

Improving Student Performance Through Teacher Evaluation: The Gallup Approach

August 2011

Gary Gordon, Ed.D.

GALLUP CONSULTING®
Education Practice

For more information about Gallup Consulting,
Gallup's Education Practice, or our solutions for
improving school performance, please visit
<http://education.gallup.com> or contact
the Education Practice at +1.800.288.8590
or education_practice@gallup.com.

Improving Student Performance Through Teacher Evaluation: The Gallup Approach

August 2011

By Gary Gordon, Ed.D.

Improving Student Performance Through Teacher Evaluation

Of all the influences schools have on student achievement, the individual teacher has the greatest effect. William Sanders and colleagues made a case for teachers as the primary difference-makers in student achievement more than 15 years ago.¹ Most Americans concur, saying that “having higher-quality, better-educated, and more-involved teachers is the best way to improve kindergarten through grade 12 education in the United States.”² Improving teacher quality requires major changes in the human capital strategy for most school districts and involves improving the recruitment, selection, development, evaluation, compensation, and retention of teachers. Teacher evaluation is just one key element in this process.

Due to several factors, however, teacher evaluation and compensation are the current focus for changing America’s schools. A series of studies clearly suggested that teacher evaluation was broken and failed to yield credible information in many — if not most — school districts. The U.S. Department of Education’s Race to the Top initiative made teacher and principal evaluation a cornerstone of education reform.³ And the Bill & Melinda Gates Foundation’s Measures of Effective Teaching project focused on identifying teacher effectiveness through new “value-added” measures of achievement and a variety of teacher observation systems that could be replicated by states and districts.⁴

Unfortunately, the rush to adopt teacher frameworks as standards for evaluation accepts without question returning to an old strategy for improving teacher performance. The conventional thinking is that someone — a principal or an outside evaluator — must be the judge of whether a teacher’s behaviors in the classroom are the correct way to teach. Making improvements to the previous practice

of teacher observations would improve the strategy. It’s important, however, to examine the underlying theory of teacher observations as a major factor in teacher evaluation and its potential for unintended consequences.

This examination reveals that there is a more effective strategy: a combination of measures that demonstrate relationships to student achievement can provide the feedback needed to help teachers grow and increase student achievement. Combining measures of student growth, assisting colleagues, and student hope and engagement results in a performance index that reflects real outcomes for individual teachers. This approach shifts the task of evaluation from subjective judgments of what is right and wrong to meaningful and individualized teacher development that results in improved student learning. Rather than treating every teacher the same, this approach acknowledges a teacher’s individual strengths and weaknesses and uses those strengths to improve performance.

Two common purposes for teacher evaluation include providing evidence for employment decisions and improving instructional performance. However, today we understand that teacher evaluation must also aim to improve student success in school. We should keep these goals in mind as we consider how teacher evaluation has been used in the past and how to improve the teacher evaluation process in the future.

What’s Wrong With Teacher Evaluation?

Changes in the current teacher evaluation system are needed because the current system has not worked. A study of 140 Midwestern school districts suggests that teacher evaluation was not taken very seriously. Fewer than half of the district evaluation-policy statements indicated how teacher evaluation results would be used. When a purpose

was mentioned, evaluations were used to provide evidence for retaining probationary teachers or dismissing tenured teachers.⁵ No consideration was given to improving teacher performance or student achievement.

Principals in these districts received little direction except to identify and remove the worst teachers. Most teacher evaluations were done by principals and assistant principals, and these administrators received district-policy guidance limited to the required aspects of the evaluation process (who would conduct the evaluation, and when and how often it would take place) rather than the content of the observation (what should be evaluated, what standards should be used, and what would be done with the results). The study found that only two in five districts specified the criteria by which teachers should be rated. About half the districts with documented criteria used external standards as the basis for rating teacher observations. Only 8% of the districts' policies called for training or certification of raters on how to conduct evaluations.⁶ In the remaining districts, principals or raters were left to their own devices or perhaps were armed with only a checklist, leaving teacher evaluation open to charges of subjectivity and favoritism. The time allotted to the evaluation process also suggests its low priority, which severely limits the effect and credibility of teacher evaluation.

The same study of Midwestern districts found that two or more principal observations were typical for beginning teachers, while experienced teachers were evaluated once every two or three years.⁷ Another study, which included 12 districts in four states, found that two or fewer observations per year was the norm, with evaluations averaging a total of 76 minutes or less.⁸ Even in school systems rated as high quality by researchers, teachers may be evaluated only in years one and four, then every fifth year.⁹

A series of studies indicates that principals tend to use only the highest categories in a scale and fail to differentiate between successful and unsuccessful teachers. In a study including districts in four states, half the districts used

only "satisfactory" or "unsatisfactory" categories for the ratings; in those settings, 99% of the teachers were rated as satisfactory. The remaining districts — those with broader rating categories — didn't see much difference in results, though, as 94% of their ratings fell into one of the top two rating categories.¹⁰ A separate study in a large urban district found that 90% of the district's teachers received evaluations in the top two rating categories. In this same district and over a four-year period, 88% of the schools did not issue an unsatisfactory rating for any of the teachers.¹¹

As a result of this approach — the lack of objective criteria, a small number of observations, and the nearly universal use of the highest ratings — it appears that these districts failed to remove teachers who were not helping students learn. In one district, over a two-year period, a study found that even in schools that were viewed as failing, no unsatisfactory ratings were issued for teachers in 79% of the schools.¹² In another study across 12 districts, 41% of the principals surveyed had never recommended nonrenewal of a probationary teacher. Yet, in the same study, 81% of administrators and 57% of teachers admitted that at least one tenured teacher in their school would be rated as delivering poor instruction.¹³

Not surprisingly, in studies where teachers were asked about the evaluation process, the majority indicated that it failed to improve their performance. In one district's study, slightly more than one-third of teachers surveyed could "strongly agree" or "agree" that the process improved their performance.¹⁴ A larger study found that during the evaluation process, nearly three of every four teachers received no feedback on how to improve; during their first four years of teaching, only 43% of teachers received feedback identifying areas in which they needed developmental improvement. Even when areas of improvement were cited, fewer than half of these teachers believed they received support to improve.¹⁵

Does every school district's teacher evaluation program suffer from these problems? Not necessarily. But the assorted problems — the lack of criteria for making

judgments in teacher observations, inadequate training of principals or raters, the infrequency of observations, nearly all teachers being rated highly, and the failure to provide information for improvement — reduce teacher evaluation to a required but perfunctory exercise in far too many school districts.

Supporters of Change

In fall 2009, two major initiatives emerged that promised to change the process and outcomes of teacher and principal evaluation. The U.S. Department of Education's Race to the Top program provided \$4.35 billion in incentives for competitive applications from states, and a significant piece of the program focused directly on the problems with teacher evaluation. The grants were designed to stimulate reform and innovation through statewide efforts. The Race to the Top selection criteria for "Great Teachers and Leaders" specifically required state evaluation systems to 1) evaluate teachers and principals annually, 2) use multiple rankings, 3) consider student academic growth, and 4) provide constructive critical feedback to teachers. In the participating states, the evaluations were required to be used for development, compensation, tenure or certification, and job action for ineffective teachers or principals. In addition, states were required to develop measures of individual student growth.¹⁶

The Bill & Melinda Gates Foundation's Measures of Effective Teaching (MET) project also was announced in fall 2009. The MET project intends to provide research-proven descriptors and measures of teacher effectiveness for feedback and evaluation. Three thousand teachers in six predominantly urban school districts are taking part in an intensive research project. The MET project is described as based on three premises: a measure of student achievement growth should be part of teacher evaluations; other evaluation processes should demonstrate a relationship to student achievement; and measures should provide feedback to teachers to support their growth and development. The MET project examines the relationships between measures of value-added student achievement,

the Balanced Assessment of Mathematics, SAT 9 Reading Open-Ended Test, and seven observation systems. Trained evaluators will score videos of teachers and the specialized observation systems for language arts, mathematics, and science.¹⁷

The Race to the Top initiative and the MET project are different in their scope but share some characteristics regarding teacher evaluation. Race to the Top focuses on policy changes and initiatives in participating school districts that have statewide implications on teacher evaluation. The MET project focuses on initiatives to improve student learning in six school districts. The Race to the Top states will provide data on various teacher evaluation systems adopted at the state level. The MET school districts will provide data from a selected group of teacher observations systems. In both Race to the Top states and MET school districts, value-added measures of student achievement will be used to judge the effectiveness of the various approaches.

The two initiatives intend to develop models of evaluation that lead to professional growth and that result in improved student performance. A major change in the criteria for evaluation, value-added or other student growth measures are part of the equation for evaluating teachers and using overall school-level student performance for evaluating principals. In addition, both initiatives aim to replace a highly subjective system that is wholly dependent on observations of teachers with a more objective approach that uses structured observation systems and trained evaluators.

What, then, are some of the major structured teacher observation systems? And how are they related to student achievement?

Teacher Evaluation Systems

A number of systems exist that improve on previous practices in teacher evaluation. Most of the systems follow the familiar practice of teacher observation.¹⁸ These

approaches differ from the past by using teaching standards as the criteria for evaluation, a rubric for assessing levels of accomplishment, and training of principals or outside evaluators in using the standards and rubric. While not intended to be inclusive of all available systems, the following approaches represent some of the most widely used or contain unique differences in approaches.

Charlotte Danielson Framework for Teaching

The Charlotte Danielson Framework for Teaching is perhaps the most widely used teacher observation model in the United States, and it will be evaluated as part of the MET project. The Danielson framework begins with four domains (planning and preparation, classroom environment, instruction, and professional responsibilities), then breaks them into 22 components. A scoring rubric serves as the standard for rating teachers in each of the components, and teachers are rated at one of four levels ranging from “unsatisfactory” to “distinguished.” The framework stems from work done by the Educational Testing Service in developing Praxis III assessments for beginning teacher licensing, and it is based on the Praxis III criteria. The Danielson framework is referred to as a standards-based observation approach because of its relationship to Praxis III and the Interstate New Teacher Assessment and Support Consortium standards.¹⁹

Studies of the Framework. The Measures of Effective Teaching (MET) project is expected to provide the most comprehensive examination of the relationships between the Danielson framework and a group of subject-specific approaches and student achievement. However, two previous studies indicate students of teachers with higher ratings on the framework performed at higher levels.

A study by Milanowski, Kimball, and White involved three districts; two used the Danielson framework and rubrics as the structure for the evaluation system, and the framework shaped the third district’s evaluation system. The study found a relationship between the evaluation ratings of teachers and their students’ value-added achievement

growth. In all three districts, effect sizes were higher in math than reading and were negative in the one district with science data.²⁰

More recently, Taylor and Tyler studied the Teacher Evaluation System (TES) in Cincinnati Public Schools, one of the districts included in Milanowski, Kimball, and White’s work. In Cincinnati, the Teacher Evaluation System gathers data from four observations — three by trained evaluators and one by the principal — and a portfolio of work products (such as teacher lesson plans and professional development activities). Teachers in Taylor and Tyler’s sample ranged from five years to 19 years in experience. The researchers found that participation in TES improved the effectiveness of the teacher in promoting student growth in math, most importantly, in the years following evaluation. No significant effect emerged for reading achievement. The least-skilled teachers benefited the most from the evaluation process.²¹

This level of evaluation, which requires increased observations by trained evaluators and principals, increases a district’s financial investment. Taylor and Tyler report that Cincinnati Public Schools budgeted \$7,000 to \$7,500 per teacher evaluated. This cost may explain why teachers in the district are evaluated in only years one and four, then every fifth year.²²

Unfortunately, we don’t have longitudinal studies that show whether the relationship found by these two studies between the Danielson framework ratings and the student performance increases continues in the years between the teachers’ evaluations. We don’t know if the improvement of teacher practice, which presumably comes from the evaluation process, produces short-lived increases in teacher effectiveness as measured by the value-added scores. If the effect of evaluating the teacher with the framework does not last — or if it fails to build on previous evaluations — the frequency of evaluation may need to be increased. Additionally, both studies included only teachers of students in grades 4 through 8, leaving the utility of the framework at the high school level unknown.

Teacher Advancement Program

Developed by the Milken Family Foundation, the Teacher Advancement Program (TAP) addresses the need for more observations, use of standards and a rubric, training for evaluators, and measures of student achievement. In addition to improving these evaluation components, TAP measures student achievement gains and provides incentives for student achievement. Jerald and Van Hook maintain that TAP was designed to measure teacher effectiveness and to improve performance over time through feedback and professional development.²³

TAP addresses many of the problems that districts encounter with teacher evaluation. Principals, mentor teachers, and master teachers observe teachers from four to six times during the school year, using three of Danielson's four domains — planning and preparation, the classroom environment, and instruction — and 19 of the framework's 22 components.²⁴ Evaluators must complete four days of training, a certification test, and an annual recertification, and meet monthly to check for inter-rater reliability and other issues. Teachers receive training in the rubric by master and mentor teachers.²⁵

After the observation, efforts focus on helping teachers improve. Jerald and Van Hook maintain that TAP evaluators provide feedback tied to indicators in the rubric in 40- to 60-minute sessions. Professional development is targeted to meet the needs identified in student achievement data and teacher observations, connecting teacher observations with teacher development. This can include coaching, model teaching, and team teaching by the mentor and master teachers. Groups of teachers meet weekly as part of the school schedule to review student data or learn new techniques. TAP teacher incentives are awarded based on evaluation ratings (50%), classroom student achievement growth (30%), and school-wide student achievement growth (20%).²⁶

Studies of TAP. An early study compared the performance of TAP schools to control schools in Arizona

and South Carolina. In Arizona, the sample of 13 TAP schools served as a third-year study of TAP schools; in South Carolina, the sample served as a first-year review for TAP. Student achievement data were available for grades 2 through 8 in Arizona and grades 3 through 8 in South Carolina. The state departments of education in both states identified control schools, but only one control school was identified for each of the six South Carolina TAP schools. Over three years, nine of the 13 Arizona TAP schools exceeded the performance of the control schools in reading, mathematics, and language. Four of the six South Carolina TAP schools performed better than their individual control schools in mathematics, and three of the six TAP schools out-performed their control schools in reading/language arts.²⁷

Springer and colleagues at the National Center on Performance Incentives²⁸ raised questions about the results in this initial review of TAP schools and a follow-up study.²⁹ Springer and colleagues conducted their own evaluation of TAP schools, using a different student achievement database that contained fall and spring mathematics scores in grades 2 through 10 and spanned four school years. In initial analysis, the TAP schools demonstrated greater student growth from the fall to spring measures than comparison schools.

However, the researchers noted that membership in TAP is selective, requiring support at the school and district levels. Controlling for the selective aspect of TAP schools, the researchers saw continued positive effects in elementary grades, but effects in grades 6 to 10 were either negative or insignificant. Springer and his colleagues noted their study's limitations, including a small sample of TAP schools, the different tests used, and the inability to know if the TAP sample schools are representative of all TAP schools.³⁰ The selectivity issue of TAP schools raised by Springer and colleagues may mean that these schools are different in other unmeasured features that explain part or all of the differences.

A more recent review of TAP schools focused on the relationship between teacher evaluation rankings and value-added scores of student achievement growth. Data from the 2006-2007 and 2007-2008 school years for 1,432 teachers in 104 schools in 10 states were used. The researchers found that a one-point increase (on a five-point scale) in evaluation ranking resulted in more than a half-point increase (on a five-point scale) in value-added scores on average. Moreover, the data suggested that teachers improved in their rankings over the two years included in the study.³¹

The emphasis on teacher development, using student growth measures, and including pay incentives for student growth place TAP more in line with the Race to the Top guidelines. But by adopting a modified version of the Danielson framework, similar questions exist about TAP's long-term effects. With the MET project testing the Danielson framework, TAP's position will be strengthened if the study finds positive relationships with student growth. The question raised by Springer and colleagues — whether schools that volunteer and are selected to be TAP schools are already performing differently than control schools — should also be explored by the TAP researchers or through independent studies.

Extensive training of evaluators, increased observations, professional development following evaluation, and performance incentives for student growth come at a price. Toch states that TAP can cost from \$250 to \$700 per student or up to 6% of per-pupil expenditures. This amounts to \$6,250 to \$14,900 per teacher, assuming 25 students per class.³²

Teacher Performance Assessment

The Teacher Performance Assessment (TPA) was developed from a partnership between the American Association of Colleges of Teacher Education and Stanford University. Linda Darling-Hammond and Raymond Pecheone led a group of Stanford researchers in the project, using the Performance Assessment for

California Teachers, a teacher licensure process, as a model. TPA was aligned with the Interstate Teacher Assessment and Support Consortium (InTASC) and Common Core State Standards initiative.³³

Though teacher observation serves as the keystone for most other teacher evaluation systems, TPA takes a different approach. Using a portfolio method, evidence of teaching competency is generated through subject-specific "Teaching Events," with separate forms for elementary teachers (multiple subjects) and secondary teachers (single subject). Different measures record teaching and learning during three- to five-day learning periods with one class of students. Evidence includes lesson plans, videos of teaching, student work, and teacher reflections, which form a teacher portfolio that is scored by trained evaluators.³⁴ The Performance Assessment for California Teachers uses a four-level rubric that is subject specific for five domains: context and planning, instruction, assessment, reflection, and academic language.³⁵ In addition, Embedded Signature Assessments allow individual assessments to be added locally. Examples of embedded assessments could include student case studies, analyses of student work, or observations. TPA aims to provide teacher preparation institutions, states, and school districts with evidence to make licensure, tenure, and employment decisions.³⁶

Studies of TPA. No performance data or cost information for TPA is currently available, though a study of the Performance Assessment for California Teachers has been done. A TPA pilot was launched in the 2010-2011 school year, and more testing is scheduled for 2011-2012. A partnership between Pearson and Stanford University, announced in March 2011, will result in a Web-based platform that will allow the electronic transmission of the Teaching Event data and scoring of the assessment.³⁷ An estimated overall pass rate of 85.4% is reported in the technical report for the Performance Assessment for California Teachers.³⁸ Unknown at present is how practicing teachers will rank on TPA's overall rubric scoring and relationships to student achievement.

District of Columbia Public Schools: IMPACT System

The District of Columbia Public Schools system has implemented teacher evaluation that combines a structured teacher observation system with value-added student growth measurement and other measures. The IMPACT system for teacher evaluation includes five components for Group 1 teachers (reading and mathematics teachers in grades 4 through 8) where value-added scores can be calculated. The five measures include Individual Value-Added Student Achievement Data (50% of the total), Teaching and Learning Framework observation ratings (35%), Commitment to the School Community (10%), and School Value-Added Student Achievement Data (5%). The fifth measure, Core Professionalism, comes at the end of the process and rates attendance, tardiness, policy compliance, and interactions with the school community. Points may be deducted from the total evaluation score if the teacher is rated below standard in Core Professionalism. A final score for each of the measures is totaled, and the score ranges are converted to Highly Effective, Effective, Minimally Effective and Ineffective ratings.³⁹

Group 2 includes more than 80% of all teachers in the District of Columbia Public Schools.⁴⁰ The IMPACT scores for Group 2 teachers, whose students are not tested in the district's assessment program, have four components. There is a greater emphasis on teacher observation, with the Teaching and Learning Framework accounting for 75% of the total evaluation score. Teacher-Assessed Student Achievement Data (10%) is used for these teachers, who have no value-added data. Commitment to the School Community (10%), School Value-Added Student Achievement Data (5%), and Core Professionalism round out the evaluation measures.⁴¹

The multiple measures are noteworthy in the District of Columbia Public Schools' IMPACT approach. An evaluator who conducted focus groups with teachers, principals, evaluators, and district office staff about the

IMPACT evaluation system noted that the use of multiple measures is viewed as the system's greatest strength, even by those critical of the system. The researcher noted that in initial conversations, teachers had strongly requested a number of measures to be included in the new evaluation system.⁴²

Teachers, principals, and central office staff reviewed several systems and developed the system's standards, its Teaching and Learning Framework, during the 2008-2009 school year. Three domains — Plan, Teach, and Increase Effectiveness — make up the framework. The focus in 2009-2010 was on the Teach domain, with professional development for principals, master educators (district evaluators), and teachers. Revisions were made to the framework as a result of feedback from evaluators and teachers. The 2010-2011 school year marked the rollout of the IMPACT teacher evaluation with teachers only rated on the Teach domain.⁴³

Another strength of the IMPACT system is the multiple observations done by principals and master educators, who serve as district evaluators. Principals conduct three teacher observations per year, and the master educator observes a teacher two times, providing two perspectives of the teacher's performance. Teachers are rated on the Teaching and Learning Framework's nine standards in the Teach domain with a rubric incorporating a one-to-four scale. The observation ratings are averaged to create a single score for Teaching and Learning. This one-to-four rating is then given 35% of the weighting for teachers with value-added scores, and a 75% weight for teachers without value-added scores, in the final evaluation score.⁴⁴

Studies of IMPACT. While an evaluation of IMPACT implementation in 2010-2011 is coming, two evaluations were done in the 2009-2010 phase. Curtis concludes that the Teaching and Learning Framework stimulated discussions among teachers, principals, master educators, instructional coaches, and district administrators. But she adds that those conversations do not appear to delve deeply into the standards, ways in which teachers may improve,

and principals' support of teachers. She notes that this may be due to the early stage of implementation or the pressure for teachers to discover how to be rated as 3 or 4 on the standard.⁴⁵

Headden conducted focus groups with teachers, principals, evaluators and district representatives. She noted that while the value-added measure raised controversy, teacher critics were most concerned about ratings on the Teaching and Learning Framework. Once continued employment and compensation make evaluation a high-stakes activity, unintended consequences begin to appear. Headden quoted an instructional coach (a teacher hired to assist in teacher development) as saying:

“Teachers aren’t stupid. Do you think they are really doing these things? They do them only for the 30 minutes they are being observed.... They pull out a new lesson plan they have in their drawer for an occasion just like this. They say [about whatever they are doing], ‘Oh kids, never mind. I think we are going to learn about the planets today.’”⁴⁶

At the same time, Headden noted that teachers intensely want feedback from the master educators. Written feedback and a conference are required following observations to serve this purpose. But teachers expressed a desire for more feedback and a tighter connection between support and evaluation. In Headden’s view, feedback to teachers is reduced in part because the master educator’s job is seen as overwhelmingly evaluative rather than developmental.⁴⁷

Is This the Best Strategy?

Though the MET project specifically tests the association of some teacher observation systems with student growth — and many Race to the Top states are adopting observation systems that will be tested — crucial questions remain about whether intensive teacher evaluation is the best answer to improving teachers’ instruction or students’ growth. Is the fundamental assumption underlying teacher

observation systems sound? Does an annual evaluation that is focused on teachers’ weaknesses — and remediation that is identified by an observation system — improve student growth? Without large infusions of federal funds, will states or individual school districts financially support the intensive efforts required of teachers, principals, and other evaluators to observe every teacher every year with multiple observations?

Admittedly, structured teacher observation systems have noteworthy assets. The structure comes from basing the desired teacher behaviors on standards developed by experts to describe good teaching and a rubric to identify varying levels of performance. This gives these systems a good deal of face validity. Moreover, teacher observation is the method by which teacher evaluation has always been done. Teachers and principals are familiar with the process, but varying levels of comfort may exist with the process in a high-stakes environment. An observation system has a language that everyone can use to describe behaviors. As indicated previously, there is some evidence of a relationship between teacher ratings in structured observation settings and value-added scores for teachers. Yet difficulties arise when teacher observation systems are put into practice.

What’s Wrong With This Strategy?

The underlying theory behind all teacher observation systems is that there is a best way to teach that results in student learning. This best way to teach is represented by the standards, which are created by experts and — we hope — supported by a positive statistical relationship with teacher value-added scores. As noted in the District of Columbia Public Schools, the greatest asset of the observation system is the common language to describe the desired standards. The rubric for each observational system differentiates teachers by measuring the gap between observed teacher behaviors and the desired behaviors exemplified in the teaching standards. The process rests on the belief that reducing the gap between desired and observed teacher behaviors improves performance; by

improving performance as measured by the rubric, the inference is that student learning will increase.

But the asset of focusing on specific, ideal teaching behaviors is also the systems' Achilles heel. At base, structured observation systems suggest that every teacher should improve in the same way — but perhaps in different areas — by exhibiting all the model's behaviors. The systems ignore all other teacher behaviors that could be meaningful to teacher effectiveness and place a high value on the observation system's standards. The argument suggests that if teachers develop more of the desired behaviors — and if those behaviors are connected to student learning — student achievement will increase.

This argument ignores the teacher. Teacher observation results in increased student learning only if two criteria are met: 1) teachers must accept, commit to, and internalize the behaviors as important to them; and 2) the standards must capture a large enough sample of the teacher behaviors that relate to student learning in a meaningful way. One method for gaining teachers' attention and forcing acceptance of the evaluation system is to tightly couple evaluation results to compensation. But the unintended consequences of treating teaching like assembly line work can be destructive, driving some teachers to leave the field and creating a general mind-set that teaching is just a job, not a profession.

When the purpose of structured observation systems moves from professional development to evaluation, retention, and compensation, those systems become high-stakes propositions. Teachers will learn to play the evaluation game, just as they did during the last wave of teacher observation in the 1980s when Madeline Hunter's model was in vogue. As the pressure increases, a teacher's focus tends to shift from teaching students to displaying the desired behaviors on demand, as suggested in the quote from the District of Columbia Public Schools instructional coach. When teacher observation systems become attached to continued employment or compensation, they tend to become a one-size-fits-all, mechanistic, prescriptive

approach to improving teaching. Everyone involved tends to lose sight of the goal, which is student learning.

Teacher observation systems often become "gotcha" schemes in which there is always a deficiency to remedy. In short periods of 30 to 50 minutes, a teacher must demonstrate as many of the required elements as possible for the evaluator. An evaluator always will be able to find areas in which even the best teachers don't measure up and need to improve. In an evaluation system, the observer's job is to find deficiencies to differentiate teachers. Though positive aspects are shared with teachers by evaluators, the focus most often becomes a teacher's weaknesses and the areas requiring remediation. And there is always something to remedy.

But the areas of the model in which the teacher is seen as deficient and asked to correct might not have strong ties to student learning, even if the model as a whole does. Not all four domains and 22 components of the Danielson framework, for example, are equally related to student learning in real classroom practice, and particularly not for every teacher. This is not a severe limitation when the model is used for development. But in an evaluation system, it can create uncertainty and distract from elements that are crucial to student learning. A teacher who produces solid student achievement may lose focus on what is important when trying to improve in areas that have a lesser impact on student growth in order to score well on the system. Yet all of the observation system's components are equally weighted as evaluators assess, provide feedback to teachers, and identify areas of improvement.

Lastly, much of the discussion around teacher evaluation assumes a simple equation: teacher effectiveness = achieving greater-than-anticipated gains on state tests. Though a focus on student achievement is crucial, it narrowly defines teacher effectiveness. It overlooks the importance of experiences with the teacher that result in higher levels of student engagement and hope for future success in the subject and in school. This focus also ignores the contribution that a number of teachers have on the

student's development, particularly in middle school and high school, where students may have six or seven teachers every day. It discounts whether the student is prepared for the next step in school or a career. It fails to count whether the student is able to take greater personal responsibility for his or her behaviors.

Laura Goe reviewed a broader definition of teacher effectiveness suggested by a variety of researchers. From this work, she proposes a definition that includes the following:

1. Effective teachers have high expectations for all students and help students learn, as measured by value-added or other test-based growth measures, or by alternative measures.
2. Effective teachers contribute to positive academic, attitudinal, and social outcomes for students, such as regular attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior.
3. Effective teachers use diverse resources to plan and structure engaging learning opportunities; to monitor student progress formatively, adapting instruction as needed; and to evaluate learning using multiple sources of evidence.
4. Effective teachers contribute to the development of classrooms and schools that value diversity and civic-mindedness.
5. Effective teachers collaborate with other teachers, administrators, parents, and education professionals to ensure student success, particularly the success of students with special needs and those at high risk for failure.⁴⁸

The current trend in using value-added scores and teacher observation systems attempts to address the first and third points in Goe's definition, but this strategy doesn't go far enough. These two measures ignore other factors that contribute to positive student outcomes, including school

climate and culture, school leadership, school resources, and parent engagement.⁴⁹

An Alternative: Multiple Measures and Teacher Performance Coaching

Efforts to improve student learning will be more successful if expectations and outcomes are clear and there is flexibility in the ways teachers achieve these goals. Intrinsic motivation and Type I behaviors, as Daniel Pink describes them, include self-direction, seeking excellence in what we do, and serving a larger purpose. Those behaviors lead to higher performance and greater physical and mental wellbeing. But as Pink suggests, "Ultimately, Type I behavior depends upon three nutrients: autonomy, mastery, and purpose."⁵⁰ Appraisal strategies must focus on the purpose of student growth, provide autonomy when earned, and offer a road to mastery for each teacher.

As a result, an alternative strategy uses multiple outcome measures rather than value-added scores and enforced compliance to teacher observation standards through evaluation. The adopted outcome measures should incorporate the broader elements of Goe's definition of teacher effectiveness and relate to student achievement. This is similar to what the MET project's goals suggest.

Additionally, an alternative strategy should achieve other objectives. From a practical view, the measures should stand as effective and efficient metrics in administration, in report distribution, and in use by teachers, principals, and other evaluators. Evidence of student learning should be clear. School-wide measures should be incorporated because an individual teacher is one of a number — growing to six or seven during the day in middle school or high school — of teachers contributing to a student's learning environment. Lastly, the alternative strategy should provide a way to recognize and reward teachers who exhibit outstanding performance; provide for an objective means to remove teachers who are seen as ineffective even after intervening with assistance; and enable the large number of teachers in the middle to grow and improve.

Just as teachers are asked to differentiate instruction for students, teacher appraisal should differentiate development for teachers. Beginning teachers need intensive assistance because the first years of teaching can be overwhelming. The development of teachers during their first five years in teaching is crucial because those years seem to be related to student growth, while later years of experience are not.⁵¹ Helping beginning teachers to master their own learning curve is justified alone by the benefits that accrue to students.

In addition, between 40% and 50% of new teachers will leave their first district or the profession in the first five years of teaching.⁵² The price of that turnover is far more costly than districts realize. Though turnover costs are typically buried in most budgets, one study of five school districts found that the cost of turnover ranges from \$8,000 to \$13,650 for every person who left the organization.⁵³ Turnover is preventable, a topic addressed in other research.⁵⁴ Teachers with less-than-expected growth also need intensive assistance that identifies specific problems, gives feedback, and provides assistance and resources for the teacher. The issue here is clear: Students aren't learning with these teachers the way they are with other teachers. If these teachers can develop and produce expected growth, we must help them. If they can't, we must arrive at a dismissal decision, allowing time for improvement but without protracted delays. Principals complain about the extensive documentation and years required to document a dismissal recommendation, and this process can be simplified.

On the other hand, teachers who demonstrate student achievement gains that are greater than expected should be allowed to make process decisions in their classrooms without labor-intensive and expensive annual observations. A key management discovery is that great managers concentrate on outcomes and avoid the temptation to try to control the steps that effective workers use to accomplish goals. First, there are typically many different ways to address any task; often, when it comes to

capitalizing on human talent, trying to enforce a single approach as "the best one" reduces the organization's efficiency.⁵⁵ Second, endowing people with autonomy in the way they do their jobs sends a powerful message that the organization's leadership respects those individuals. In the 2011 Phi Delta Kappa/Gallup Poll, nearly three-fourths of Americans indicated that teachers should be allowed the "flexibility to teach in ways they think best" rather than "follow a prescribed curriculum."⁵⁶ These effective and highly effective teachers are those in whom Americans place their trust — and so should those who evaluate them.

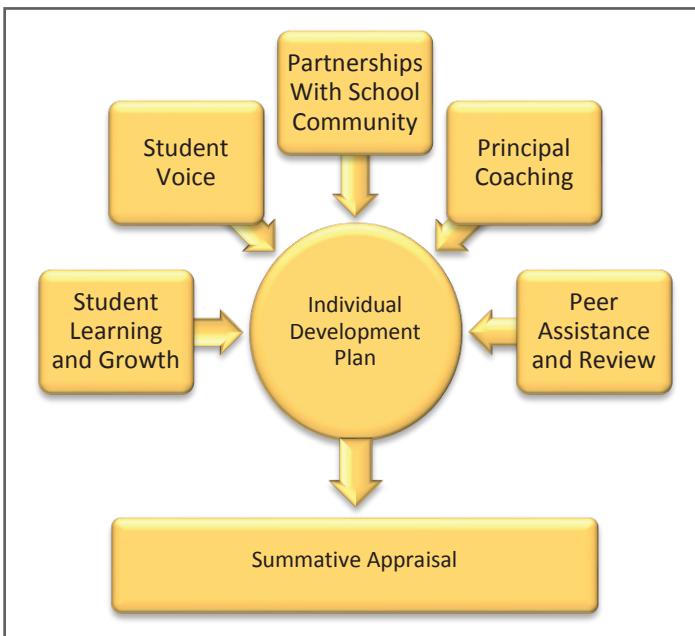
The structure of a proposed evaluation strategy includes multiple measures of teacher effectiveness as outlined in the figure below and the explanation that follows. The measures include individual and school-wide examples to create a more complete picture of a teacher's contribution to students and the school community. Four points summarize the approach:

1. Value-added scores, which are the best measure at this time of a teacher's contribution to student achievement, become an important outcome measure where available. Evidence of student growth can be provided by other teachers when value-added scores are not available.
2. Teacher observation should be used with beginning teachers and with teachers whose students are not showing sufficient student growth. The observations inform an individual development plan written specifically to the individual teacher's needs. Experienced teachers whose students show expected or greater-than-expected growth also need individual plans for their own learning and growth. They may request expert teachers to observe their classrooms to obtain feedback.
3. Peer assistance and review provides the greatest availability of expertise for beginning or struggling

teachers and experienced teachers whose students show expected or greater-than-expected growth who want additional feedback. Expert teachers conduct the teacher observations, then provide written documentation from the observation and give feedback to the teachers. As subject matter or grade level experts, these teachers can provide a broader perspective from across the district and a higher level of training.

- As coaches, principals should be required to visit their teachers' classrooms to understand how teachers are performing. If an experienced teacher who previously demonstrated effective or highly effective results is not performing well, the principal should recommend an appropriate intervention. As the best research on principals suggests, principals should create opportunities for collective leadership with staff members and focus on influencing teachers' motivation and working conditions.⁵⁷

Figure 1. Teacher Performance Index



Student Learning and Growth

The questions about student performance and the tests to measure performance have changed over the last 30 years. Following the publication of *A Nation at Risk* in 1983, nearly all states created course and grade-level standards and tests to measure achievement of those goals. The question became, "Are students proficient?" The passage of the No Child Left Behind Act of 2001 added subgroup breakouts, such as minority and special education students, to compare the performance of the subgroups with other students.⁵⁸ Yet it was apparent that students could make substantial learning gains (growth) but not reach proficiency (achievement). The movement toward using growth measures accelerated with the Race to the Top initiative's requirement that states develop growth measures.

While growth and value-added measures are sometimes considered as synonymous, there are significant differences. A growth model measures achievement scores for a student in two consecutive years in the same subject, such as fourth-grade math and fifth-grade math, and compares the change in achievement. Some growth models incorporate additional years of data, but the process still measures the achievement in one year compared to the preceding year's achievement level.⁵⁹

Value-added models use a more complex approach to measure growth for individual teachers. Test scores in multiple years in the same subject are used to predict a student's score in the next year using complex statistical techniques. As an example, math scores in the third, fourth, and fifth grades would be used to predict expected growth in the sixth grade. The student's growth is viewed as the difference between the predicted score based on past performance and the score based on actual performance. The average of the student gain scores for a specific teacher is used as a measure of the teacher's impact on the students' learning. Value-added models may incorporate a number of factors to improve reliability, including factors that may

influence student performance, such as gender, race, or socioeconomic status.⁶⁰

Teachers can be compared with other teachers in the district or state based on learning impact. Ranking teachers by their students' growth becomes the value-added score for the teacher. A value-added ranking for a teacher whose students grew as expected equates to average or effective; student growth that is greater than expected is considered above average or highly effective; and a ranking of less than expected is seen as ineffective.⁶¹

Though controversial,⁶² the MET project's report of its initial findings, as well as other researchers' views, supports the use of value-added scores. Though value-added scores have limitations and some variability from year to year, they offer a good assessment of student gains and a prediction of future performance.⁶³ Benefits exist for teachers, principals, and mentors by providing a broader view for comparing the teacher's performance to others outside the school or school district. Value-added scores are the best approach available at this time for estimating the teacher's influence on student learning.⁶⁴

Value-added scores are not available for all teachers. States vary considerably in their testing programs. In some states, this group of teachers may include grades 4 through 8 in reading and math and one grade in high school, but in other states, the subject areas tested are much broader and include more teachers. End-of-course testing is coming at the high school level and in subject areas beyond reading and math in elementary and middle schools. The Common Core State Standards initiative continues to move ahead in developing common assessments.⁶⁵ Consequently, the likelihood is fairly high that more teachers will have test data available in their subjects in the near future.

Teachers who do not have test data from which value added scores are calculated and school organizations without value-added scores need samples exhibiting student achievement. The goal is for all teachers to find

ways to demonstrate that their students are learning and growing.

Student achievement evidence can take many forms. Examples include end-of-course tests used in common by grade levels or departments, advanced placement test scores, individualized education program data, samples of student work, contest ratings, or the percentage of students moving to higher levels of a grade or subject. At the beginning of the year, teachers and administrators should agree to the measures to be used. Teachers' submissions can be ranked based on student achievement or growth on the assessments.

Student Voice

Perhaps the most neglected of all measures of teacher effectiveness is the voice of students. More than 1.2 million students drop out of high school every year. The personal decision to leave school develops through a slow process in which students lose hope and engagement in school.⁶⁶ Individually, and collectively as a staff, teachers are most effective when they raise both the hope and engagement of students in the learning program. Students are in the classroom every day and best understand the learning conditions and climate in individual classrooms and the school as a whole.

Three elements — hope, engagement, and wellbeing — provide a critical insight for teachers and administrators. *Hope* taps into the ideas and energy students have for the future. Student success and hope are intertwined. When students achieve successful outcomes, especially early in their school careers, hope grows. *Engagement* reflects the commitment to, involvement in, and enthusiasm for the work at hand in classrooms and the school. Engagement leads to higher levels of effort and persistence in the classroom and school. *Wellbeing* reflects how students think about and experience their lives. Wellbeing tells us how students are doing today and predicts their success in the future. Students with thriving wellbeing think about their present and future lives in positive terms.

- **Hope.** C. R. Snyder developed a psychological theory and model of hope that is based on goal-directed thinking. Hope theory describes a person's ability to visualize a future with goals, create strategies to reach those goals, and maintain the energy and drive to use those strategies.⁶⁷ Hopeful students have better grades in core subjects in middle school⁶⁸ and higher scores on achievement tests.⁶⁹ High school students⁷⁰ and beginning college students⁷¹ with higher levels of hope have higher overall grade point averages. The ability of measures of hope to predict achievement remained when controlling for prior achievement⁷² and entrance examination scores.⁷³
- **Engagement.** Student engagement encompasses three crucial factors in student learning: behavior, emotion, and cognition. Behavioral engagement can range from merely doing the required work and complying with rules to active participation in school activities. Emotional engagement can vary from disliking the school to valuing and identifying with it. Cognitive engagement ranges from memorization for a test to developing a love for the subject and increased knowledge and skill.⁷⁴ Moreover, engagement, like hope, is malleable; engagement levels can be changed with intervention.⁷⁵

A number of studies show positive relationships between student engagement and achievement outcomes. These studies include elementary, middle, and high schools at the student level.⁷⁶ Discipline problems and higher absenteeism are related to lower levels of engagement.⁷⁷ Moreover, lack of engagement and participation by students begins as early as first grade and can predict high school dropouts.⁷⁸ Engagement, as emotional commitment to the school, schoolwork, and a sense of belonging, is significantly related to academic achievement.⁷⁹

- **Wellbeing.** Wellbeing, how students think about and experience their lives, tells us how our students are

doing today and predicts their success in the future. Nobel Laureate Daniel Kahneman distinguishes between evaluative and experienced wellbeing. Evaluative wellbeing is the way people remember their experiences at a later time; experienced wellbeing focuses on emotions in the immediate moment.

Research is being done around the world on adult wellbeing. Recent studies indicate that wellbeing is important for success in school⁸⁰ and work.⁸¹ As with hope and engagement, wellbeing is malleable⁸² and can be fostered by focusing on students' thoughts and feelings.

The student's voice can be measured at the school level, for a school-wide measure for all staff members; it can also be measured at the individual teacher level. Assessing the levels of student hope, engagement, and wellbeing can give us a noncognitive measure that is a leading-edge indicator of student achievement. Schools that measure hope, engagement, and wellbeing have a diagnostic tool to identify strengths and weaknesses in the school climate that influence how students feel, behave, and learn in the school.

Partnerships With School and Community

Partnerships with key stakeholders provide a broader measure of the individual teacher's contribution to student and school success. These partnerships may be with other teachers, parents, administrators, and other school district or community stakeholders, and they focus on particular student or school goals and initiatives. These partnerships may be related to grade-level, departmental, or school-wide initiatives; participation in learning communities; or school or school district committees. Attendance, on-time promotion, and high school graduation, for example, have direct connections to short- and long-term student success and pose challenges that can only be met with partnerships across the school community.⁸³

Standards can be written to identify and measure the different ways and the extent of participation

that individual teachers contribute to the school and community. The principal is in the best position to evaluate the teacher's overall contribution to the school.

Peer Assistance and Review

Peer Assistance and Review (PAR) programs generate firm support from more than 70 school districts across the country currently using the program.⁸⁴ In part, PAR hasn't been adopted widely because it differs from the norm in what teachers and principals do, and the program requires high levels of collaboration and trust between the teachers' union and administrators. Most PAR programs have some common elements, including programs for beginning teachers and for experienced teachers who are facing issues in the classroom or low student achievement. A joint union-administration committee typically manages the program and selects expert teachers to serve as "consulting teachers" to support and assess beginning and struggling teachers.⁸⁵

In most situations, the responsibility for evaluation and development in PAR shifts from principals to consulting teachers. Consulting teachers conduct extensive observations, maintain comprehensive written reports of the teacher's progress, report teacher progress, and make employment recommendations to the PAR panel. They provide individualized assistance to meet the teacher's needs, such as assisting with lesson planning, conducting demonstration lessons, facilitating observations of other teachers, and giving specific feedback. For their work, consulting teachers receive stipends and release time.⁸⁶

Consulting teachers play a crucial role in teacher evaluation and development for beginning and struggling teachers. In most districts, PAR serves as the induction program for beginning teachers. The consulting teacher serves as an information source, mentor, and expert consultant as beginning teachers work through the first year of teaching. Formative and summative reports of the teacher's progress are presented to the PAR panel with employment recommendations. When an experienced teacher is

recommended for intervention, a consulting teacher assesses the teacher's situation and reports to the PAR panel and principal. If a consulting teacher is assigned, assistance is individualized to the particular problems.⁸⁷

If the performance of a beginning teacher or of an experienced teacher who is recommended for intervention does not improve, the consulting teacher initiates a review process. Intervention is a serious matter. If the consulting teacher does not believe the teacher is meeting the district's teaching standards, a recommendation is made to the PAR panel. A PAR panel review is based on the reports provided to the teacher by the consulting teacher. The PAR panel normally consists equally of union and administrative representatives, and the panel can recommend dismissal to the superintendent or continued assistance for the next school year.⁸⁸

A PAR strategy is expensive, but it has many benefits. Cost estimates range from \$4,000 to \$7,000 per teacher participant in the program. Districts reported, however, that PAR saved the district money eventually through lower turnover. By comparison, the study of five school districts cited previously indicated the total cost of turnover ranged from \$8,000 to more than \$13,000 per teacher. Additionally, ineffective teachers can be dismissed without excessive delay and extensive legal costs, including lawsuits. PAR was seen by district representatives as effective in improving instruction for students, serving as superior professional development for consulting teachers when they return to the classroom, and creating a culture that attracts, develops, and retains teachers.⁸⁹

PAR systems have been criticized for failing to dismiss teachers any more frequently than other evaluation systems. Two of the 12 schools involved in The New Teacher Project's 2009 study use PAR programs. Additionally, administrators, union representatives, and consulting teachers in some districts indicate that failing teachers are not reported to PAR for intervention.⁹⁰ With value-added scores for some teachers and evidence

of student achievement for other teachers, the need for intervention by PAR should be apparent to administrators.

At the same time, other districts cite very positive results. Jerry Weast, former superintendent in Montgomery County, Maryland, maintains that over an 11-year period, the PAR panels recommended 200 teachers for dismissal, and another 300 teachers left the district rather than participate in a PAR intervention. By comparison, in the 10 years preceding PAR, five teachers were dismissed in Montgomery County.⁹¹

The specific teaching standards to be used may include those adopted by the state or the school organization. General models like the Danielson framework, modifications to Danielson as in the Teacher Advancement Program, National Board for Professional Teaching Standards, Teacher Performance Assessment, the Marzano Evaluation Model, CLASS, any of the subject-specific models, or standards and a rubric developed by the school organization can all be used as the focus of — and the language to describe — what teachers should be able to do in their classrooms.

Regardless of the specific model adopted, the model's purpose as a developmental tool is to structure and provide a language for observations and consistent input for individual development plans and teacher improvement. Though the greater specificity of some of the models may be seen as desirable, it leads to highly prescriptive recommendations when these models are used as development tools. A more holistic observation model and rubric provides guidance but allows for individualized approaches to similar goals. In the end, it's not what we do to teachers that will change behaviors and increase student learning; it's what we do with and for teachers that will make the difference for students.

Principal Coaching

The principal or another school administrator serves in the critical role as a coach, providing direction, arranging

resources, and consulting with teachers. The principal is in a unique position to see the big picture of the school and know the individual teachers. By considering each teacher's strengths, weaknesses, and areas of contribution to the school community, a principal can guide the teacher in collaboratively designing an individualized development plan. Arranging and coordinating resources for teachers becomes an important coaching function because of the principal's understanding of resources available within the school and district. Principals can help teachers reflect on their classroom work and how to contribute more broadly to the school community.

As a coach, the principal's role moves from judging teachers' performance to assisting teachers in continuous improvement. Formative conferences provide checkpoints and opportunities for adjustment. After the teacher's individual development plan is approved, periodic principal and teacher conferences allow for progress checks and updates of the action plans and resource needs. While the teacher assumes responsibility for bringing evidence to the formative and summative meetings, the principal and a consulting teacher, as part of the PAR program, bring additional evidence to the conferences. In this way, the principal stays informed of the teacher's progress and a participant in the process.

Individual Development Plan

All teachers should have individual development plans that contain specific and meaningful growth goals with measurable outcomes and provide assistance in meeting those outcomes. The specific goals of the individual development plans must aim at meaningful student success, using specific standards developed by the school district or state. These development plans should include the teacher's strengths and weaknesses as measured by research-based assessments, the insights of a school administrator who directs and coordinates the plan's development, and support from a highly effective teacher in the content area with training as an observer and coach for beginning and struggling teachers. Individual development plans should

focus on self-awareness, reflection, and input and should be targeted to the specific needs of the teacher.

At the same time, we should be wary of the old adage that admonishes us to “hire good people and get out of their way.” Like their best students, effective and highly effective teachers are motivated learners and need to continue to grow. We should encourage and support their growth through plans — developed jointly with principals and content evaluators — that are based on the teacher’s unique talents and strengths. These teachers have demonstrated that they produce student growth. Our challenge is how to make them even better at what they do well. Self-awareness and application of this awareness is the first step.

The individual development plan incorporates assessments of the multiple measures as data for goal setting and action planning. In the absence of past measures, a first-year teacher could use the basic components of an induction program to form a portion of the developmental goals and action plans, while other portions would be individualized. Beginning teachers with more than one year’s experience, teachers whose students grew less than expected, and teachers with expected or greater-than-expected growth would use the previous year’s data and summative evaluation to formulate the current year’s plan. The final individualized plan for the current year should include developmental goals and action plans that identify activities and resources needed that are jointly developed by the teacher and the administrator.

Periodic, formative assessments of the plan’s goals and an annual summative assessment should inform the completion level for each of the goals. The teacher assumes the responsibility for bringing evidence to the formative and summative meetings with the principal and a teacher who is an expert in the content field. With evidence contributed by the teacher being evaluated, the expert teacher and the principal contribute to a summative review of the development plan. This ranking would then be incorporated with the other Teacher Performance Index

metrics to create a final ranking for the appraisal.

Incorporate Strengths. The individual development plan includes feedback and a growth map, but part of the growth can occur through awareness of the individual teacher’s strengths and the strengths of others. The key to improving performance through evaluation is to avoid the weakness trap. Our society generally, and much of education particularly, seems intent on pointing out weaknesses. All too quickly, we see differences in people as weaknesses. Behaviors that have a negative impact on performance can’t be ignored, but an incessant drumbeat of weakness and remediation doesn’t lead to excellence. Weakness-fixing may take a teacher from -10 to -4; a strengths-based approach can take a teacher from +10 to +40.⁹² The goal is to help teachers manage their weaknesses and build their strengths.

As part of evaluation as growth, individual teachers must become aware of their strengths. The adage that “you can do or be anything you want” is pervasive and wrong. It suggests that we are all “blank slates” with identical potential, differentiated only by how hard we try. People are not completely moldable through conditioning. In fact, current brain research suggests that each of us carries unique potential when we are born — potential that is activated and developed by the opportunities we have and the choices we make throughout life.⁹³

Four elements are crucial to developing teacher strengths. The first element, *talent*, is “a natural way of thinking, feeling, or behaving.” Talents are the dispositions that come naturally and make an individual teacher unique from others despite similar education, experience, or professional development. Teaching talents include a passion for teaching, developing solid relationships, and working with individual student differences. *Knowledge* comprises two components: what you know (subject matter content) and experiential learning. *Skills* are the basic abilities (stating the objective of the lesson, making a PowerPoint). *Practice* enables teachers with the talent, knowledge, and skills

to fine-tune their work with students. Strength can be understood as consistent excellence in performance.⁹⁴

Combining talents with knowledge, skills, and practice results in a multiplier of the elements of great teaching. Examples of consistent and near-perfect performance include the ability of a teacher to win over a difficult group of students, detect the patterns in student achievement data, match a learning activity to a struggling student, or persevere after a particularly trying day.

So how do we identify individual teachers' talents? That's more difficult than it sounds. Though talents come naturally to us, our recognition of them often doesn't. Because our greatest talents are so instinctual to us, we often assume they're commonplace. Whether it's an innate ease in learning mathematics or the inherent need to take action when confronted with a problem, our greatest talents are what we do naturally; they are how we react and behave when we are at our best. Others may marvel at our ease in an area of talent, but we often fail to recognize how our own talents contribute to high performance. The areas of our lives in which we apply those talents tend to go smoothly; thus, we spend little time consciously thinking about them. Fortunately, there are research-driven psychometric tools to help identify talents.⁹⁵

However, identifying individual talents is only the first step in consciously using them. Feedback, instruction, reflection, and mentoring activities move development from a one-time event to continuing growth. They also enable teachers to intentionally apply their talents to achieve higher performance in their work. An understanding of talents launches an ongoing discussion with colleagues, mentors, supervisors, friends, and family that focuses on what is right with the teacher.⁹⁶

Launching this discussion in an organization yields positive results. Using strengths as a common language to talk about individual differences in a positive way contributes to a more constructive and productive working and learning climate. In mentor programs, discussions

about talents and strengths provide insights for teachers and their mentors that build stronger relationships. Once teachers know the talents of team members, they begin to have clearer understandings and deeper appreciation of their colleagues, including those talents that may not seem positive initially. With an appreciation of talents and strengths, administrators develop a clearer understanding of teachers and how they can make greater contributions in teacher teams.

The importance of understanding talents and developing strengths is underscored by a series of studies. A Gallup poll asked adults from across the United States whether their manager primarily focused on their weaknesses, focused on their strengths, or ignored them. Gallup also asked a series of items that measure the respondents' level of engagement at work. Engaged employees work with passion and feel a profound connection to their company. Actively disengaged employees, in contrast, tend to be negative, actively spreading their negativity and posing obstacles to change. The results are displayed in Table 1.

When we look at the relationship between a manager's focus and employee engagement, we see that among U.S. employees who indicated their manager primarily focused on their weaknesses, 22% were actively disengaged. Strikingly, only 1% of U.S. employees who indicated that their manager primarily focused on their strengths were actively disengaged. For employees, a manager who provides no feedback at all is far worse than one who focuses on weaknesses; 40% of U.S. employees who said that managers primarily ignored them were actively disengaged.⁹⁷

Table 1

The Manager's Focus and Its Effect on Active Disengagement	
If your manager primarily:	Your chances of being actively disengaged are:
Ignores you	40%
Focuses on your weaknesses	22%
Focuses on your strengths	1%

Studies of strengths development found strong relationships to engagement. A review of 896 workgroups found that workgroups whose managers received some strengths feedback increased the groups' engagement level significantly compared with workgroups whose managers did not receive strengths feedback. This is noteworthy because only the managers received the strengths intervention — not the employees — yet engagement increased for the workgroup overall. At the individual level, data for 12,157 employees showed that among employees who received strengths feedback, engagement improved significantly compared with employees who did not receive feedback.⁹⁸

Strengths programs also assist in reducing turnover. A study of turnover data for 65,672 employees shows that among employees who received some strengths feedback, turnover rates were 14.9% lower than for employees not participating in a strengths program.⁹⁹

Two other studies looked at productivity after a strengths intervention. Of the 530 workgroups studied, workgroups whose managers participated in the strengths program demonstrated 12.5% greater productivity after the program when compared to those workgroups whose managers did not participate. In examining the productivity data for 1,874 employees, researchers found that the productivity of employees receiving the strengths intervention increased 7.8% as compared with employees who did not receive the strengths intervention.¹⁰⁰

Successful strengths programs that achieve results like those described here are not “feel good” or “fluff” initiatives. Private-sector companies that embrace a strengths-based approach to development seek a solid return on investment, and so should school organizations. A study in a large business organization that had invested in strengths over time provides insights. A group of components explained most of the performance range across the 606 business units and seemed crucial to the success of the strengths program:

- employees knew their strengths and the strengths of colleagues;
- employees intentionally applied their strengths and created feelings of success;
- employees saw a commitment to strengths by coworkers and managers; and
- employees sensed a shared commitment to strengths from the leadership of the company.¹⁰¹

The payoff of strengths as part of individualized development can be considerable in measurable ways, as the studies show. In addition to improving performance measures, a strengths-based program results in other positive aspects for the organization. The benefits of strengths-based development:

- apply to all individuals in the organization regardless of their position
- provide a common language that helps employees understand the ways in which each person most naturally thinks, feels, and behaves
- promote growth and top performance based on what each employee naturally does best
- provide a common language and conceptual framework through which individuals and teams can more fully develop, perform, and partner
- cultivate a positive culture focused on growth¹⁰²

Next Steps

The focus on teacher effectiveness and teacher evaluation is encouraging because both have been neglected for far too long. Measurement does improve performance, but care should be taken in what is measured. Unfortunately, the rush to adopt teacher frameworks for evaluation accepts teacher observation as one of sometimes just two measures of teacher effectiveness. Instead, a combination of measures

provides more objective measurement and a focus on the development of teachers.

Success, as measured by student growth, becomes more likely if expectations and outcomes are clear and teachers have autonomy, when earned, in how to achieve their goals and an opportunity to attain mastery. Multiple measures provide a more complete picture of how teachers influence student learning and development. The metrics can do much of the difficult work of differentiating teacher performance, enabling principals and expert teachers to concentrate on growing teachers' capacity.

If we begin with the belief that individual teachers comprise the biggest influence schools have on student achievement, successful teachers deserve an appraisal system that is based on their needs, aimed at developing mastery, and building strengths. New teacher appraisal systems are an important step forward. But new systems for recruiting, selecting, and compensating teachers are also needed.

Notes

- 1 Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67.
- 2 Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center. Retrieved December 20, 2007, from <http://www.mccsc.edu/~curriculum/cumulative%20and%20residual%20effects%20of%20teachers.pdf>
- 3 Jordan, H. R., Mendro, R. L., & Weerasinghe, D. (1997, July). *Teacher effects on longitudinal student achievement: A report on research in progress*. Dallas, TX: Dallas Public Schools. Retrieved September 30, 2001, from <http://www.dallasisd.org/cms/lib/TX01001475/Centricity/Shared/evalacct/research/articles/Jordan-Teacher-Effects-on-Longitudinal-Student-Achievement-1997.pdf>
- 4 Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.
- 5 Jones, J. (2009, August 25). Public says better teachers are key to improved education. Retrieved September 1, 2009, from <http://www.gallup.com/poll/122504/Public-Says-Better-Teachers-Key-Improved-Education.aspx>
- 6 U.S. Department of Education (2010, September 23) U.S. Department of Education announces \$442 million in Teacher Incentive Fund grants. Press release. Retrieved November 1, 2010, from <http://www.ed.gov/news/press-releases/department-education-announces-442-million-teacher-quality-grants-62-winners-27>
- 7 Bill & Melinda Gates Foundation. (No date). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Retrieved June 1, 2011, from <http://www.gatesfoundation.org/Programs/Educational-Improvement/Measures-of-Effective-Teaching.aspx>

http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf

5 Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Issues & Answers Report, REL 2007-No. 030). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved June 15, 2011, from <http://ies.ed.gov/ncee/edlabs>

6 Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Issues & Answers Report, REL 2007-No. 030). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved June 15, 2011, from <http://ies.ed.gov/ncee/edlabs>

7 Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Issues & Answers Report, REL 2007-No. 030). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved June 15, 2011, from <http://ies.ed.gov/ncee/edlabs>

8 Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on teacher effectiveness*. New York: The New Teacher Project. Retrieved June 25, 2010, from http://tntp.org/files/TheWidgetEffect_2nd_ed.pdf

9 Taylor, E. S. & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student*

achievement data of mid-career teachers (NBER Working Paper No. 16877). Cambridge, MA: National Bureau of Economic Research. Retrieved July 10, 2011, from <http://www.nber.org/papers/w16877>

10 Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on teacher effectiveness*. New York: The New Teacher Project. Retrieved June 25, 2010, from http://tntp.org/files/TheWidgetEffect_2nd_ed.pdf

11 The New Teacher Project. (2007). *Hiring, assignment, and transfer in Chicago Public Schools*. New York: Author. Retrieved June 1, 2011, from <http://tntp.org/files/TNTPAnalysis-Chicago.pdf>

12 The New Teacher Project. (2007). *Hiring, assignment, and transfer in Chicago Public Schools*. New York: Author. Retrieved June 1, 2011, from <http://tntp.org/files/TNTPAnalysis-Chicago.pdf>

13 Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on teacher effectiveness*. New York: The New Teacher Project. Retrieved June 25, 2010, from http://tntp.org/files/TheWidgetEffect_2nd_ed.pdf

14 The New Teacher Project. (2009). *Human capital reform in Cincinnati Public Schools: Strengthening teacher effectiveness and support*. New York: The New Teacher Project. Retrieved June 22, 2011, from http://tntp.org/files/TNTP_Cincinnati_Report_Doc09.pdf

15 Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on teacher effectiveness*. New York: The New Teacher Project. Retrieved June 25, 2010, from http://tntp.org/files/TheWidgetEffect_2nd_ed.pdf

- 16 U.S. Department of Education. (2009). *Race to the top program: Executive summary*. Washington, D.C.: Author. Retrieved April 15, 2010, from <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>
- 17 Bill & Melinda Gates Foundation. (No date). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Retrieved June 1, 2011, from http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf
- 18 Additional examples of teacher observation systems include:
- Marzano's Teacher Evaluation Framework, <http://www.marzanevaluation.com/>
 - Pianta's Classroom Assessment Scoring System, <http://www.teachstone.org/about-the-class/>
 - Other observation systems are designed for specific subject areas, including mathematics, science, and language arts.
- 19 Danielson, C. (1996). *Enhancing professional practice: A framework for teaching* (2nd ed). Alexandria, VA: Association for Supervision and Curriculum Development.
- 20 Milanowski, A. T., Kimball, S. M., & White, B. (2004). *The relationship between standards-based teacher evaluation scores and student achievement: replication and extensions at three sites* (CPRE-UW Working Paper Series TC-04-0). Madison, WI: Consortium for Policy Research in Education. Retrieved May 1, 2011, from http://www.cprewisconsin.com/papers/3site_long_TE_SA_AERA04TE.pdf
- 21 Taylor, E. S. & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (NBER Working Paper No. 16877). Cambridge, MA: National Bureau of Economic Research. Retrieved July 10, 2011, from <http://www.nber.org/papers/w16877>
- 22 Taylor, E. S. & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (NBER Working Paper No. 16877). Cambridge, MA: National Bureau of Economic Research. Retrieved July 10, 2011, from <http://www.nber.org/papers/w16877>
- 23 Jerald, C.D., & Van Hook, K. (2011). *More than measurement: The TAP system's lessons learned for designing better teacher evaluation systems*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved June 17, 2011, from http://www.tapsystem.org/publications/eval_lessons.pdf
- 24 Toch, T. (2008, October). Fixing teacher evaluation. *Educational Leadership*, 66(2), 32-37.
- 25 Jerald, C.D., & Van Hook, K. (2011). *More than measurement: The TAP system's lessons learned for designing better teacher evaluation systems*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved June 17, 2011, from http://www.tapsystem.org/publications/eval_lessons.pdf
- 26 Jerald, C.D., & Van Hook, K. (2011). *More than measurement: The TAP system's lessons learned for designing better teacher evaluation systems*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved June 17, 2011, from http://www.tapsystem.org/publications/eval_lessons.pdf
- 27 Schacter, J., Thum, Y. M., Reifsneider, D., & Schiff, T. (2004). *The teacher advancement program report two: Year three results from Arizona and year one results from South Carolina TAP schools*. Santa Monica, CA: Milken Family Foundation. Retrieved June 15, 2011, from http://www.tapsystem.org/pubs/tap_results_azsc2004.pdf
- 28 Springer, M. G., Ballou, D., & Peng, A. (2008). *Impact of the Teacher Advancement Program on student test score gains: Findings from an independent appraisal*. Nashville, TN: National Center on Performance Incentives. Retrieved June 15, 2011, from <http://>

- www.performanceincentives.org/data/files/news/PapersNews/Springer_et_al_2008.pdf
- 29 Solmon, L.D., White, J.T., Cohen, D., & Woo, D. (2007). *The effectiveness of the Teacher Advancement Program*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved June 15, 2011, from http://www.tapsystem.org/pubs/effective_tap07_full.pdf
- 30 Springer, M. G., Ballou, D., & Peng, A. (2008). *Impact of the Teacher Advancement Program on student test score gains: Findings from an independent appraisal*. Nashville, TN: National Center on Performance Incentives. Retrieved June 15, 2011, from http://www.performanceincentives.org/data/files/news/PapersNews/Springer_et_al_2008.pdf
- 31 Daley, G., & Kim, L. (2010). *A teacher evaluation system that works*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved June 15, 2011, from http://www.tapsystem.org/publications/wp_eval.pdf
- 32 Toch, T. (2008, October). Fixing teacher evaluation. *Educational Leadership*, 66(2), 32-37.
- 33 Pearson. (2011, March 17). Stanford University and Pearson collaborate to deliver the Teacher Performance Assessment (TPA). Press Release. Retrieved May 24, 2011, from <http://www.pearsoned.com/2011/03/17/stanford-university-and-pearsn-collaborate-to-deliver-the-teacher-performance-assessment-tpa/>
- 34 American Association of Colleges of Teacher Education. (2011, March). Teacher performance assessment consortium. Retrieved May 24, 2011, from <http://www.aacte.org/index.php?/Programs/Teacher-Performance-Assessment-Consortium-TPAC/teacher-performance-assessment-consortium.html>
- 35 Performance Assessment for California Teachers. (No date.) Retrieved May 24, 2011, from http://www.pacttpa.org/_main/hub.php?pageName=Supporting_Documents_for_Candidates#Rubrics
- 36 American Association of Colleges of Teacher Education. (2011, March). Teacher performance assessment consortium. Retrieved May 24, 2011, from <http://www.aacte.org/index.php?/Programs/Teacher-Performance-Assessment-Consortium-TPAC/teacher-performance-assessment-consortium.html>
- 37 Pearson. (2011, March 17). Stanford University and Pearson collaborate to deliver the Teacher Performance Assessment (TPA). Press Release. Retrieved May 24, 2011, from <http://www.pearsoned.com/2011/03/17/stanford-university-and-pearsn-collaborate-to-deliver-the-teacher-performance-assessment-tpa/>
- 38 Pecheone, R. L. & Chung Wei, R. R. (2007). PACT technical report: Summary of validity and reliability studies for the 2003-2004 pilot year. Stanford, CA: PACT Consortium. Retrieved May 24, 2011, from http://www.pacttpa.org/_files/Publications_and_Presentations/PACT_Technical_Report_March07.pdf
- 39 District of Columbia Public Schools. (2010a). IMPACT: The District of Columbia Public Schools effectiveness assessment system for school-based personnel 2010-2011. Group 1 Guidebook. Washington, D.C.: Author. Retrieved November 15, 2011, from [http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+\(Performance+Assessment\)/IMPACT+Guidebooks](http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/IMPACT+Guidebooks)
- 40 Curtis, R. (2011). *District of Columbia Public Schools: Defining instructional expectations and aligning accountability and support*. Washington, D.C.: The Aspen Institute. Retrieved July 1, 2011, from http://www.aspeninstitute.org/sites/default/files/content/docs/education%20and%20society%20program/AI_DCPS_teacher%20evaluation.pdf
- 41 District of Columbia Public Schools. (2010b). IMPACT: The District of Columbia Public Schools effectiveness assessment system for school-based personnel 2010–2011. Group 2 Guidebook.

- Washington, D.C.: Author. Retrieved November 15, 2011, from [http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+\(Performance+Assessment\)/IMPACT+Guidebooks](http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/IMPACT+Guidebooks)
- 42 Headden, S. (2011). *Inside IMPACT: D.C.'s model teacher evaluation system*. Washington, D.C.: Education Sector. Retrieved July 21, 2011, from <http://www.educationsector.org/publications/inside-impact-dcs-model-teacher-evaluation-system>
- 43 District of Columbia Public Schools. (2010a). IMPACT: The District of Columbia Public Schools effectiveness assessment system for school-based personnel 2010-2011. Group 1 Guidebook. Washington, D.C.: Author. Retrieved November 15, 2011, from [http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+\(Performance+Assessment\)/IMPACT+Guidebooks](http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/IMPACT+Guidebooks)
- 44 District of Columbia Public Schools. (2010a). IMPACT: The District of Columbia Public Schools effectiveness assessment system for school-based personnel 2010-2011. Group 1 Guidebook. Washington, D.C.: Author. Retrieved November 15, 2011, from [http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+\(Performance+Assessment\)/IMPACT+Guidebooks](http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/IMPACT+Guidebooks)
- 45 Curtis, R. (2011). *District of Columbia Public Schools: Defining instructional expectations and aligning accountability and support*. Washington, D.C.: The Aspen Institute. Retrieved July 1, 2011, from http://www.aspeninstitute.org/sites/default/files/content/docs/education%20and%20society%20program/AI_DCPS_teacher%20evaluation.pdf
- 46 Headden, S. (2011). *Inside IMPACT: D.C.'s model teacher evaluation system*. Washington, D.C.: Education Sector. Retrieved July 21, 2011, from <http://www.educationsector.org/publications/inside-impact-dcs-model-teacher-evaluation-system>
- 47 Headden, S. (2011). *Inside IMPACT: D.C.'s model teacher evaluation system*. Washington, D.C.: Education Sector. Retrieved July 21, 2011, from <http://www.educationsector.org/publications/inside-impact-dcs-model-teacher-evaluation-system>
- 48 Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, D.C.: National Comprehensive Center for Teacher Quality. Retrieved January 15, 2010, from <http://www.tqsource.org/publications/teacherEffectiveness.php>
- 49 Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, D.C.: National Comprehensive Center for Teacher Quality. Retrieved January 15, 2010, from <http://www.tqsource.org/publications/teacherEffectiveness.php>
- 50 Pink, D. H. (2009). *Drive: The surprising truth about what motivates us*. New York: Riverhead Books.
- 51 Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The market for teacher quality*. (NBER Working Paper No. w11154). Cambridge, MA; National Bureau for Economic Research. Retrieved June 1, 2010, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=669453
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2), 247-252.
- 52 National Commission on Teaching and America's Future. (2003, January). *No dream denied: A pledge to America's children*. Washington, D.C.: Author. Retrieved March 24, 2006, from http://www.nctaf.org/documents/no-dream-denied_summary_report.pdf

- See Bridgeland, J. M., DiJulio Jr., J. J., & Morison, K. B. (2006). *The silent epidemic: Perspectives of high school dropouts*. Washington, D.C.: Civic Enterprises, LLC. Retrieved June 10, 2007, from <http://www.civicenterprises.net/pdfs/thesilentepidemic3-06.pdf>
- Hammond, C., Linton, D., Smink, J., & Drew, S. (2007). *. Clemson, SC: National Dropout Prevention Center/Network, and Communities In Schools, Inc. Retrieved June 12, 2010, from http://www.dropoutprevention.org/sites/default/files/uploads/major_reports/DropoutRiskFactorsandExemplaryProgramsFINAL5-16-07.pdf*
- Epstein, J. L., & Sheldon, S. B. (2002). Present and accounted for: Improving student attendance through family and community involvement. *Journal of Educational Research*, 95(5), 308-319.
- Alexander, K. L., Entwistle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70(2), 87-107.
- 53 Barnes, G., Crowe, E., & Schaefer, B. (2007). *The cost of teacher turnover in five school districts: A pilot study*. Washington, D.C.: National Commission on Teaching and America's Future. Retrieved July 15, 2011, from http://www.nctaf.org/resources/demonstration_projects/turnover/documents/CTTFullReportfinal.pdf
- 54 Wagner, R. & Harter, J. K. (2006). *12: The elements of great managing*. New York: Gallup Press.
- Harter, J. K., Schmidt, F. L., Killham, E. A., & Asplund, J. W. (2006). *Q¹² meta-analysis*. Omaha, NE: The Gallup Organization.
- 55 Buckingham, M., & Coffman, C. (1999). *First, break all the rules: What the world's greatest managers do differently*. New York: Simon & Schuster.
- 56 Bushaw, W. & Lopez, S. J. (2011, September). Betting on teachers: the 43rd annual Phi Delta Kappa/Gallup Poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 93(1), 8-26. Retrieved September 23, 2011, from http://www.pdkintl.org/poll/docs/pdkpoll43_2011.pdf
- 57 Seashore Louis, K., Leithwood, K., Wahlstrom, K. L., Anderson, S. E. (2010). *Learning from leadership: Investigating the links to improved student learning: Final report of research findings*. Minneapolis, MN: University of Minnesota . Retrieved December 10, 2010, from <http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/Pages/Investigating-the-Links-to-Improved-Student-Learning.aspx>
- 58 Battelle for Kids. (2011). *Selecting growth measures: A guide for education leaders*. Columbus, Ohio: Author. Retrieved July 28, 2011, from <http://www.edgrowthmeasures.org/home/home.html>
- 59 Battelle for Kids. (2011). *Selecting growth measures: A guide for education leaders*. Columbus, Ohio: Author. Retrieved July 28, 2011, from <http://www.edgrowthmeasures.org/home/home.html>
- 60 Goe, L. (2008). *Key issue: Using value-added models to identify and support highly effective teachers*. Washington, D.C.: National Comprehensive Center for Teacher Quality. Retrieved January 15, 2011, from <http://www2.tqsource.org/strategies/het/UsingValueAddedModels.pdf>
- 61 Goe, L. (2008). *Key issue: Using value-added models to identify and support highly effective teachers*. Washington, D.C.: National Comprehensive Center for Teacher Quality. Retrieved January 15, 2011, from <http://www2.tqsource.org/strategies/het/UsingValueAddedModels.pdf>
- 62 For critiques of value-added measures in teacher evaluation, see the following:

- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., & and Shepard, L.A. (2010) *Problems with the use of student test scores to evaluate teachers* (Briefing Paper #278). Washington, D.C.: Economic Policy Institute. Retrieved September 2, 2010, from <http://www.epi.org/publication/bp278/>
- Goldhaber, D., & Hansen, M. (2008). *Assessing the potential of using value-added estimates of teacher job performance for making tenure decisions*. Seattle, WA: Center on Reinventing Public Education. Retrieved April 27, 2009, from http://www.crpe.org/cs/crpe/download/csr_files/brief_crpe_badclass_nov08.pdf
- McCaffrey, D. F., Sass, T. R., & Lockwood, J. R. (2008, June 27). *The intertemporal stability of teacher effect estimates*. Retrieved January 15, 2011, from http://www.wcer.wisc.edu/news/events/VAM%20Conference%20Final%20Papers/IntertemporalStability_McCaffreySassLockwood.pdf
- McCaffrey, D.F., Koretz, D., Lockwood, J.R., & Hamilton, L.S. (2004). *The promise and peril of using value-added modeling to measure teacher effectiveness* (Research Brief No. RB-9050-EDU). Santa Monica, CA: RAND Corporation. Retrieved January 15, 2011, from http://www.rand.org/pubs/research_briefs/2005/RAND_RB9050.pdf
- Goe, L. (2008). *Key issue: Using value-added models to identify and support highly effective teachers*. Washington, D.C.: National Comprehensive Center for Teacher Quality. Retrieved January 15, 2011, from <http://www2.tqsource.org/strategies/het/UsingValueAddedModels.pdf>
- Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service. Retrieved January 15, 2011, from <http://www.ets.org/Media/Research/pdf/PICVAM.pdf>
- 63 Bill & Melinda Gates Foundation. (No date). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Retrieved June 1, 2011, from http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf
- 64 Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., Whitehurst, G. (2010, November). *Evaluating teachers: The important role of value-added*. Washington, D.C.: The Brookings Institution. Retrieved August 15, 2011, from http://www.brookings.edu/~/media/Files/rc/reports/2010/1117_evaluating_teachers/1117_evaluating_teachers.pdf
- 65 Partnership for Assessment of Readiness for College and Careers (2011). PARCC content frameworks. Retrieved August 5, 2011, from <http://www.parcconline.org/parcc-content-frameworks>
- 66 Bridgeland, J. M., DiJulio Jr., J. J., & Morison, K. B. (2006). *The silent epidemic: Perspectives of high school dropouts*. Washington, D.C.: Civic Enterprises, LLC. Retrieved June 10, 2007, from <http://www.civicenterprises.net/pdfs/thesilentepidemic3-06.pdf>
- 67 Snyder, C. R. (1994). *The psychology of hope: You can get there from here*. New York: Free Press.
- 68 Marques, S. C., Pais-Ribiero, J.L., & Lopez, S. J. (2009). Validation of a Portuguese version of the Children's Hope Scale. *School Psychology International*, 30(5), 538-551.
- 69 Snyder, C. R., Hoza, B., Pelham, W. E., Rapoff, M., Ware, L., Danovsky, M., Highberger, L., Rubinstein, H., & Stahl, K. J. (1997). The development and validation of the Children's Hope Scale. *Journal of Pediatric Psychology*, 22(3), 399-421.
- 70 Gallup. (2009). [Hope, engagement, and wellbeing as predictors of attendance, credits earned, and GPA in high school freshmen.] Unpublished raw data. Omaha, NE.

- Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L., Bibb, J., Langelle, C., & Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, 60 (4), 570-585.
- Worrell, F. C., & Hale, R. L., (2001). The relationship of hope in the future and perceived school climate to school completion. *School Psychology Quarterly*, 16(4), 370-388.
- 71 Gallagher, M. W., & Lopez, S. J. (2008, August). Hope, self-efficacy, and academic success in college students. Poster session presented at the annual convention of the American Psychological Association, Boston, MA.
- Snyder, C. R., McDermott, D., Cook, W., & Rapoff, M. (2002). *Hope for the journey* (revised ed.). Clinton Corners, NY: Percheron Press.
- 72 Gallagher, M. W., & Lopez, S. J. (2008, August). Hope, self-efficacy, and academic success in college students. Poster session presented at the annual convention of the American Psychological Association, Boston, MA.
- Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L., Bibb, J., Langelle, C., & Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, 60 (4), 570-585.
- Snyder, C. R., McDermott, D., Cook, W., & Rapoff, M. (2002). *Hope for the journey* (revised ed.). Clinton Corners, NY: Percheron Press.
- 73 Gallagher, M. W., & Lopez, S. J. (2008, August). Hope, self-efficacy, and academic success in college students. Poster session presented at the annual convention of the American Psychological Association, Boston, MA.
- Snyder, C. R., McDermott, D., Cook, W., & Rapoff, M. (2002). *Hope for the journey* (revised ed.). Clinton Corners, NY: Percheron Press.
- 74 Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109.
- 75 Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, 82(2), 221-234.
- 76 Connell, J. P., Spencer, M. B., & Aber, J. L. (1994). Educational risk and resilience in African-American youth: Context, self, action, and outcomes in school. *Child Development*, 65(2 special number), 493-506.
- Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, 47(3), 633-662.
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in elementary, middle, and high school years. *American Educational Research Journal*, 37(1), 153-184.
- Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, 82(2), 221-234.
- Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology*, 82(1), 22-32.
- 77 Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, 82(2), 221-234.

- Alexander, K. L., Entwistle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70(2), 87-107.
- 78 Alexander, K. L., Entwistle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70(2), 87-107.
- 79 Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, 47(3), 633-662.
- Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement: The influence of school- and individual-level factors on academic achievement. *Education and Urban Society*, 40(1), 179-204.
- Wilms, J. D. (2003). *Student engagement at school: A sense of belonging and participation*. Paris: Organisation for Economic Co-operation and Development. Retrieved December 27, 2010, from <http://www.oecd.org/dataoecd/42/35/33689437.pdf>
- 80 Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin*, 131(6), 803-855.
- 81 Boehm, J. K., & Lyubomirsky, S. (2008). Does happiness promote career success? *Journal of Career Assessment*, 16(1), 101-116.
- Judge, T. A., & Hurst, C. (2008). How the rich (and happy) get richer (and happier): Relationship of core self-evaluations to trajectories in attaining work success. *Journal of Applied Psychology*, 83(4), 849-863.
- 82 Sin, N. L., & Lyubomirsky, S. (2009). Enhancing wellbeing and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *Journal of Clinical Psychology*, 65, 467-487.
- Suldo, S. M., Huebner, E. S., Michalowski, J., & Thalji, A. (in press). Promoting subjective well-being. In M. Bray & T. Kehle (Eds) *Oxford handbook of school psychology*. New York: Oxford University Press
- 83 Bridgeland, J. M., Dilulio Jr., J. J., & Morison, K. B. (2006). *The silent epidemic: Perspectives of high school dropouts*. Washington, D.C.: Civic Enterprises, LLC. Retrieved June 10, 2007, from <http://www.civicenterprises.net/pdfs/thesilentepidemic3-06.pdf>
- 84 National Comprehensive Center for Teacher Quality. (No date). Guide to teacher evaluation products. Retrieved July 12, 2011, from <http://www3.learningpt.org/tqsource/gep/GEPTool.aspx?gid=23>
- 85 Harvard Graduate School of Education Project on the Next Generation of Teachers. (2008). *A user's guide to peer assistance and review*. Cambridge, MA: Author. Retrieved July 15, 2011, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf
- 86 Harvard Graduate School of Education Project on the Next Generation of Teachers. (2008). *A user's guide to peer assistance and review*. Cambridge, MA: Author. Retrieved July 15, 2011, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf
- 87 Harvard Graduate School of Education Project on the Next Generation of Teachers. (2008). *A user's guide to peer assistance and review*. Cambridge, MA: Author. Retrieved July 15, 2011, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf
- 88 Harvard Graduate School of Education Project on the Next Generation of Teachers. (2008). *A user's guide to peer assistance and review*. Cambridge, MA: Author. Retrieved July 15, 2011, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf
- 89 Harvard Graduate School of Education Project on the Next Generation of Teachers. (2008). *A user's guide to peer assistance and review*. Cambridge, MA: Author.

- Retrieved July 15, 2011, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf
- 90 Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on teacher effectiveness*. New York: The New Teacher Project. Retrieved June 25, 2010, from http://tntp.org/files/TheWidgetEffect_2nd_ed.pdf
- Harvard Graduate School of Education Project on the Next Generation of Teachers. (2008). *A user's guide to peer assistance and review*. Cambridge, MA: Author. Retrieved July 15, 2011, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf
- 91 Winerip, M. (2011, June 5). Helping teachers help themselves. *New York Times*. Retrieved July 31, 2011, from http://www.nytimes.com/2011/06/06/education/06oneducation.html?_r=1
- 92 Clifton, D. O., & Harter, J. K. (2003). Investing in strengths. In K. S. Cameron, J. E. Dutton, & R. E. Quinn (Eds.), *Positive Organizational Scholarship: Foundations of a New Discipline* (pp. 111-121). San Francisco, CA: Berrett-Koehler Publishers, Inc.
- 93 Ridley, M. (2003). *Nature via Nurture: Genes, experience, & what makes us human*. New York: HarperCollins.
- 94 Rath, T. (2007). *StrengthsFinder 2.0*. New York: Gallup Press.
- 95 Asplund, J., Lopez, S. J., Hodges, T., & Harter, J. (2007). *The Clifton StrengthsFinder 2.0 technical report: Development and validation*. Princeton, NJ: Gallup.
- 96 Hodges, T. D., & Asplund, J. (2009). Strengths development in the workplace. In Linley, P. A., Harrington, S., & Garcea, N. (Eds.). *Oxford handbook of positive psychology and work*. New York: Oxford University Press.
- 97 Rath, T. (2007). *StrengthsFinder 2.0*. New York: Gallup Press.
- 98 Asplund, J., Lopez, S. J., Hodges, T., & Harter, J. (2007). *The Clifton StrengthsFinder 2.0 technical report: Development and validation*. Princeton, NJ: Gallup.
- 99 Hodges, T. D., & Asplund, J. (2009). Strengths development in the workplace. In Linley, P. A., Harrington, S., & Garcea, N. (Eds.). *Oxford handbook of positive psychology and work*. New York: Oxford University Press.
- 100 Asplund, J., Lopez, S. J., Hodges, T., & Harter, J. (2007). *The Clifton StrengthsFinder 2.0 technical report: Development and validation*. Princeton, NJ: Gallup.
- 101 Asplund, J., & Blacksmith, N. (2011, August). Making strengths-based development work: Effective implementation and support are vital to a program's success. *Gallup Management Journal*. Retrieved August 11, 2011, from <http://gmj.gallup.com/content/148691/Making-Strengths-Based-Development-Work.aspx>
- 102 Asplund, J., & Blacksmith, N. (2011, August). Making strengths-based development work: Effective implementation and support are vital to a program's success. *Gallup Management Journal*. Retrieved August 11, 2011, from <http://gmj.gallup.com/content/148691/Making-Strengths-Based-Development-Work.aspx>

GALLUP®

World Headquarters

The Gallup Building
901 F Street, NW
Washington, D.C. 20004

t +1.877.242.5587
f +1.202.715.3045

www.gallup.com