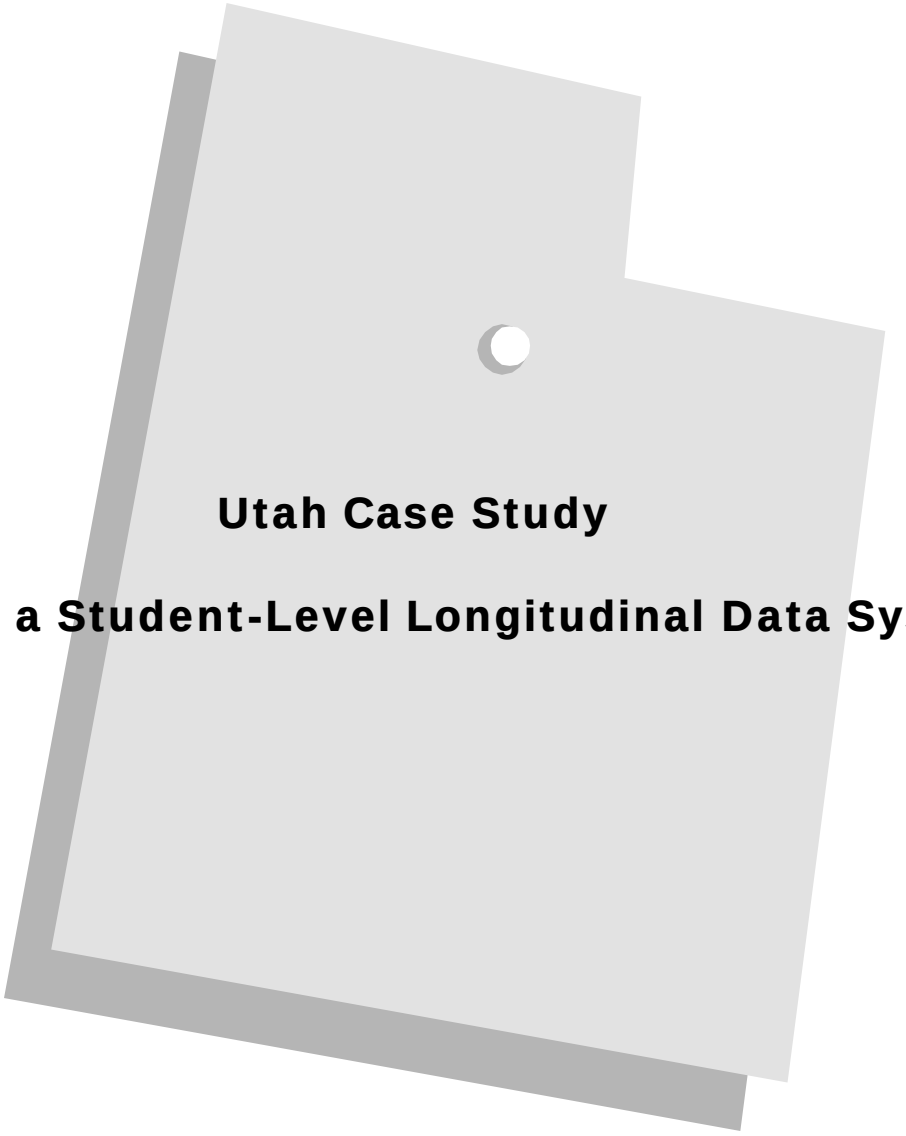
DATA QUALITY CAMPAIGN

Using Data To Improve Student Achievement



Utah Case Study

Building a Student-Level Longitudinal Data System

Data Quality Campaign

August 2006

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

To these ends, four site visits were conducted in the spring of 2006 to state education agencies (SEAs) to gather information on their experiences in developing statewide longitudinal data systems: Florida, Utah, Virginia, and Wisconsin.

This publication was produced by the Data Quality Campaign/National Center for Educational Accountability for submission under contract with the Bill & Melinda Gates Foundation. The report was written by Cherry Kugle and Nancy Smith and was reviewed and approved by the Utah State Office of Education.

Utah: Not Waiting for Mandates

The Utah State Office of Education (USOE) oversees 40 public school districts and 52 charter schools in which 510,000 students were enrolled in the 2005-06 school year. Although per pupil expenditures in Utah have historically been among the lowest in the nation, average scores on the National Assessment of Educational Progress (NAEP) for Utah students are higher than national averages.

History of Development

Data Collection

Utah has not bought a package data system, choosing instead to build almost everything from scratch. The USOE began to collect integrated student data, replacing "stovepipe collections," in 1994-95. At that time, input from program areas was solicited to ensure that all data needs were met. The data were collected as large, flat, fixed-format files that programmers imported into a relational database and accessed to meet reporting requirements. This effort evolved to the Clearinghouse, a statewide system that collects student-level data at three points throughout the year from district-level student information systems.

A statewide licensing system for teachers has been in place since the 1970s. Prior to 1999 teachers' social security numbers (SSNs) were used as identification (ID) numbers; these have now been replaced with unique SSN-independent IDs. A student-teacher linkage was built into the course membership record as part of the Clearinghouse submission.

Legislation

In 1998-99 the state legislature mandated an effort that became the Data Warehouse. They provided an initial $150,000 for hardware and appropriated $350,000 a year later, plus $200,000 a year for ongoing costs for maintenance. Clearinghouse data are passed to the Data Warehouse for integration with other data systems, including the results of state assessments. Although the development of the Data Warehouse was met with skepticism in some program areas, it has become key to developing the state accountability system enacted in 2000 (Utah Performance Assessment System for Students, U-PASS) and meeting the federal requirements of No Child Left Behind (NCLB).

Partnership

In 1999-2000, the department that is now Assessment and Accountability purchased a contract with Cognos, a producer of business intelligence software, that allows all districts in the state to use the software at no charge locally. In 2005 the contract was renewed through 2013. The intent was to provide assessment data to district personnel, principals and teachers in a format they could use to inform instruction.

Unique Identifiers

Planning for a statewide student identifier (SSID) system began in October 2004, with development continuing through March 2005. Testing and training took place in July 2005, and implementation in September 2005. Although districts had local student IDs and the state had internal student IDs for many years, the state education agency (SEA) was unable to track students across years until recently. Under the SSID system, districts request ID information from the state, either by retrieving an existing SSID for a transferring student, or getting a new SSID for a student new to Utah public education. All student level data

(*e.g.*, assessment, Clearinghouse) submitted to the USOE must contain the SSID for each student.

Essentially, state and national mandates to calculate individual student growth measures, assign accountability categories, and create district and school performance reports, came after the USOE had already put most of the pieces in place to implement them. A notable exception was the SSID that became operational in the 2005-2006 school year. The emphasis early on was to get data into the hands of teachers to inform instruction and to ensure the data would be accurate so it would be worthwhile and trusted. The state mandate provided the authority to collect data at the student and teacher level, after it had been collected for years.

Implementation Issues

The state legislature has been very involved in accountability development in Utah and wishes to see the state continue their system (U-PASS). As is true in other states that already had accountability systems in place when NCLB was enacted, the federal accountability system requirements have created conflicts with the state system. Currently the state has two accountability systems with some overlap.

The state legislature meets annually for 45 days, with interim committees that can meet year round. The USOE sometimes cannot take action until the interim committee gives approval. New requirements and frequent changes to statute create difficulties in terms of sustaining all projects.

Utah's SSID system is in large part based on similar systems implemented in other states (*e.g.*, California, Colorado, Florida and Oregon). Staff from USOE used Personnel Exchange funds available from the National Center for Educational Statistics (NCES) to make site visits to Oregon and Colorado to learn about their systems.

Districts were involved early in the planning to determine "how it would hurt and what they needed." This allowed some potential conflicts to be resolved early in the process. USOE staff stated that districts frequently provide ideas and solutions that they would not have thought of at the state level. The majority of issues for the districts center on funding (or rather, the lack thereof), not design. Weekly Data Warehouse meetings are held at the state department and districts are welcome to come to those at any time. On a more formal basis, the USOE conducts monthly meetings with districts to elicit their input in order to avoid imposing too costly a set of changes and to get their feedback on how to do things more efficiently and effectively. There are also semi-annual data meetings that provide an opportunity for consistent communication from the state department and facilitate districts' integrating changes on a regular basis.

When the Cognos contract was implemented, the state did not have the funds to hire trainers. Over time, districts began to see the benefits of the software and there is now a vibrant user group which facilitates communication with the state and shares knowledge and practices among districts.

When hiring contractors, USOE staff are clear in their specifications and hire senior-level consultants, often working with consultants who had previously contracted with USOE. There is a high degree of productivity between the state staff and the consultants, who work under both cooperative agreements and fixed contracts.

Other factors that facilitated the work were attendance at national meetings (such as the annual NCES MIS Conference and Council of Chief State School Officers committees) that created awareness of states that were further along in developing their systems, and

longevity of staff which helps create a long-term vision of what is needed and provides consistency over time in implementing the vision.

Costs

Estimated costs to the state:

- Because legislation frequently changes the scope of work and projects frequently intersect and overlap, it is difficult to determine firm costs for specific projects. In 2005-06, it is estimated that all state level IT costs for all systems related to longitudinal student data (SSID, Data Warehouse, Clearinghouse, U-PASS and NCLB reporting, as well as CRT scanning and scoring) totaled approximately $800,000 per year.
- There are 50 staff members in the IT department at USOE. Approximately ten FTEs spend the majority of their time on state-level data systems and on federal reporting to the Electronic Data Exchange Network (EDEN), with an additional two FTEs working on teacher licensing.
- USOE scans and scores the answer documents for the assessment program and hires temporary employees to produce a four-day turnaround. Results are provided back to districts down to the objective level.
- In 1998-99 the legislature provided an initial $150,000 for hardware and appropriated $350,000 a year later, plus $200,000 a year for ongoing costs for maintenance.
- Approximately $200,000 in contract funds, including the development of a user manual. The contractors were those who helped developed the Data Warehouse so they were familiar with Utah's data systems and needed very little supervision. USOE has since hired a fulltime developer to maintain the system and provide enhancements over time. One technical support specialist is also assigned to provide part-time SSID support.
- Cognos: $2.2 million dollars over 14 years, which comes to approximately 37¢ per student in Utah per year.

Estimated costs to districts:

- Districts did not receive financial help or staff to implement the system. Some technical training is provided. Data are submitted via a website so hardware and software costs to the districts are minimized.
- In general, districts did not add staff because there were no additional appropriations to fund a new position. Work is generally absorbed as "other duties as assigned" and as lower priority projects are displaced. District representatives interviewed variously estimated the staff impact of managing the SSID system as:
  - using ½ to ¾ of a high-level FTE (District Size: 24,000 students)
  - increasing a ½ time FTE to full time (District Size: 62,000 students)
  - using two FTEs to maintain after using eight FTEs to implement (District Size: 78,000 students)
- A district representative described student mobility as a big issue for verifying the SSIDs. Described as a human, intuitive process that is currently an add-on to existing work loads, it was estimated that it will take at least a tenth of an FTE to do the work, once the high-level staff member learns how to do it efficiently and can train someone else to interface with the state to resolve ambiguities.
- Cognos: training is being done on district time by one individual. In many districts, the Educational Technology Directors have become the contact and training sources for Cognos use.

Benefits and Uses of System

Benefits

- In general, USOE staff feel that the state is now getting higher quality data and more comprehensive data, but cannot assign specific costs and savings values to these factors. There have been reductions in paper collections and time spent doing data entry. There have definitely been improvements in efficiency as the data collection interface is a huge labor-saving device.
- In particular, implementing an SSID has resulted in improvements in data quality for high-stakes testing, longitudinal studies, and NCLB reporting.
- The implementation of NCLB has motivated districts in Utah to focus on data accuracy. The quality and timeliness of the data collection has improved every year since Adequate Yearly Progress (AYP) has been calculated and publicized for each school and district.

Uses

- It is anticipated that the SSID system will help give insight and oversight into how programs and practices are working.
- Districts receive their own subset of data (with value added assessment indicators) which can be integrated with additional student and teacher data at the district level both within and across years. The Cognos application provides an educator-accessible mechanism for integrating state and local data.
- Data are being used more often to inform decisions at the state and local levels.
- *Ad hoc* data are not generally provided to researchers. A goal of the USOE is to provide public-use datasets capability for research purposes, or to create a research agenda to take to the higher education community. One area of current interest is the Novice Teacher Program which will examine how training of teachers affects instruction. The goal is to provide the data in sufficient detail that analyses and reports can be created by researchers.

Lessons Learned

Design

- Build only what is necessary and required and add “nice to know” data later.
- Don’t design a new system without determining what is already available.
- While it is easy to look for an external resource, it is ultimately more powerful and useful to build capability within your agency.
- Before investing in business intelligence software, think about what the actual needs are in your state.

Staffing

- Bring in policy and program area staff to help define data and changes needed over time. Each program area should have the capability to extract data and produce reports.
- Program area staff and IT staff need to communicate with one another to ensure they are providing consistent definitions and instructions about processes to districts.
- Require districts to resolve ID conflicts and duplicates unless there are sufficient staff at the state agency to manage that effort.
- Promote the designation of a data coordinator position in each district, similar to a testing coordinator position.
- The data retrieved are only as good as the people collecting and coding them, who must share and understand common definitions. Now that high-stakes testing can

affect people's lives, the transactions between students or parents and the person coding the data are critical.
- Due to turnover in staff at the local level, districts must constantly work to maintain and sustain the data quality cycle.

Maintenance/Change Control
- Student attributes are not pre-coded or bubbled onto assessment answer documents at the time of testing. Instead, these are assigned from the SSID system during the scanning and scoring process.
- The capability to match students to teachers, and teachers to classes taught exists in the data collection structure.
- The capability to match students from the public education system to higher education institutions within the state will be available in 2007. The need for concurrent enrollment data has led to the development of a central warehouse for higher education data, eliminating the need to request data from each individual college or university.
- Educational Technology has emerged as an area providing training to administrators to examine their data and link them to teachers in their building in order to improve instruction.
- Districts may not want to depend on the statewide ID for local systems because other users in the state can access the statewide ID data and modify them.
- Create and document business rules regarding how the data submitted are used for accountability purposes.

Partnership
- Build a self-sustaining coalition among the districts. Solicit participation so system implementation can be seen as helpful and collaborative, rather than an edict being imposed on the districts.
- To assure quality data, state education agencies must have the authority, resources, and freedom to direct and manage education data collection and reporting.

Communication
- Although data awareness is increasing among district staff, skill in using and interpreting the data have been slower to develop.
- District representatives described allocating resources to "data coaches" who help teachers understand and evaluate their data. They described more positive effects from this approach compared to simply providing the data to teachers and expecting them to use them.
- After analyzing their data, one district put all their high schools on a common schedule to accommodate the mobility of their students. "When the delivery of the curriculum is uniform across schools, formative assessments have greater power because the students are being assessed in real time with comparable data." This district is now harvesting data out of grade books on a weekly basis to target students who are struggling. However, these efforts come with a significant staff development cost.
- Introducing data to the schools "so they know what is available and what the data can tell you is a very important effort. You have to be able to ask the right question, as the way you ask it is important. Everyone in the education system has a role to play and data help to tell the story of where we are and where we need to go."
- Marketing and public relations directed internally to state department and externally to school districts, policymakers, and the public is important. Although it takes a different skill set to do marketing than build the system, invest the time to sell the vision.

- If promoted properly, districts will see some benefit from it, especially in terms of more accurate accountability data and faster turnaround of student information.
- Although *accountability* is the current driving force of data collection activities, progressive districts see data collection as useful for *informing instruction* and targeting *intervention*.

Other Considerations

- In terms of implementing a change from inception to quality data, a realistic time frame is at least three years. Even then, you will find that as policy and reporting needs change the system will require continual modifications and enhancements.
- Technology is not the issue now. What is required to do this well is professional development and the articulation of polices and standards rather than a mandate of how it should be done.
- Flexibility in the interpretation of data definitions and business rules is inversely proportional to data quality.
- Do not underestimate the magnitude of the job. There will always be maintenance and ongoing changes that must be managed.

Recommendations for Future Development

- Continue to provide training throughout the state to create a culture of using data to make instructional decisions.
- Continue to promote a good single-point of contact in each district to serve as the data steward.
- Continue work that will enable matching students from P-12 public education system to higher education data.
- Provide a public-use datasets capability for research purposes, or create a research agenda to take to the higher education community.
- Monitor and provide input into new rules and regulations so new data collection and reporting requirements can be successfully implemented.