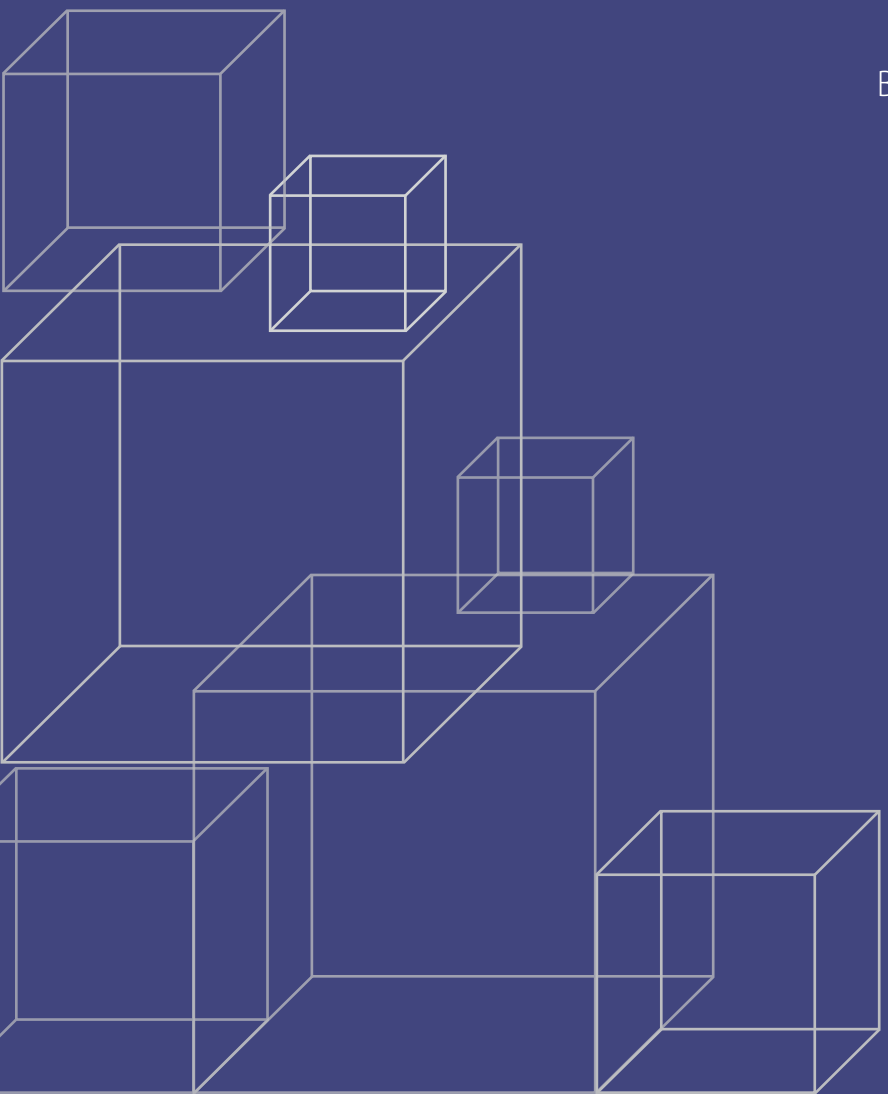
DATA QUALITY CAMPAIGN
Using Data To Improve Student Achievement

www.DataQualityCampaign.org

They Can Pass, but Are They College Ready?

Using Longitudinal Data To Identify College and Career Readiness Benchmarks on State Assessments

By Chrys Dougherty, Ph.D.



October 2008

DATA QUALITY CAMPAIGN

Using Data To Improve Student Achievement

Managing partners of the Data Quality Campaign include:

- Achieve, Inc.
- Alliance for Excellent Education
- Council of Chief State School Officers
- Education Commission of the States
- The Education Trust
- National Association of State Boards of Education
- National Association of System Heads
- National Center for Educational Achievement
- National Center for Higher Education Management Systems
- National Governors Association Center for Best Practices
- Schools Interoperability Framework Association
- Standard & Poor's School Evaluation Services
- State Educational Technology Directors Association
- State Higher Education Executive Officers

Endorsing partners of the Data Quality Campaign include:

- ACT
- Alliance for Quality Teaching
- American Association of Colleges for Teacher Education
- American Association of State Colleges and Universities
- American Board for Certification of Teacher Excellence
- American Youth Policy Forum
- APQC
- Business-Higher Education Forum
- Center for Teaching Quality
- College Summit
- Consortium for School Networking
- Educational Policy Institute
- ETS
- GreatSchools
- Institute for a Competitive Workforce
- Institute for Educational Leadership
- James B. Hunt, Jr. Institute for Educational Leadership and Policy
- Jobs for the Future
- Knowledge Alliance
- League of Education Voters Foundation
- Learning Point Associates
- Midwestern Higher Education Compact
- National Alliance for Public Charter Schools
- National Association of Secondary School Principals
- The National Center for Public Policy and Higher Education
- National Council for Accreditation of Teacher Education
- New England Board of Higher Education
- Pathways to College Network
- Postsecondary Electronic Standards Council
- Pre–K Now
- Roads to Success
- Southern Regional Education Board
- Western Interstate Commission for Higher Education

The Data Quality Campaign is a national, collaborative effort to encourage and support state policymakers to improve the collection, availability and use of high-quality education data and to implement state longitudinal data systems to improve student achievement. The campaign aims to provide tools and resources that will assist state development of quality longitudinal data systems, while also providing a national forum for reducing duplication of effort and promoting greater coordination and consensus among the organizations focusing on improving data quality, access and use.

This publication was produced by the Data Quality Campaign/National Center for Educational Achievement. It was written by Chrys Dougherty, Ph.D., senior research scientist, National Center for Educational Achievement, and produced with support from the Bill & Melinda Gates Foundation.

Table of Contents

Introduction

As of fall 2007, 22 states had set minimum standards for high school graduation on their state tests.[1] But does that mean that students who meet those requirements are ready to succeed in college or in other postsecondary learning opportunities that lead to skilled careers? States that create longitudinal student databases containing the right information will be in a better position to answer this question.

Many state education leaders would like for a high school diploma to mean that a graduate is ready for college and other opportunities. But at the same time, they don't want to set standards at a level that will deny high school diplomas to a high percentage of disadvantaged students. This tension has led to states setting graduation standards that are well below the college readiness level. For example, only 20 percent of graduating high school seniors in Colorado, 21 percent in Illinois and 17 percent in Michigan met ACT's college readiness benchmarks in all four subjects,[2] and a recent report found that more than half of first-time Florida college students attending state public universities and community colleges required remediation, despite the fact that students in that state must pass a state assessment to graduate.[3]

Because meeting state graduation standards alone does not signify college and career readiness, states can use other approaches to identify whether students are college ready. For example, they can require every student to take a college readiness exam, as four states currently do.[4] They can augment the state high school exit test with additional items to measure college readiness, as California does.[5] On the state test, they can identify separate college and career readiness benchmarks — scores that indicate that a student is college and career ready in English, mathematics or science. Or they can use a combination of these approaches.

[1] Center on Education Policy, *State High School Exit Exams: Working to Raise Test Scores*, Washington, DC, September 2007.

[2] See www.act.org/news/data/08/statemenu.html. ACT data are used in this report due to availability; similar analyses can be done with SAT data. Colorado, Illinois and Michigan were chosen as examples because almost all of the graduating seniors in the state in 2008 took the ACT as part of state policy. In Texas and California, where passing the state high school exit exam is required to graduate but fewer students took the ACT, the corresponding percentages were 20 and 28 percent.

[3] Office of Program Policy Analysis and Government Accountability, *Half of College Students Needing Remediation Drop Out; Remediation Completers Do Almost as Well as Other Students*, Tallahassee, FL, 2007. Under these conditions, requiring that all graduates be "college ready" is likely to result in watering down the definition of college readiness.

[4] As of spring 2008, Colorado, Illinois and Michigan administered the ACT to all public high school students in grade 11 statewide, while Maine required the SAT.

[5] California has added to the state's graduation test a voluntary section of additional items that students take to show readiness for the California State University (CSU) system. A standard-setting committee identified percentages of these items that a student should answer correctly to demonstrate partial or full college readiness. Students who demonstrate full college readiness are exempt from placement exams at CSU campuses. See www.cde.ca.gov/ci/gs/ps/eapindex.asp.

Data Requirements for Identifying College Readiness Benchmarks on State Tests

This resource guide focuses on data that states must collect and analyze to identify scores on the state test that provide strong assurance that the student is college ready. To inform the setting of these benchmarks, states need information on the relationships among state test scores and other indicators of student readiness for college and workforce training programs. For example, with a properly designed longitudinal student information system, the state can match individual students' state test scores to three kinds of information on the same students:

- **Scores on national exams designed to assess college readiness**, such as SAT and ACT — element 7 of the Data Quality Campaign's (DQC) 10 essential elements of longitudinal data systems (see box, page 3);
- **Scores on college course placement exams**, particularly those that assess students' readiness for credit-bearing college courses in specific subjects — element 9; and/or
- **Students' actual college enrollment and their completion of college degrees or workforce certifications** within a specified number of years — element 9.

Connecting student ACT and SAT scores to state test results requires the state to obtain the student-level scores from ACT, Inc., and the College Board and to match those records to the state test score records as accurately as possible. This matching can be done more accurately if the state collects student Social Security numbers. These numbers also are included in the ACT and SAT databases and can be used to verify, for example, that "Billy Jones" in the 11th grade at XYZ High School is indeed the same student whose name appears as "Charles William Jones" in a second database and "Chas W. Jones" in a third database.

Determining the relationship between how well students perform on the state test and the same students' success in higher education requires that a single entity acquire and match the P–12 and higher education records for the same students. This matching requires pulling the P–12 test records for individual students together in a single database with the enrollment and remediation records for all of the state's public two- and four-year colleges.

Elements and Components of Longitudinal Data Systems

The DQC has identified *10 essential elements* that states must include to build a highly effective longitudinal data system:[6]

1. A unique statewide student identifier that connects student data across key databases across years
2. Student-level enrollment, demographic and program participation information
3. The ability to match individual students' test records from year to year to measure academic growth
4. Information on untested students and the reasons they were not tested
5. A teacher identifier system with the ability to match teachers to students
6. Student-level transcript information, including information on courses completed and grades earned
7. Student-level college readiness test scores
8. Student-level graduation and dropout data
9. The ability to match student records between the P–12 and higher education systems
10. A state data audit system assessing data quality, validity and reliability

As outlined in *Creating a Longitudinal Data System: Using Data To Improve Student Achievement*,[7] a white paper by the DQC, a longitudinal data system also needs to include the following components:

- **A technology infrastructure.** Schools, districts and state agencies have access to computers, servers, networks and the Internet to collect, transfer and use data.
- **A data architecture** that defines how data are coded, stored, managed and used. Data definitions are important. When everyone uses standard definitions, different systems can share information, staffing resources and process time are minimized, and data are provided to users when they need them. Privacy protection measures allow unique student identifiers to be used without revealing the data associated with a specific student when the data are shared with other organizations. Security protocols, like encryption, allow the secure transmission of data among systems.
- **A data warehouse** that stores, organizes and links student, school and district information — over time. Warehouses are designed to make it easy for users to "query" the database and produce standard or customized reports for different stakeholders. Researchers can use the data warehouse to answer questions such as the value-added of schools, identify which programs work for which students or identify which schools are closing the achievement gap — without violating student privacy.
- **Ongoing professional development** for those who are charged with collecting, storing, analyzing and using the data. Training ranges from how data are input locally to how teachers access and use the data for school and instructional improvement to how state education leaders use the system to make policy changes. Professional development continues as the system is refined and gains capacity for data-driven decisionmaking.

[6] Data Quality Campaign, *Creating a Longitudinal Data System: Using Data To Improve Student Achievement*, 2006, www.DataQualityCampaign.org/files/Publications-Creating_Longitudinal_Data_System.pdf.
[7] Ibid.

How Many States Meet These Requirements?

According to the annual survey of state data systems conducted by the DQC, only 15 states reported that they were collecting student-level ACT and/or SAT scores (element 7), and of those, only six collected the student Social Security numbers needed to make the most accurate matches. Meanwhile, 22 states reported that they could link individual students' P–12 and higher education records (element 9), although these states may face varying levels of difficulty in consolidating all of their student higher education records.[8]

Using all of this information to create benchmarks still requires judgment on how high to set the benchmark. How high an SAT or ACT score, or how low a probability that the student will need remediation in college, is sufficient? The data can provide information on the relationship between early measures and later success but cannot directly answer the question that must be addressed by educators and policymakers, "How good is good enough?"

In addition, the state test must cover sufficiently challenging content to make a college readiness benchmark on that test meaningful. For example, if the state test covers content only through 7th or 8th grade, a college readiness benchmark would exceed the ceiling on that test.[9] A state assessment system that adequately covers the content of high school courses traditionally treated as college preparatory should meet this requirement. This goal is best accomplished with a statewide end-of-course exam for each core high school course.[10]

[8] See www.DataQualityCampaign.org/survey_results/elements.cfm.

[9] An alternative approach for a state with a weak high school exam but a relatively strong 8th grade exam would be to match the 8th grade data at the student level with results on a national 8th grade exam, such as EXPLORE, that has been linked to college readiness standards on the ACT. This approach requires that a sufficiently representative sample of students in the state take EXPLORE.

[10] Given that many subjects such as algebra don't differ that much across states, end-of-course exams for use in multiple states have been developed by Achieve for Algebra II and by ACT for 14 courses through its QualityCore program.

Identifying College Readiness Benchmarks — A Case Study

The example in this resource guide is taken from Texas, which has had a statewide longitudinal data system since 1990, has collected student-level SAT and ACT test scores, and has been matching P–12 and higher education student records since the late 1990s. However, any state that develops these capabilities as part of its longitudinal student data system could do the following type of analysis using its own data.

Since 1989, all Texas two- and four-year public colleges and universities have administered placement exams to identify whether students are ready to take credit-bearing courses. Students may be exempted from these exams based on their SAT, ACT or state exam scores, using criteria that have varied over the years. This information on placement and the need for remediation has been collected on all Texas public college students by the Texas Higher Education Coordinating Board (THECB).

When the Texas Assessment of Knowledge and Skills (TAKS) exam was introduced in 2003, researchers from THECB and the Texas Education Agency (TEA) sought a benchmark score on the TAKS test that could be used to indicate that the student is college ready, eliminating the need to give the student a placement exam when entering college. Researchers from the two agencies linked TAKS scores with the same students' scores from the SAT; the ACT; and the Texas Higher Education Assessment (THEA), the most common placement exam given in Texas public colleges and universities. The results from this analysis are summarized in Table 1.

The information in the table can be interpreted as follows: A student with a score of 2100 on the TAKS English Language Arts (first row of data) would be predicted to earn a score of 17.7 on the English ACT and 461 on the Verbal SAT and would have a 57 percent chance of meeting the THEA standard indicating that he or she was

Table 1

Relationship of 11th Grade TAKS Scores to College Readiness Measures Based on Linking Student-Level Data

	11th grade TAKS score	Predicted ACT score	Predicted SAT score	Approximate probability of THEA score > 230**	Approximate probability of THEA score > 270**
English Language Arts*	2100	17.7	461	57%	n/a
	2200	20.1	502	77%	n/a
	2300	22.5	543	90%	n/a
	2400	24.9	584	100%	n/a
Mathematics	2100	19.5	472	67%	5%
	2200	21.9	521	90%	26%
	2300	24.3	570	100%	77%
	2400	26.7	618	100%	100%

*The 11th grade TAKS English Language Arts covers both reading and writing.
**The state used a score of 230 or higher on the THEA to indicate that the student was ready to take credit-bearing college courses and did not need remediation, while a score of 270 on the mathematics THEA was treated informally as an indicator of readiness for college algebra.
Source: Texas Higher Education Coordinating Board, 2003

ready to take credit-bearing English courses and did not need remediation in that subject. A student with a 2100 on the TAKS Mathematics had a two-thirds chance of not needing remediation in mathematics but only a 5 percent chance of being ready for college algebra. In 2005, 2100 on the TAKS test became the state passing standard required for graduation.

To set a college readiness benchmark on the state test using this information, THECB had to weigh the competing advantages of higher and lower standards. Set the standard too high, and aspiring students might get the word that they are "not college ready" and be discouraged from attending college. Set the standard too low, and students might struggle in credit-bearing courses in college because they were not identified as needing remediation in advance and did not receive help. As in many other situations, policymakers were concerned about both false positives (identifying a student as needing remediation when he or she does not need it) and false negatives (failing to identify a student who does need remediation).[11] With these considerations in mind, the agency set a college readiness benchmark of 2200 on both the TAKS English Language Arts and Mathematics exams.

[11] False positives and false negatives are sometimes called Type I and Type II errors, respectively.

Using Multiple Years of Data To Validate College Readiness Benchmarks

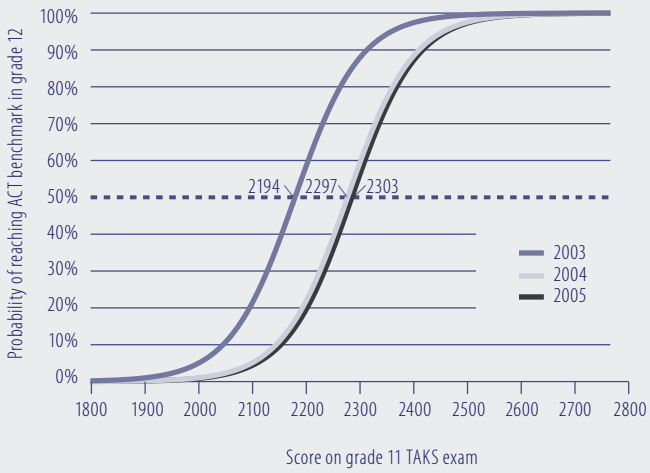
Texas set its benchmark based on the first year of administration of the high school exam, when it did not count for graduation. Because state test scores often rise rapidly in the second year of administration (due to increased familiarity with the test) and when the exam starts to count for graduation (due to increased student effort on the test), the information in Table 1 is likely to understate the TAKS score a student would need to receive in subsequent years to achieve the ACT, SAT and THEA outcomes predicted in the table.

Therefore, states and researchers need access to the data to check for shifts over time. Researchers from the National Center for Educational Achievement (NCEA), a nonprofit organization that manages the DQC and is involved in school improvement efforts, performed such an analysis using three years of matched TAKS and ACT data.

As shown in Figure 1, the relationship between TAKS and ACT mathematics results did shift between the 2003 and 2004 grade 11 TAKS administrations but stabilized with the TAKS administration in spring 2005. In 2003, a TAKS Mathematics score of 2194 corresponded to a 50 percent probability that a student would meet the ACT mathematics college readiness benchmark in 2004.[12] Yet in 2004 and 2005, the relationship shifted so that higher TAKS Mathematics scores of 2297 and 2303, respectively, were required to have the same odds of meeting the ACT benchmark.[13] This shift is shown by the curve in Figure 1. NCEA researchers viewed this analysis as supporting a choice of 2300 as the TAKS Mathematics college and career readiness benchmark.[14]

Figure 1

Relationship of TAKS Mathematics Performance to the Probability of Reaching ACT's College Readiness Benchmark in Mathematics



Source: NCEA analysis, 2008

[12] ACT established its college readiness benchmarks in 2005 by matching student-level ACT test records with information on the same students' grades in freshman credit-bearing courses in college. The benchmarks correspond to a 50 percent probability that a student will receive a B or better in the appropriate credit-bearing course in the "median" college and a 75 percent probability that the student will receive a C or better. This analysis produced college readiness benchmarks of 18, 21, 22 and 24, respectively, on the ACT English, Reading, Mathematics and Science exams. See J. Allen and J. Sconing, 2005, *Using ACT Assessment Scores to Set Benchmarks for College Readiness*, www.act.org/research/researchers/reports/index.html.

[13] 2297 and 2303 were essentially the same score because meeting or exceeding these scores required students to get the same number of items correct.

[14] For a fuller discussion of NCEA's reasoning in preferring the higher standard, see C. Dougherty, L. Mellor and N. Smith, *Identifying Appropriate College Readiness Standards for All Students*, 2006, www.just4kids.org/en/files/Publication-Identifying_Appropriate_College-Readiness_Standards_for_All_Students-05-03-06.pdf.

Identifying Benchmarks for Ensuring Students in Earlier Grades Are on Track to College and Career Readiness

Once college readiness benchmarks on the state exam have been identified in the upper grades, similar benchmarks may be identified in the earlier grades to indicate whether a student is on track to being college and career ready by the time he or she leaves high school. If test score records for the same students are available for grades 8 and 11, for example, the probability that a student will meet the benchmark on the state test in grade 11 based on the student's 8th grade state test score can be estimated by doing an analysis similar to the one illustrated in Figure 1 on page 7.

However, because of changes in state tests and the newness of longitudinal data systems, states typically do not have 3rd grade and 11th grade scores for the same students on the same state test. The state would have had to administer a test very similar to the current 3rd grade state test eight years earlier, and those 3rd grade scores would have to be matched to scores for the same students eight years later.

A second and more feasible approach is to pick a base year (usually the year after a test was first introduced) and to treat scores in the same place on the achievement distribution in different grades in that year as comparable. Thus, for example, a student scoring a certain distance above the average in the 3rd grade test score distribution would be treated as on track to earn a score the same distance above the average in 11th grade.[15]

[15] Scores at the same percentile (e.g., the 84th percentile) or the same number of standard deviations above the average in 3rd grade and 11th grade can be treated as "the same distance above the average." The approach using standard deviations, called "statistical moderation," is used in Dougherty, Mellor and Smith, 2006, as well as in two major studies — one comparing scores on state tests with those on the National Assessment of Educational Progress (NAEP) by the Institute of Education Sciences, 2007, and one comparing state average NAEP scores with those of various countries on the Trends in International Math and Science Study by Gary Phillips of American Institutes for Research, 2007.

Recommendations

As educators and state policymakers consider the importance of preparing students for college and skilled careers, they should be identifying whether students are on track to college and career readiness, starting in the early grades. States with longitudinal data systems can do this identification in the upper grades by directly matching state test scores to scores on the SAT, ACT and college placement exams that are used to determine whether students need remediation. This matching requires the use of a statewide longitudinal data system that can match state test, SAT, ACT and higher education records for the same students. The state then should extrapolate these benchmarks down into the lower grades.

The fact that college and career readiness benchmarks may not match up with the proficiency standards used for NCLB and state accountability systems should not deter states and school systems from paying attention to those benchmarks. As educators increasingly focus on academic growth, they should monitor student growth across multiple achievement levels (e.g., below basic, basic, proficient, college and career ready, and possibly an advanced standard for college and career readiness). Statewide longitudinal student data systems not only will help educators and policymakers set appropriate benchmarks, but they also will make it possible to identify and learn from the schools and school systems that are doing the best job of producing student growth toward these benchmarks.

References

J. Allen and J. Sconing, *Using ACT Assessment Scores to Set Benchmarks for College Readiness,* 2005, www.act.org/research/researchers/reports/index.html.

California Department of Education, Early Assessment Program, www.cde.ca.gov/ci/gs/ps/eapindex.asp.

Center on Education Policy, *State High School Exit Exams: Working to Raise Test Scores,* Washington, DC, September 2007, www.cep-dc.org.

Data Quality Campaign, *Creating a Longitudinal Data System: Using Data To Improve Student Achievement,* 2006, www.DataQualityCampaign.org/files/Publications-Creating_Longitudinal_Data_System.pdf.

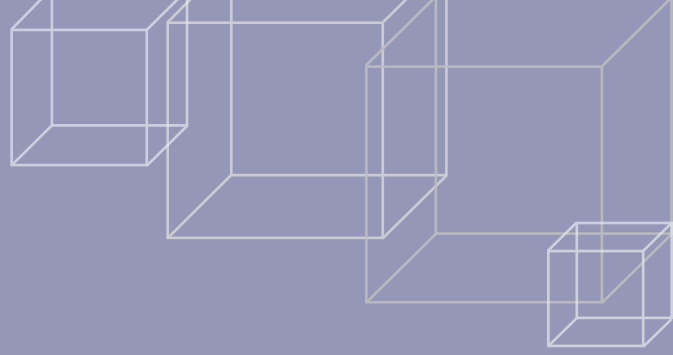
C. Dougherty, L. Mellor and N. Smith, *Identifying Appropriate College Readiness Standards for All Students,* 2006, www.just4kids.org/en/files/Publication-Identifying_Appropriate_College-Readiness_Standards_for_All_Students-05-03-06.pdf.

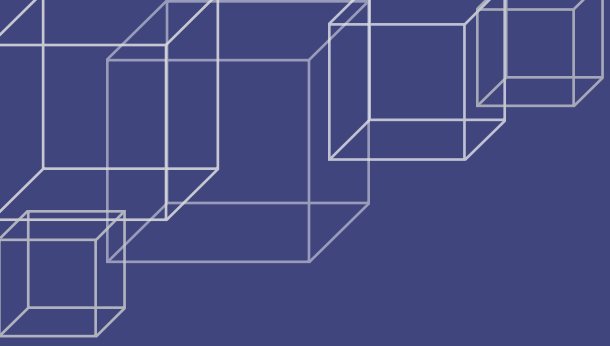
C. Dougherty, L. Mellor and N. Smith, *Identifying College and Career Readiness Benchmarks on State Assessments: An Update*, forthcoming.

Institute for Educational Sciences, *Mapping 2005 State Proficiency Standards onto the NAEP Scales,* NCES, 2007, http://nces.ed.gov/nationsreportcard/pdf/studies/2007482.pdf.

Office of Program Policy Analysis and Government Accountability, *Half of College Students Needing Remediation Drop Out; Remediation Completers Do Almost as Well as Other Students,* Tallahassee, FL, 2007, www.oppaga.state.fl.us/reports/pdf/0731rpt.pdf.

G.W. Phillips, *Expressing International Achievement in Terms of U.S. Performance Standards: Linking NAEP Achievement Levels to TIMSS,* American Institutes for Research, Washington, DC, 2007, www.air.org/news/documents/naep-timss.pdf.





DATA QUALITY CAMPAIGN

Using Data To Improve Student Achievement

www.DataQualityCampaign.org
Phone: 512.320.1816
Fax: 512.320.1877