Creating a Longitudinal Data System

Using Data To Improve Student Achievement

DATAQUALITY
CAMPAIGN

www.DataQualityCampaign.org

Update 2006

Does your state collect the data you need to answer these questions?

▶ Which schools produce the strongest academic growth for their students? *(23 states report having the data to answer this question)*

▶ What achievement levels in middle school indicate that a student is on track to succeed in rigorous courses in high school? *(5 states)*

▶ What is the state's graduation rate, according to the calculation agreed to in the 2005 National Governors Association compact? *(26 states)*

▶ What high school performance indicators (e.g., enrollment in rigorous courses or performance on state tests) are the best predictors of students' success in college or the workplace? *(4 states)*

▶ What percentage of high school graduates take remedial courses in college? *(14 states)*

▶ Which teacher preparation programs produce the graduates whose students have the strongest academic growth? *(9 states)*

THE DATA QUALITY CAMPAIGN AT YEAR ONE

Creating the Will and Understanding To Use a Statewide Longitudinal Data System

A consistent refrain is heard in state houses, in district offices and in school staff meetings: "We need quality education data to do our jobs."

Good information is critical to both the processes and the outcomes of a high-performing education system, and increasing numbers of voices — inside and outside the system — are calling for this issue to be a national priority. In November 2005, 10 organizations launched a national campaign *to improve the collection, availability and use* of high-quality education data. The Data Quality Campaign (DQC) aims not only to have longitudinal data systems in place in every state by 2009 but also, equally important, to change the culture surrounding data use in education.

Shining the Spotlight on Essential Data and on State Data System Progress
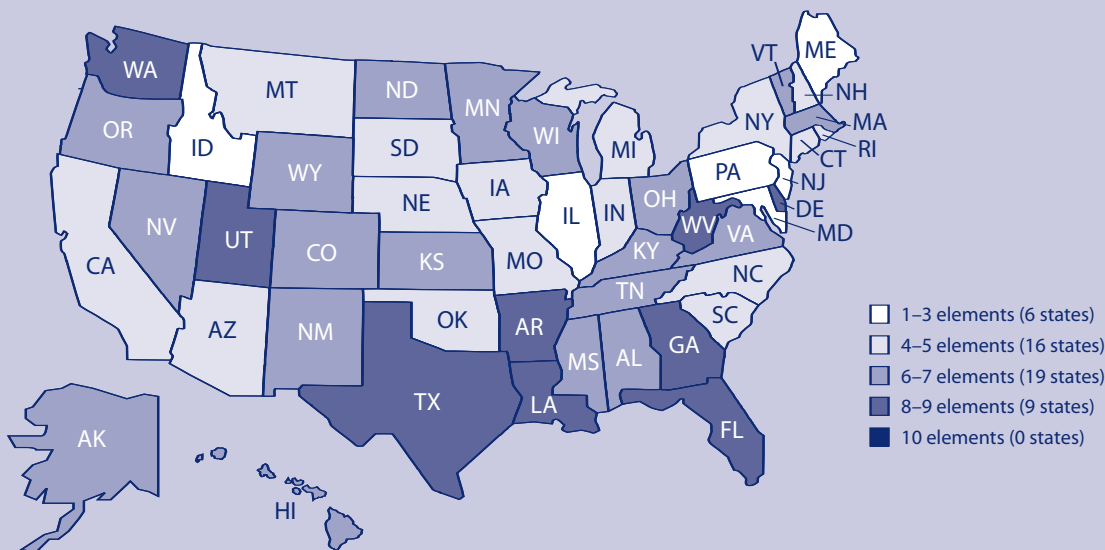
Over the past year, the DQC has generated power behind the issues of data collection, availability and use. A growing network of committed partners at the national, state and institutional levels is shining a brighter and wider spotlight on the pragmatic ways to build and use these longitudinal unit record data systems. We have evidence that this intense focus is bearing results. It is widely acknowledged that vital policy conversations now under way — conversations about increasing the rigor and relevance of high school, improving teacher quality, promoting higher graduation rates and reducing achievement gaps among student populations — cannot be successful unless they are informed by reliable longitudinal data.

Just as more education leaders are recognizing the need for better data, more states are doing the hard work of addressing that need by putting in place the DQC's 10 essential elements of a longitudinal data system. To date:*

▶ 42 states report having a unique student identifier (an integral part of a longitudinal data system) in place, up from 36 last year.

▶ Nine states have eight or nine of the essential elements (no state reports having all 10), while only six states have three or fewer.

More States Are Putting Essential Data System Elements in Place, 2006

Legend:
- ☐ 1–3 elements (6 states)
- ☐ 4–5 elements (16 states)
- ☐ 6–7 elements (19 states)
- ☐ 8–9 elements (9 states)
- ■ 10 elements (0 states)

▶ 26 states indicate they have or are working on building data warehouses.

▶ 36 states (up from 19 states last year — the most significant increase among the 10 elements) have put in place audit systems to ensure quality data.

▶ 28 states have some form of Web-based data and analysis tools available for local educators.

The progress we have made in the past year is encouraging, but there is still much work to do. This campaign is more than a checklist of elements in a data system. We must build the will and understanding to use this information for policy and program decisions. Policy conversations will be far better informed if they are based on information from longitudinal, student-level data. And teachers and administrators will be better able to tailor instruction and programs to individual student needs if they can access and use longitudinal data. Therefore, states must continue to build and maintain these functional, informative data systems to address such pressing issues as:

▶ Identifying which schools produce the strongest academic growth for their students. **(23 states report having the data systems to do so)***

▶ Knowing what achievement levels in middle school indicate that a student is on track to succeed in rigorous courses in high school. **(5 states)**

▶ Calculating the state's graduation rate, according to the 2005 National Governors Association graduation compact. **(26 states)**

▶ Determining which high school performance indicators (e.g., enrollment in rigorous courses or performance on state tests) are the best predictors of students' success in college or the workplace. **(4 states)**

▶ Identifying the percentage of high school graduates who take remedial courses once enrolled in college. **(14 states)**

▶ Improving communication and collaboration between postsecondary and P–12 systems. (State education staff respondents to the DQC survey indicated that **18 states** *had the ability* to link postsecondary and P–12 data systems. But according to a National Center for Higher Education Management Systems survey of postsecondary data managers, **11 states** report that they *actually have linked* student postsecondary data with high school records, and **10 states** use their higher education student unit record systems to regularly produce feedback reports to high schools by district on such topics as need for remediation, credits enrolled for and grade performance.)

▶ Identifying which teacher preparation programs produce the graduates whose students have the strongest academic growth. **(9 states)**

DQC: A Catalyst To Advance the Power of Longitudinal Data

For years, states have built their data systems and acknowledged the relationship of data to sound education decisionmaking. Over the past 12 months, the DQC has raised the conversation to a national priority and provided a sense of urgency to the states' work. DQC is aggressively fostering the demand for better, more "actionable" data among all constituencies, while also making the case for increased investments to build effective systems that can meet this demand. How are we doing this?

Raising the chorus of voices to champion the use of longitudinal data. We continue to expand our partnership to ensure that a broad array of national organizations embrace and incorporate the DQC goals into their state-focused initiatives. In the past, an organizer for a policy conference most likely would have minimized a session about building education data systems, but in the past year, most of the major national organizations have addressed this topic in their plenary sessions. They also have featured the campaign, and the importance of building and using these

*Although states report having the data elements to address this issue, the National Center for Educational Accountability/DQC survey did not capture whether states actually were using the data to this end.

systems, on their Web sites and in scores of publications. Data are no longer a sleeper topic that "someone else will deal with." A growing number of organizations view quality data systems as an integral component of their education reform agendas.

Marshalling the experience and expertise of state data directors. States are at various stages of building and using education data systems, as evidenced by the map on page 1. The DQC provides a forum for states to learn from each other as they continue to improve their systems. Our campaign is informed by a Task Force on Longitudinal Student Data Systems managed by the Council of Chief State School Officers' Education Information Management Advisory Consortium to ensure that we are being responsive and helpful to the needs of states. We have captured the lessons from states that have built longitudinal data systems, disseminated case studies and resource briefs, and convened forums for states to learn from their peers.

Building a national forum to reduce duplication and increase efficiency. DQC provides cohesiveness and leadership for the previously fragmented and disconnected conversations among data managers and policymakers nationwide. We have taken the initiative to develop a coordinated data request from states that can be used by national organizations and research entities. This effort has been applauded by the states, and some have reported that they already see a reduction in the number of "data asks" made by national groups. These national organizations — thanks to the network of the DQC — are now regularly sharing their plans for data requests and surveys of the states to work together rather than duplicating each others' efforts.

Expanding the "bully pulpit" for data issues. The DQC provides a unified voice on data issues on behalf of its 14 managing partners and growing number of endorsing partners. The DQC partners are increasingly seen as the authoritative resource for information on longitudinal data systems and are called on to make presentations, offer commentary, publish articles and serve on advisory committees related to cutting-edge data topics. To promote the essential role of longitudinal data in today's key policy conversations, the DQC convenes quarterly gatherings of national leaders, produces practical information and action items, and conducts Webcasts to spread these messages nationwide. A growing stock of products and tools (see box) makes the case for policymakers and stakeholders that the value of longitudinal data outweighs the challenges of implementing such systems and that help is available.

Promoting, Developing and Using Longitudinal Data Systems: DQC Resources

To download full copies of these documents, go to www.DataQualityCampaign.org.

- *Aligning P–12 and Postsecondary Data Systems*

- *Building Student-Level Longitudinal Data Systems: Lessons Learned from Four States*

- *Creating a Longitudinal Data System: Using Data To Improve Student Achievement*

- *Every Student Counted: Using Longitudinal Data Systems To Calculate More Accurate and Useful High School Graduation Rates To Improve Student Success*

- *Maximizing the Power of Education Data while Ensuring Compliance with Federal Student Privacy Laws: A Guide for State Policymakers*

- *Measuring What Matters: Creating a Longitudinal Data System To Improve Student Achievement*

- *Using Data in the Central Office and the Classroom To Improve Student Achievement*

Driving the Use of Longitudinal Data: A Look Ahead at Year Two

In year two, the DQC will focus on promoting the use of longitudinal student-level data for accountability purposes and for tailoring instructional programs and policies, while continuing to support state efforts to build longitudinal systems. Many educators harbor negative perceptions of data because in the past the data have been incomprehensible, unhelpful or used solely for compliance purposes. Our goal is to demonstrate to teachers, administrators and other education partners that high-quality, easily accessible and functional longitudinal data are among the most powerful tools they can use in their efforts to prepare every student for success in an increasingly knowledge-based world.

As the DQC begins its second year, this focus on promoting the powerful and indispensable use of data to all education stakeholders will be pervasive. The following actions will guide the campaign's work:

▶ **Build longitudinal data systems with end users in mind.** As states design and construct these systems, it is vital that they do so with the needs in mind of the people who will be using the data — teachers, parents, administrators, students, researchers and policy-makers. We cannot wait until these systems have been completed to start building the knowledge and capacity of individuals to use the information.

▶ **Create toolkits for education stakeholders that demonstrate the power of longitudinal data** to inform their policy and program-matic decisions and that guide them on how to begin using these data effectively. *(The first toolkit targets principals, teachers and school-based leaders.)*

▶ **Advocate for continued investments in state data systems** — not just to build and maintain longitudinal data systems, but to ensure all users are trained to understand how these data can inform decisions in teaching and management. We will continue to promote the crucial need to invest in building and using longitudinal data systems with both national and state policymakers.

▶ **Generate opportunities for states to learn from one another** as they build, maintain and use state longitudinal data systems. The DQC will continue to develop resource and implementation materials for building successful longitudinal data systems and will highlight the efforts of various states that are leaders in the field. We will reach out to states that are further behind in establishing these systems to provide targeted assistance.

As we begin our second year of the campaign, all of the DQC partners commit to reducing duplication and promoting efficiencies not only in our organizational work but also in data collection initiatives. This partnership and campaign, by their very existence, reinforce the integral role of data in the increasingly aligned and complementary agendas of each of our individual organizations. We look forward to working in partnership with each other and with state leaders to meet our goal of having longitudinal data systems completed, accessible and used to inform the important discussions under way about improving America's schools. Our nation and our children cannot afford to wait for better information.

Data Quality Campaign Managing Partners
November 2006

Introduction

A rapidly changing global economy and concerns about our ability to create a competitive workforce have focused national attention on the education systems of America's states, highlighting their critical role in ensuring a prosperous future for our country. As a result, American schools are expected to perform better than ever before in preparing all students to meet rigorous educational requirements for postsecondary education and the workplace.

Educators and policymakers are recognizing the value of *better information* as an essential tool for improving schools. They understand that when states collect the most relevant data and are able to match individual student records *over time,* they can answer the questions that are at the core of educational effectiveness. Longitudinal data (data gathered on the same student from year to year) make it possible to:

▶ follow students' academic progress as they move from grade to grade;

▶ determine the value-added and effectiveness of specific schools and programs;

▶ identify consistently high-performing schools so that educators and the public can learn from best practices;

▶ evaluate the effect of teacher preparation and training programs on student achievement; and

▶ focus school systems on preparing a higher percentage of students to succeed in rigorous high school courses, college and challenging jobs.

Policymakers and educators need longitudinal data systems capable of providing timely, valid and relevant data. Access to these data gives teachers the information they need to tailor instruction to help each student improve, gives administrators the resources and information to effectively and efficiently manage, and enables policymakers to evaluate which policy initiatives show the best evidence of increasing student achievement.

Essential Elements of a State Longitudinal Data System

Although each state's education system is unique, 10 elements are essential in a longitudinal data system:

1. A unique statewide student identifier.

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.

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 postsecondary systems.

10. A state data audit system assessing data quality, validity and reliability.

In building a statewide data system with each of these components, the state should ensure that student records can be easily transferred, student privacy is protected, data definitions and requirements are clear to all concerned, and the data system is organized in ways that facilitate data use and user-friendly reporting.

The 10 Essential Elements in Detail

1   A Unique Statewide Student Identifier

36 ▷ 42   *Growth in states that report having this element, 2005 to 2006* [1]

A unique statewide student identifier is a single, non-duplicated number that is assigned to and remains with a student throughout his or her P–12 career. A student who leaves the state and returns should be assigned his or her original number.

A student identifier will allow the state to follow the progress of each student over time, from prekindergarten though grade 12, and across campuses or districts within the state while ensuring his or her privacy. It also makes it possible to identify information about a single student across various data sources (e.g., enrollment, program participation, demographics and assessment) to evaluate the relationship between program participation and performance and to study student mobility patterns and evaluate the effect of mobility on performance.

A statewide student identifier can help policymakers and educators know, among other things:

▸ The academic value-added of a school or program.

▸ The achievement levels in early grades that indicate that a student is on track to succeed in subsequent grades.

▸ The test scores in early grades that should be thresholds for intervention.

Action Steps for Policymakers*

◾ Every student in the state is assigned a unique student identifier that is consistent throughout his or her public education in the state.

2   Student-Level Enrollment, Demographic and Program Participation Information

38 ▷ 46   *Growth in states that report having this element, 2005 to 2006*

Accurate information on student enrollment, demographics and program participation (e.g., student participation in special education or the free and reduced-price lunch program, which is the most common indicator of student poverty status) is essential to evaluate the effects of schools and programs on student achievement and to assess the impact of student mobility and continuous enrollment on learning.

Correct student demographic and program participation information also is critical for the accurate disaggregation of test scores, assuming that a student identifier can connect these data to test results. For example, test scores

---

[1] In September 2006, the National Center for Educational Accountability, with support from the Bill & Melinda Gates Foundation and The Broad Foundation, administered a survey of all 50 states, territories and the District of Columbia to update the results of the 2005 survey, which captured how many states already have the 10 essential elements in place. Similar surveys were conducted in 2003 and 2004. Fifty states responded to the survey, and results can be viewed at www.DataQualityCampaign.org.

*See Appendix for examples of State Education Agency Actions associated with each essential element.

cannot be correctly disaggregated for special education, English language learner, or free and reduced-price lunch students unless accurate information on these students is connected to the test database.

Attendance data also can be an important indicator of student motivation and the likelihood that students will drop out. Collecting information on the percentage of days that students attend school (and specific classes if a state collects attendance at that level) in a given six- or nine-week period and matching those data to results on subsequent statewide assessments or end-of-course exams can help identify students who fail those exams even though their attendance rates are high.

With student-level enrollment, demographic and program participation information, policymakers and educators will know:

▶ The extent to which free and reduced-price lunch enrollment drops off in high school, and how that might affect measures of each high school's poverty rate.

▶ How the percentage of minority students in gifted and talented programs compares with that of white students.

▶ The rate at which English language learners are entering the state for the first time in high school, and how they are doing on the state's high school exams.

Action Steps for Policymakers

■ Enrollment, demographic and program participation data are collected and permanently stored at the state education agency so they can be matched across years. At a minimum, these data are collected at least one time during the year other than when the state assessment is administered. Ideally, states will collect this information continually or at multiple times during the year.

■ Attendance data are collected and permanently stored at the state education agency. At a minimum, these data are collected over six- or nine-week increments.

3 The Ability To Match Individual Students' Test Records from Year to Year To Measure Academic Growth

32 ▷ 41 *Growth in states that report having this element, 2005 to 2006*

A statewide database of individual student results on state exams and state-mandated local exams should be maintained with the ability to disaggregate the results by individual item and objective. This will allow the state to provide good diagnostic information to teachers. Though most states do have annual test records for individual students, only some of these

states have created the ability to match records for individual students across time and with other databases (e.g., enrollment, course completion and graduation databases).

Combining a student-level statewide test database with the ability to match individual student records over time allows states to monitor student academic

growth and provide valuable diagnostic information from the student's academic history to teachers and principals. Having this information stored in a statewide database makes it easier for the information to follow the student as he or she changes grades, schools or districts.

With the ability to match individual students' test scores from year to year, policymakers and educators will know (by grade and subject):

▶ The percentage of students who were below proficient last year but who met the state's proficiency standard this year.

▶ Whether or not proficient and advanced students are achieving at least a year's growth every year.

4 Information on Untested Students

25 ▷ **30** *Growth in states that report having this element, 2005 to 2006*

Too often students "fall through the cracks" in educational systems. To address this problem, the No Child Left Behind Act of 2001 requires states to keep track of the number and percentage of students who do not take the state tests. However, states need to go one step further to find out *why* these students are not tested and then match their records to separate enrollment and program participation databases. This makes it possible to identify patterns associated with specific student populations (e.g., special education students or English language learners) and ensure that all students are held to high expectations.

With information on untested students, policymakers and educators will know:

▶ Which students were not tested by grade and subject and why.

▶ Whether there are trends over time in the number and percentage of untested students from each student group (e.g., English language learners, special education students, different ethnic groups).

▶ Whether or not particular schools and districts have excessive absences on test day or questionable patterns of absences and exemptions across years (these measures can be used in a state's data audit system to ensure data quality).

5 A Teacher Identifier System with the Ability To Match Teachers to Students

13 ▷ 16 *Growth in states that report having this element, 2005 to 2006*

Many states collect data on teacher education and certification, but matching teachers to students by classroom and subject is critical to understanding the connection between teacher training and qualifications and student academic growth. As with students, teachers should be given a unique statewide identifier that follows them over time. This makes it possible to determine which students and which courses are being taught by teachers with different levels and types of preparation or certification and which forms of teacher training and certification have the greatest effect on students' academic growth in the classroom. Such a match makes it possible to evaluate the effectiveness of teacher preparation programs, including traditional and alternative certification programs, based on students' academic progress. Combining this information with student demographic information also allows states to determine the experience level of the teachers teaching low-income or special needs students.

With a teacher identifier and the ability to connect teacher and student data, policymakers and educators will know:

▶ The teacher preparation programs that produce graduates whose students have the strongest academic growth.

▶ How the experience levels of the teachers in the district's high-poverty schools compare with those of teachers in the schools serving affluent students, and how these experience levels are related to the academic growth of the students in their classrooms.

▶ The relationship between the performance of the district's low-income students on the state math exam and teacher preparation in that subject.

Action Steps for Policymakers

■ Every public education teacher in the state is assigned a unique teacher identifier that is consistent throughout his or her teaching career in the state.

■ Teacher preparation information is maintained by and/or shared with the state education agency.

■ Teacher data can be matched to data on students in each of the teachers' classes.

6 Student-Level Transcript Information, Including Information on Courses Completed and Grades Earned

7 ▷ 12 *Growth in states that report having this element, 2005 to 2006*

Many states are encouraging students, particularly low-income and minority students, to take rigorous courses in high school so that they are better prepared for success in postsecondary education and the job market. In most states, however, course-taking data are not collected at the state level, making it impossible to monitor the effects of these policies. To fill in the missing information, states should collect student-level transcript information from middle and high schools, including courses taken and grades earned. These data can be aggregated to the school and district levels to track patterns in overall student course-taking and the relationship between those courses and student readiness for college and work.

Because titles for the same courses may not be consistent, and some titles may not match the material taught in the classroom, states should explore creating a statewide course classification system (numbering and course description/naming). This system should be based on the state standards that are covered in the course. For example, the title Algebra I could be applied to courses designed to cover certain content; courses that cover only part of this content would receive a different title.

To make sure that students are actually learning the content implied by the course titles (e.g., students actually learn algebra in courses titled Algebra I), information on students' course completion and grades should be linked to other measures of learning, such as scores on state end-of-course tests and college readiness exams, and to data on students' need for remediation in college.

With student-level transcript information, including courses completed and grades earned, policymakers and educators will know:

▶ The number and percentage of students who are enrolling in and completing rigorous courses in high school, disaggregated by ethnicity and income status.

▶ The middle schools that are doing the best job of preparing students for rigorous courses in high school.

▶ Whether students in more rigorous courses in high school have been more successful in college or in the workplace.

▶ Whether there is evidence of grade inflation (e.g., students with the same test scores receiving dramatically higher grades in the same course in certain schools or districts).

Action Steps for Policymakers

■ A standardized course classification (numbering and course description/naming) system is developed and used by the state.

■ Course completion records and grades for all courses taken in middle and high school are collected for all students.

Student-Level College Readiness Test Scores

7 ▷ 9 *Growth in states that report having this element, 2005 to 2006*

To ensure that students make a successful transition from high school to postsecondary education, it is important for states to collect and report student performance data on college admissions, placement and readiness tests. Student performance on SAT, SAT II, ACT, Advanced Placement (AP) and International Baccalaureate (IB) exams is an important indicator of students' college readiness; states should collect and report these data annually. Currently, only nine states maintain this information from year to year at the student level.

Some states are even building college readiness tests into their statewide assessment systems so that all students — not just a self-selected group — are tested on these skills. This will provide an even clearer picture of how well schools are preparing students for postsecondary education.

Matching students' college readiness test information from high school with the same students' test scores in middle school makes it possible to analyze the effectiveness of high schools for students with varying levels of preparation.

With student-level college readiness test scores, policymakers and educators will know:

▶ How participation rates and scores on SAT, ACT, AP and IB exams change over time for low-income and minority students.

▶ The percentage of students who met the proficiency standard on the state 8th grade test who also take AP or IB courses in high school and pass the corresponding exams.

▶ The percentage of low-income students who met the proficiency standard on the state high school test who also take the SAT and ACT exams and earn a score that indicates college readiness on those exams.

Action Steps for Policymakers

■ Student-level college-ready assessment data, including SAT, SAT II, ACT, AP, IB and end-of-course test data, are collected and maintained for all students who take the tests.

■ Cooperative discussions are entered into with testing agencies such as ACT and the College Board to establish processes for obtaining and using student-level assessment results.

8 | Student-Level Graduation and Dropout Data

34 ▷ 40 *Growth in states that report having this element, 2005 to 2006*

A majority of states currently collect annual records on individual graduates and dropouts. But the National Governors Association (NGA) compact signed by all states aims to create a more valid, reliable and consistent graduation rate that tracks students from 9th to 12th grade. Based on National Center for Educational Accountability analyses, only 26 states (up from 14 last year) currently can calculate the graduation rate defined in the NGA compact.[1]

The calculation of accurate graduation rates also requires being able to accurately account for what happens to students who leave public education. For example, states must be able to distinguish between students who drop out or get a GED and students who transfer to another school.

States can engage in the following activities to maintain accurate records on students who are not accounted for:

▶ Maintain a statewide enrollment database.

▶ Match student records over time.

▶ Merge this information with student records on completed diplomas and GEDs.

▶ Establish a statewide coding system for the various reasons why students leave a school and clear guidelines on how those codes should be applied.

▶ Provide training and hold school districts accountable for using the statewide coding system.

▶ Spot-check or audit districts that are likely to have problems.

With accurate graduation and dropout data in place and the ability to match records to other databases, policymakers and educators will know:

▶ When and why students leave the state's public education system.

▶ The percentage of first-time 9th graders in a given year who graduate from high school within four, five or six years.

▶ The schools and school systems that are doing the best job of reducing the dropout rate.

▶ The characteristics of high school dropouts and whether or not there are early warning signs that schools can look for in elementary and middle school.

Action Steps for Policymakers

■ Graduation and dropout data are collected at the student level for all students in grades 7–12.

■ A standardized system for identifying why a student leaves the public education system (e.g., moves to another district, is home schooled or leaves the state) is developed and used.

■ Consequences are applied to school districts that do a poor job of accounting for missing students.

---

[1] In early 2005, the NGA convened a task force charged with helping states develop a high-quality, comparable high school graduation measure. The result was a compact signed by all governors to develop a standard, four-year, adjusted-cohort graduation rate. To calculate the agreed-upon rate, states need a way to track students over time. According to responses to the National Center for Educational Accountability survey of state longitudinal data systems, only 26 states (up from 14 last year) have the necessary elements (numbers 1, 2, 8 and 10) in place to calculate the graduation rate defined by the NGA graduation compact.

$12 \triangleright 18$ *Growth in states that report having this element, 2005 to 2006*

As states and school systems work to align expectations in high school with the demands of postsecondary education, they need better data on student success when they leave the P–12 system and enter college. Most states today do not have data systems that enable this two-way communication. There are often two separate data systems, and they rarely can exchange information.

In fact, in a survey conducted in spring 2006 by the National Center for Higher Education Management Systems (NCHEMS), the disconnect between P–12 and postsecondary systems was evident. Although 18 state education agency data managers responded to the DQC survey that the state *has the ability* to connect these two data systems, postsecondary data managers reported on the NCHEMS survey that 11 states *actually have linked* student postsecondary data with high school records to conduct studies.

States must move toward a more integrated system. Postsecondary institutions should provide annual feedback reports to individual high schools on the success of their graduates in their first year of college-credit coursework. (The NCHEMS survey found that only 10 states use their higher education student unit record systems to regularly produce feedback reports to high schools by district on such topics as need for remediation, credits enrolled for and grade performance.) These data would give high schools valuable information for improving the rigor and effectiveness of high school curricula and instruction.

It is advantageous for the P–12 and public postsecondary systems to agree on a common data set to include in the high school transcript state assessment data that may be used in the placement process. The application to public colleges should include a request for the unique statewide student identifier so that student records can be electronically exchanged among high schools, community colleges and four-year colleges. Although federal privacy laws place some restrictions on the exchange of individual records, they do not prohibit states from sharing student records. Several states have worked out ways to make this exchange possible while protecting student privacy and remaining within the bounds of federal law.

Just as P–12 information needs to be collected in a central database so that it does not have to be requested from each district, so too should analysts be able to retrieve student-level postsecondary records centrally without having to request those records from each institution.

With the ability to match student records between P–12 and postsecondary systems, policymakers and educators will know:

▶ The percentage of each district's high school graduates who enrolled in college within 15 months after graduation.

▶ The percentage of last year's graduates from each high school or school district who needed remediation in college, and how this percentage varied by student income and ethnicity.

▶ The percentage of students who met the proficiency standard on the state high school test and still needed remediation in the same subject in college.

▶ How students' ability to stay in and complete college is related to their high school courses, grades and test scores.

Action Steps for Policymakers

■ A mechanism for sharing data between the P–12 and postsecondary systems is developed and used.

■ The state investigates the possibility of developing and maintaining a central repository for all of these data for all the state's public postsecondary institutions.

10 A State Data Audit System Assessing Data Quality, Validity and Reliability

19 ▷ 36 *Growth in states that report having this element, 2005 to 2006*

The decisions made in education are only as good as the data on which they are based. A poorly organized data management system, limited staff and unclear rules for data entry all lead to less-than-valid information. Invalid or careless reporting by some schools and districts is a problem in a number of states — one that is likely to continue in the absence of checks on the quality and accuracy of the data submitted by schools and districts. Without a well-designed and well-implemented state data audit system, the public cannot have confidence in the quality of the information coming out of the state's public education system.

At the outset, states need to ensure that the data elements they request are clearly and unambiguously defined, as are any rules or interpretations concerning the entry or reporting of those data. Similarly, states should provide materials or professional development opportunities for school and/or district staff to ensure that they understand state rules, regulations, definitions and protocols. States also need a well-developed system to identify data submitted by school districts that are likely to be in error, to spot-check other information on a random basis and to conduct site visits as necessary to audit the validity of the data. In addition, states must be prepared to impose consequences on districts for submitting incomplete or incorrect information.

With a robust data audit system, policymakers and educators will know:

▶ Whether or not the disaggregated student information used to rate schools for Adequate Yearly Progress (AYP) is valid.

▶ The districts that do the best job of accurately reporting their dropout data.

▶ Whether or not districts are reporting valid numbers of untested students and reasons for not testing the students.

▶ The amount and type of data quality problems identified by districts and how those problems are being addressed.

Action Steps for Policymakers

■ The state adopts and clearly communicates a set of data definitions and standards to all entities submitting data to the state.

■ A thorough data audit system is developed and used to ensure that data received by the state education agency are accurate and match the data definitions and standards adopted by the state.

■ Thorough training and professional development on data definitions and standards are provided to all education units in the state.

Fundamentals in Designing State Longitudinal Data Systems

In addition to the 10 essential elements, states need to ensure that they take into account the following fundamental concepts in the construction of their longitudinal systems.

Privacy Protection: One of the critical concepts that should underscore the development of any longitudinal data system is preserving student privacy. An important distinction needs to be made between applying a "unique student identifier" and making "personally identifiable information" available, for example. It is possible to share data that are unique to individual students but that do not allow for the identification of that student. It also is critical to put in place encryption and data security protocols to secure the transmission or transaction of data between and among systems. States should ensure that they bring privacy considerations into the development of each repository and the exploration of each protocol or report.[1]

Data Architecture: Data architecture defines how data are coded, stored, managed and used. Good data architecture is essential for an effective data system. Many states are in the process of improving their data architecture so that they can clearly communicate with all entities with which they share and from which they receive data. Districts need to know specifically how data elements are defined (e.g., what a "dropout" is), how they should be formatted, and how and when the data should be transferred to the state education agency. Without these standard definitions and dictionaries, state education agencies will have an extremely difficult time making sense of the data received from their districts. With standards in place that are used by everyone, staffing resources and processing or cycle time can be greatly reduced, data can be made available to users when they need them, and reports can be based on clear and common definitions.

Data Warehousing: Twenty-six states have designed and built or upgraded their data warehouses or are in the process of doing so. Policymakers and educators need a data system that not only links student records over time and across databases but also makes it easy for users to query those databases and produce standard or customized reports.

A data warehouse is, at the least, a repository of data concerning students in the public education system; ideally, it also would include information about educational facilities and curriculum and staff involved in instructional activities, as well as district and school finances. The warehouse should ensure student and teacher confidentiality, allow longitudinal analyses, and include analytical capabilities for its users. Examples of the capabilities that should be available in a data warehouse include, but are not limited to, trend analyses; tracking of students over time and across campuses and/or districts; queries designed and conducted by different users (with different levels of access to detailed data, depending on user classification); and standard summary reports at the campus, district or state level for policymakers and educators. The key to effective data warehousing is the timely and efficient use and reporting of data.

---

[1] For more information on how states can ensure that they are safeguarding personally identifiable information while building and using longitudinal data systems, see the DQC issue brief *Maximizing the Power of Education Data While Ensuring Compliance with Federal Student Privacy Laws: A Guide for State Policymakers.*

Interoperability: Data interoperability entails the ability of different software systems from different vendors to share information without the need for customized programming or data manipulation by the end user. Interoperability reduces reporting burden, redundancy of data collection, and staff time and resources. It allows for better, faster and clearer reporting of data. It depends on systems having common data standards and definitions (as addressed in element 10 above). Organizations such as the Schools Interoperability Framework Association work to ensure the creation of platform-independent, vendor-neutral open standards that can be used by educators and vendors to design and implement interoperable data systems.

Portability: Data portability is the ability to exchange student transcript information electronically across districts and between P–12 and postsecondary institutions within a state and across states. Portability has at least three advantages: it makes valuable diagnostic information from the academic records of students who move to a new state available to their teachers in a timely manner; it reduces the time and cost of transferring students' high school course transcripts; and it increases the ability of states to distinguish students who transfer to a school in a new state from dropouts. The large interstate movement of students in the wake of Hurricane Katrina made the value of such a system obvious. Data portability is supported by the implementation of interoperable systems, but it requires states that use these systems to have a set of common definitions or protocols.

Professional Development around Data Processes and Use: Building a longitudinal data system requires not only the adoption of key elements outlined in this paper but also the ongoing professional development of the people charged with collecting, storing, analyzing and using the data produced through the new data system. The local school person who inputs course grades needs to understand fully how his/her work fits into the broader data system, the principal needs to understand how data can effect daily school management — both facilities and academic decisions — and policymakers need to understand how their decisions are limited or expanded based on the quality of the data available. For these changes in culture and management to occur, states need to make it a priority to rethink and possibly reorganize how education data are managed throughout the system, increase training and professional development for staff — both managers and users — and assist all employees and stakeholders of the state education system to be active consumers of the longitudinal data system.

Researcher Access: Research using longitudinal student data can be an invaluable guide for improving schools and helping educators learn what works. These data are essential to determining the value-added of schools, programs and specific interventions. States are developing ways to make student-level data available to researchers while protecting the privacy of student records under the Family Education Rights and Privacy Act. Because state education agencies and local school districts usually do not have the resources to conduct this research themselves, providing access to the data to outside researchers with appropriate privacy protections allows critical research to be done at no cost to the state or school districts.

Future Directions of State Data Systems

These 10 elements and design fundamentals are essential but not sufficient. States need to plan for a series of next-generation improvements — in fact, some states are already working on them. In the future, data systems can make it possible to do the following:

Connect school performance to spending: Building a longitudinal data system will go a long way toward determining how much value schools and districts are adding to student outcomes. This system will help identify which school systems do particularly well and might serve as models for lower-performing peers. However, the question that such identification begs is, why are these school systems so effective? How do these school systems allocate their resources — time, staff and money — to achieve their success? How can less successful school systems reallocate their existing resources or allocate new resources to achieve the same success?

The existing financial data reported within states are limited. For example, most systems do not report beyond the district level, making it impossible to tell how resources are allocated within individual schools. Twenty-nine states report the ability to connect financial data to student-level data at the state, district and/or campus levels. To really harness the power of a longitudinal data system, states need to collect financial data at the school level, and ultimately at the program level, so they can match initiatives or interventions with any subsequent changes in outcomes. In fact, another complete set of "essentials" would be required to build a high-quality data system that links resource allocation with the longitudinal data system described in this paper.

That said, states could use existing financial data to better understand how the functional resource allocations of the most successful schools (or their districts) compare to those that are not as successful. For example, most states could use existing financial data to answer the following questions:

▶ Are the most successful schools more likely to be located in districts that allocate more of their money toward instruction than are less successful schools?

▶ Do districts with schools that have increased their success with student cohorts over time spend more money on staff development than other districts?

▶ Do patterns of improvement in student outcomes correspond to changes in overall spending levels or specific spending allocations?

▶ How does the overall spending level of districts with the most successful schools compare with the amount spent by less successful districts?

These are questions that policymakers grapple with regularly in almost every state. Creating a comprehensive data system that links longitudinal student outcome data to financial data at the district level — and ultimately, the school or even classroom level — can help decisionmakers get to the next level and figure out not only what works but also what it takes to deliver what works.

Connect school performance to employment: The P–12 education system is a precursor to employment either directly (straight from high school) or indirectly (after postsecondary education). Educators and poli-

cymakers need to know whether schools are preparing students for long-term success in the workplace. Obtaining this information requires matching the P–12 and postsecondary academic records and employment records of individual students. Because employment records use social security numbers (SSNs), the educational data systems must be able to collect students' SSNs under suitable privacy protections to make the

matching of school and employment records possible. This requires some states to lift restrictions on the collection of social security numbers by the P–12 and postsecondary systems.

States also should consider incorporating into their education data systems as needed records from other social service agencies that have information relevant to students' health and safety.

Call to Action

"Without data, you're just another person with an opinion." This aphorism underscores the importance of making educational decisions at every level based on valid and reliable information.

Only a handful of states have data systems with most of the 10 essential elements, and no state has all 10 elements. The number of states with three or fewer elements has dropped from 12 last year to six this year. Progress is being made, but there is work to be done across states. (To see how your state stacks up, visit www.DataQualityCampaign.org.) The partner organizations of the Data Quality Campaign are eager to support states in the development and maintenance

of data systems that include all 10 essential elements. The partners challenge every state to put these systems in place by 2009. The campaign also continues to provide tools, resources and support as states develop and use quality longitudinal data systems.

There has never been a more unique opportunity or urgent need for every state to create and embrace a longitudinal data system. As a nation, we must take advantage of this unique moment and work together to ensure that states have the data foundation and infrastructure they need to support and enrich the hard work under way to strengthen our schools and improve student achievement.

Appendix: State Education Agency Actions

This paper has dealt primarily with actions and issues that have state policy implications related to each of the 10 essential elements. Each of those, however, translates to multiple specific actions that need to occur at the state education agency level. Examples of specific state education agency actions associated with each element include:

1. A Unique Statewide Student Identifier

■ The state assigns each student a unique statewide student identifier that can be used to match records accurately across databases and years.

■ The state develops procedures to ensure that two identifiers of the same type are not assigned to the same student (e.g., when the student moves, she/he keeps the same identifier) and that two students are not assigned the same identifier.

■ The state assigns an identifier that will follow each student from kindergarten (or prekindergarten when applicable) through 12th grade as he or she moves across campuses or districts and leaves and re-enters the state's public education system.

2. Student-Level Enrollment, Demographic and Program Participation Information

■ The state collects information at least annually on each student's:

  • campus of enrollment
  • grade level
  • gender
  • ethnicity
  • economic status
  • English language learner status
  • participation in bilingual or English as a Second Language program
  • special education status
  • migrant status
  • Title I status
  • gifted and talented status

■ At least one enrollment data collection takes place in the fall.

■ At least one data collection for each of these items occurs at a different time from when the state test is administered.

■ The information is stored permanently in a state database. (If the information is used to populate the test database and then discarded, the state loses track of enrollment over time.)

■ The state collects student attendance data either daily or over a small period of time (e.g., over a six- or nine-week period) that includes at a minimum campus of attendance, number of days absent and number of days present.

3. The Ability To Match Individual Students' Test Records from Year to Year To Measure Academic Growth

■ The state updates its student test database with demographic and program participation information collected earlier in the school year rather than updating it at the time of the test administration. (This assumes a student identifier is available to connect the two databases.)

■ The state collects and permanently stores information on each student's test score in each subject for year-to-year comparisons. The information may be disaggregated by skill or skill area for each student (e.g., reading comprehension/ability to identify the main idea).

■ The state makes the data available and/or uses the statewide database to conduct research and program evaluation activities (e.g., the measurement of year-to-year student academic growth).

■ At the beginning of the year, the state makes available to each teacher student test score information on state exams that can be broken out by specific skill areas within each subject for each of the teacher's students.

4. Information on Untested Students

■ The state maintains a record for each untested student in a tested grade, including information on why the student was not tested.

■ The state matches information on untested students to demographic, program participation and attendance information.

■ The state conducts analyses of patterns among untested students across campuses and districts.

5. A Teacher Identifier System with the Ability To Match Teachers to Students

■ The state assigns each teacher a unique statewide identifier that can be used to match records accurately across databases and years.

■ The state develops procedures to ensure that two identifiers of the same type are not assigned to the same teacher and that two teachers are not assigned the same identifier.

■ The state collects information on each teacher's college major, graduate school degrees by degree type and subject, types of certification or credential, certification exam scores, salary, and experience.

■ The state collects data from each school district that match each teacher to the students taught in each of the teacher's classes, by teacher and student identifier.

6. Student-Level Transcript Information, Including Information on Courses Completed and Grades Earned

- The state adopts or develops and maintains an ongoing electronic course classification system (including standard course numbers, titles and descriptions).

- The state collects individual course completion records for all courses taken in middle and high school. These include:

  - courses taken during the regular fall or spring semesters
  - courses taken in summer school
  - courses taken in middle school for high school credit (e.g., Algebra I)
  - courses taken at local colleges for dual credit
  - credits transferred from private high schools or home school
  - credit received for distance learning

- The state also collects the grade the student earned in each course and the student's overall grade point average.

- Individual student records in the course completion database can be connected to the same students' records in the enrollment, demographic, program participation and test databases.

7. Student-Level College Readiness Test Scores

- The state acquires from the College Board, ACT and the IB Organization and permanently stores student-level results by individual exam for each of the following: SAT, SAT II Subject Tests, ACT, AP, IB.

- The state is able to connect the student-level test data mentioned above to the enrollment, demographic, program participation and test databases.

8. Student-Level Graduation and Dropout Data

- The state collects and stores graduation and dropout data at the student level.

- The state collects and stores student-level graduation data by diploma type (e.g., Recommended Graduation Plan, New York Regents Diploma).

- For students in grades 7–12 who were enrolled in one year, not enrolled the next year and did not graduate, the state collects information from local school districts on where each departing student went. The evidence on departing students can be used to determine whether students graduated; dropped out; transferred to another school, district or state; earned a GED; or are missing (they cannot be located, and no evidence exists on where they went).

- The state makes every effort to track reported dropouts back to other schools in the state via the enrollment, test and/or attendance databases.

- The state has standards for the types of evidence that may be used to determine where departing students went.

- The state has standards for the percentage of departing students that school districts should be able to locate.

- The state applies consequences to school districts that do a poor job of accounting for missing students (e.g., lower accountability rating).

9. The Ability To Match Student Records between P–12 and Postsecondary Systems

- The state works with the postsecondary system to match student-level records among all institutions of the state's P–12 and public higher education systems.

- The postsecondary information to be matched includes, but is not limited to, student records on:

  - enrollment
  - course completion
  - graduation
  - degrees and certificates received
  - performance on mandated state tests administered by the postsecondary system

10. A State Data Audit System Assessing Data Quality, Validity and Reliability

- The state develops a clear set of data standards and definitions that apply to all data received by the state education agency.

- The state provides training on these data standards to local school district personnel.

- The state performs statistical checks on data submitted by school districts.

- The state has criteria established for determining when data submitted by school districts are likely to be in error.

- The state has a system for investigating the accuracy of data that are flagged by the statistical checks.

- The state has a system for occasionally spot-checking the accuracy of data in cases that are not flagged by statistical checks.

- The state has a system of selecting districts for on-site audits and performs on-site audits in the selected districts.

- The state imposes consequences on school districts that do a poor job of collecting and submitting accurate and complete information.

Find Out More

▶ **the 10 essential elements** and the state policy actions required to establish, maintain and use a quality longitudinal data system;

▶ the results of NCEA's **2006 update of its annual survey** that show where your state stands on the 10 essential elements;

▶ **the tools, materials, meetings and information** that can aid states and interested organizations seeking to ensure increased quality, accessibility and use of data; and

▶ how your organization can **partner with DQC** to generate the understanding and will to build and use state longitudinal data systems.

This white paper was produced by the Data Quality Campaign. It is based on work originally written by Chrys Dougherty, Ph.D., NCEA director of research, and Nancy J. Smith, Ph.D., DQC deputy director.

DATAQUALITY
CAMPAIGN
Using Data To Improve Student Achievement

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

▶ implement state longitudinal data systems to improve student achievement.

The campaign provides tools and resources that states can use as they develop quality longitudinal data systems and also serves as a national forum for reducing duplication of effort and promoting greater coordination and consensus among like-minded organizations.

Data Quality Campaign Partner Organizations

Managing Partners

▶ 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 Accountability
▶ 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

▶ ACT
▶ Alliance for Quality Teaching
▶ American Association of Colleges for Teacher Education
▶ American Association of State Colleges and Universities
▶ American Board for Certification of Teaching Excellence
▶ APQC
▶ Center for Teacher Quality
▶ College Summit, Inc.
▶ Consortium for School Networking
▶ Educational Policy Institute
▶ GreatSchools
▶ Jobs for the Future
▶ League of Education Voters Foundation
▶ National Alliance for Public Charter Schools
▶ National Association of Secondary School Principals
▶ National Education Knowledge Industry Association
▶ Postsecondary Electronic Standards Council
▶ Roads to Success
▶ Southern Regional Education Board

The campaign is managed by the National Center for Educational Accountability and supported by the Bill & Melinda Gates Foundation.