COULD A NEW WAY OF COLLECTING DATA TRANSFORM EDUCATION IN CALIFORNIA?

The state's K–12 public schools are about to experience the launch of two new data systems designed to track student academic information and teacher qualifications

Technological advancements have dramatically changed the way the world collects and stores data, providing us with a surfeit of information literally at the touch of a keyboard. Businesses and government agencies—including many of the nation's public school systems—manage vast databases that help us quickly access the facts and figures we need. And now California is making inroads toward improving the databases of its public schools.

The state plans to launch a more sophisticated process for collecting educational data—acquiring a type of data known as longitudinal data—that can measure change, such as how student graduation rates change, given various factors.

The California Longitudinal Pupil Achievement Data System (CALPADS) is the name of the state's new system that will help maintain data on public school students in Kindergarten through 12th grade, and the California Longitudinal Teacher Integrated Data Education System (CALTIDES) is the data system that will maintain information on public school teachers.

Together these two state data programs have been in development for several years: CALPADS is scheduled to be implemented during the 2009–10 instructional year, and CALTIDES is expected to launch in the fall of 2011.

Learning More About California's Schoolchildren
A new way for California to compile academic data on public K–12 students is currently in development. Known as CALPADS, the data system is expected to launch during the 2009–10 instructional year and will generate education information that has never before been offered in this state.

What Is CALPADS?

The California Longitudinal Pupil Achievement Data System—CALPADS—was created by California Senate Bill 1453 (Alpert, Chapter 1002, Statutes of 2002), which establishes various goals for this data system.[1] Once CALPADS launches, the following information for each student[2] attending a K–12 public school[3] in California will become available:

> **Demographic data.**[4] A student's legal name, gender, date of birth, race/ethnicity, socioeconomic status,[5] primary language, and, optionally, city and state of birth.

> **Pupil achievement data.** A student's grades, credits earned for certain courses taken in grades 7–12,[6] and a student's scores from the California Standards Tests, California High School Exit Exam, and California English Language Development Test (for English learners).[7]

> **Pupil enrollment data.** A student's school, grade level, enrollment status, course names (and whether the courses satisfy University of California or California State University admission requirements), and graduation status.

> **Program participation data.** Whether a student is a special education student, migrant student, Gifted and Talented Education Program participant, English-language learner, and/or a participant in other state or federal programs, such as the California Partnership Academies and Title I (for economically disadvantaged students) and Title X (education for homeless children) of the federal No Child Left Behind Act.

> **Instructional services data.** Whether a student receives independent study, federal Title I services (for economically disadvantaged students), and English Language Development services (including the type of language-development service, such as bilingual education or Specially Designed Academic Instruction in English).

> **Student discipline data.** Whether a student was truant during the school year, suspended, and/or expelled, and the reason for disciplinary actions.

CALPADS also will include data on some of the state's K–12 public school employees:

> **Employee data.** For credentialed employees assigned to specific classes, this data will include the employee's job classification (teacher, administrator, counselor, librarian, nurse, speech therapist); gender; race/ethnicity; years of service in the district and in the profession; highest education level attained; and, for teachers, the courses they are teaching. (See "What Is CALTIDES?" on page 4.)

CALPADS and CALTIDES: Why Are They Important?

These new data-gathering systems will provide parents, educators, school administrators, and educational policymakers with an unprecedented volume of information that is expected to aid in the enhancement of California's public education system.

Both CALPADS and CALTIDES were initiated, in part, to provide data required by the federal No Child Left Behind Act (NCLB), a sweeping 2001 federal law that requires states to improve student achievement and teacher quality. Under NCLB, state and local school districts must report graduation rates and achievement scores by certain categories, such as those who are economically disadvantaged, of a particular ethnic background, or are English-language learners.

NCLB also requires states to ensure that public school teachers have the appropriate credentials for the classes they teach—and CALPADS and CALTIDES will provide the data needed to meet this requirement.

In addition, longitudinal data—data that reflects the same type of information that has been collected on the same subject (or person) at multiple points in time—will be collected by these new programs. For example, this longitudinal data could show how specific students' grades have changed over a period of several years.

In the past, school districts have submitted aggregate reports summarizing some student and teacher statistics for a particular point in time and for a particular group of students. These aggregate reports have provided useful snapshots of data, but the data from one year's report typically cannot be compared accurately to the data from another year's report because each snapshot is of a different (and constantly changing) group of students.

For instance, under the previous system, when California compared the achievement of its English-language learners from one year to the next, progress could not be distinguished because these comparisons often were not of the same students. After a period of time, some students who had been studying English as a second language may have learned enough to depart the English learner category, while at the same time new students would have entered into this category. Furthermore, aggregate data does not show how students progress over time, whereas longitudinal data will provide those results.

Collecting data on an individual level enables the state to respond better to reporting requirements, which often change, and to produce custom-made reports without having to ask school districts to re-collect and re-submit data for each report.[8] For example, if the federal government were to change the reporting requirements for a program that accounts for students grouped by gender to one that instead groups them by race, the state would not have to request additional reports from school districts to

What Is CALTIDES?

The California Longitudinal Teacher Integrated Data Education System—CALTIDES—was initiated in the 2005 Budget Act and subsequently authorized by California Senate Bill 1614 (Simitian, Chapter 840, Statutes of 2006), which establishes the following goals for this data system:

> **Identify how many teachers** will be needed in California in future years, as well as trends in the public K–12 teaching workforce, including retention, attrition, and mobility.

> **Provide high-quality program evaluations,** including analyses of teacher-preparation programs and professional-development programs.

> **Provide the state and school districts** with the information needed to better monitor whether teachers are actually teaching the classes they have been authorized to teach (a teaching requirement mandated by state and federal law).

> **Offer a repository of information** to help develop and review state policy.

> **Provide age ranges of teachers** in the workforce, teacher retirement projections for the next 10 years, subject areas with severe teacher shortages, geographic distributions of teachers by type of credential, and patterns of in-service education for teachers.

CALTIDES will use teaching-credential information from a database maintained by the Commission on Teacher Credentialing and link it to the teacher-assignment data in CALPADS. This step will allow the state to report on, among other things, compliance with federal requirements, including ensuring that those who teach certain core courses are "highly qualified."

The following data will be included in CALTIDES for credentialed personnel[9] in California's K–12 public schools:

> **The type of credential** an employee holds and what courses he or she is authorized to teach or what services he or she is authorized to provide.

> **How a credential was obtained** (for instance, was it given after the completion of a teacher-preparation program, the passage of an exam, or participation in a nontraditional credential program, such as a school-district internship program).

> **Whether a teacher participated** in the Beginning Teacher Support and Assessment Program.

(Note: It is prohibited by law to use CALTIDES data "either solely or in conjunction with data from CALPADS, for purposes of pay, promotion, sanction, or personnel evaluation of an individual teacher or groups of teachers, or of any other employment decisions related to individual teachers.")

satisfy this change in federal requirements; rather, the new CALPADS program could be used to access this data instantly.

Other benefits of the CALPADS and CALTIDES data systems include:

> **Readily accessible information that will help students when they transfer to a new school district.** When a student moves to a different school district, it is important for the new district to have basic information about the student (such as test scores and special-education status) as soon as possible to ensure the student is given appropriate services promptly. CALPADS will make it easier for a new school district to quickly access basic student information instead of, for example, waiting for the child's records to be delivered in the mail.

> **An "early warning system" to help students who are behind in their education.** A longitudinal data system provides information that allows the state and local school districts to develop their own early warning system that can identify students who are at risk of failure in middle school and high school based on certain data, such as their performance on state tests. For example, one might ask whether elementary school students with lower-than-average test scores and/or discipline problems are more likely to fail the California High School Exit Exam than their peers who don't struggle with such issues. CALPADS will allow for such analyses.

> **Better data to identify what works best for students and teachers.** A comprehensive educational-program evaluation includes

Gaining a Better Understanding of the State's K–12 Teachers and Other Public School Staff
Evaluating teacher-training programs, monitoring how long teachers stay in the teaching profession, and identifying trends in the public K–12 workforce are just some of the areas that could become easier to analyze once the new informational data system called CALTIDES—the companion data system to CALPADS—is launched.

how a program's participants perform over time. For example, did students who participated in a particular behavioral-modification program experience a reduction in suspensions? Or did students who participated in middle-school college-awareness programs take more college preparation courses in high school? To obtain the data necessary to answer such questions, longitudinal data is needed so a comparison can be made between groups of students who do not participate in a program with groups of students who do participate in that program. Longitudinal data also allows for comparisons of a student's performance *before* he or she attended a particular program and *after* attending it.[10]

Similarly, the longitudinal data provided by CALTIDES is expected to help with the evaluation of various teacher-training programs, as well as provide information on where program participants are getting teaching jobs and how long their teaching careers generally last.

> **Improved data reporting.** CALPADS will allow state and local school districts to more efficiently and effectively meet state and federal No Child Left Behind Act reporting requirements. For instance, the state will be able to calculate and report drop-out and graduation rates, since the new data system will enable school districts to track students as they progress through middle and high school. CALPADS also will facilitate reporting on the achievements of groups of students, such as English-language learners or those who are economically disadvantaged.

> **A way to monitor teacher assignments.** CALTIDES will allow the state and school districts to monitor more easily whether teachers have the appropriate credentials to teach the courses they are assigned to teach, in adherence with current law and the *Williams* lawsuit settlement.[11] The data system also will allow the state to report on its compliance with No Child Left Behind's "highly qualified"[12] teaching requirement.

Linking CALPADS and CALTIDES Data From Year to Year

> **CALPADS**
Upon entering California's public school system, each student is assigned a unique 10-digit student identification number (known as an identifier), which is randomly assigned and stays with a student throughout his or her attendance in California's K–12 public education system. This identifier will be the "hook" upon which individual student information is submitted and maintained from year to year in CALPADS, and the identifier will not contain a code or any other hidden information that might reveal personal data about the student. (Social security numbers will not be submitted to CALPADS.)

Each school district will have access to CALPADS information that pertains to each of its enrolled students.

> **CALTIDES**

Each credentialed employee working in the K–12 public school system—and aspiring educators applying for their credentials—also will be assigned a 10-digit state-educator identification number (known as an identifier, just like those described under CALPADS) that will follow him or her throughout his or her public school career in California. These identifiers also will not contain a code that could reveal the holders' identity or other personal information, nor will they include social security numbers, home addresses, telephone numbers, or e-mail addresses; however, they will provide school districts with the ability to match a teacher's credential and authorization data in CALTIDES with his or her class-assignment data in CALPADS, as well as access to other relevant employee data.

Implementation Status

> **CALPADS**

State funding has been provided for the development of CALPADS in the state's annual budgets since the authorizing legislation went into effect in 2002. After several years of required research and planning, the California Department of Education (CDE) selected a private vendor in January 2008 to build and develop the system, and local school districts (with the assistance of a $30 million grant) are now preparing their data systems for CALPADS,[13] which is expected to launch during the 2009–10 instructional year. Nearly all students already have been assigned an individual student identifier.

> **CALTIDES**

Since 2005, the state's annual budgets also have provided funds for the development of CALTIDES, which will be managed jointly by the California Department of Education (CDE) and the Commission on Teacher Credentialing (CTC). To date, CDE and CTC have managed most of the technological planning for CALTIDES and now are in the process of selecting a vendor for the program.[14] CTC has assigned identifiers to nearly all teachers and credentialed staff currently working in California's public schools,[15] and county education offices and school districts are in the process of inserting those identifiers into their local data systems. CALTIDES is expected to launch in the fall of 2011.

CALPADS and CALTIDES: What They Won't Deliver

While there will be many benefits of the CALPADS and CALTIDES programs, there also will be limitations. For example, the state's existing student testing records are not designed to be used as longitudinal data, which means the tests are not designed to allow a comparison between a student's test scores from one year to the next.

To address this limitation, California's 2007 Budget Act included funds for the California Department of Education (CDE) to investigate options for developing a way to track test scores so they can be compared from year to year. In spring 2008, as required by the 2007 Budget Act, CDE's test contractor provided various options to

California's Students May Someday Reap Benefits From an Interlinked Family of Data Systems
Linking numerous existing databases—ranging from the data programs at preschools and universities to those at health and social service departments—with the upcoming CALPADS and CALTIDES K–12 data systems could provide the information needed to enrich the academic experience of every public school student in the state, while also offering benefits to school administrators and staff.

accomplish this task, which would allow parents and school staff to see how the test scores of an individual student—or a group of students—compare over the years.

Other limitations: some student data (such as attendance data) initially will not be included in CALPADS. However, CALPADS is being designed to accommodate additional data in the event the state decides (via legislation) to increase the types and quantity of data gathered by this system and assuming additional funds are provided to pay the state and local costs of adding more data.

In addition, the state does not yet have a formal plan to link CALPADS and CALTIDES to systems outside of the K–12 public schools, such as the data systems of California's public higher-education institutions. Yet on September 29, 2008,

Governor Arnold Schwarzenegger signed Senate Bill 1298 (Simitian, Chapter 651, Statutes of 2008), which requires the state to map out, in a strategic plan, the technical details for linking CALPADS and CALTIDES to other systems.

Linking these databases to other state systems would enable educators to more easily answer numerous important questions, such as: What percentage of a school district's high-school graduates enroll in college within 15 months of graduation? What percentage of the state's high-school graduates need remediation once they get to college? How is a student's ability to successfully complete college related to his or her high-school courses, grades, and test scores? Do students who attend a state-funded preschool have a better chance of meeting state academic expectations in elementary school? And in high school?

What We Don't Know Now But Will Know Soon…

California's existing system of compiling annual education reports does not provide the data needed to answer some fundamental questions about the academic performance of the state's students. With the implementation of CALPADS and CALTIDES, however, this critical data about K–12 public education finally will become available, which will help education policymakers and researchers answer important questions, including the following:

> Can the academic performance of students in the 3rd or 4th grade predict the likelihood that they will pass the California High School Exit Exam?

> How do English-language learners who recently have moved to the United States tend to score on the California High School Exit Exam? And how does their academic achievement compare to English learners who have been attending the state's public schools for many years?

> How do students who fail to meet the state's academic expectations in 3rd grade tend to perform in 4th grade? In 5th grade? In 8th grade?

> What percentage of 9th graders drop out of high school before they graduate? Would identifying such factors as below-average test scores, suspensions, or expulsions for students in elementary school help predict whether those students might be candidates for dropping out of middle school or high school?

> How long do teachers typically stay in their jobs in high-poverty schools?

> Are there particular teacher-preparation programs that tend to produce instructors who make teaching a longer-than-average career?

> How many new math and science teachers received their credentials last year?

Beyond CALPADS and CALTIDES: The McKinsey Recommendations

In 2008, the California Department of Education and the office of Governor Arnold Schwarzenegger announced they had hired McKinsey and Company, a private consulting firm, to recommend how the state should develop and implement a comprehensive statewide data system that integrates existing data systems, including the soon-to-be-launched CALPADS and CALTIDES programs. The consulting firm's findings were published in a December 2008 report.[16]

McKinsey and Company analyzed businesses and schools that have "continuous-learning systems," that is, systems that help organizations make small improvements toward their long-term goals using data-collection processes.

To set up continuous-learning systems in California's public education system the report recommends continuing with the development of CALPADS and CALTIDES and eventually expanding and combining these data systems to create a statewide system that could accomplish the following:

> Provide a one-stop-shop web-enabled education portal, in which teachers and students could access information to help them improve.

> Allow school districts to design their own student assessments and parent surveys and use the statewide system to process and track the results.

> Provide more comprehensive data about teachers, students, and schools (for example, there would be more information on student absences and the amount of state and private funding received by schools).

> Allow teachers and administrators to use the system to share effective practices and collaborate with their peers throughout the state.

> Provide teachers easy access to professional-development programs and allow school districts to survey the participants about the usefulness of such programs.

> Link CALPADS and CALTIDES to other data systems such as those used by preschool, higher education, criminal justice, and health and social-service programs.

To achieve this vision of a comprehensive statewide data system the McKinsey report recommends improving the quality, accessibility, and use of its current K–12 data systems; adding about 30 more data elements to CALPADS and CALTIDES; and creating interagency links that would connect information in the K–12 system to other parts of a student's academic experience.

The Road Ahead

Accurate, accessible information is important for any institution that wants to improve its outcomes and this is particularly true of California's public school system, which has the added burden of scarce financial resources for the foreseeable future. Collecting and securely maintaining data can be costly but also can yield many benefits that have the potential to impact the education of California's school children in a positive way for many future generations.

As the state embarks on the implementation of CALPADS and CALTIDES, there will continue to be significant challenges, from balancing the long-term costs to explaining the long-term benefits of expanding these databases. And throughout this process the state also must ensure that the privacy of student and teacher data will forever be protected, another essential ingredient to maintaining the long-term viability of CALPADS and CALTIDES.

Endnotes

1. The goals for CALPADS established by Senate Bill 1453 include: provide school districts and the California Department of Education with the data needed to comply with the reporting requirements of the federal No Child Left Behind Act; improve the evaluation of "educational progress and investments" over time; provide school districts with information that can be used to improve pupil achievement; and create an efficient, flexible, and secure way to maintain statewide longitudinal pupil data.

2. The following types of students are required to have individual student identifiers: all public school students enrolled in grades K–12, certain special-education preschool students, and some adult students who intend to take the California High School Exit Exam.

3. This includes charter and alternative schools operated by county offices of education.

4. Demographic data is collected in part to help school districts distinguish students from one another when the student identification numbers are being generated—an important part of any longitudinal student data system (see "Linking CALPADS and CALTIDES Data From Year to Year" on page 6). Only the pertinent school district will have access to its students' names; other districts may have limited access to students' names when new students are transferring to their districts. In addition, student addresses are collected by districts that opt to directly certify their students for participation in the National School Lunch Program; these addresses are then matched with the appropriate data files maintained by the California Department of Health Services.

5. In this case, socioeconomic status is determined by whether a student participates in the National School Lunch Program and the educational level of the student's parents or guardians.

6. This will include course completion for departmentalized courses in traditional (not alternative) schools.

7. Test scores will include performance levels and scale scores. After students take the California Standards Test they receive a raw score, and that score will fall into one of five categories: far below basic, below basic, basic, proficient, and advanced. The California State Board of Education has established the proficient category as the minimum performance expectation for all students for purposes of compliance with the federal No Child Left Behind Act.

8. Longitudinal data will not be available through CALPADS until it has been collected and maintained by the system for at least two years.

9. Credentialed personnel includes teachers, school administrators (such as principals), counselors, librarians, nurses, and speech therapists.

10. Unless changes are made to California's educational testing system, CALPADS will not show year-to-year growth in test scores for individuals or groups of students because state tests are not designed to provide test scores that are statistically comparable from year to year. See "CALPADS and CALTIDES: What They Won't Deliver" on page 7.

11. The *Eliezer Williams, et al., v. State of California, et al. (Williams)* case was settled in 2004. The settlement requires California to monitor the following conditions in low-performing schools: adequate school facilities, adequate textbooks, and adequately trained teachers.

12. No Child Left Behind defines "highly qualified" as one who has a bachelor's degree, a state-issued authorization to teach, and demonstrated competence in the subject matter he or she is authorized to teach.

13. Prior to the conception of CALPADS, previous California state budgets allotted funds for CALPADS' predecessor, California School Information Services (CSIS), which was initiated in California's 1997 Budget Act. The CSIS program assists school districts in implementing data systems, among other things.

14. California has a defined procurement process for all information-technology projects, which is controlled by the California Department of General Services.

15. Teachers and credentialed staff who hold a lifetime credential and have their credential information only available on microfilm have not yet been assigned identifiers.

16. The research was funded by private foundation grants.