Data Warehouses and Student Achievement

This report is comprised of two sections. The first section profiles eight school districts across the country that have implemented data warehouse/data management software district-wide. State assessment statistics from before software implementation and after implementation are examined to determine if changes in student achievement have occurred and to what extent. The second section is a brief review of literature pertaining to data warehousing at school districts.
The Impact of Data Warehouses on Assessment Outcomes

This section will examine the standardized test scores of a number of districts before and after implementation of a data warehouse/data management system, in order to determine if there is evidence whether the use of data warehouses at the district, school, and teacher levels affect standardized test achievement. Districts that are examined in this section were selected based on two criteria: if data warehouse software is currently being used by the district and its schools, and if the year of software implementation could be determined by looking at previous literature on the topic, software websites (case studies and client lists), or school district websites. A variety of data warehouse software products are also represented in the sample of districts examined in this section.

States and districts use different assessments with different performance standards; therefore, standardized assessment scores before and after implementation of data warehouse software at each district are examined to account for differences between tests and standards. If differences in scores before and after implementation do exist, there is reason to believe that district use of data warehouse software contributed to the change, whether it is positive or negative. However, there may be other district-wide or school-wide changes independent of data warehouse use—changes in policy, instructional approaches and philosophies, state assessment standards, and student demographics to name some examples—that can affect standardized assessment scores from year to year. Therefore, definitive conclusions about the causal effect of data warehouse usage on assessment scores should not be drawn based on the findings presented here. Although the degree of causality is difficult to determine, the example districts in this section, each using a different type of data warehouse software, are intended to provide evidence that data warehouses may be an integral part of a district’s efforts to improve learning outcomes.

Methodology

State assessment results from eight school districts using data management software were gathered. As was described above, assessment results from the year prior to and two years following implementation of data software were determined. For all districts, the statistic used to measure district-wide student achievement on the assessments is the percent of students scoring at or above a specific cutoff score determined by the respective state department of education. For example, many districts determine the percentage of students achieving a “satisfactory” or “proficient” score, and a separate percentage of students achieving an “advanced” level score. For the purposes of this report, whenever possible, these two percentages were summed into one percentage for each grade level and subject for the school years prior to and following implementation. For example, the percentage of students earning either satisfactory or advanced scores for 4th grade math and separately for 4th grade reading in the year prior to implementation and two years after implementation.
could be determined for most districts presented in this section. For most districts, assessment results for math and English/language arts portions of state tests are used because those subjects are tested most frequently.

In order to determine the amount of change in assessment results before and after data software implementation, the difference in the percent of students achieving satisfactory or advanced scores for each grade and subject was calculated and is presented for each district in the following profiles.

**Summary of Results**

- Among all eight districts, the average numerical change in the percentage of students achieving satisfactory or advanced scores for all subjects was **an increase of 7.8 percent from the year before implementation compared to the second year after implementation**.
  - The average numerical change in the percentage for **reading/language arts** portion of the test among all districts was an **increase of 7.5 percent**.
  - For **math**, the average numerical change was an **increase of 9.0 percent**.

- These findings indicate that **the implementation of data warehouse technology is associated with improved student achievement**.
Poway Unified School District (CA)

According to the Consortium for School Networking’s case studies of districts engaged in data-driven decision making, Poway USD implemented its Total Information Management System (TIM) in fall 2002. The system allows users to view classroom and student data that can filtered by period, course, or any of the NCLB categories such as ethnicity, gender, and so on. The data warehouse pulls data from the student information system, human resources database, and student assessments.1

California Standards Test Scores Comparison

The following table shows the percent of students scoring at the “proficient” or “advanced” level for grades 2-11 in 2001 and 2004, one year before implementation and two years after. California changed from the Stanford 9 test to the CAT/6 test between 2001 and 2004; the Stanford test had previously tested both math and reading for grades 2-11 but was not administered in 2004 and so cannot be used for comparison of achievement between years. Therefore, results from the California Standards Test are examined here. Achievement data from the English/Language Arts version of the CST test were available for 2001, but unfortunately data for the 2001 math test were not.2

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English/Language Arts (2001)</strong></td>
<td>67%</td>
<td>63%</td>
<td>68%</td>
<td>60%</td>
<td>58%</td>
<td>59%</td>
<td>61%</td>
<td>58%</td>
<td>57%</td>
<td>52%</td>
</tr>
<tr>
<td><strong>English/Language Arts (2004)</strong></td>
<td>66%</td>
<td>63%</td>
<td>72%</td>
<td>72%</td>
<td>68%</td>
<td>71%</td>
<td>62%</td>
<td>67%</td>
<td>60%</td>
<td>58%</td>
</tr>
<tr>
<td><strong>Percent Change</strong></td>
<td>-1%</td>
<td>0%</td>
<td>+4%</td>
<td>+12%</td>
<td>+10%</td>
<td>+12%</td>
<td>+1%</td>
<td>+9%</td>
<td>+3%</td>
<td>+6%</td>
</tr>
</tbody>
</table>

Source: California Department of Education

As can be observed, the percent of students achieving a level of “proficient” or “advanced” on the English/Language Arts version of the California Standards Test increased from 2001 to 2004 except for grades 2 and 3. Three grade levels witnessed an increase of at least 10 percent more students achieving proficient or advanced scores in 2004 compared to 2001.

---

Firebaugh-Las Deltas Unified School District (CA)

The district began using DataDirector software in 2004. The district has data clerks at each school to input various classroom and benchmark test and assessment data so that teachers have access to test results as soon as possible.3

California Standards Test Scores Comparison

The following table shows the percent of students scoring at the “proficient” or “advanced” level for grades 2-11 in 2003 and 2006, one year before implementation of DataDirector and two years after. Results for the CST math test are available for grades 2-7, whereas the English Language Arts test results are available for grades 2-11.4

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>English/Language Arts (2003)</td>
<td>14%</td>
<td>14%</td>
<td>25%</td>
<td>13%</td>
<td>13%</td>
<td>17%</td>
<td>12%</td>
<td>7%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>English/Language Arts (2006)</td>
<td>28%</td>
<td>22%</td>
<td>46%</td>
<td>27%</td>
<td>26%</td>
<td>26%</td>
<td>27%</td>
<td>28%</td>
<td>18%</td>
<td></td>
</tr>
<tr>
<td>Percent Change</td>
<td>+14%</td>
<td>+8%</td>
<td>+21%</td>
<td>+14%</td>
<td>+13%</td>
<td>+9%</td>
<td>+14%</td>
<td>+20%</td>
<td>+16%</td>
<td>+8%</td>
</tr>
<tr>
<td>Math (2003)</td>
<td>23%</td>
<td>19%</td>
<td>40%</td>
<td>9%</td>
<td>16%</td>
<td>22%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Math (2006)</td>
<td>33%</td>
<td>34%</td>
<td>60%</td>
<td>37%</td>
<td>36%</td>
<td>29%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+10%</td>
<td>+15%</td>
<td>+20%</td>
<td>+28%</td>
<td>+20%</td>
<td>+7%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Source: California Department of Education

The data show that the district has seen significant gains in the percent of students achieving proficient or better scores on state assessments since 2003, one year before DataDirector was implemented. Most grade levels witnessed double-digit increases in this percentage for both subjects between 2003 and 2006.

---

http://star.cde.ca.gov/star2006/viewreport.asp
Boston Public Schools (MA)

A study by the Annenberg Institute for School Reform at Brown University examines the effect of data warehousing at public school districts. One district examined is Boston Public Schools. According to the study, the district implemented the MyBPS web portal to provide school and classroom-level student data to educators. Attendance, grades, class schedules, and state assessment data are all available through the portal. Teachers are able to filter assessment results by student subgroup, examine questions that were most often missed, and view individual students’ responses.⁵

The MyBPS system was implemented in the 2003-2004 school year.⁶ The portal has been continuously improved, including software upgrades and the integration of a formative assessment system into the MyBPS system to help teachers use formative assessment to refine instruction methods.⁷

**Massachusetts Comprehensive Assessment System (MCAS) Results**

The following table shows the percent of students scoring at the “proficient” or “advanced” level for grades 4, 6, 7, 8 and 10 on the English/Language Arts (ELA) and math versions of the MCAS assessment in 2002 and 2005 (one year before MyBPS implementation and two years after). Since students do not take both subject tests every year, test data for ELA and math are not available for every grade level; however, comparisons can still be made between 2002 results and 2005 results for any given grade level and subject.⁸

**Table 3: Boston Public Schools State Assessment Scores - Percent Proficient or Advanced (2002 & 2005)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24% 40% 35%</td>
<td>25% 44% 38%</td>
<td>+1%</td>
<td>15% 16% 19% 24%</td>
<td>21% 23% 23% 39%</td>
<td>+6%  +7%  +4%  +15%</td>
</tr>
</tbody>
</table>

Source: Massachusetts Department of Elementary and Secondary Education

---

⁸ Massachusetts Department of Elementary and Secondary Education. “MCAS Performance Results.” http://profiles.doe.mass.edu/state_report/mcas.aspx
As the data demonstrate, **there was a modest increase in the percent of students achieving proficient or advanced level scores** on the MCAS language arts and math tests between 2002 and 2005; 10th grade math results witnessed the largest improvement, with 15 percent more students achieving proficient or advanced scores compared to 2002.
Brockton Public Schools (MA)

Brockton Public Schools began using eScholar Complete Data Warehouse in fall 2005. The software integrates all of the district’s data and allows educators to view reports online. To ease transition to the new system, the district created staff positions to facilitate implementation and work with users to help them create relevant reports, as well as create support documentation such as reference guides to ensure that the data warehouse is effectively used by district and school staff. An example of how the warehouse has been used by the district is in identifying and monitoring students who are at risk of dropping out of school in order to better focus resources and interventions.9

Massachusetts Comprehensive Assessment System (MCAS) Results

The following table shows the percent of students scoring at the “proficient” or “advanced” level for grades 4, 6, 7, 8 and 10 on the English/Language Arts (ELA) and math versions of the MCAS assessment in 2004 and 2007 (one year before implementation of eScholar and two years after). Since students did not take both subject tests every year as of 2004, data for ELA and math tests are not available for every grade level for 2004; however, comparisons can still be made between available 2004 and 2007 results for any given grade level and subject.10

Table 4: Brockton Public Schools State Assessment Scores - Percent Proficient or Advanced (2004 & 2007)

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>4</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>English/Language Arts (2004)</td>
<td>38%</td>
<td>-</td>
<td>41%</td>
<td>-</td>
<td>54%</td>
</tr>
<tr>
<td>English/Language Arts (2007)</td>
<td>39%</td>
<td>-</td>
<td>51%</td>
<td>-</td>
<td>60%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+1%</td>
<td>-</td>
<td>+10%</td>
<td>-</td>
<td>+6%</td>
</tr>
<tr>
<td>Math (2004)</td>
<td>21%</td>
<td>26%</td>
<td>-</td>
<td>16%</td>
<td>37%</td>
</tr>
<tr>
<td>Math (2007)</td>
<td>33%</td>
<td>41%</td>
<td>-</td>
<td>24%</td>
<td>50%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+12%</td>
<td>+15%</td>
<td>-</td>
<td>+8%</td>
<td>+13%</td>
</tr>
</tbody>
</table>

Source: Massachusetts Department of Elementary and Secondary Education

The percentage of students achieving proficient or advanced level scores on state assessments increased significantly since 2004, one year prior to implementation of the eScholar software. Double-digit changes occurred for four out of seven administrations of the MCAS assessment between 2004 and 2007.

9 eScholar, LLC. “Brockton Public Schools Increases Efficiency and Drives Student Outcomes with the eScholar Complete Data Warehouse Solution.” www.escholar.com/documents/BrocktonPS_CaseStudy.pdf
Western Heights Public School District (OK)

Western Heights Public School District implemented its data management system using the Schools Interoperability Framework (SIF) Specification, a set of rules and definitions established by various software vendors, school districts, and departments of education that enable software from different vendors to share data without additional programming. Under this framework, the district’s various software applications can communicate with each other, allowing educators to aggregate, analyze, and report data more effectively. Beginning in 2003, the district began to integrate its software applications into the framework in a staged manner, gradually adding its student information system, nutrition, library, and grade book applications.

A data warehouse and custom reporting tools, developed by Mizuni, Inc., were deployed in 2005.11

Oklahoma Core Curriculum Test Results

The following table shows the percent of students scoring at the “satisfactory” or “advanced” level for grades 3, 4, 5, 8 and at the end of instruction on the reading and math versions of the Priority Academic Student Skills assessment. Results for 2004 and 2007 (one year before the implementation of a data warehouse and two years after) are provided.12

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>End of Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading (2004)</td>
<td>70%</td>
<td>80%</td>
<td>62%</td>
<td>66%</td>
<td>48%</td>
</tr>
<tr>
<td>Reading (2007)</td>
<td>72%</td>
<td>87%</td>
<td>72%</td>
<td>67%</td>
<td>54%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+2%</td>
<td>+7%</td>
<td>+10%</td>
<td>+1%</td>
<td>+6%</td>
</tr>
<tr>
<td>Math (2004)</td>
<td>62%</td>
<td>68%</td>
<td>65%</td>
<td>60%</td>
<td>16%</td>
</tr>
<tr>
<td>Math (2007)</td>
<td>61%</td>
<td>68%</td>
<td>76%</td>
<td>76%</td>
<td>76%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>-1%</td>
<td>0%</td>
<td>+11%</td>
<td>+16%</td>
<td>+60%</td>
</tr>
</tbody>
</table>

Source: Oklahoma State Department of Education

In general, a larger percentage of students in all tested grades achieved satisfactory or advanced scores on the state assessment in 2007 compared to 2004, except for a slight decrease in achievement on the 3rd grade math assessment and no change in achievement for 4th grade math. Large increases in achievement between the two years occurred for 5th grade reading and math, 8th grade math, and the end of instruction math assessment.

Rockwall Independent School District (TX)

Rockwall ISD sought a single, integrated database solution to replace its outdated software and began implementing Skyward’s School Management System software in 2006 to handle the district’s student data and financial data. Implementation of the student data software package was completed in January 2007; the software is comprised of various modules that track attendance, demographics, current and next year class scheduling, discipline, grades, progress towards graduation requirements, and health records. The software also enables parents to view this information online. The financial software package tracks budget, payroll, inventory, and employee data.13

Texas Assessment of Knowledge and Skills (TAKS) Results Comparison

As implementation of the student data software was completed in January 2007, state assessment data from 2006 and 2009 will be examined. The table below displays the percentage of students whose assessment scores far exceeded state standards. For performance that far exceeds assessment standards, Texas uses the term “commended performance.” Results for the math and reading portions of the test for grades 3-11 are provided.14

Table 6: Rockwall ISD State Assessment Scores - Percent Achieved
Commended Performance (2006 & 2009)

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading (2006)</td>
<td>62%</td>
<td>33%</td>
<td>39%</td>
<td>63%</td>
<td>37%</td>
<td>52%</td>
<td>34%</td>
<td>20%</td>
<td>29%</td>
</tr>
<tr>
<td>Reading (2009)</td>
<td>65%</td>
<td>40%</td>
<td>45%</td>
<td>61%</td>
<td>41%</td>
<td>60%</td>
<td>28%</td>
<td>23%</td>
<td>45%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+3%</td>
<td>+7%</td>
<td>+6%</td>
<td>-2%</td>
<td>+4%</td>
<td>+8%</td>
<td>-6%</td>
<td>+3%</td>
<td>+16%</td>
</tr>
<tr>
<td>Math (2006)</td>
<td>39%</td>
<td>46%</td>
<td>62%</td>
<td>57%</td>
<td>27%</td>
<td>24%</td>
<td>36%</td>
<td>20%</td>
<td>26%</td>
</tr>
<tr>
<td>Math (2009)</td>
<td>53%</td>
<td>50%</td>
<td>62%</td>
<td>56%</td>
<td>27%</td>
<td>38%</td>
<td>36%</td>
<td>26%</td>
<td>46%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+14%</td>
<td>+4%</td>
<td>0%</td>
<td>-1%</td>
<td>0%</td>
<td>+14%</td>
<td>0%</td>
<td>+6%</td>
<td>+20%</td>
</tr>
</tbody>
</table>

Source: Texas Education Agency

Overall, the district has seen increases in student achievement on the TAKS assessment since 2006. The percent of students achieving the “commended” performance level increased the most for 3rd and 8th grade math, and for 11th grade math and reading. Modest decreases in achievement since 2006 also occurred for some tests.

---

13 Skyward, Inc. “Case Studies: Rockwall ISD.” www.skyward.com/Page.ashx/Home
14 Texas Education Agency. “TAKS Aggregate Web Reports.”
   http://ritter.tea.state.tx.us/cgi/sas/htmSQL/student.assessment/reporting/taksagg/index.htm
Northside Independent School District (TX)

Northside ISD implemented its SchoolNet Instructional Management Suite in spring 2005. A large and diverse district, the software package enables teachers and administrators to disaggregate and monitor academic data for different student subgroups in order to modify curricula and instruction and reduce gaps in achievement.

Texas Assessment of Knowledge and Skills (TAKS) Results Comparison

As implementation of the SchoolNet software was completed in 2005, state assessment data from 2004 and 2007 will be examined. Identical to Rockwall ISD in the previous subsection, the table below displays the percentage of students whose assessment scores far exceeded state standards, achieving “commended performance.” Results for the math and reading portions of the test for grades 3-11 are again provided.

Table 7: Northside ISD State Assessment Scores - Percent Achieved Commended Performance (2004 & 2007)

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading (2004)</td>
<td>34%</td>
<td>23%</td>
<td>27%</td>
<td>25%</td>
<td>17%</td>
<td>22%</td>
<td>11%</td>
<td>6%</td>
<td>15%</td>
</tr>
<tr>
<td>Reading (2007)</td>
<td>37%</td>
<td>31%</td>
<td>27%</td>
<td>52%</td>
<td>24%</td>
<td>44%</td>
<td>30%</td>
<td>17%</td>
<td>39%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+3%</td>
<td>+8%</td>
<td>0%</td>
<td>+27%</td>
<td>+7%</td>
<td>+22%</td>
<td>+19%</td>
<td>+11%</td>
<td>+24%</td>
</tr>
<tr>
<td>Math (2004)</td>
<td>26%</td>
<td>21%</td>
<td>29%</td>
<td>21%</td>
<td>8%</td>
<td>12%</td>
<td>19%</td>
<td>10%</td>
<td>18%</td>
</tr>
<tr>
<td>Math (2007)</td>
<td>27%</td>
<td>36%</td>
<td>45%</td>
<td>33%</td>
<td>18%</td>
<td>18%</td>
<td>19%</td>
<td>19%</td>
<td>25%</td>
</tr>
<tr>
<td>Percent Change</td>
<td>+1%</td>
<td>+15%</td>
<td>+16%</td>
<td>+12%</td>
<td>+10%</td>
<td>+6%</td>
<td>0%</td>
<td>+9%</td>
<td>+7%</td>
</tr>
</tbody>
</table>

Source: Texas Education Agency

The data demonstrate that the district has seen significant gains in the percent of students achieving commended performance on state assessments since 2004, one year before SchoolNet was implemented. Half of the grade levels witnessed double-digit increases in this percentage for both subjects between 2004 and 2007, the largest increase occurring for the 6th grade reading assessment.

---

Independent School District of Boise City (ID)

Seeking a web-based integrated student information system, Independent School District of Boise City chose Infinite Campus software to replace its Pearson SASIxp software. The Infinite Campus software enables the district to input data through a single point and gives teachers, students, and parents access to the data in real-time. The software was implemented in summer 2008.18

Idaho Standards Achievement Test (ISAT) Results Comparison

The following table shows the percent of students scoring at the proficient or advanced level for grades 3-8 and 10 on the reading, math, and “language usage” portions of the ISAT assessment. Results from the spring 2008 and 2010 administrations of the test (several months before implementation of the software and two years after) are provided.19

| Table 8: Boise City State Assessment Scores - Percent Proficient or Advanced (2008 & 2010) |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
|                                 | 3     | 4     | 5     | 6     | 7     | 8     | 10    |                                 |
| Grade Level                     | 3     | 4     | 5     | 6     | 7     | 8     | 10    |                                 |
| Reading (2008)                  | 86%   | 86.3% | 86.3% | 87%   | 83.1% | 90.1% | 89.2% |                                 |
| Reading (2010)                  | 89.7% | 87.7% | 88.7% | 88.4% | 88.7% | 92.7% | 86.3% |                                 |
| Percent Change                  | +3.7% | +1.4% | +2.4% | +1.4% | +5.6% | +2.6% | -2.9% |                                 |
| Math (2008)                     | 89%   | 85.2% | 78.4% | 81.5% | 75.8% | 82.5% | 81.6% |                                 |
| Math (2010)                     | 87.7% | 86.2% | 80.2% | 83.3% | 75.4% | 81.4% | 78.2% |                                 |
| Percent Change                  | -1.3% | +1%   | +1.8% | +1.8% | -0.4% | -1.1% | -3.4% |                                 |
| Language Usage (2008)           | 75.9% | 82%   | 77.1% | 77.6% | 74.1% | 72.8% | 74.8% |                                 |
| Language Usage (2010)           | 77.1% | 84.4% | 79.6% | 80.7% | 75.3% | 76.7% | 75.1% |                                 |
| Percent Change                  | +1.2% | +2.4% | +2.5% | +3.1% | +1.2% | +3.9% | +0.3% |                                 |

Source: Idaho State Department of Education

As the table displays, the district witnessed both slight increases and slight decreases in the percentage of students achieving proficient or advanced scores on the Idaho state assessment between 2008 and 2010. Although percentage changes were modest, increases were more common than decreases.

---

Data Warehouse in Public School Districts

Introduction

Data collection in public school systems is not a new phenomenon. For years, districts and schools have maintained records containing student assessment scores, student histories and demographics, and various other student data. Recent initiatives such as the No Child Left Behind (NCLB) Act and the Elementary and Secondary Education Act have only increased the importance of data-gathering for states and districts, especially in regards to student performance data. However, although districts have been accumulating vast amounts of data for accountability purposes and for state and federal requirements, there often is a gap between the amount of data that is collected and the degree to which it is effectively used by teachers and administrators to improve their schools and student learning. The previously cited Annenberg Institute study states that “while [districts] collect, store, and report data extensively, very rarely is any of it used to analyze student needs or quality of instruction, help form strategic plans or policy, or determine whether resources are distributed equitably.”

Similarly, it has been observed that initiatives intended to close achievement gaps, such as NCLB, assume that the availability of data will inform and initiate improvements in educational practice, but do not specify ways to use data at the classroom level. In order to close achievement gaps, educators need to be able to use actionable data; in other words, they need to engage in “data-driven decision making.” There are several primary reasons why analyzing, interpreting, and using data to improve student achievement have previously been difficult in the K-12 realm. The Annenberg Institute study summarizes three obstacles to effective data use in schools:

- **“Scattered and Antiquated Technology Infrastructure”** – Schools often store relevant data in separate locations using different software that do not communicate; for example, student demographic data and test score data may be stored in different databases, while the human resources department stores important data on teacher certification and participation in professional development. Additionally, some data may not be available in an electronic format. Storing important data on students, teachers, the school, and the district in different locations and on different systems makes data-driven decision making much more labor-intensive, and in some cases, impossible.

- **“Data Accessible Only Through ‘Gatekeepers’”** – For many school districts, data are only available to data analysts or other specialists, partly

---


21 Ibid., p.12-13.
because technical software and analytical expertise is required to examine the data.

- **“Inaccurate, Late, or Cumbersome Data”** – Districts often report student performance data to schools and teachers, but the data may arrive after students have progressed to the next grade or have graduated. The data can often be in unmanageable or “unwieldy” formats. Additionally, important student data are not dynamic because they are often collected infrequently, decreasing their relevance to educators trying to improve student learning based on data.

Regarding technology infrastructure, experts have noted the “lack of [data] systems providing user-friendly access to timely, relevant information” as a primary obstacle to effective data-driven decision making at the district, school, and classroom levels.\(^{22}\) A leading researcher in the use of data in education, Professor Jeffrey Wayman, states that “although schools have been ‘data rich’ for years, they were also ‘information poor’ because the vast amounts of available data they had were often stored in ways that were inaccessible to most practitioners. Recently emerging technology is changing these circumstances.”\(^{23}\)

**Data Warehouses in School Districts**

Information management systems are frequently promoted as one of the keys to organizing and presenting data in a more accessible and user-friendly format. A study by the Center on Educational Governance at the University of Southern California looked at how high-performing school systems use data to improve student learning. One of the main strategies of “performance-driven” school systems highlighted in the study was an investment in information management systems; a **technology infrastructure that is user-friendly and can present data in a comprehensible way.**\(^{24}\)

Data warehouse software is a primary component of information management systems, and is increasingly being implemented at the state and district levels. Data warehouses potentially enable districts to combine and integrate data from different systems; for example, student achievement data, grades, discipline and attendance records, and teacher information. This allows various data relationships to be examined, including longitudinal analysis of data and linkages between learning

---


outcomes and educational processes and variables. Such analyses were previously very difficult or impossible to carry out due to technological deficiencies.

The rapid adoption of data warehouses and other data management software in the education sector is reflected in recent surveys of school districts. A report by the U.S. Department of Education (DOE), which surveyed more than 400 school districts during the 2007-2008 school year, found that only 10 districts did not have any kind of electronic student data system in place or other tools to inform educational decision making. These 10 districts tended to be smaller than the rest of the survey sample. The remaining districts reported to have an electronic student data system of some sort, and 77 percent of districts reported to have a data warehouse in 2007-2008.25

Despite the increasingly widespread existence of data warehouses and other data systems in school districts, the U.S. DOE study found that districts are still in the process of taking full advantage of new technologies. Districts have electronic access to vast amounts of student data, but are less likely to be able to combine data from different types of systems or link educational variables to achievement data. The survey identified five types of data linkages that may be part of a data system: (1) school performance linked to finance data, (2) student performance linked to teacher information and characteristics, (3) student performance linked to AYP subgroups, (4) student performance linked to specific teachers, and (5) student performance linked to specific instructional programs.26 It was reported that only 15 percent of districts have systems that can perform all five data linkages, and a majority of districts cited that the main barrier to data-driven decision making was having information located in multiple databases.27

The ability to link learning outcomes to various educational processes and variables is still limited among most district data systems, making it more difficult to make practical instructional improvements based on data. However, most data systems allow educators to query and manipulate datasets for closer examination of information. Approximately three-quarters of district data systems have “drill-down” capabilities and 86 percent of systems can provide longitudinal student information such as schools attended and GPA. Drill-down capability allows educators to examine a subset of data at the grade, classroom, or student level.28

The U.S. DOE report also found that the majority of districts with data warehouses (56 percent) use commercially-developed software, and approximately a quarter of districts use a data warehouse supplied by their state.29

---

26 Ibid., p.33.
27 Ibid., p.37.
28 Ibid., p.43.
29 Ibid., p.40.
The report provides a list of the data warehouse software brands and vendors most frequently identified by districts that participated in the survey, with the number of citations in parentheses:\(^{30}\)

- Pearson SASI (20)
- Sungard Pentamation (14)
- TetraData (11)
- Pearson PowerSchool (10)
- STI (10)
- Aeries (10)
- Skyward (9)
- Educational Data Warehouse (8)
- CIMS (7)
- Cognos (7)
- JPAMS (6)
- DataDirector (5)
- Otis Ed (4)
- ProSoft (4)

It should be noted that some of the software listed above may also be categorized as student information systems (SIS), and may not be strictly defined as data warehouses. Also, there are some software brands that may have been phased out and are no longer commercially available.

**Data Warehouse Guidelines and Recommendations**

Previous literature on data warehouses has provided a number of guidelines and recommendations regarding the implementation of a data warehouse, or more generally, a data management system. The following are summaries of best practices for warehouse implementation that have been presented by experts in the field.

In a study of high-performing school districts that use data to improve instruction, the University of Southern California’s Center on Educational Governance looked at information management systems used at four districts. As the researchers state, “merely having data does not ensure that data-driven decision making will take place.” Although all four districts use different types of data systems, the report identifies two common characteristics of the districts’ data management approach that help ensure that data are used effectively by educators:\(^{31}\)

- **Effective systems are user-friendly**, and provide easily accessible and comprehensible data in a timely manner. Systems are scalable; that is, software

---

\(^{30}\) Ibid., p.135.

capabilities can be expanded as data needs change. According to the districts, the most useful data reports can quickly identify students that need extra assistance and in what academic areas.

- Most school districts have dedicated personnel to assist with data management and use at the school and district levels. Data management staff in schools are usually experienced educators as well, such as principals and head teachers, not technologists or statisticians.

As for the implementation of a data warehouse and what it requires, the Annenberg Institute study examined eight districts that had established a data warehouse, three of which had earned national recognition for effectively using technology. Based on the experiences of these districts, the study provided several key suggestions for districts looking to building a data warehouse:

- Prior to implementing a warehouse, a district must prepare by assessing the kinds of data it currently collects and how it is stored, as well as any additional data it currently does not collect but plans to include in the warehouse. The districts in the Annenberg study spent much time cleaning data—making sure data are accurate and not duplicated, as well as standardizing data labels across existing systems.

- Districts must think carefully about what they want to know; that is, discern the questions that are most important and how data can answer those questions. The Annenberg researchers recommend drawing from strategic plans and other key goals of the district to help determine what kind of data to store and use, with input from a range of stakeholders. Districts may want to expand beyond standardized test scores to include data on educational resource allocation, curricula and lesson plans, demographical information, and survey data, for example.

- It is important to open up channels of communication among various stakeholders so that all end-users of the warehouse can offer input and feedback during the planning stages and after implementation. Teachers, principals, and administrators will have different data needs and perceptions of the system. Districts in the Annenberg study used e-mail, listservs, message boards, meetings and focus groups to gather input from stakeholders.

- Potential vendors and software products should be investigated by sampling and viewing demonstrations of products, and perhaps most importantly, by speaking with other districts that are currently using the software or service to learn about their implementation process and

---

33 Ibid., p.34-36.
challenges. **Any commercially developed software should be compatible with the School Interoperability Framework (SIF),** an initiative of educational software companies to ensure that new software is compatible and can communicate with other systems.

- **The data warehouse should be rolled out as quickly as possible.** Most districts in the Annenberg study built and began rolling out the warehouse within a year. The districts recommended phasing in access for different user groups, beginning with the top-level decision makers and continuing with principals, teachers and, if desired, parents and students.

- **Fostering a data culture within a school district is just as important as the data warehouse software.** The Annenberg researchers state that the collaborative planning and communication that is necessary to implement such a technology may conflict with traditional district culture and hierarchies. Taking full advantage of the technology requires going beyond its technical capabilities and involving not only the IT department, but also the staff responsible for training and helping educators interpret and use the data in a practical way to improve instructional practice. Sufficient time and money needs to be invested to support use of the data warehouse throughout the district.

Finally, a 2004 paper by Dr. Jeffrey Wayman, Sam Stringfield, and Mary Yakimowsky of Johns Hopkins University is an excellent resource for districts planning to implement a data warehouse.\(^{34}\) The paper includes a list of important features of data analysis software that could potentially be useful when selecting a product. The following is an abridged version of the list:\(^{35}\)

**User Friendliness**
- Software is intuitive and easy to use
- Software requires little training
- Presentation is familiar to user
- Access speed is fast and efficient

**User Features**
- Comprehensive query tools available for every level of user
- Flexible drill-down capability from any form of data aggregation
- Data can be accessed from anywhere

**Information Access**
- Multiple ways to access information
- Varied methods of representing information (e.g., tables, graphs)
- Wide range of data available for analysis

---

\(^{34}\) Wayman, J., et. al. 2004. “Software Enabling School Improvement Through Analysis of Student Data.”

\(^{35}\) Ibid., p.20.
Interface provides immediate access to relevant information
- Pre-formatted reports are clear, varied, relevant, and comprehensive
- Longitudinal presentation of data available at every user level

Creating and Sustaining Quality Data
- Provides capacity to enable clean data
- Company accepts responsibility to facilitate data process with schools
- System allows for expansion past initial implementation
- System provides proper security for data transmission
- Integration of different areas of information is seamless to the user
- Software accepts many common data formats

Additional Features
- Online student work samples available
- Software exports into common programs
- Users can access electronic discussion groups
- Easy access to learning standards information
- Software offers capacity to link individual teacher data to student data
**Project Evaluation Form**

Hanover Research is committed to providing a work product that meets or exceeds member expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.


**Note**

This brief was written to fulfill the specific request of an individual member of Hanover Research. As such, it may not satisfy the needs of all members. We encourage any and all members who have additional questions about this topic – or any other – to contact us.

**Caveat**

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties which extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every member. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Members requiring such services are advised to consult an appropriate professional.