other information and links—are available at **www.highereducation.org**. Single copies of most of these reports are also available from the National Center. Please FAX requests to 408-271-2697 and ask for the report by publication number.

Measuring Up on College-Level Learning, by Margaret A. Miller and Peter T. Ewell (October 2005, #05-8). In this report, the National Forum on College-Level Learning proposes a model for evaluating and comparing college-level learning on a state-by-state basis, including assessing educational capital. As well as releasing the results for five participating states, the National Forum also explores the implications of its project findings in terms of performance gaps by race/ethnicity and educating future teachers.

The Governance Divide: A Report on a Four-State Study on Improving College Readiness and Success, by Andrea Venezia, Patrick M. Callan, Joni E. Finney, Michael W. Kirst, and Michael D. Usdan (September 2005, #05-3). This report identifies and examines four policy levers available to states that are interested in creating sustained K–16 reform: finance, assessments and curricula, accountability, and data systems. In addition, the report examines the importance of other factors—such as leadership and state history and culture—in initiating and sustaining K–16 reform.

Borrowers Who Drop Out: A Neglected Aspect of the College Student Loan Trend, by Lawrence Gladieux and Laura Perna (May 2005, #05-2). This report examines the experiences of students who borrow to finance their education but do not complete their postsecondary programs. Using the latest comprehensive data, this report compares borrowers who drop out with other groups of students, and provides recommendations on policies and programs that would better prepare, support, and guide students—especially low-income students—in completing their degrees.

Case Study of Utah Higher Education, by Kathy Reeves Bracco and Mario Martinez (April 2005, #05-1). This report examines state policies and performance in the areas of enrollment and affordability. Compared with other states, Utah has been able to maintain a system of higher education that is more affordable for students, while enrollments have almost doubled over the past 20 years.

Measuring Up 2004: The National Report Card on Higher Education (September 2004). *Measuring Up 2004* consists of a national report card for higher education (report #04-5) and 50 state report cards (#04-4) The purpose of *Measuring Up 2004* is to provide the public and policymakers with information to assess and improve postsecondary education in each state. For the first time, this edition of *Measuring Up* provides information about each state's improvement over the past decade. Visit www.highereducation.org to download *Measuring Up 2004* or to make your own comparisons of state performance in higher education.

***Technical Guide Documenting Methodology, Indicators, and Data Sources for* Measuring Up 2004** (November 2004, #04-6).

Ensuring Access with Quality to California's Community Colleges, by Gerald C. Hayward, Dennis P. Jones, Aims C. McGuinness, Jr., and Allene Timar, with a postscript by Nancy Shulock (April 2004, #04-3). This report finds that enrollment growth pressures, fee increases, and recent budget cuts in the California Community Colleges are having significant detrimental effects on student access and program quality. The report also provides recommendations for creating improvements that build from the state policy context and from existing promising practices within the community colleges.

Public Attitudes on Higher Education: A Trend Analysis, 1993 to 2003, by John Immerwahr (February 2004, #04-2). This public opinion survey, prepared by Public Agenda for the National Center, reveals that public attitudes about the importance of higher education have remained stable during the recent economic downturn. The survey also finds that there are some growing public concerns about the costs of higher education, especially for those groups most affected, including parents of high school students, African-Americans, and Hispanics.

Responding to the Crisis in College Opportunity (January 2004, #04-1). This policy statement, developed by education policy experts at Lansdowne, Virginia, proposes short-term emergency measures and long-term priorities for governors and legislators to consider for funding higher education during the current lean budget years. *Responding to the Crisis* suggests that in 2004 the highest priority for state higher education budgets should be to protect college access and affordability for students and families.

With Diploma in Hand: Hispanic High School Seniors Talk about their Future, by John Immerwahr (June 2003, #03-2). This report by Public Agenda explores some of the primary obstacles that many Hispanic students face in seeking higher education, barriers which suggest opportunities for creative public policy to improve college attendance and completion rates among Hispanics.

Purposes, Policies, Performance: Higher Education and the Fulfillment of a State's Public Agenda (February 2003, #03-1). This essay is drawn from discussions of higher education leaders and policy officials at a roundtable convened in June 2002 at New Jersey City University on the relationship between public purposes, policies, and performance of American higher education.

Measuring Up 2002: The State-by-State Report Card for Higher Education (October 2002, #02-7). This report card, which updates the inaugural edition released in 2000, grades each state on its performance in five key areas of higher education. *Measuring Up 2002* also evaluates each state's progress in relation to its own results from 2000.

***Technical Guide Documenting Methodology, Indicators, and Data Sources for* Measuring Up 2002** (October 2002, #02-8).

State Policy and Community College–Baccalaureate Transfer, by Jane V. Wellman (July 2002, #02-6). Recommends state policies to energize and improve higher education performance regarding transfers from community colleges to four-year institutions.

Fund for the Improvement of Postsecondary Education: The Early Years (June 2002, #02-5). The Fund for the Improvement of Postsecondary Education (FIPSE) attained remarkable success in funding innovative and enduring projects during its early years. This report, prepared by FIPSE's early program officers, describes how those results were achieved.

Losing Ground: A National Status Report on the Affordability of American Higher Education (May 2002, #02-3). This national status report documents the declining affordability of higher education for American families, and highlights public policies that support affordable higher education. Provides state-by-state summaries as well as national findings.

The Affordability of Higher Education: A Review of Recent Survey Research, by John Immerwahr (May 2002, #02-4). This review of recent surveys by Public Agenda confirms that Americans feel that rising college prices threaten to make higher education inaccessible for many people.

Coping with Recession: Public Policy, Economic Downturns, and Higher Education, by Patrick M. Callan (February 2002, #02-2). Outlines the major policy considerations that states and institutions of higher education face during economic downturns.

Competition and Collaboration in California Higher Education, by Kathy Reeves Bracco and Patrick M. Callan (January 2002, #02-1). Argues that the structure of California's state higher education system limits the system's capacity for collaboration.

Measuring Up 2000: The State-by-State Report Card for Higher Education (November 2000, #00-3). This first-of-its-kind report card grades each state on its performance in higher education. The report card also provides comprehensive profiles of each state and brief states-at-a-glance comparisons.

***Beneath the Surface: A Statistical Analysis of the Major Variables Associated with State Grades in* Measuring Up 2000**, by Alisa F. Cunningham and Jane V. Wellman (November 2001, #01-4). Using statistical analysis, this report explores the "drivers" that predict overall performance in *Measuring Up 2000*.

Supplementary Analysis for* Measuring Up 2000: *An Exploratory Report, by Mario Martinez (November 2001, #01-3). Explores the relationships within and among the performance categories in *Measuring Up 2000*.

***Some Next Steps for States: A Follow-up to* Measuring Up 2000**, by Dennis Jones and Karen Paulson (June 2001, #01-2). Suggests a range of actions that states can take to bridge the gap between state performance identified in *Measuring Up 2000* and the formulation of effective policy to improve performance in higher education.

***A Review of Tests Performed on the Data in* Measuring Up 2000**, by Peter T. Ewell (June 2001, #01-1). Describes the statistical testing performed on the data in *Measuring Up 2000* by the National Center for Higher Education Management Systems.

***Recent State Policy Initiatives in Education: A Supplement to* Measuring Up 2000**, by Aims C. McGuinness, Jr. (December 2000, #00-6). Highlights education initiatives that states have adopted since 1997–98.

***Assessing Student Learning Outcomes: A Supplement to* Measuring Up 2000**, by Peter T. Ewell and Paula Ries (December 2000, #00-5). National survey of state efforts to assess student learning outcomes in higher education.

***Technical Guide Documenting Methodology, Indicators and Data Sources for* Measuring Up 2000** (November 2000, #00-4).

A State-by-State Report Card on Higher Education: Prospectus (March 2000, #00-1). Summarizes the goals of the National Center's report card project.

Great Expectations: How the Public and Parents—White, African-American and Hispanic—View Higher Education, by John Immerwahr with Tony Foleno (May 2000, #00-2). This report by Public Agenda finds that Americans overwhelmingly see higher education as essential for success. Survey results are also available for the following states:

Great Expectations: How Pennsylvanians View Higher Education (May 2000, #00-2b)
Great Expectations: How Floridians View Higher Education (August 2000, #00-2c)
Great Expectations: How Coloradans View Higher Education (August 2000, #00-2d)
Great Expectations: How Californians View Higher Education (August 2000, #00-2e)
Great Expectations: How New Yorkers View Higher Education (October 2000, #00-2f)
Great Expectations: How Illinois Residents View Higher Education (October 2000, #00-2h)

State Spending for Higher Education in the Next Decade: The Battle to Sustain Current Support, by Harold A. Hovey (July 1999, #99-3). This fiscal forecast of

state and local spending patterns finds that the vast majority of states will face significant fiscal deficits over the next eight years, which will in turn lead to increased scrutiny of higher education in almost all states, and to curtailed spending for public higher education in many states.

South Dakota: Developing Policy-Driven Change in Higher Education, by Mario Martinez (June 1999, #99-2). Describes the processes for change in higher education that government, business, and higher education leaders are creating and implementing in South Dakota.

Taking Responsibility: Leaders' Expectations of Higher Education, by John Immerwahr (January 1999, #99-1). Reports the views of those most involved with decision making about higher education, based on focus groups and a survey conducted by Public Agenda.

The Challenges and Opportunities Facing Higher Education: An Agenda for Policy Research, by Dennis Jones, Peter T. Ewell, and Aims C. McGuinness (December 1998, #98-8). Argues that due to substantial changes in the landscape of postsecondary education, new state-level policy frameworks must be developed and implemented.

Higher Education Governance: Balancing Institutional and Market Influences, by Richard C. Richardson, Jr., Kathy Reeves Bracco, Patrick M. Callan, and Joni E. Finney (November 1998, #98-7). Describes the structural relationships that affect institutional effectiveness in higher education, and argues that state policy should strive for a balance between institutional and market forces.

Federal Tuition Tax Credits and State Higher Education Policy: A Guide for State Policy Makers, by Kristin D. Conklin (December 1998, #98-6). Examines the implications of the federal income tax provisions for students and their families, and makes recommendations for state higher education policy.

The Challenges Facing California Higher Education: A Memorandum to the Next Governor of California, by David W. Breneman (September 1998, #98-5). Argues that California should develop a new Master Plan for Higher Education.

Tidal Wave II Revisited: A Review of Earlier Enrollment Projections for California Higher Education, by Gerald C. Hayward, David W. Breneman, and Leobardo F. Estrada (September 1998, #98-4). Finds that earlier forecasts of a surge in higher education enrollments were accurate.

Organizing for Learning: The View from the Governor's Office, by James B. Hunt Jr., chair of the National Center for Public Policy and Higher Education, and

former governor of North Carolina (June 1998, #98-3). An address to the American Association for Higher Education concerning opportunity in higher education.

The Price of Admission: The Growing Importance of Higher Education, by John Immerwahr (Spring 1998, #98-2). A national survey of Americans' views on higher education, conducted and reported by Public Agenda.

Concept Paper: A National Center to Address Higher Education Policy, by Patrick M. Callan (March 1998, #98-1). Describes the purposes of the National Center for Public Policy and Higher Education.

www.highereducation.org