GETTING DOWN TO FACTS II

RESEARCH BRIEF | SEPTEMBER 2018

Making California Data More Useful for Educational Improvement

Meredith Phillips
University of California, Los Angeles

Sarah Reber
University of California, Los Angeles

Jesse Rothstein
University of California, Berkeley

About: The Getting Down to Facts project seeks to create a common evidence base for understanding the current state of California school systems and lay the foundation for substantive conversations about what education policies should be sustained and what might be improved to ensure increased opportunity and success for all students in California in the decades ahead. *Getting Down to Facts II* follows approximately a decade after the first Getting Down to Facts effort in 2007. This research brief is one of 19 that summarize 36 research studies that cover four main areas related to state education policy: student success, governance, personnel, and funding.

Stanford University

PACE Policy Analysis for California Education

This brief summarizes the *Getting Down to Facts II* technical report, **Making California Data More Useful for Educational Improvement**, by Meredith Phillips, Sarah Reber, and Jesse Rothstein, September 2018.

This brief draws from two additional studies about California's systems for gathering and sharing education data:

Using Data for Improvement: Learning from the CORE Data Collaborative
Heather Hough, Erika Byun, and Laura Mulfinger, September 2018.

In Need of Improvement? Assessing the California Dashboard After One Year
Morgan S. Polikoff, Shira Korn, and Russell McFall, September 2018.

These and all GDTFII studies can be found at www.gettingdowntofacts.com.

Introduction

What impact do California's publicly-financed preschool programs have on kindergarten readiness and student success?

Which schools are moving low-income, Hispanic English learners to full English proficiency most successfully?

Are smaller K-3 class sizes a smart investment for California?

Currently, the ability of California education leaders and policymakers to answer such questions is severely limited by weaknesses in the state's education data systems. Many of those weaknesses could be readily solved.

If statewide data systems were better integrated and more accessible, California leaders would have a stronger basis for determining the school and policy improvement efforts they want to implement. Such systems create the transparency that local educators and state leaders require to identify needs and areas of growth, validate intervention strategies, and monitor progress. They shed light on how well state policies and programs are working (together or separately) and identify areas where change is needed.

In recent years, California has made important improvements in the collection and availability of education data, but it remains far behind other states in its ability to use those data to understand what is working in its education system and what isn't. The state could strengthen that ability by taking advantage of modern computing technology that enables governments at all levels to better use the data they already collect.

This brief describes the current status of California's education data systems. It documents some areas of inquiry that could be strengthened by a better integrated and more accessible approach to education data, providing examples—from other states and from within California—of what is possible. The brief also includes a discussion of the technical and political levers that could help make those improvements.

KEY FINDINGS

- California has improved its data systems through the development of the California Longitudinal Pupil Assessment and Data System or CALPADS.
- Data availability and usefulness still fall short of what other states have developed.
  - CALPADS is not integrated with other data systems.
  - CALPADS and other systems are underutilized due in part to access issues.
- California has already made a big investment in data systems, and with some additional investment, these data could be made more widely available and more useful.
- Developing a comprehensive statewide data system requires leadership and commitment.

Summary of Key Findings

California has improved its data systems through the development of CALPADS

During the past decade, California has integrated multiple K-12 education data sets into the California Longitudinal Pupil Assessment and Data System. Maintained by the California Department of Education (CDE), CALPADS rolls up some, but not all, of the student- and staff-level information that school districts collect. It includes individual student data on school attendance, special program participation (e.g., English learner and special education programs), course enrollments, and state test scores. Each school district in the state reports these data to the department, which compiles them in a statewide data system that follows students across time.

The development of this data system has expanded the possibilities for how data can be analyzed and reported, and it has supported policy development that would not have been possible otherwise. For example, the state uses CALPADS data to develop an "unduplicated count" of high-needs students to determine Local Control Funding Formula (LCFF) supplemental and concentration grant amounts for school districts.

CALPADS data have also been tapped to create the school- and district-level accountability reports provided in the California School Dashboard (see box). The Dashboard fulfills the accountability requirements of LCFF and the federal Every Student Succeeds Act (ESSA).

.

THE CALIFORNIA SCHOOL DASHBOARD—AN EXAMPLE OF WHAT IS POSSIBLE

The California School Dashboard is the state's consolidated accountability system. The state's Local Control Funding Formula lists eight priorities, which range from pupil achievement to parent engagement. Districts are required to describe how they will address each of these priorities in their Local Control and Accountability (LCAP) plans. The state-provided indicators in the Dashboard provide data for five of these eight priorities. The Dashboard is intended to:

1) Support local education agencies (LEAs) in identifying strengths, weaknesses, and areas for improvement;

2) Assist in determining whether LEAs are eligible for assistance; and

3) Assist the state superintendent of public instruction in determining whether LEAs are eligible for more intensive state support/intervention.

Under the federal ESSA, the state must develop a method for identifying low-performing schools for intervention. The Dashboard will be used for that purpose as well.

The Dashboard's website (caschooldashboard.org) offers detailed ratings for California's public schools and districts. Each school and district Dashboard report presents performance on an array of indicators such as student results on state assessments, the progress of English learners, and graduation and suspension rates. The Dashboard also provides information about the performance of student subgroups. Schools and districts are rated in a way that takes into account both the current status and change on each indicator.

In addition to providing data for the California School Dashboard, CALPADS data can be accessed by pulling reports from the Education Data Partnership (ed-data.org) and DataQuest (data1.cde.ca.gov) sites. Local test coordinators can access student assessment data from the California Assessment of Student Performance and Progress (CAASPP) portal (caaspp.org). Researchers working independently or in collaboration with school districts or other policymakers can, at least in theory, obtain individual-level data for studies that align with CDE's research priorities. These arrangements must protect the confidentiality of the data, and access remains very limited.

CDE collects other K-12 data that can be used to inform policy and practice decisions. For example, the Education Data Partnership website offers data based on districts' annual financial reports.

All of these systems serve up "frozen data" that provide information about the status of the system at a given point in time. As such, they are primarily useful for:

- describing system performance, such as the portion of students who graduated this past year, and
- measuring the impact of local policies, such as determining what proportion of students have access to arts courses.

This is in contrast to "real time" data local educators can use to evaluate how well individual students have mastered their multiplication tables or whether a new approach to reading instruction is working. The report by Heather Hough and associates, *Using Data for Improvement: Learning from the CORE Data Collaborative,* goes into more detail about this use of data, as does a separate brief in the *Getting Down to Facts II* project, *Toward a Common Vision of Continuous Improvement in California*, which discusses continuous improvement.

Data availability and usefulness still fall short of what other states have developed

During the past decade, California has made considerable progress in compiling and reporting on K-12 education data. However, the state still lacks the kind of single statewide education data system found in some other states.

CALPADS is not integrated with other data systems

CALPADS data cannot be linked to databases maintained by California state agencies with responsibility for other aspects of children's lives (e.g., social services) or by other segments of the education system (e.g., higher education). Examples include:

- the state's three higher education systems (the community college, University of California, and California State University systems),
- the California Student Aid Commission,
- the early education sector,
- the California Commission on Teacher Credentialing,
- the Employment Development Department,
- social service agencies, and
- the criminal justice system.

The state of California or its agents already own each of the above data systems and each already contains a great deal of information in a readily usable form. Integrating these databases and providing better access could help educators, policymakers, and researchers better understand the K-12 system.

In recent years, efforts to integrate data across the various systems have largely come from the bottom up. Individual districts or other education providers saw how better data could add to their ability to provide needed services. These projects have mostly been localized, limited to a single pair of data systems (e.g., a single district's records merged to the local community college district) or to a particular geographic area. The bottom-up approach limits the opportunity to improve systems statewide. Nevertheless, these local projects, along with examples from other states, demonstrate the value that could be generated through more intentional, comprehensive efforts to integrate data across the state (see the box on the following page for examples).

INFORMATION THAT INTEGRATED AND ACCESSIBLE DATA SYSTEMS CAN PROVIDE

Understanding how K-12 schools and other systems interact

- Matching students from K-12 schools to the community colleges they attend has made it possible to examine what measures besides test scores could inform course placement for new community college students. **Cal-PASS Plus**
- Researchers look at how well the K-12 and higher education systems are aligned to ensure that K-12 students are prepared for postsecondary success. **California Ed Lab at the University of California, Davis**
- A data repository connects academic and human-services data to help school staff understand students' mental health and child welfare involvement, while helping social workers and caseworkers understand children's school performance, attendance, and disciplinary history. **Allegheny County, Pennsylvania**
- School districts now have information about which of their students are foster youth and/or are eligible for free meals under the National School Lunch program. **CDE/CALPADS and the California Department of Social Services**

Comparing schools and districts to identify successes and challenges among those that are similar

- Districts participate in a data collaborative that makes the state data more valuable for district leaders by (a) adding additional data sources, (b) using these data to create additional measures of student progress, and (c) creating data displays and tools that enable district and county leaders to look at data more easily and in different ways. **CORE Data Collaborative**
- Publicly available estimates of earnings by discipline and college enable researchers to examine the returns to career-technical certificates and degrees. **California Community College system and the Employment Development Department (EDD)**

Evaluation of the effectiveness of K-12 policies and state investments

- State leaders can evaluate the effectiveness of policies designed to reduce teacher shortages in math, science, and special education through the use of longitudinal data on teachers and students. **State of Florida**
- By linking birth and public school records, officials can measure the impact of statewide early childhood policy initiatives—including funding to improve child care quality and expand preschool slots—on children's math and reading achievement by the end of elementary school. **State of North Carolina**

CALPADS and other systems are underutilized due in part to access issues

CALPADS data could be of tremendous value in evaluating whether a specific state education policy is effective and in identifying schools and districts that have similar challenges and different outcomes. However, CALPADS data are not as accessible to interested researchers as they might be, and are therefore underutilized.

CDE has developed standardized procedures by which agencies and researchers can gain access to CALPADS data for specified research projects. But a shortage of resources at CDE and a lack of clarity about permissible uses prevents analysts from using the data to nearly the extent that similar data systems are used in some other states to understand and improve educational practices.

Moreover, the CALPADS access procedures are the exception, not the rule: most other California data systems have no regular processes or procedures by which analysts from other agencies, policymakers, or researchers can obtain access.

Access to large databases can facilitate the work of state agencies, large school districts, and the higher education community, to name a few. But simply providing access assumes a higher level of analytical skill than is present in the majority of local education agencies. CALPADS data are the backbone of the California School Dashboard and other state data systems, but these source data are not made available to districts or those who support them in ways that are easy to use.

The CORE districts—eight urban districts in California that banded together to improve education practice—set out to address that issue by creating the CORE Data Collaborative. The collaborative has developed systems and tools to make the state's data system more valuable and user-friendly, including offering better data displays, calculating innovative metrics, and adding new sources, such as surveys and links to higher education. The collaborative's experiences illustrate the importance of making it easier for local education leaders to access information from a statewide system and put those data to work.

CORE's data system is available to any district that is able to pay a small fee to join the collaborative. Now that an infrastructure has been developed by CORE for California data, the cost for adding additional districts is relatively low. This existing system represents an opportunity for the state to get a head start on a major improvement in data accessibility.

California has already made a big investment in data systems, and with some additional investment, these data could be made more widely available and more useful

California's current data systems provide limited support for the variety of uses that are possible. Many agencies collect data for themselves, or go without, because they do not have access to the data systems that already house the information. This leads to a costly duplication of effort, ineffective service delivery, and holes in the web of education services that the state provides.

The locally-initiated efforts described above, and many others, illustrate the value of linked data for informing education policy and practice. But they are not a substitute for a statewide system. Each effort involved large investments of time and resources, typically over several years. In some cases, the investment created a one-off data system that could not—due to legal or technological limitations—be used for other purposes or by other agencies.

A second important limitation is the spotty coverage provided by this patchwork of local data-linking efforts. Some districts or colleges participate, while their neighbors do not. This patchwork approach limits the value of the data, as some students disappear from a longitudinal database simply because they moved to a nonparticipating district or college. It is also difficult to know how the results would generalize to the state as a whole. In addition, a system based on voluntary initiatives will tend to exclude smaller and rural districts. A more comprehensive, statewide data system would enable these districts to benefit from the same data access that larger, urban districts have.

Statewide data systems also make possible a range of useful comparisons, analyses, and reports:

- Agencies can access relevant educational, human services, and health data even when students move across district or county lines.
- K-12 districts can more accurately understand students' outcomes in higher education and the workforce even if their former students have moved to a different part of the state.
- Districts and counties can compare their outcomes to other districts and counties that serve similar populations, making it possible to identify a good comparison group for program evaluation purposes.
- Policymakers can more easily determine the impact of state policies and understand how that impact varies across localities.

With statewide systems, agencies and regions benefit from a centralized resource that links data and develops protocols for sharing and displaying data. These benefits are likely to be especially important for smaller agencies and localities that typically have fewer resources for data management and analysis than their larger counterparts. To ensure that data and analyses remain relevant to local needs, agencies can augment the statewide linked data with their own local data, as desired.

Developing a comprehensive statewide data system requires leadership and commitment

Developing data systems that serve teachers, school leaders, policymakers, and California's parents and children will require the state to overcome a variety of technical and political barriers. The experiences in other states suggest that these barriers can be surmounted, but that political leadership is required.

An integrated data system requires obtaining data from many sources, linking them together, storing the linked data, ensuring data quality, sharing them with authorized users, and keeping data secure from those who should not have access. These steps have been accomplished in many other states, and the technical challenges are well understood. California could draw on these states' experiences, the guidance of organizations such as the Data Quality Campaign, and expertise residing in agencies, nonprofit organizations, and universities to determine the best solution to these issues. In fact, in the K-12 domain, California has already overcome these technical challenges to develop CALPADS.

The CALPADS experience, however, also points to the importance of creating governance rules that allow access to the data while specifying the conditions under which the data can be used. States that provide differing examples of such governance rules include Florida, Kentucky, Maryland, North Carolina, Texas, and Washington. (Details are available in the full report.)

Indeed, the technical barriers to developing a more comprehensive education data system in California are fairly easily solved, at least compared with the political considerations. Experiences in other states point to three factors that are important for maintaining political support:

1) Strong leadership is critical; successful efforts in other states have typically required both political leadership and individuals inside the Department of Education and other agencies who understand the value of the work and are committed to developing better data systems and to improving the effectiveness of the education system.

2) State-level actors need to be consistently engaged with stakeholders, particularly county and district officials, throughout the process.

3) Advocates for improved data systems—including agency leaders, outside allies, and political leaders—should clearly articulate the value of the data system, how it will be used, and how privacy will be protected.

In developing support for a comprehensive data system in California, it seems particularly important to emphasize that data can be used for a wide range of purposes other than accountability, and that more comprehensive data systems will make it easier for agencies to serve Californians well.

Conclusion

A more integrated and accessible California education data system could enable a wealth of important education questions to be addressed, substantially reduce costs, increase efficiency, and allow students to be better served.

California has made progress with the creation of its CALPADS data system, but that investment is underutilized, in part because the data are mostly available in aggregate form. That limits their utility for districts that lack the capacity to compile large data sets, conduct analyses, and provide the information to their constituents in usable formats. Building local capacity and encouraging use of these data, as well as integrating education data with other state data systems, would improve access to information in California and bring California's system closer to the systems seen in some other states. A more integrated system would also further the goals of transparency often pursued by the many independent, ground-up data projects that have proliferated in California given the absence of a comprehensive statewide education data system.

California's existing political, technological, and organizational barriers to the creation of improved statewide data systems are not insurmountable. Other states have overcome these barriers, and their experiences demonstrate the substantial value of better data systems. Enormous technical projects are not required—significant progress can be made at a relatively low cost, given the political will to overcome bureaucratic and organizational inertia.

Lead Author Biographies

Meredith Phillips is associate professor of Public Policy and Sociology at the University of California, Los Angeles. She studies the causes and consequences of educational inequality. In 2012, she cofounded the Los Angeles Education Research Institute (LAERI), a research-practice partnership with the Los Angeles Unified School District.

Sarah Reber is associate professor of Public Policy at the UCLA Luskin School of Public Affairs and a California Policy Lab affiliate. She is an economist who studies school segregation, school finance, and college access.

Jesse Rothstein is professor of Public Policy and Economics at University of California, Berkeley. His research includes studies of education finance, teacher evaluation, school and neighborhood segregation, and higher education admissions. He is also cofounder and codirector of the California Policy Lab, founded in 2017 to facilitate close working partnerships between state and local public agencies in California and University of California researchers. The lab works with agencies to analyze and improve public programs through empirical research and technical assistance.