



CALIFORNIA  
DEPARTMENT OF  
EDUCATION

1430 N STREET  
SACRAMENTO, CA  
95814-5901

**JACK O'CONNELL**

State Superintendent of  
Public Instruction

PHONE: (916) 319-0800

October 10, 2007

Little Hoover Commission Testimony  
October 25, 2007

Presented by  
Keric Ashley, Director  
Data Management Division  
California Department of Education

### **Background of California Education Data**

For many years, the California Department of Education (CDE) has collected student and teacher data from schools and districts. The majority of the data was collected on Information Day (the first Wednesday in October) as part of a data collection known as the California Basic Educational Data System (CBEDS). Schools and districts reported to CDE, aggregated counts of students by school, grade level, gender, and race/ethnicity. The data was not submitted with a student identifier of any kind and, therefore, was not longitudinal. Along with enrollment counts, schools and districts also submitted counts of dropouts, graduates, graduates who met UC and CSU requirements, and enrollment in a few selected higher level math and science courses. With the enactment of the federal No Child Left Behind Act of 2001, dropout and graduate data was collected by certain subgroups (i.e., Migrant Education, Limited English Proficient, Special Education and Socioeconomically Disadvantaged). Because these counts were aggregated at the school level, none of the data could match students from year to year.

CBEDS also collected teacher and staff counts by school, gender, race/ethnicity, education level, years of service, and authorized teaching areas. A data submission on each individual teacher also included the specific courses taught by the teacher. Although each teacher was reported separately, there was no statewide identifier assigned to each teacher and, therefore, the data was not longitudinal.

None of the student data could be matched with the teacher data.

Over the years, as federal and state categorical programs have increased, the requirement for student participation data has also increased. Program staff, across multiple divisions within CDE, were required to initiate separate data collections to meet

specific federal and state reporting requirements. More student level data were also collected through the answer documents related to separate statewide assessments. Federal and state accountability systems required subgroup reporting and, therefore, each answer document collected grade level, gender, race/ethnicity, and special program participation.

The result was multiple data collections collecting similar, but not identical data and no way in which to match the data across the different collections.

### **Background of California School Information Services**

With an acknowledgement that student and teacher level data reporting could be streamlined with electronic reporting, the California School Information Services (CSIS) program was created in 1997 under the oversight of the Fiscal Crisis Management and Assistance Team (FCMAT), operated from the Kern County Office of Education.

As stated in Education Code Section 49080, the mission of CSIS is to do all of the following:

- Build the capacity of local education agencies to implement and maintain comparable, effective and efficient pupil information systems that will support their daily program needs, assist local education agencies in improving the outcomes of pupils, and promote the use of information for educational decision making by school site, district office, and county staff.
- Enable the accurate and timely exchange of pupil transcripts between local education agencies and to postsecondary institutions.
- Assist local educational agencies (LEAs) to transmit state and federal reports electronically to the State Department of Education, thereby reducing the reporting burden of LEA staff.

School districts received incentive funding to participate in a voluntary program to participate in capacity building activities and submit student and teacher data electronically through CSIS to the CDE. As part of this submission of data, students in these "voluntary CSIS Districts" were given individual student identifiers or "CSIS IDs" as they were called at the time. Through CSIS, the CDE today receives student-level data from approximately 250 school districts representing a majority of the state's student enrollment. The other 750 school districts still submit aggregate data to CDE to meet reporting requirements.

At this same time, the Department of Finance sponsored a review of data management practices within the CDE. The intent of the study was to identify strategies to improve

the efficiency and effectiveness of data management within the CDE. As a result of that study, funding was made available and CDE established an Education Data Office. The CDE adopted guiding principles for collecting and managing data and began the process to catalogue all of its separate data collections. Superintendent Jack O'Connell supported the elimination of any duplicate or non-essential data collections. The CDE was able to eliminate over 10 percent of the current data collections, but even more collections will be discontinued, once the state has a longitudinal system that collects data based on individual student identifiers.

### **National Focus On Student and Teacher Level Data Reporting**

As previously mentioned, the No Child Left Behind Act of 2001 increased the amount of data states must collect from schools and districts. States have responded by authorizing longitudinal data systems for student and teacher level data. Federal grants have also been made available to states for initiating or enhancing the development of these longitudinal systems. The Data Quality Campaign (DQC) managed by the National Center for Educational Accountability has three goals:

- Longitudinal education data systems in 50 states by 2009
- Increased understanding by policymakers and educators of how to use longitudinal and financial data in their efforts to improve student achievement
- Promotion of data standards and efficient data transfer and exchange

The DQC has also published a list of ten essential elements critical to a state's longitudinal data system:

- A unique statewide student identifier
- Student-level enrollment and demographic and program participation information
- The ability to match individual student's test records from year to year to measure academic growth
- Information on untested students
- Student-level graduation and dropout data
- Student-level transcript information, including information on courses completed and grades earned
- A state data audit system assessing data quality

- Student-level college readiness scores
- The ability to match student records between K-12 and postsecondary systems
- A teacher identifier system with the ability to match teachers to students

In 2006, the National Governors' Association sponsored a compact signed by all 50 governors to begin implementing a standard four-year adjusted cohort graduation rate that is only possible through a longitudinal data system that receives student level data through a unique student identifier.

### **California's Response For Longitudinal Student and Teacher Data**

In 2002, Senate Bill 1453 (Alpert) established the California Longitudinal Pupil Achievement Data System (CALPADS). It stated, "In order to comply with the federal No Child Left Behind Act of 2001, California must have access to longitudinal pupil data to assess the long-term value of its educational investments and programs and provide a research basis for improving pupil performance." SB 1453 requires all schools and districts (including charter schools) to acquire and maintain a Statewide Student Identifier (SSID) for each of their K-12 public school students. It also requires the creation of a longitudinal student data system using the SSID that includes demographic, program participation, assessment data (CAHSEE, STAR and CELDT) and highly qualified teacher data. The system must be able to meet all federal reporting requirements, including four-year graduation and dropout rates.

Along with meeting federal reporting requirements, the goals of the system are to:

- Provide a better means of evaluating educational progress and investments over time
- Provide LEAs information that can be used to improve pupil achievement
- Provide an efficient, flexible, and secure means of maintaining longitudinal statewide pupil level data in a manner that promotes good data management practices.

CALPADS will be beneficial to both state policymakers and school districts.

Benefits to the state include:

- streamlined data collection and reporting system with better data quality
- Rich data for research and evaluation
- Ability to calculate more accurate drop out and graduation rates

- Longitudinal data which may be used for accountability measures

Benefits to schools and districts include:

- Immediate provision of basic student information to facilitate decisions (e.g. CELDT, CAHSEE, Special Education)
- Access to longitudinal data
- Reduction in aggregate reporting to the state and student level data supplied to test vendors
- Ability to update data on an ongoing basis

### **CALPADS Timeline for Implementation**

- June 2005 SSIDs were assigned to all students by CSIS
- Fall 2006 CDE used the SSID data to certify enrollment counts
- Fall 2007 CDE to use SSID data for graduates and dropouts
- Fall 2007 Special Project Report submitted for approval
- Winter 2007 Contract awarded and vendor begins development
- 2008-09 Development of CALPADS completed & pilot testing
- 2009-2010 Statewide implementation of CALPADS

### **CALPADS Preparation**

The 250 school districts that have participated in the CSIS program are better-equipped to make the transition to reporting student and teacher level data to CALPADS. Their reporting infrastructure is already in place and they only need to maintain their current CSIS reporting efforts until CALPADS is implemented. For the remaining 750 school districts and a couple of hundred direct-funded charter schools, funding was made available (beginning in 2006-07) to prepare them for CALPADS. This preparation program, operated by CSIS, is called the Best Practices (BP) Cohort Project. This project will help eligible school districts and charter schools implement sustainable local data management practices that will contribute to improved student achievement

through better local data-driven decision making and will prepare them to submit data to CALPADS. Funding to participate in the project is based upon enrollment counts. Schools and districts must learn to collect, manage, and report data differently through submitted individual student-level and teacher-level files.

### **The Challenges Ahead Of Us**

There are five major challenges that need to be addressed as we move toward having longitudinal data systems:

- The contract for building CALPADS will begin in the next few months. Building a longitudinal data system that meets the requirements of NCLB and makes useful data available for LEAs, policymakers, researchers, and the general public is a difficult and complex task. Sufficient resources must be available to successfully build such a system.
- Building a successful system is a separate issue from having quality data to put into such a system. The data reporting that comes out of CALPADS will only be as good as the data that is submitted by approximately 10,000 schools in California. We need high quality data, if it is to be used for high-stakes decision-making. Some LEAs have great capacity and infrastructure to supply quality data while many others need to build their capacity. Issues of resources, hardware, software and ongoing training must be addressed. Quality data won't just happen because we build a system.
- LEAs must be trained to make the most use of longitudinal data. CALPADS will have statewide assessment data, but LEAs will also need to learn how to link the data with local formative assessments to better impact student achievement.
- California must address how we will make use of individual student level and individual teacher level data within the current restrictions of the federal Family Educational Rights and Privacy Act (FERPA). There is a growing expectation that individual level data will be made readily available to LEAs, policymakers and researchers. While LEAs will have access to their own data and aggregate reports can be made available to policymakers and researchers, there are current federal and state restrictions from making available individual level data.

The California Legislative Analyst's Office (LAO) is currently working on recommendations of how California might address these restrictions.

- We must give consideration to the short-term and long-term expansion of CALPADS, once it is operational. CALPADS is being built to meet the requirements of NCLB. There are current CDE data collections that could be added to CALPADS, but are not currently in the scope of the project. Adding some of these collections may require legislative authority. Although some additions may require additional resources, some collections could be added without any additional resources. Some of these collections would not create mandated cost issues because the data is already being collected and collecting the data through CALPADS will only change the method of collection. There must also be a process to determine when new data elements are collected by CALPADS. New data elements may require additional resources or initiate the submission of mandated cost claims. There are already bills before the legislature that would require adding new data elements to CALPADS. In the long-term, there is considerable interest from many in the education and business to community to either expand CALPADS into a preschool through post-college data system or make CALPADS data "linkable" to data from other educational institutions or agencies.

### **Meeting The Challenges**

The CDE is already engaged in efforts to meet the challenges and opportunities in the development of our longitudinal data system, CALPADS. Working groups and technical advisory groups have been meeting to discuss the specifications of such a system. School district representatives and other stakeholders from across the state have been solicited for advice. The Superintendent has called for a Data Policy Advisory Committee consisting of educators, policymakers, parents and the business community to discuss the development, implementation and future expansion of longitudinal data.