Rethinking School Evaluation and Accountability
to Get Credit for All that is Being Done

Along with expanding the school improvement policy framework, there is a need to rethink and reframe approaches to evaluation and accountability. ESSA does too little in this respect. For example, perhaps it's evident, but in case it's not: *Just adding one or two nonacademic indicators to accountability requirements will do little to move transformation of student/learning supports forward.*

School accountability is a policy tool with extraordinary power to reshape schools – for good and for bad. Systems are driven by accountability measures. This is particularly so under "reform" conditions.

Policy makers demand that schools show that their efforts are effective. But effective in what way? To what degree? At what cost? It is commonplace for there to be dissatisfaction over whatever is mandated. A major concern is that, too often, the data have little relationship to improving schools.

It is evident that accountability and all other evaluation activity can be difficult, troublesome, and controversial. In choosing what to look at, how to gather and interpret the data, and what to report, schools reflect prevailing policy makers' values and priorities. As a result, major disagreements are ongoing about which indicators to collect and how to use the amassed data.

Most of us agree that schools should constantly strive to improve and must be accountable. Every educator is aware of the importance of having data on results. We know that evaluation can be the door to a better future. But it is complicated. And it is more than ironic when prevailing policies and practices close rather than open that door.

Needed at this juncture is policy that reframes evaluation and accountability to ensure that schools get credit for all they are doing and are guided to what they can improve. With specific respect to accountability for addressing barriers to learning and teaching, this requires an expanded accountability framework that encompasses a whole person and whole system focus. This includes:

    (1) cognitive development and engagement (e.g., academics),
    (2) social and personal development and behavioral engagement,
    (3) direct intervention efforts essential for dealing with barriers to learning and teaching and re-engaging disconnected students,
    (4) evaluating system development, performance, and impact within the political-social-economic context (e.g., the neighborhood) in which a school is enmeshed.

About Evaluation

Evaluation involves determining the worth or value of something. It is one of several basic assessment functions used at schools. In that context, evaluation is defined as a systematic process designed to describe and judge the overall impact and value of an intervention for purposes of making decisions and advancing understanding. Properly developed, formative and summative evaluations can aid efforts to (1) assess efficiency, effectiveness, costs, and impact, (2) make good decisions for improving schools, and (3) advance knowledge in ways that improve policy, practice, and personnel preparation.

Given that many more indicators can be proposed than can be feasibly used to gather data, decisions must be made about what will be evaluated. In addition to matters highlighted above, these include decisions about (1) general concerns of interest (e.g., concerns about students, teachers, support staff, administrators; classroom and schoolwide conditions and climate; intervention antecedents/inputs, immediate objectives, intermediate goals, long-range aims), (2) specific facets to be evaluated, (3) level of specificity used in designating indicators, (4) measures and methods for gathering data on designated indicators, and (5) standards to be used in analyzing the data and arriving at judgments. In making such decisions, concerns arise because what can be evaluated currently is far less than what schools assert as their mission. Furthermore, all such decisions are influenced by various sources of bias.

In discussing evaluation, Yankelovich sagely cautioned:

> *The first step is to measure whatever can be easily measured. This is okay as far as it goes.*
>
> *The second step is to disregard that which can't be measured .... This is artificial and misleading.*
>
> *The third step is to presume that what can't be measured easily isn't very important. This is blindness.*
>
> *The fourth step is to say what can't be measured really doesn't exist. This is suicide.*

Appreciating the Many Functions of Assessment at a School

Schools need to pursue *assessment* related to various functions. Decisions about indicators begin with clarity about such functions. And, most of the functions call for more than indicators of results.

Besides generating data for system management, assessment plays a key role in:

1. *Identification* – indicators that help find and label matters of specific interest. The focus may be on person variables, environmental factors, or both, and on problems, strengths, or both (e.g., data to help identify effective teachers and effective schools; data to help identify gifted and talented students and those who are not doing well at school – including those needing special education).

2. *Selection* – indicators that help make decisions about general changes in status (e.g., data to inform decisions about moving teachers and principals to different schools, choosing schools for special intervention, placing students in specific programs).

3. *Planning for specific changes* – indicators that help make decisions about immediate and short-term objectives and procedures for accomplishing long-term goals (e.g., data to inform school improvement planning, professional development, specific student interventions – including data from response to intervention efforts and IEP assessments).

4. *Evaluation of School Results* -- indicators that help make decisions about effectiveness based on positive and negative outcomes and related costs (e.g., focus may be on impact on students, particular subgroups, society as a whole). These data play a major role in system improvement and policy decisions (e.g., accountability).

Expanding the School Accountability Framework

Under the No Child Left Behind Act, the only measures that really counted were achievement test scores. Such scores drove school accountability, and what such tests measured became the be-all and end-all of what is attended to by many decision makers. This produced a growing disconnect between the realities of what it takes to improve academic performance and the direction in which many policy makers and school reformers were leading the public.

Prevailing accountability pressures marginalize almost every effort not seen as directly and quickly producing higher achievement scores

The disconnect was especially evident in schools serving what often are referred to as "low wealth" families. Such families and those who work in schools serving them have an in-depth understanding of a range of barriers to learning and teaching that must be addressed so students can benefit from the teacher's efforts to teach. These stakeholders long-stressed that, in many schools, major academic improvements are unlikely until the impact of such barriers are reduced effectively.

At the same time, it is evident to anyone who looks that there is no *direct* accountability for how schools address factors interfering with student success at school. To the contrary, efforts essential for addressing barriers to development and learning often are devalued, further marginalized, and cut when achievement test scores do not reflect an immediate impact.

Thus, rather than building a system that effectively addresses barriers to learning and teaching, prevailing accountability measures mainly pressure schools to pursue "better" instruction. The underlying assumption is that students are both motivated and able each day to benefit from instruction. The reality, of course, is that in too many schools this is not the case for a large group of youngsters.

Students confronted with a host of interfering factors usually are not in a position to benefit even from significant instructional improvements. The result is enduring low test scores and an achievement gap.

As was the case with the No Child Left Behind Act, the Every Child Succeeds Act (ESSA) provides a fragmented vision for a range of student/learning supports intended to promote equity of opportunity. As a result, there remains a fundamental disconnect between ESSA and the realities of what it takes to improve academic performance.

Ensuring a Focus on Barriers to Learning and Teaching

Logically, well designed, systematic efforts are essential for addressing interfering factors. However, current accountability pressures override the logic and marginalize almost every effort not seen as directly and quickly leading to higher achievement scores. Adding a couple of "nonacademic" accountability indicators clearly is not a solution. Doing so will likely contribute to the trend to drive student/learning supports in ways that deemphasize essential work that is not a specified accountability indicator. For example, efforts to improve attendance often stress rounding up and bringing truants back to school, but do little to help teachers re-engage these students in classroom instruction. This is a recipe for a revolving door.

Those calling for "multimetric" accountabilty capture the essence of the problem of emphasizing only one or a few nonacademic indicators. However, they have yet to face up to developing an accountability framework that effectively encompasses a systemic focus on addressing barriers to learning and teaching and re-engaging disconnected students. If the goal of a school accountability system is to improve schools so that they increasingly enhance equity of opportunity, these factors must be included in a comprehensive and systemic manner.

All this leads to an appreciation of the need for an expanded framework for school accountability – a framework that includes direct measures of achievement *and much more*. We view this as a move toward what has been called *intelligent accountability.* The Exhibit on the following page highlights such an expanded framework.

As illustrated, there is no intent to deflect from the laser-like focus on meeting high academic standards. Debate will continue about how best to measure academic outcomes, but clearly schools must demonstrate they effectively meet goals for *cognitive development* (especially higher order learning).

It is self-defeating not to hold schools accountable for enhancing personal and social growth

At the same time, policy must acknowledge that schools also are expected to pursue high standards in promoting positive *personal and social functioning*, including enhancing civility, teaching safe and healthy behavior, and some form of "character education." Every school we visit has specific goals related to this facet of student development and learning. Yet, it is evident that there is no systematic evaluation or reporting of the work. As would be expected, then, schools direct few resources and too little attention to these unmeasured concerns. Yet, society wants schools to attend to these matters, and most professionals understand that personal and social functioning are integrally tied to academic performance. From this perspective, it seems self-defeating not to hold schools accountable for student improvements related to such goals.
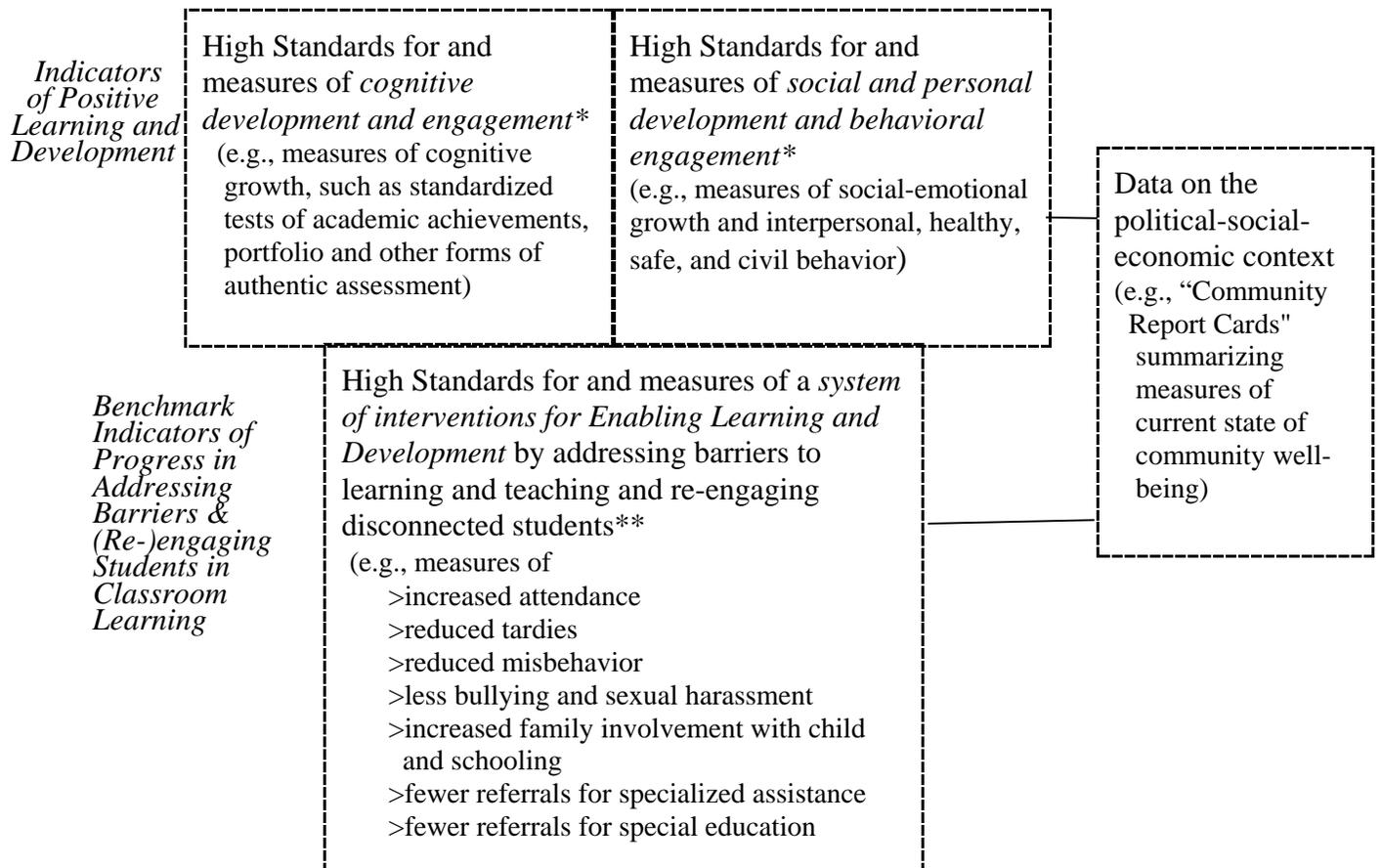
It is also self-defeating not to attend to benchmarks related to learning supports

And, for schools where a large proportion of students are not doing well, it is also self-defeating not to attend to benchmark indicators of progress in addressing barriers to learning and teaching. Schools cannot teach children who are not in class. Therefore, increasing attendance always is an expectation (and an important budget consideration). But there are also other basic indicators of school improvement and precursors of enhanced academic performance (e.g., reductions in tardiness, problem behavior, suspensions, dropouts, inappropriate referrals for special education). Given this, the progress of school staff related to such matters should be measured and treated as a significant aspect of school accountability, with credit given for all progress.

Exhibit

Expanding the Framework for School Accountability

Indicators of Positive Learning and Development

High Standards for and measures of *cognitive development and engagement**
(e.g., measures of cognitive growth, such as standardized tests of academic achievements, portfolio and other forms of authentic assessment)

High Standards for and measures of *social and personal development and behavioral engagement**
(e.g., measures of social-emotional growth and interpersonal, healthy, safe, and civil behavior)

Data on the political-social-economic context (e.g., "Community Report Cards" summarizing measures of current state of community well-being)

Benchmark Indicators of Progress in Addressing Barriers & (Re-)engaging Students in Classroom Learning

High Standards for and measures of a *system of interventions for Enabling Learning and Development* by addressing barriers to learning and teaching and re-engaging disconnected students**
(e.g., measures of
>increased attendance
>reduced tardies
>reduced misbehavior
>less bullying and sexual harassment
>increased family involvement with child and schooling
>fewer referrals for specialized assistance
>fewer referrals for special education

*Results of interventions for directly facilitating development and learning.

**Results of interventions for addressing barriers to learning and development.

Context Matters

Finally, school outcomes, of course, are influenced by the well-being of the families and the neighborhoods in which they operate. Therefore, performance of any school should be judged within the context of the current status of indicators of community well-being, such as economic, social, and health measures. If those indicators are not improving or are declining, it is patently unfair to ignore these contextual conditions in judging school performance.

Of Course it's Complex

How effective is the intervention? Do you have data to support that approach? Where's your proof?

The questions are so logical and simple to ask, and they can be so devastating in their impact. One problem is that such questions imply that relevant data are easy to gather, and so if data aren't available, the intervention must be ineffective or else those in charge are being irresponsible. Usually ignored by the questioners are the many complexities associated with valid and ethical evaluation.

Indeed, under reform conditions, policy makers often want a quick and easy recipe to use. This leads to measures aimed at holding administrators and staff accountable for specific, short-term results. Little thought is given to the negative effects such a limited focus can have on achieving more complex desired long-term results.

Methodologically, for instance, evaluation and accountability must be carried out with tools that are technically limited (e.g., many lack adequate standardization, have poor reliability, have poor validity). Interpretations of findings are made using different and sometimes biased perspectives. Consequently, what is reported often is controversial.

Moreover, almost everyone has experienced negative consequences from evaluation. Those evaluated often are harmed, and consumers of evaluation reports frequently are misled. Evaluations create tensions and dilemmas and can be misused to create undesirable degrees of uniformity and conformity. Ethically, we should be as concerned with the consequences of evaluation as we are with improving our evaluation capability.

It is important to remember that choices about what data to gather and exclude are guided by policy decisions, and major decisions about education involve considerations that go well beyond the availability of valid data. Profound and conflicting social-political-economic-philosophic agenda are at play; so no one should be surprised that relevant data often are ignored, and some data are manipulated during policy debates and at decision making tables. As Rutkowski cautions, "Through educational indicators a set of 'truths' is arguably produced. However, these 'truths' are very open to interpretation."* And as Planty and Carlson stress, "Indicators of poor quality certainly distort and misguide decision making and policy."**

Furthermore, two unfounded presumptions are at the core of most current formal and informal evaluations in education. One premise is that an intervention in widespread use has arrived at a relatively evolved stage of development and, therefore, warrants the cost of summative evaluation. The other presumption is that major conceptual and methodological problems associated with evaluating intervention are resolved. The truth is that interventions are frequently introduced prior to adequate development, with a view to evolving them based on lessons learned. This is the case for many science-based practices brought to schools.

It is well to remember that empirical support indicating *efficacy* does not predict *effectiveness* when school personnel implement the practice under common school conditions.) Moreover, many well-institutionalized approaches remain relatively underfunded and underdeveloped. Finally remember that every review of the literature on formal evaluations outlines major unresolved concerns.

> Given this state of affairs, the nature and scope of too many accountability demands often are unreasonable and chronically reflect a naive view of the evidence base.

*Rutkowski, D. (2008). Towards an Understanding of Educational Indicators. *Policy Futures in Education, 6*, 470-481.
**Planty, M. & Carlson, D. (2010). *Understanding education indicators: A practical primer for research and policy.* New York: Teachers College Press.

School Related Assessments:
A Few Concluding Comments

> *Today's enthusiastic embrace of data has waltzed us directly from a petulant resistance to performance measures to a reflexive and unsophisticated reliance on a few simple metrics.... The result has been a nifty pirouette from one troubling mind-set to another; with nary a mistep, we have pivoted from the "old stupid" to the "new stupid."*
>
> Frederick Hess

Gathering good data to evaluate schools clearly contributes to school improvement. Doing so, however, involves more than amassing limited data on results. Evaluations can as easily reshape schools in negative as in positive directions.

Current practices must be rethought. A particular concern is ensuring that accountability pressures do not inappropriately narrow the mission of public education.

To these ends, policy makers must expand the framework for school accountability beyond indicators of cognitive development to ensure that systems are driven in ways that

- promote an equal opportunity for all students to succeed at school (e.g., include indicators that evaluate direct efforts to address barriers to learning and teaching and re-engage disconnected students)

- facilitate students' personal and social development

- evaluate system performance and development in the context of the surrounding community's current status.

Moreover, policy makers need to invest in supporting development of district and school information management systems that enable gathering and reporting data in aggregated and disaggregated ways (with data on individuals appropriately safeguarded).

For more, see

> *>Evaluation and Accountability: Getting Credit For All You Do!*
> http://smhp.psych.ucla.edu/pdfdocs/evaluation/evaluation.pdf
> *>Standards & Quality Indicators for an Enabling or Learning Supports Component*
> http://smhp.psych.ucla.edu/summit2002/qualityindicators.pdf

Folks who work in schools deserve more credit.

Sure, but they wouldn't need it if we paid them better.

Feel free to share this with concerned others.