Indicators Supporting School Quality:
Lessons Learned from the United States

Joan L. Herman | CRESST/UCLA

Introduction

This paper shares lessons learned about the design and use of indicators to monitor and support the improvement of school quality. Drawing largely on the experience of the United States, the paper starts with assumptions about the nature and purpose of such indicators and criteria for judging their validity. After briefly reviewing indicators used in international comparisons, the paper concentrates on common and emerging indicators used in the United States to support school accountability and improvement, indicators that focus on student learning outcomes. The paper concludes with lessons learned about the use of indicators in school improvement.

Purpose of School Quality Indicators

Educational indicators are statistics that enable educators, education leaders and other stakeholders, policymakers, and/or the public to monitor the condition of schools by providing aggregate measures of important system components. They enable users to understand and better judge the current status of schools, to analyze trends, and to forecast future changes (Shavelson, McDonnell & Oakes, 1991). Not very useful in isolation, individual indicators are usually considered as part of a system measuring distinct components of the education system that are central to understanding how well it is functioning and potentially provide evidence on how well critical components are working together to achieve educational success. Although educational indicators may address how well schooling is functioning at any level of the educational system—national, regional, district or school—this paper focuses on school-level indicator systems, that is, indicators that can be used to monitor the quality of individual schools. Moreover, it concentrates on indicators of student performance.

National Center for Research on Evaluation, Standards, & Student Testing
UCLA | Graduate School of Education & Information Studies

School level-indicators, of course, provide a general barometer that serves decision-making purposes at multiple levels of the system. They support accountability for individual schools and provide data to inform improvement. Likewise, at the district and state levels, educational leaders and administrators can use indicator data to identify schools that are struggling as well as those that may embody promising practices, establish improvement and other policy priorities, and allocate resources, to name just a few purposes. In the United States, school report cards provide wide access to school indicators, and parents can use them to select schools for their children.

Validity Criteria

At the same time, not all indicators are created equal, and not all provide meaningful and useful data for decision-making. *Standards for Educational and Psychological Measurement* (AERA, APA, & NCME, 1999) and CRESST criteria for assessment and accountability systems (see Linn, Baker, & Dunbar, 1991; Baker, Linn, Herman, & Koretz, 2002) establish validity benchmarks that can be applied to school indicator systems as well. As a set, school indicator systems should characterize central features and processes that are thought to make a difference in student learning and other important outcomes, be composed of multiple indicators of short- and longer-term institutional outcomes, and include major contextual variables essential to interpreting school and student success. Contextual factors that are beyond schools' control yet influence student outcomes are important considerations in making fair comparisons. For example, comparisons between schools need to take into account differences in school populations that influence students' entering abilities and that may likewise influence student outcomes. Concerns for fairness and validity of inferences from indicator systems also has led to a dual focus on both the current status of student learning and progress in learning from one year to the next.

> **Starting in the 1990's, schools increasingly have made available school report cards to provide the public access to key indicators of performance.**

Presumably, actors at each level use indicators not only to monitor the status and progress of schools within their purview, but also, when necessary, to take action—to change policies and/or practices that will improve quality. Moreover, in addition to providing data for evidence-based decision-making, research also clearly shows that indicators serve signaling and motivational purposes. Public reporting of quality indicators, at least in the United States, are highly visible, garner significant attention and tend to energize educators and schools to improve their results—even more so when incentives are tied to performance, as is the case in US accountability systems. "What gets measured gets managed" and "what you test is what you get" are popular axioms in the education and business literatures in the United States. Moreover, startling results—for example, low performance on the Program for International Student Assessment (PISA)—can spark wide public attention and provide an important platform for public dialogue and for leveraging changes in policy and practice.

School indicator systems that support improvement, moreover, will focus primarily on indicators that are actionable (i.e., that reflect constructs that schools can influence and for which it is reasonable to hold schools accountable). To the extent possible, indicators also should have diagnostic value in not only identifying relative system

strengths and weaknesses, but also in helping to clarify the source of these strengths and/or challenges. School indicator systems intended to support improvement must also be comprehensible and credible to intended users and reflect measures that are technically adequate; that is, they must be reliable, accurate, comparable, and aligned with intended constructs, as well as valid and fair for intended purposes. Feasibility is a final issue: Indicators should be worth the time and energy required to develop, analyze, and use.

In summary, individually and as a system, sound indicators will incorporate the following criteria:

- Aligned with and provide accurate data on essential education system elements, processes, and outcomes
- Fair
- Valid
- Reliable
- Comparable
- Credible and meaningful
- Comprehensible and understandable to intended users
- Actionable
- Feasible

International Indicator System

International indicator systems incorporate many of these criteria, although their purpose is to monitor progress and support policy and action at the national, rather than the school level, which of course has strong implications for system design, intended use and users, and for what constitutes actionable data.

The European Commission (2000), for example, focused on 16 indicators organized into the categories of: learning attainment, organized by subject area; success and transition, including drop out and completion rates plus participation in post secondary education; monitoring, including indicators of school evaluation and parent participation; and finally resources and structure, including indicators of expenditures, teacher training, preschool participation, and the number of students per computer.

Approximately a decade later, the appetite for and wider availability of indicators and data are evident in the Organisation for Economic Co-operation and Development's (OECD) most recent Education at a Glance (2012). While still organized into four general indicators, OECD shows attention to a much broader set of indicators and burgeoning attention to equity, including:

1. Educational outputs, including indicators of educational attainment, graduation and learning achievement, gender and equity outcomes and labor market and social outcomes of education.
2. Financial and human resource investment, including national and per-student spending on education, higher education costs and support and how financial resources are allocated.
3. Access to education, participation, and progress, which include indicators of access to early childhood education, primary, secondary and tertiary education, transitions from schools to work, and adult learning.
4. Learning environment and organization of schools, which provides indicators on teacher characteristics, salaries, teaching time, class size, school decision-making, and examinations.

School Report Cards in the United States

School report cards in the United States show these same trends: An explosion of interest in data on schools, increasing availability of data and sophisticated data management tools, and increasing interest in accountability, particularly accountability for equity outcomes. Starting in the 1990's, schools increasingly have made available school report cards to provide the public access to key indicators of performance. Starting in 2001, federal (No Child Left Behind [NCLB], 2001) legislation

required states, districts, and schools to provide annual school reports on their status and progress in achieving NCLB goals, including those related to academic proficiency for *all* students (including subgroup performance), student attendance, drop out data, and teacher qualifications. The contrast in reporting and media options between the late 1990's and today is striking.

As Figure 1 shows, report card models from the earlier period attempted to provide in a single page a picture of school performance. Because the reports were intended for parents and the public, they were designed to be user-friendly and simple to interpret: Figure 1, for example, uses colors—green, yellow, and red—to support appropriate interpretation of the indicators. At a glance, users can get an overall picture of achievement results by grade and subject, student attendance, graduation, completion of college admissions coursework, student discipline, and teacher credentialing, among others.

Figure 1 was set in a time when paper was the dominant model of delivery and simply providing access to such reports via the Web was considered innovative, as was appealing graphical design and color printing.

In contrast today, report cards for *every* public school are routinely available digitally through each state's website. Providing access to a wealth of data, the reports are interactive, provide multiple representations of each indicator, and multiple points of comparison to aid in indicator interpretation. For example, indicators of current levels of student achievement for a school may be compared to the district and state averages, as well as to prior year performance. One simply goes to any state's website for access to report cards for all districts and schools in the state. Common categories for parent reporting are:

- *School profile* data, including data on *student* enrollment and demographic characteristics, average class size, teacher qualifications, teacher turnover rates, staff counts.
- *School climate* data such as attendance, suspensions, and expulsions.

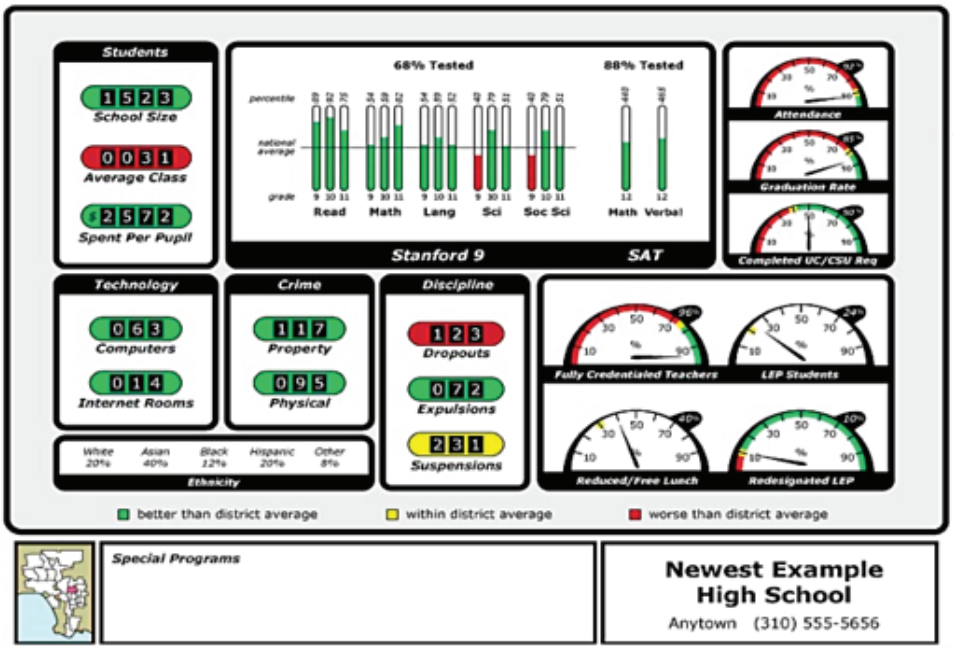

Figure 1. CRESST School Report Cards: High School Example.

National Center for Research on Evaluation, Standards, & Student Testing

- *Student outcome* data, including performance in academic subjects, school completion, and drop out rates—for all students combined and by subgroup.
- *School accountability* results, which are derived from student outcome data to indicate schools' overall performance and the extent to which they have met their annual state goals.

Some states also provide data on school finances, including average instructional expenditures per student, average total expenditures, average teacher salary, the status of school facilities, curriculum and instructional resources—whether there are sufficient texts for each student, currency, etc., and students' post-secondary success (e.g., college attendance, completion).

Schools and districts often customize and augment base reports to emphasize additional priority goals. For example, the Los Angeles Unified School District (2012), the second largest school district in the country, organizes its indicators for high schools (grades 9-12) into four categories that underscore district priorities and district definitions of each. Performance on each indicator is compared to prior year performance and to that of the district as a whole.

- *Student progress:* Percentage of students who meet state proficiency standards in English Language Arts and in Mathematics; percentage of students who have been promoted from one grade to the next; and school growth in performance—the extent to which student scores from the prior year grew more than expected, or less than expected, based on value-added modeling.

- *Student readiness for college and career:* Percentage of graduates who passed courses required for college admissions (termed A-G requirements) with a C or better; percentage of students who, given their grade level, are "on track" to pass required college courses with a C or better; percentage of students who take and achieve minimally acceptable scores on college entrance tests (SAT or ACT); and the percentage of students who take and achieve at least a "C" in Advanced Placement courses.

- *Learning environment*: student, parent, and teacher perceptions of the learning environment and student expectations for college, based on survey responses; student attendance and suspension rates; teacher attendance and turnover rates.

- *Subgroup performance*: for all students and each numerically significant subgroup at the school, where subgroups are defined by race/ethnicity, socioeconomic status (SES), language background, and disability.

Summary accountability indicators can be very complicated.

Accountability Indicators

School report cards represent a part of the data that are regularly available to schools and educators. It's a wealth of data, and the public can get lost in its details. School accountability targets represent one, crucial indicator by providing a summary judgment of how a school is doing relative to state and/or federal accountability goals. Summary accountability indicators can be very complicated, because they try to incorporate goals for both excellence and equity and include variables that discourage obvious gaming of results.

Accountability targets established by the No Child Left Behind legislation are the most familiar. Schools must demonstrate that they are making progress toward all students scoring proficient on annual state achievement

tests—for students overall and for every numerically significant subgroup, including those defined by race/ethnicity, poverty (SES), language status (limited English proficient), and students with disabilities. To eliminate schools trying to increase their scores by encouraging some students to stay home on test day, test results must be based on at least 95% of eligible students. To guard against schools trying to improve their scores by pushing some low scoring students out, drop out rates also are part of the equation.

Some states have initiated more sophisticated indices and have provided data on how schools are performing relative to the state overall and relative to demographically similar schools. For example, California's Academic Performance Index (API) draws on annual individual student scores in a range of subject areas and drop out rates to measure the academic performance and progress of every public school in the state. The API scale ranges from a low of 200 to a high of 1000, with a score of 800 established as the minimum performance goal for every school. Schools also have annual progress goals for moving toward and surpassing the 800-mark target, which represents an average student score between basic and proficient achievement.

The API index is weighted to provide differential value for different subject areas and for moving students to higher levels of achievement. In general (leaving out some of the technical details and safeguards), the process is as follows: First, student scale scores for each subject are used to assign students to different levels of proficiency—from *advanced, proficient, basic, below basic*, and *far below basic*. Students are assigned score values as follows:

- Advanced: 1000 points
- Proficient: 875 points
- Basic: 700 points
- Below basic: 500 points
- Far below basic: 200 points

Schools make the most progress from year-to-year by moving students from the lowest categories to higher ones, providing an incentive to attend to the lowest performing students.

Table 1

Content weights for API 2011-2012

Content Area	Elementary K-5	Middle School 6-8
English-Language Arts	56.5	51.4
Mathematics	37.6	34.3
Science, grades 5 & 8	5.9	7.1
History/Social Science, grade 8	NA	7.1

(See CDE, http://www.cde.ca.gov/ta/ac/ap/documents/infoguide12.pdf)

The next step is to apply weights to the average scores for each subject. For example, Table 1 shows the weights applied per subject in typical elementary and middle schools. The weights show that English Language Arts is weighted nearly 50% more than mathematics, and that science and history are given relatively little weight, even given that they are tested far less often than ELA and mathematics. These weights provide a strong signal for relative school attention. Many schools listen.

Total scores are then computed and used to categorize all schools in the state, by school level, on a 1-10 scale, where 10 reflects the highest performing schools and 1 the lowest performing schools. Further, because background

characteristics are highly related to school performance, every school is assigned a similar school rating. Every school is compared to the 100 most demographically similar schools in the state, based on poverty, language status, ethnic and racial composition, percent of students with disabilities, etc. Schools are classified into 10 categories relative to the performance of schools that are demographically similar, with 10 being the highest rating relative to similar schools and 1 the lowest. Thus, schools and their public audiences can compare themselves to the state as a whole and to schools serving similar school populations. This is an attempt to hold all schools to the same high standards, yet also provide realistic comparisons by which each school may compare its accomplishments. Therefore, it is possible that a school would score in the 1-5 range relative to all schools in the state and still receive a high rating (e.g., 7-10) relative to similar schools. In addition to these relative ranks, schools also are expected to make specified progress toward an overall score of 800 or beyond.

There are rules to try to assure that school scores are legitimately attained and not as a result of gaming, for example, by assuring that scores are based on the vast majority of students (even with parent opt out provisions) and by incorporating drop out rates. The system emphasizes equity goals by establishing specific, annual API growth targets for each numerically significant subgroup at a school, including ethnic/racial minorities and socio-economically disadvantaged students. The intent is to drive attention to closing historical achievement gaps for traditionally underperforming students.

Indices for Identifying Distinguished Schools

States also are continuing to create additional, new indices for identifying and officially recognizing excellent schools—and school that are the most at risk. These indices combine school performance over a variety of indicators. For example, the state of Washington has created an accountability index based on performance on each of four indicators across five outcome areas. The outcome areas are the subject areas assessed by the state assessment program—reading, writing, math, and science, in addition to drop out or graduation rate. The indicators for each subject area address:

- Achievement of non-low income students,
- Achievement of low income students,
- Achievement vs. peers, and
- Improvement from previous years.

For each outcome area and indicator, a 1-7 rating is computed, representing the percentage of students who meet established standards. A rating of 7, for example, means that 90.1-100% of students meet established standards, a rating of 6 means that 80.1-90% of students meet established standards, etc. As shown in Table 2, the overall index is a simple average of the ratings over all

Table 2

Washington State Accountability Index: Hypothetical Example

Indicators	Outcomes					
	Reading	Writing	Math	Science	Grad Rate	Average
Achievement of non-low income students	7	6	6	2	5	5.20
Achievement of low income students	5	4	4	2	7	4.40
Achievement vs. Peer	6	6	7	4	6	5.2
Improvement from previous year	4	4	5	6	7	5.2
INDEX SCORES	**5.5**	**4.25**	**5.5**	**3.5**	**6.25**	**5.00**

(see http://www.sbe.wa.gov/documents/2010.03.15%20recognition%20flyer.pdf)

categories. While the overall index scores give equal weight to each outcome area and indicator, the figures in the table provide easy comparison for identification of relative strengthens and weaknesses. For example, the index scores by outcome area in Table 2 reveal science as a relatively weak area; and the difference between average ratings for the achievement of non-low income and that of low-income students reveals a school achievement gap, 5.20 vs. 4.40. The state defines seven tiers of school quality based on specified index ranges. Schools scoring between 5.5 and 7, for example, are designated exemplary, while those in the range of 1.00-2.49 are designated struggling, and those below 1.00 are designated as high priority for state improvement.

Statewide, individual student-level longitudinal data bases are becoming more and more common, opening up a wealth of additional indicators based on growth modeling and value added analyses. The state of Maine, for example, has just proposed a system that establishes individual performance and growth targets for every school, based on their prior student performance.

Summary and Conclusions

The United States has been actively using school level indicators and related public reports for more than a decade to support school improvement. Indicators have become progressively more complicated and sophisticated in an attempt to reflect multiple goals for improving equity and excellence in student learning and to try to create indicators less susceptible to gaming. This experience reinforces a number of conclusions about the use of indicators:

- Any indicator is only as good as the measure(s) underlying it. Indicators in the United States have focused on measures of student learning, based predominantly on standardized, multiple-choice tests. Research indicates that these measures focus largely on basic and lower level, cognitive skills at the expense of complex thinking and problem solving. Critics also charge that these tests are not sensitive to instruction (see for example, Herman, 2007; Popham, 2009).
- The indicators themselves, rather than what they are supposed to represent, may become the focus of attention. Research shows that when schools are held accountable for student performance and/or when school performance indicators receive high visibility, educators tend to focus their attention on the test that underlies the indicator, rather than the significant learning goals that the test performance is intended to represent. In the United States, a focus on improving performance indicators has too often led to a curriculum of test preparation rather than one focused on meaningful learning goals (for example, see Hout & Elliot, 2011, for a review of evidence).
- The experience in the United States underscores the truth of Campbell's Law (1976): "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor." When test scores become the goal of schools, they both lose their value as indicators of learning and distort the teaching and learning process.

Even so, indicator systems can provide important evidence on the status and progress of schools that can help to motivate and provide data for improvement. However, at the same time, the data alone cannot foster improvement. Moving from indicators to effective improvement strategies requires a host of socio-political, technical, and pedagogical capabilities, for example, school leadership and culture, effective infrastructure and processes supporting indicator use, technical and pedagogical expertise, and commitment to continuous improvement (see, for example, Winters & Herman, 2012; CRESST data use website, 2012). Along with attention to quality, indicator systems need policies and practices that support the effective use of these data for improvement.

References

American Educational Research Association, American Psychological Association, and National Council for Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: APA.

Baker, E., Linn, R., Herman, J., & Koretz, D. (2002). Standards for educational accountability. *Policy Brief* #5. Los Angeles, CA: CRESST.

California Department of Education. (2012). 2011-12 *Academic performance index reports: Information guide.* Available on line at http://www.cde.ca.gov/ta/ac/ap/documents/infoguide12.pdf

Campbell, D. T. (December 1976). *Assessing the impact of planned social change.* The Public Affairs Center, Dartmouth College, Hanover, NH, USA.

CRESST. (2012). *Data use website.* See http://datause.cse.ucla.eduEuropean Commission Directorate General for Culture and Education. (2000). *European report on the quality of school education.* Retrieved September 17, 2012 from http://ec.europa.eu/education/policies/educ/indic/rapinen.pdf

Herman, J. L. (2007). *Accountability and assessment: Is public interest in K-12 education being served?* CRESST Technical Report #728. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). See www.cresst.org

Hout, M., & Elliot, S., (Eds.). (2011). *Incentives and test-based accountability in education.* Washington, DC: National Academies Press.

Linn, R., Baker, E., & Dunbar, S. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher,* 20(8), 16-21.

Los Angeles Unified School District (2012). *School report card*. Retrieved at http://www.getreportcard.lausd.net/reportcards

No Child Left Behind Act of 2001 (Vol. Pub. L. No. 107-110, 115 Stat. 1425).

OECD (2012). *Education at a Glance 2012: OECD Indicators*. Retrieved September 17, 2012 from http://www.oecd.org/edu/eag2012.htm

Popham, W. J. (2009). *Instruction that measures up*. Alexandria, VA: ASCD.

Shavelson, R. J., McDonnell, L., & Oakes, J. (1991). What are educational indicators and indicator systems? *Practical Assessment, Research & Evaluation,* 2(11). Retrieved September 9, 2012 from http://PAREonline.net/getvn.asp?v=2&n=11

Winters, L., & Herman, J. (2012). *Turn-around toolkit: Managing rapid, sustainable school improvement.* Thousand Oaks, CA: Corwin Press.