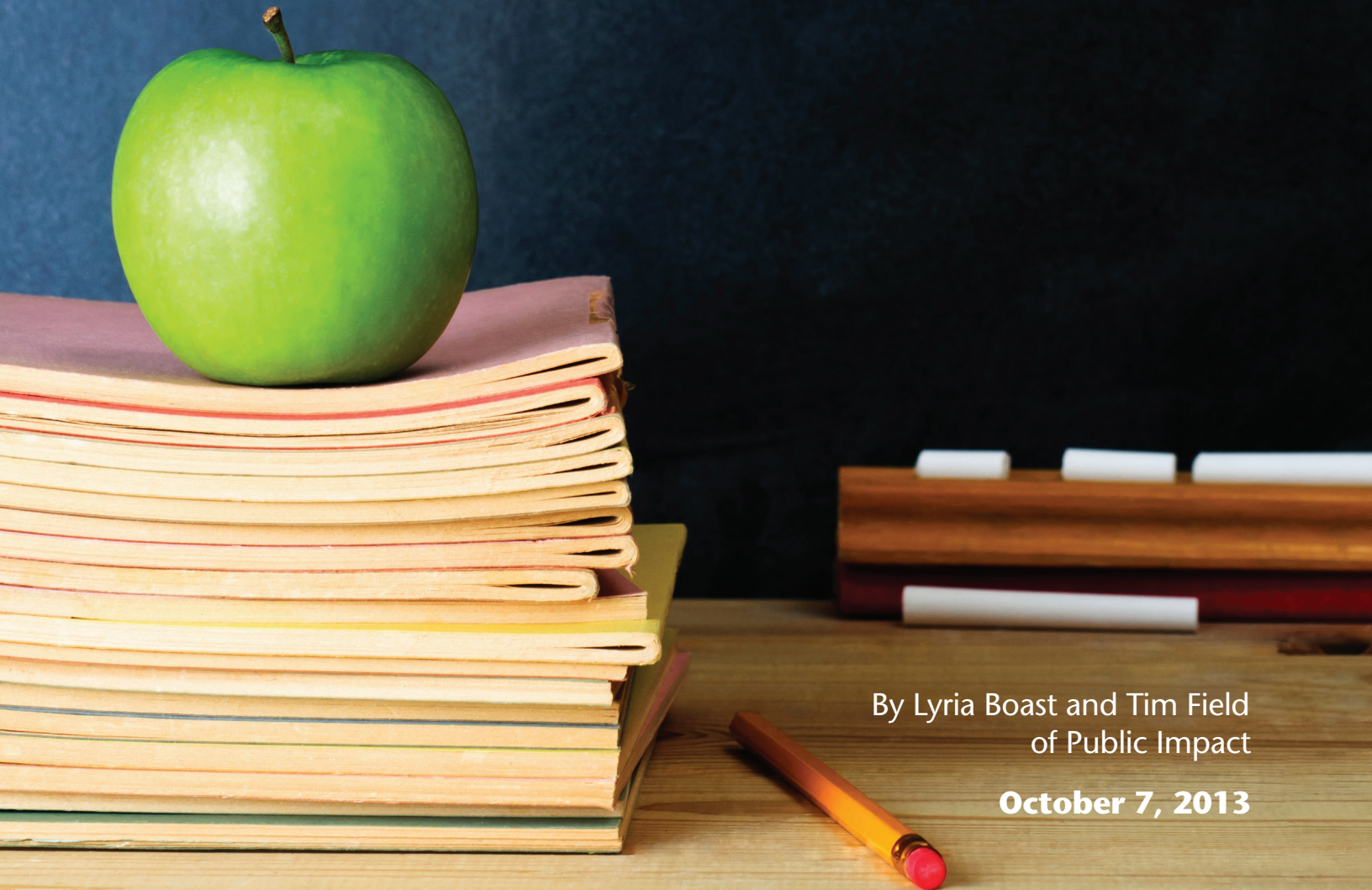National Alliance for Public Charter Schools

QUALITY SCHOOL RATINGS:
Trends in Evaluating School Academic Quality

By Lyria Boast and Tim Field
of Public Impact

October 7, 2013

INTRODUCTION

Parents, policymakers, and educators are committed to identifying great schools and calling attention to failing ones, but they often lack the tools to distinguish adequately between the two. The proliferation of available student data, advancements in analytic methods, and increased public demand for school quality information has spurred innovation and diversification in school quality rating systems. But with this rapid evolution comes the challenge to understand and track these diverse systems' characteristics. This report conducts an inventory of state and district school rating systems, as well as independent rating systems from charter and education reform organizations. Specifically, we seek to answer:

- What are the primary characteristics of school quality rating systems across the spectrum of rating systems and organizations?

- What are the key trends in the design and implementation of school quality rating systems?

- What challenges exist in developing accountability systems that are applicable across states?

These are especially relevant questions for the public charter school sector, with its explicit intention to provide innovative, autonomous schools that are held accountable to rigorous and meaningful quality measures.

The increasing sophistication of state accountability systems and frameworks that charter school authorizers use to evaluate charter school quality is providing the public with better information to evaluate and influence the growth of a high-quality charter school sector. However, it remains difficult to compare charter school quality across states. There does not yet exist a robust, national rating system on charter school quality, and several challenges inhibit establishing such a system—most significantly, the lack of comparability across state assessments, and the inability to collect student-level data across states for growth measures. A national system could offer the primary advantage of providing a consistent and comprehensive measure of charter school quality to inform parent choice and authorizer decisions. States and charter authorizers that have not yet established robust school quality rating systems to use in charter school accountability especially need a national system. Although the National Association of Charter School Authorizers (NACSA)[1] and the Building Charter School Quality project[2] have published recommendations for evaluating charter school quality, no national system evaluates charter schools, or traditional schools, with a consistent set of metrics across states.

While this report does not attempt to recommend a national charter school quality rating system, the analysis of existing systems can provide useful insights into school quality measures, methods, and reporting formats that should be considered in a national system.

BACKGROUND: SCHOOL QUALITY RATING SYSTEMS

Since the reauthorization of the Elementary and Secondary Education Act (ESEA) through the No Child Left Behind Act (NCLB) in 2001, the United States has focused more on using quantitative measures of school quality in K–12 public education.

NCLB mandates annual assessments of all students in grades 3 through 8, and at least once in high school grades, to measure student proficiency against state learning standards in English language arts, math, and science, and mandates the disaggregation and reporting of assessment results across student subpopulations.[3]

More frequent assessments and wider inclusion of students have led to a significant increase in the amount of data reported and available for evaluating school performance. Increased data access and rapid advancements in data infrastructure and analytics have led states and districts to design and implement more sophisticated systems to track student academic performance, not just by district or school, but by subgroups and at the individual level.[4] Since 2006, the federal government has invested half a billion dollars to help 41 states build longitudinal data systems that look beyond isolated proficiency rates by grade and subject, and track students' progress throughout their elementary and secondary careers.[5] In addition, NCLB reporting mandates required states to produce annual school and individual student reports that provide students, families, and educators with detailed information on student academic performance by subject and/or skill in a timely manner.[6]

Limitations of Existing School Quality Measures

One of the most visible components of NCLB school accountability was the strengthening of Adequate Yearly Progress (AYP) requirements in order to measure of the progress of all students, including major subgroups of students, towards reaching state proficiency standards. Due to the statewide AYP public reporting requirements, and the consequences of missing AYP targets in the NCLB accountability system, AYP quickly became a *de facto* measure of public school quality. Yet the increased scrutiny on school performance, as measured by AYP and proficiency levels on state assessments, also highlighted significant gaps in the adequacy of the existing school evaluation and accountability system, including the inconsistency in assessment rigor across states and insufficient data access for parents and communities.

Inconsistent rigor of state standards

Because NCLB did not establish national proficiency expectations, each state sets its own proficiency standards for students in every assessed subject and grade. As a result, student expectations vary widely across states. In some cases, the gap between expectations from one state to another is "so large that it is more than twice the

size of the national black-white achievement gap."[7] A lack of transparency about the assessments' rigor, coupled with accountability systems focused on meeting proficiency targets, actually gives states the incentive to decrease the rigor of assessments and proficiency standards, and raises questions about how meaningful state proficiency rates are, when "the word 'proficiency' means whatever one wants it to mean."[8]

Inaccessible and non-actionable data for parents and communities

School quality data is not consistently reported in an accessible and actionable way to families and communities. Parents often do not have easy access to transparent and useful information that will help them choose quality schools for their children. In 2012, only 26 states promoted awareness of available data, and only 31 states provided training for parents to use the data they received.[9]

Emergence of Post-AYP Measures of School Quality

In light of the gaps and consequences of the NCLB accountability framework, states, districts, educators, and families began to seek more rigorous, meaningful, and sophisticated techniques for evaluating school quality in order to identify and focus interventions on the lowest performing schools. Given the law's 100 percent proficiency target for 2014, an increasingly unmanageable number of schools were on track for being identified for school improvement interventions under AYP. In the absence of legislative changes to NCLB through reauthorization, the U.S. Department of Education announced plans in September 2011 to allow states to apply for ESEA waivers from NCLB mandates. These waivers, which included flexibility to redesign state school accountability systems and incorporate new measures of school quality, accelerated the movement toward a new generation of school accountability systems and post-AYP measures of school quality. Many of these new systems incorporate measures of individual student academic growth, a broad range of college readiness indicators, variations in achievement gap metrics, and more transparent school quality rating labels.[10]

Parents across the country are looking for better options for their children: the National Alliance estimates that there are nearly one million student names on charter school waiting lists.[11] Nearly 70 percent of Americans favor charter schools, according to the recent survey results.[12] In response to the success of charter schools in urban settings, reform-oriented urban districts are increasingly pursuing a "portfolio" model in order to give families more educational options among a diverse pool of high-performing schools that have autonomy over staff and funding.[13]

In response to parent demand for choice, districts have created their own rating systems to help parents, community members, and educators evaluate their schools. Denver and Philadelphia, for example, rate schools on a variety of factors, including student growth.[14] New York City and Chicago provide progress report cards that give schools an overall performance rating, but also include information on standardized assessments, growth measures, and school environment.[15]

A new tier of education organizations such as charter associations, authorizers, and investors has also emerged, which aggregate, report, and use school quality data to advocate for policy changes.[16] Organizations such as the National Association of Charter School Authorizers have developed principles and standards that provide guidance for authorizers in evaluating public charter schools against rigorous academic expectations, while still preserving autonomies that allow charters to pursue specific missions, instructional designs, and student populations.[17]

As the Internet becomes the primary medium for accessing school information, parents are also turning to private ranking systems such as GreatSchools.org, *U.S. News* Best High School Rankings, *Newsweek's* America's Best High Schools, and *The Washington Post's* Challenge Index. In addition, numerous state and city-based organizations have been established over the past decade to help families navigate school choice decisions.[18]

Current School Quality Rating Systems

To examine trends in evaluating school quality, we reviewed 25 rating systems used by state departments of education, large public school districts, charter associations and authorizers, and private news and advocacy organizations (Table 1).

TABLE 1

State Accountability Systems	Arizona California Colorado Florida	Indiana Louisiana Maryland Massachusetts	Minnesota Mississippi New Mexico Oklahoma	Rhode Island South Carolina Utah Wisconsin
City School Districts	Chicago Denver	New York City Philadelphia		
Charter Authorizers and Organizations	Washington, D.C. Public Charter School Board California Charter School Association			
Private Ratings	*U.S. News & World Report* (high schools only) *The Washington Post* Challenge Index (high schools only) GreatSchools			

We reviewed only systems that result in a single, overall rating or grade for schools. In order to have access not only to published descriptions of adopted rating systems, but also to view results and reports, we limited our review of state and district systems to those that published ratings for all schools for the 2011–12 school year. Many states and districts have adopted rating systems that will be implemented in 2012–13 or later; these are not included. During the review of systems, we identified the following types of performance metrics used for evaluation:

TABLE 2

Rating Components	Rating Component Variables	
Student Achievement Measures	■ Proficiency rates ■ Comparison with district or state performance ■ Advanced proficiency rates ■ Point systems for performance at different proficiency levels ■ Controls for differences in student population ■ Trend in overall school proficiency	
Student Growth Measures	■ Student Growth Percentiles (SGP) model ■ Value-added growth model ■ Value table growth model ■ Growth to proficiency/Adequate Growth Percentiles (AGP)	
College- and Career-Readiness Measures	■ Graduation rate ■ Disaggregated graduation rate ■ Extended graduation rate (5- or 6-year) ■ Diploma quality ■ Advanced coursework (e.g., AP, IB) ■ College readiness examinations (e.g., ACT, SAT)	■ Industry certification ■ College remediation ■ College attendance/acceptance ■ Dual credit
Subgroup Performance and Achievement Gap	■ Subgroup performance ■ Consolidated subgroup performance (use of "supergroups") ■ Growth of lowest-performing students ■ Reduction of achievement gap	
Measures of Student Engagement	■ Dropout, retention, or re-enrollment rates ■ Attendance, absenteeism rates ■ Parent and student satisfaction surveys	
Rating Format	■ A–F grades ■ Performance labels ■ Ranking	
Reporting Format	■ School report card ■ Interactive online interface	

To provide detailed examples of how organizations calculate school ratings, Appendices A-G presents profiles of systems. While there are similarities among the systems, they vary in the measures chosen for inclusion, the weighting of those measures, the method used for calculating an overall score or grade, and the presentation of the final results to the public. Two systems may have similar measures that are weighted differently, or very different measures that result in the same type of public reporting, such as an A–F grade. The profiles were selected to provide examples of the range of measures, weighting, methodology, and reporting formats, and are referred to throughout the report for illustration. We also conducted interviews with national experts on school accountability systems and selected practitioners of school rating systems, so we could identify additional insights and trends.

COMMON COMPONENTS AND TRENDS IN SCHOOL RATING SYSTEMS

Our review of systems produced five broad categories of components: student proficiency, student growth, measures of subgroup performance or achievement gap, postsecondary readiness and success, and measures of student engagement. Although the methods for assessing these components differed across systems, interesting trends exist in each area.

We also reviewed the methods for categorizing performance levels and reporting these results to the public. Table 3 (pages 8-9) highlights how each rating system incorporates metrics into these categories. Highlights and trends for each category are discussed in detail below.

Student Proficiency

Student proficiency on standardized state assessments served as the central measure of school quality under NCLB accountability systems. Although student growth and college readiness metrics have become increasingly important measures in new state accountability systems authorized under ESEA flexibility waivers, student proficiency remains a critical measure of school quality, and continues to represent the largest single factor in most school rating systems. It is a particularly useful metric for school quality rating systems because academic proficiency is widely recognized by educators, parents, and communities as a primary goal for public education. Furthermore, the general public can easily understand the metric of "meeting proficiency," and because results of state standardized assessments are publicly available for all public schools, proficiency information can be incorporated into school rating systems that do not have access to student-level information.

Use of Student Proficiency Measures in School Rating Systems
In most systems, academic proficiency remains the most important factor for evaluating school quality, and this report identified several trends used by school evaluation systems to augment the relevance of proficiency as a measure of school quality.

■ **Proficiency relative to the state or local district.** Several school quality rating systems compare school proficiency rates to district and state averages or decile performance levels. This approach is particularly useful for evaluation systems that are designed to compare school performance across states, because it partially controls for large variations in proficiency levels that are associated with differences in assessment rigor. Some evaluation systems also incorporate subgroup performance in comparisons with statewide performance. GreatSchools utilizes this methodology to create a 1 to 10 rating system for public schools in all states. The 1 to 10 score is based on decile performance versus the state, and includes an overall evaluation score,

TABLE 3[19]

Sample of Components of School Rating Systems 2011-12	Student Achievement/Proficiency					Student Growth					College and Career Readiness										
	Percent of Students Meeting Proficiency	Separate Evaluation of Advanced Proficiency	Trend over time in Proficiency Rates	Index - Proficiency Levels	Controlling for differences in student populations	Student Growth Percentiles	Value-Added	Value Table	Adequate Growth Percentiles/ Growth to Proficiency	Other	Cohort Graduation Rate	Disaggregated Graduation Rate	Extended Graduation Rates (5-year or 6-year)	Diploma Quality	Advanced Coursework	Readiness Exams: ACT/SAT/Accuplacer…	Industry Certification	College Remediation	College Attendance /Acceptance	Dual Credit	Assessment of Readiness at Elementary or Middle School
States																					
Arizona	■					■					■										
California	■																				
Colorado	■					■			■		■	■				■					
Florida	■							■			■	■	■		■	■	■			■	■
Indiana	■					■					■				■		■			■	
Louisiana				■			■				■		■	■	■	■	■			■	■
Maryland	■		■							■	■		■		■		■				
Massachusetts	■	■	■			■					■		■						■		
Minnesota	■							■			■	■									
Mississippi				■			■					■			■						
New Mexico	■	■	■		■	■	■				■		■		■	■	■			■	
Oklahoma				■							■		■	■	■	■	■	■		■	
Rhode Island	■	■	■			■					■		■								
South Carolina	■										■	■									
Utah	■					■					■										
Wisconsin	■					■			■		■	■				■					■
Districts																					
Chicago	■						■						■		■				■		
Denver	■	■			■	■			■		■				■	■		■		■	
New York City	■				■	■					■	■	■	■	■	■	■		■	■	■
Philadelphia	■						■				■				■	■					
Charter Authorizers and Charter Associations																					
CA Charter School Association	■		■		■																
D.C. Charter Board	■	■				■					■				■	■			■		■
Private Ratings																					
GreatSchools	■																				
US News&World Reports (HS)	■			■	■										■						
Washington Post Challenge Index (HS)											■				■						

TABLE 3[19] (CONTINUED)

Sample of Components of School Rating Systems 2011-12	Subgroup Performance				Student Engagement			Reporting Format		
	Disaggregated Subgroup Performance	Consolidated sub-group performance ("supergroups")	Growth of lowest-performing students	Reduction of achievement gap	Dropout/Retention/Re-enrollment Rates	Absenteeism/Attendance	Parent/Student Satisfaction Surveys	A/F Grades	Performance Labels	Ranking
States										
Arizona			■		■			■		
California	■									
Colorado	■		■		■				■	
Florida			■					■		
Indiana			■					■		
Louisiana			■		■	■		■		
Maryland					■				■	
Massachusetts		■			■				■	
Minnesota	■								■	
Mississippi								■		
New Mexico			■			■	■	■		
Oklahoma			■		■	■	■	■		
Rhode Island		■							■	
South Carolina	■							■		
Utah			■							■
Wisconsin		■			■	■			■	
Districts										
Chicago					■	■	■		■	
Denver	■				■	■	■		■	
New York City	■		■			■	■	■		■
Philadelphia	■					■	■			■
Charter Authorizers and Charter Associations										
CA Charter School Association									■	
D.C. Charter Board					■	■			■	
Private Ratings										
GreatSchools										■
US News & World Reports (HS)	■								■	■
Washington Post Challenge Index (HS)										■

as well as separate scores for each student subgroup at the school. A state decile approach is also used by private organizations, such as the Charter School Growth Fund and the Broad Foundation, to identify high-performing schools or districts that achieve exceptional academic results for students.

- **Advanced proficiency rates.** While most school quality rating systems consider the percentage of students meeting proficiency on state assessments, some systems include an additional evaluation of the percentage of students achieving advanced proficiency. Focus on advanced proficiency rates underlines a concern that educators are not merely focusing efforts on moving students into proficiency, but are continuing to challenge students at different performance levels. Furthermore, "proficiency" on many state assessments does not necessarily correlate to college-readiness levels, and "advanced" can be viewed as a more accurate measure of true college readiness.[20] Advanced proficiency rates are included in rating systems used by Rhode Island, the District of Columbia Public Charter School Board, and *U.S. News & World Report.* These systems evaluate schools based on the percentage of students achieving advanced proficiency in addition to evaluating the percentage of students who meet proficiency benchmarks.

- **Point systems for state assessment performance levels.** While most school quality rating systems evaluate proficiency by assessing the percentage of students achieving proficiency, some systems focus on the percentage of students at each performance level on state assessments (e.g., far below basic, below basic, basic, advanced). This approach gives a more detailed and balanced view of student performance; schools receive more credit for having students at advanced levels, while schools with most students far below proficiency would be rated differently than schools with most of their students nearing proficiency.

- **Controlling for differences in student populations.** Some systems include a consideration of student populations when evaluating school-level proficiency rates.

  - A number of systems, such as those for Rhode Island and the California Charter School Association, use regression on student characteristics in an attempt to remove the effects of student characteristics from the assessment of school proficiency rates. Although there has been a national debate about this practice, some would argue that adjusting for student characteristics is only fair when comparing schools in affluent communities to those serving poor students in disadvantaged urban communities.

  - New York City uses a similar-schools comparison. Each school is compared to schools serving similar student populations. This approach provides actual schools for comparison, but can often generate debate about the appropriateness of specific schools chosen for the analysis, because schools are rarely exactly similar.

- **Trend in proficiency rate.** The main component of the NCLB AYP reports was an evaluation of the improvement in overall school proficiency rates over time. Although this evaluation is easy to calculate and understand, many states have moved to incorporate individual student growth measures as better models to calculate growth. The "trend in proficiency rate" is still used in newly adopted systems in Rhode Island, Massachusetts, and South Carolina.

Considerations for the Use of Student Proficiency

There are important limitations of student proficiency as a measure of school quality. Most notably, proficiency measures fail to account for different academic starting points—thus attributing "credit" for a school's low or high performance to the characteristics of its students when they entered the school, rather than the academic progress they achieved at the school. The significant variety of state assessments' rigor and proficiency cut scores also create large variations in state proficiency levels, even as performance on NAEP and other national student assessments reflect relatively modest variations in academic performance across states.[21] In addition, as noted above, "proficiency," as defined by state assessments, does not always correspond to academic performance levels that are an accurate assessment of college- and career-readiness standards.

Student Academic Growth

The increasing demand for alternatives to proficiency measures launched the proliferation of student academic growth models. States and districts usually adopt value-added models, student growth percentiles, and growth from pre- to post-tests. Value-added models predict how much growth a student will make based on the historical growth of similar students, and then determine whether a school's students achieved growth that fell short of, met, or exceeded those predictions.[22] Student growth percentiles measure a student's growth in comparison to his peers, and can show whether schools have met expected growth for their students.[23] Pre- and post-tests evaluate student progress toward growth goals set by students and educators.

The inclusion of student growth measures is perhaps the most significant advancement in the evolution of state accountability systems. Of the states that have received ESEA waivers, nearly all have included measures of student academic growth on state assessments, in addition to academic proficiency.[24] In advance of the waivers, several school districts and charter authorizers had already begun to incorporate growth measures into school accountability systems that are used for evaluating the relative performance of schools to inform school intervention and replication strategies.

By incorporating growth measures, schools can be evaluated based on the academic progress of each student, regardless of where a student started academically. This prevents schools from being "punished" for educating students who started school behind academically or "rewarded" for students who started out highly proficient, but failed to achieve expected growth during the school year.

Types of Growth Models

The systems use various methodologies to calculate student growth.[25] Chris Domaleski and Marianne Perie from the Center on School Assessment provided an overview in 2012 of the most common methodologies that state accountability systems use: the student growth percentile model, value-added evaluations of growth, and value tables.

- **Student growth percentile** (SGP) models calculate individual student progress in comparison to the progress of academic peers—students with similar performance on previous assessments. Simply put, each student's growth in assessment results is ranked against all students who had the same baseline assessment result. A student with an SGP of 50 demonstrated higher growth than half of his academic peers with similar performance history across the state. A school median SGP of 60 indicates that at least half of the students in the school showed more growth than 60 percent of their academic peers with similar performance across the state.

- **Value-added models**, such as the Tennessee Value-Added Assessment System (TVAAS), control for student characteristics such as poverty status, to isolate the amount of growth for which the school may take credit.

- **Value tables, or transition models**, assign points for each student who shows a defined amount of growth or moves from one performance category to another. For example, the Florida growth model assigns "learning points" for each student who moves from one performance category to another on two consecutive state assessments. More points are given for more desired transitions, such as the move from non-proficiency to proficiency.

- **Growth-to-proficiency models**, and the adequate growth percentile (AGP) model, assess whether students' individual growth is sufficient to achieve or maintain proficiency within a certain time. The AGP evaluates whether students are on track to reach proficiency within three years. In contrast, the SGP and value-added models compare individual students' growth with the growth of their peers to evaluate the typical or average growth shown by students. Although it is important to understand average growth patterns, typical growth may be insufficient to bring low-performing students up to grade level. Although consistently achieving average growth for an "advanced" student would suggest that the student will graduate college-ready, average growth for a "basic" or low-performing student likely means that the student is not making progress toward proficiency.

- **Other models:** Some additional growth models used in the ratings include gain scores and improvement models.

  - Gain score models evaluate the change from year to year in individual student test scores. Though gain score models are simple to calculate and understand, they do not take into consideration differences in

growth expected at different starting points of performance.[26] For example, students far below proficiency show higher growth in scale score points, on average, than students far above proficiency.

■ Improvement models compare changes in overall proficiency rates, as opposed to change in individual student performance. NCLB is an example of an improvement model, with annual targets for overall school proficiency. Improvement models consider only the overall proficiency rate for all students in the school; they do not take into consideration changes in student population.

Use of Growth Measures in School Rating Systems

In addition to differences in the methodology for measuring student academic growth, rating systems vary in their use and weighting of student growth to calculate an overall quality score. Although many use fairly equal weightings for proficiency and growth, some systems give more weight to growth, while others place more importance on status or proficiency (for example, Colorado gives growth a 75 percent weight in elementary and middle schools, while Rhode Island puts it at just 25 percent).

In addition, the rating systems consistently apply different weights to growth measures for high schools versus middle and elementary schools. For high schools, growth is either not calculated, due to the limited number of annual state assessments administered in grades 9 through 12, or it is de-emphasized in favor of college- and career-readiness metrics. Alternative metrics are often developed to evaluate growth at the high school level. Rhode Island and Arizona evaluate the change from year to year in overall proficiency rates on high school end-of-course tests as a proxy for growth. Louisiana is using the ACT series of college readiness assessments (EXPLORE, PLAN, ACT) to calculate student growth for high school students.

■ Some state rating systems use multiple student growth models to evaluate school performance. Colorado and Wisconsin include both SGP and AGP models to provide a comprehensive assessment of student growth that answers two important questions: 1) Are students showing typical growth? and 2) Is that growth sufficient to bring all students to grade level?

The growth methodologies used in school evaluation systems vary in complexity and are often tailored to the features of the underlying data from the state assessments. Several studies compare the pros and cons of these various growth models.[27] Although there is not yet a "gold standard" for the most effective, it is important that growth model selection is guided by a clear theory of action about the purpose of the school evaluation system.

Considerations for the Use of Student Growth Measures

Despite the advantages of incorporating growth measures into school rating systems, the validity of these models can be compromised by the limitations of

the underlying state assessment. Criterion-referenced state tests often focus on a narrow band of academic performance clustered around a proficiency standard for each grade level. For students who are performing at or around that standard, the assessment might be more relevant. But for students performing well above or well below the proficiency level, it is difficult to gauge how much they have learned during the year.[28]

Some districts, schools, and charter organizations have addressed this limitation by using national norm-referenced tests to evaluate student growth.[29] Rather than measuring academic growth based on a set of grade-specific criterion measures, norm-referenced tests evaluate student growth in comparison with students who began the academic year at a similar starting point.

The development of computer-based, adaptive student assessments provides another solution to the limitations of state assessments. Adaptive assessments dynamically select (adapt) the difficulty of questions based on correct and incorrect student responses.[30] Adaptive assessments can thus assess a much broader range of student knowledge, providing a more accurate measure of student growth. The norm-referenced Measures of Academic Progress (MAP) assessment, developed by Northwest Evaluation Association (NWEA), and the planned Smarter Balanced student assessments for the Common Core are two examples. Delaware, Hawaii, and Oregon have also begun to incorporate computer adaptive testing into their state assessment systems.

College and Career Readiness

Rating systems are also focusing on measures that evaluate schools' ability to prepare students for success in college and the workplace. The 2008 Higher Education Opportunity Act encouraged high schools to monitor college enrollments, and provided support to use this data to improve schools. In 2009, the American Recovery and Reinvestment Act gave states funding to report college enrollment and credit accumulation rates. Race to the Top grants included college enrollment and attainment rates as measures of success for their grantees.[31]

This broader perspective on student achievement has encouraged and enabled states to use existing data to monitor how schools are preparing students over time for success beyond the K–12 system. Most states produce individual student reports that include measures of academic growth and college/career readiness.[32] Forty-three states link K–12 and postsecondary data, and 17 states link postsecondary and workforce data.[33]

Use of College- and Career-Readiness Indicators in School Rating Systems
The proliferation of postsecondary readiness metrics represents a significant evolution from earlier state accountability systems, and highlights a growing emphasis on college and career readiness as an ultimate goal for public schools. Table 3 (pages 8-9)

highlights the range of postsecondary readiness indicators that are being incorporated into state, district, and private school rating systems, including:

- **Graduation rates.** Though all states reported graduation rates under NCLB, states used different methods to calculate the percentage of students successfully earning a high school diploma. However, in the 2012–13 school year, all states will calculate and report graduation rates using the National Governor's Association four-year cohort method, which follows a cohort of students from ninth grade though graduation, reporting the percentage who graduate. Some rating systems also include five-, six-, or seven-year graduation rates, giving schools credit for students who take longer to finish. Additionally, some systems, such as Wisconsin and Oklahoma, further assess the graduation rates of at-risk students or students in subgroups.

- **Diploma quality.** In states with a range of available diplomas, some rating systems evaluate schools based on the overall quality of their diplomas. Louisiana, for example, awards more points in the state rating system for higher-quality diplomas and fewer points for students earning a GED. Mississippi, Oklahoma, and New York City also give more credit in ratings to schools with higher rates of advanced diplomas.

- **College readiness assessments.** Many rating systems include some combination of ACT, SAT, AP, or IB participation and performance results. They provide results that are comparable across states, and the ACT and SAT have established benchmarks for college readiness. Participation rates, however, may not be sufficient everywhere to warrant inclusion in an accountability system. States such as Colorado that fund the ACT or SAT for all students are well positioned to include assessment results in school accountability frameworks. In cities or states with low participation rates, however, schools should not be assessed based on the results of a small proportion of students, often those with the means to pay test fees. Similarly, AP and IB results are appropriate only when all students have access to test participation and there are no incentives for school leaders to discourage low-performing students from participating.

- **Industry certification.** In order to address the preparation of students who may not go to college, many states and cities, including Florida, Indiana, Louisiana, Maryland, New Mexico, Oklahoma, and New York City, have included industry certification completion rates in rating systems.

- **College attendance.** The goal of many traditional and charter schools is to prepare students to get into and graduate from college. Until recently, the inability to collect college attendance and retention data has hindered efforts to evaluate the success of individual schools in preparing students for college. The development and expansion of state systems and the National Student Clearinghouse now allow most schools to see college attendance rates for their students, providing a valuable assessment tool. While only one of the state systems we reviewed, Maryland, included college attendance rates, we

expect that more states will include them in the future. NACSA recommends the use of college attendance rates, and many charter school authorizers use them, including the D.C. Public School Charter Board and Ball State University.

- **College remediation.** Although growing national attention has focused on high rates of required participation in college remedial classes, only one of the systems reviewed, Denver, included college remediation rates in its rating systems. Remediation rates are generally collected state by state through statewide college reporting databases. Inclusion in accountability systems requires the linking of K–12 and postsecondary data systems.

- **Elementary and middle school readiness.** If postsecondary success is the ultimate goal, then one could argue that postsecondary readiness should be measured starting in middle school or even elementary school. A growing number of rating systems have added "early indicators" of readiness to K–8 school rating systems. Common measures of early readiness include performance on fourth-grade reading and eighth-grade math state assessment results, or assessments such as EXPLORE, which is administered by ACT in eighth or ninth grade and gives both an assessment of college readiness and the ability to measure growth when administered with the ACT test.

Considerations for the Use of College- and Career-Readiness Measures

Although there is increased nationwide focus on college and career readiness, state and national data systems cannot always provide the necessary data to evaluate schools on readiness. Many states are prioritizing efforts to link K–12 with postsecondary and state data systems in order to follow students from high school through college or employment. In addition to state data sources, college- and career-readiness data may be accessed from testing agencies such as ACT and the College Board, and from the National Student Clearinghouse, which calculates college attendance rates.

Even as additional data sources become available, opinions differ on whether the goal of having all students attend college is appropriate. Depending on the objectives of the rating system, or the mission of the school being evaluated, measures of career readiness can augment or substitute for college readiness.

Subgroup Performance and Achievement Gap

Nearly every school quality rating system reviewed for this report measured the difference in academic performance among student demographic subgroups. Under ESEA flexibility waivers, states must set annual measureable objectives (AMOs) for each student subgroup, but are not required to include all subgroup performance levels in their school rating systems. With ESEA waivers, states may adjust how subgroup performance is evaluated in their accountability systems, and may incorporate student growth rates (as opposed to proficiency rates) as a measure of the achievement gap.

As mentioned above, many states have used the waivers to redefine subgroups to avoid double-counting students who are classified under multiple subgroups. For example, a Hispanic, low-income student receiving English language learner (ELL) services would be included in all three subgroups if each group is evaluated separately for accountability purposes. Although all state accountability systems continue to report disaggregated student subgroup proficiency levels, many school rating systems have adopted new approaches to incorporating subgroup performance into school quality measures.

Use of Subgroup Performance and Achievement Gap Measures in School Rating Systems

The improvement of the lowest-performing students is an important priority for most educators. The rating systems reviewed use several methods to evaluate the performance of students in subgroups.

- **Disaggregated subgroup performance.** Most of the systems reviewed for this report continue to evaluate the performance of disaggregated subgroups. Though students may be double-counted if they belong to more than one group, evaluating each subgroup separately avoids the possibility that the performance of one subgroup may be masked by the higher performance of other students.

- **Consolidated student subgroups.** Some rating systems evaluate the performance of all students who are part of any traditionally low-performing subgroup. This "supergroup" may include any combination of minority, free and reduced-price lunch, ELL, and special education students. Including these students in one group for evaluation avoids double-counting students who fall into more than one subgroup. Massachusetts, South Carolina, and Wisconsin, for example, evaluate supergroup performance.

- **Lowest-performing students as a subgroup.** Some argue that educators should focus on the performance of the lowest-performing students (e.g., students in the bottom 25 percent for the school), regardless of whether they fit a defined subgroup. Arizona, Florida, Indiana, New Mexico, Oklahoma, and Utah have all included a measure of growth of the lowest-performing students in their state rating systems. Lowest-performing students are usually identified based on the prior year's assessment results.

- **Achievement gap reduction.** To directly address the question of an achievement gap between high- and low-performing student groups, some rating systems include a comparison of groups and evaluate the magnitude of the gap or progress toward reducing the gap. In Rhode Island, for example, the performance of subgroups within each school is compared to the highest-performing subgroup in the district.

Achievement gap measures continue to be a factor in non-state rating systems. In part because student subgroup performance data is publicly available, several private rating organizations use these data to evaluate school quality. GreatSchools.org

produces a 10-point rating system that does not directly include an achievement gap metric in the overall score, but does provide separate student subgroup scores, using a 10-point scale, that compare school subgroup performance to the overall state student population. *U.S. News & World Report's* Best High Schools does not publish performance levels for student subgroups, but includes less-advantaged subgroup categories (black, Hispanic, and low-income) to adjust overall school performance based on statistical expectations and to compare performance levels of these subgroups against state averages. Similarly, the California Charter School Association evaluates California charter schools as "meeting" or "not meeting" expectations based on an adjusted calculation of the state's API score that controls for differences in student demographics.

Student Engagement

Many state and district rating systems, as well as the D.C. Charter School Board, include various measures of student engagement to evaluate school quality. Student engagement metrics are less pervasive than metrics related to student growth, achievement gap, and college readiness indicators. The most common metrics include student dropout rates, attendance, and parent/student satisfaction surveys. These are a small component of the overall rating system, generally accounting for a small percentage of "points" in a school's aggregate score. Because parent and student satisfaction data requires a consistent process for administering surveys, this data is most commonly included in district systems and by charter authorizers that can require and oversee survey administration, although New Mexico and Oklahoma included survey results in new state accountability ratings.

It can be argued that student engagement metrics represent important leading indicators of school quality, and in our interviews with accountability experts, several recognized the desire to incorporate measures of student engagement and school climate/culture into rating systems. However, experts also pointed to the real challenges of developing measures that are reliable (not subject to the reporting bias found in parent satisfaction surveys, for example) and which do not create incentives for counterproductive school management decisions (e.g., underreporting student discipline incidents).

School Quality Reporting Formats and Rating Methodology

Most of the reporting formats reviewed for this report require the rating system to combine multiple metrics into a point system. This approach is described in the literature as a "compensatory" model, wherein low performance on a single metric can be *compensated for* by high performance on separate measures (and vice versa). This compensatory model is a sharp departure from the "conjunctive" model of AYP, under which schools must meet minimum targets for every subgroup in order to receive the "quality rating" of "Made AYP."

Rating systems using a compensatory model must assign different weights to each measure, reflecting values about what metrics are most important about school quality. For example, if growth is valued more highly than proficiency, growth may be given a higher weight in a rating model. The approach for weighting individual measures must be considered carefully—a single set of underlying performance metrics can result in very different assessments of quality depending on the weights given to individual measures.

We have focused on school rating systems that assign a grade, score, or performance labels to schools. Under NCLB accountability systems, relative performance levels across schools were articulated through school improvement status designations (e.g., Making AYP, School Improvement, Corrective Action), and solely based on meeting or not meeting AYP proficiency targets. Rating systems have evolved to include more sophisticated measures of school performance intended to provide families, communities, and public officials with more informative and actionable data on school quality, resulting in school ratings that include both detailed information about school performance as well as easy-to-interpret overall ratings or grades for individual schools. The ESEA waiver process accelerated this evolution, and there is now a diverse range of school quality reporting formats that have been established by districts, states, and private organizations.

Table 3 (pages 8-9) shows the reporting formats used by the systems reviewed for this report; the rating results are presented as A–F grades, performance labels, or rankings. Even while school rating systems may use different methodologies for evaluating school quality metrics, many share similar approaches to combining metrics into a single performance classification. For example, multiple states use an A–F grading system, even though the components of the grading systems and the methods for calculating a final score can differ significantly. Most grading systems with a rating classification have an underlying point system to assign classifications; however, not all of these grading systems publicly display points with the school rating classification.

Regardless of reporting format, if school rating systems are to provide actionable information, the performance classifications should be assigned in a manner that provides rigorous distinctions in relative school performance. Systems in which more than 50 percent of schools are in one performance category (as in Colorado), or systems whose targets for the highest target category are below state average performance levels (as in Virginia), fail to distinguish adequately between schools. A school performing at the state average is not on par with a school in the top 5 percent of schools statewide, and the performance classification should reflect these differences.

A–F Systems

States are increasingly adopting A–F school reporting formats, following the example first set by Florida in 2001 (see Florida profile on page 42). Although underlying methodologies for calculating an A–F grade may be complex, the summative grade provides a very transparent and intuitive label of school quality. The widespread adoption of the A–F reporting in multiple states has been supported, in part, by

the Foundation for Excellence in Education, which Governor Jeb Bush established in 2008 to support education policy reform across the country, including a specific policy agenda of promoting A–F grading systems.

Although many states have adopted an A–F reporting format, their metrics and methodologies can vary significantly. For example, Florida, Louisiana, South Carolina, and Arizona all report A–F grades for schools, yet there are significant differences in the components included in these systems. Florida, Louisiana, and Arizona use the results of three growth models (value table, value-added, and student growth percentiles) and assign different weights to these growth results, while South Carolina does not use a student growth model in school evaluation. Louisiana applies a range of postsecondary measures to evaluate high schools (graduation rate, diploma quality, completion of advanced coursework, performance on college readiness exams, industry certifications, and dual credit) while Arizona and South Carolina include only graduation rate.

Performance Labels

Many systems use performance labels or categories to classify relative school performance, rather than assigning a grade, ranking, or number to schools. Examples of performance labels include:

- Colorado: Accredited, Improvement, Priority Improvement, or Turnaround;

- Rhode Island: Commended, Leading, Typical, Warning, Focus, Priority; and

- Wisconsin: Significantly Exceeds Expectations, Exceeds Expectations, Meets Expectations, Meets Few Expectations, Fails to Meet Expectations.

In some rating systems, the labels correspond to specific sets of rewards and consequences, while in others they are illustrative and intended to provide a description of the school's relative performance. For Rhode Island, the labels of "Focus" and "Priority" specifically align to the state's strategies for identifying and supporting low-performing schools as articulated in its ESEA flexibility waiver. In some systems, the performance label language is somewhat similar to the NCLB performance classifications. However a primary difference is that NCLB classifications are based on successive years of achieving (or not achieving) AYP, versus assigning performance labels based on an annual, aggregate score derived from a combination of performance metrics.

In order to provide clear and transparent information about school quality, performance labels should use language that effectively communicates school performance levels. Although the Colorado ratings include many worthy features, the labels of "Accredited," "Improvement," and "Priority Improvement" provide a less intuitive description of school quality than an A–F or numerical rating system.

Ranking

Six of the rating systems reviewed used a ranking to report final assessments. Schools are first assigned a performance index, or score, based on performance across the multiple performance measures. The score is used to order schools, which are then assigned a percentile or decile ranking. A school that outperforms 55 percent of schools would have a percentile ranking of "55" and a decile ranking of "6." Utah and Philadelphia present both a score and a ranking. New York City school reports provide a grade, score, and ranking, while the three private ratings in our review all use rankings to report results of their assessments.

KEY CONSIDERATIONS FOR DEVELOPING AN EFFECTIVE SCHOOL RATING SYSTEM

Aligning School Rating Systems to a Theory of Action

As organizations develop a school quality rating system, they will undoubtedly grapple with constraints related to data quality and accessibility, as well as political considerations and their capacity to reliably calculate and disseminate school performance data. However, the essential starting point for every organization is to first establish a clear *theory of action* for their rating system. That is, what is the purpose of the school evaluation system, and what actions is it designed to guide and motivate? Although all models should be technically sound and account for the strengths and limitations of the assessment data, the system's effectiveness ultimately depends on how well it aligns to the organization's theory of action and the purpose of the rating system. Organizations should begin and end their development of a rating system with this in mind.

Aligning a school rating system to an organization's theory of action goes beyond the calculation of the school rating or performance classification, extending to the manner in which school quality information is communicated to target communities. (See Data Transparency and Public Accessibility on page 22).

For example, it is helpful to compare characteristics of different rating systems and contemplate what their methodologies say about the underlying theory of action. As we discussed, state accountability systems have made different decisions about the relative weight of student growth measures. The Colorado model weights growth as 75 percent of the overall school rating for elementary and middle schools, whereas the Rhode Island system puts it at 25 percent. Although leaders from both states might declare that their objective is to ensure that all students graduate from high school prepared for postsecondary success, the differences in their methodology might suggest different views on how their school accountability systems and

state assessments support that objective. The Colorado model implies a theory of action that holds that by focusing accountability on academic growth, schools will maximize the potential learning gains of each student so that he or she is more likely to be prepared for postsecondary success. In contrast, the Rhode Island model implies a theory of action holding that by focusing accountability on academic achievement, schools will raise the rigor of student instruction to meet a high standard of academic performance.

The school quality rating system developed by the California Charter School Association (CCSA) provides another example of how a rating system was designed to align to a specific theory of action. Beginning in 2009, the CCSA began focusing on the challenges posed by having a state charter sector that included a high concentration of charter schools that were distributed at both the low and high ends of the performance spectrum for public schools. The CCSA determined that perpetuating low-performing charter schools hurt the success of the charter sector and was not in the best interest of students. Rather than simply wait for charter authorizers to close low-performing schools, the CCSA decided to be play an active role in identifying the highest- and lowest-performing charter schools in the state. The CCSA determined that the primary state accountability measure, API, was not sufficient to accurately and reliably differentiate the lowest- and highest-performing schools, and thus developed a charter school quality rating system that included enhanced methodologies to use in addition to the API. (See CCSA profile on page 30 for more details.) The CCSA theory of action in this case was to use their school quality rating system to provide actionable data to facilitate closing low-performing charter schools (and to highlight the state's most successful charter schools).

Data Transparency and Public Accessibility

A primary motivation for developing school quality rating systems is to more effectively communicate the concept of school quality to the public. The systems reviewed for this report use various levels of methodological complexity to create their ratings. Some systems, such as Arizona and Florida, use a fairly simple approach of assigning points based on the percentage of students meeting proficiency or growth targets (e.g., 60 points for 60 percent proficiency), while other systems, such as New Mexico, use a relatively complex rating formula that includes two growth models—SGP and value-added—and a range of postsecondary measures, yet results in an easily interpreted A–F grade.

Regardless of formula complexity, effective rating systems summarize information in an easily understood reporting format. This report asked practitioners who have been involved in designing and assessing school quality rating systems whether there is a point at which the complexity of the rating methodology interferes with its ability to provide transparent and accessible information, but they generally did not see this as a major concern.

Based on the review of existing rating systems in this report, several factors contribute to a transparent and accessible school quality rating system:

- **Clear, intuitive school reporting format.** Many states have adopted the A–F partly because it communicates a very clear message about school quality. This level of clarity can also be achieved through well-designed numeric scores or performance labels. But, as discussed in the "School Reporting Format" section of this report, performance labels do not always articulate a clear and meaningful distinction in relative school performance.

- **Informative descriptions of rating methodologies and reports.** Beyond providing intuitive and clear rating formats for schools, effective rating systems are supported by informative and accessible information about how to understand and interpret the system. This often includes web-accessible school report cards, FAQs, and information guides that summarize the school rating methodology, along with additional information about the school that is relevant to families and school communities. This level of information is particularly important to support informed school choice, when parents are navigating complicated choices about where to enroll their children. The D.C. Public School Charter Board is a good example of a charter authorizer that provides useful information about its performance management framework. In 2012, it began publishing a "Parent Guide to Public Charter School Performance" that summarizes the framework methodology and provides guidance on how to use the report cards to make informed decisions about school selection.

- **Interactive data-reporting applications.** A select number of reporting systems have developed websites and reporting applications that go beyond describing rating methodologies and results to allowing parents and educators to create customized school quality data reports. Colorado has been a pioneer in this area, adopting the "School View" web-based software that allows users to create customized school reports. Additionally, parents and students may use secure passwords within School View to view individual student reports showing achievement and growth results.

  In the design and rollout of School View, the Colorado Department of Education did much work to educate school districts and communities about the rating systems and reports. The department conducted focus groups with teachers, principals, and parents to ensure that the system was user-friendly and that educators, parents, and students could use it to track and discuss academic performance and progress.

  The Colorado School View system is also a unique example of cross-state sharing and collaboration, as Colorado is now licensing the School View system to other states to use with their respective school accountability systems.[34]

Challenges of Cross-State Analysis

A few rating systems attempt to evaluate schools across states. The two great challenges to such comparisons are: 1) the differences in state standards and assessments that form the backbone of school quality rating systems, and 2) the challenge of collecting student-level data needed for academic growth metrics. Adopting Common Core assessments (see below) may partly address these formidable obstacles, but they will continue to hinder the feasibility and value of a robust national school rating system. However, some organizations have developed cross-state rating systems that use a variety of approaches to compensate for these data limitations.

Within-State Ranking Systems

Several organizations have developed systems that use publicly available state assessment data to conduct within-state comparisons of school performance. GreatSchools uses a 1–10 ranking system that compares schools' performance with the state average for each combination of grade/subject at the school. "Performance" is based on the percentage of students meeting proficiency targets, and the 1–10 rankings are based on decile performance versus statewide performance. The rating system also provides 1–10 rankings by student subgroups to show how subgroups in a school compare to subgroup performance in the state. (See the GreatSchools profile in Appendix A). Although school ratings are not comparable across states, the methodology is consistent, and the results provide a snapshot of school performance relative to all other schools in the state.

Within-state comparisons are also used by the Charter School Growth Fund (CSGF) and the Broad Prize competition to identify higher-performing charter management organizations (CMOs) and school districts.[35] The Broad Prize for Public Charter Schools evaluates charter school proficiency levels (and proficiency improvement) on state assessments in comparison to the state and local district, using publicly available state-reported data.[36] CMO performance is based on school-level data aggregated across schools, and CMO comparisons to the state are expressed as a decile ranking, and calculated for each subgroup at the elementary, middle, and high school levels. Math, reading, and science decile ranking data are provided to show performance (e.g., percent of students meeting standard) and improvement at "proficient" and "advanced" levels. The Broad Prize analysis also uses state comparisons on the reduction and magnitude of the achievement gap, graduation rates, and participation rates and scores on SAT/ACT and AP/IB exams.

However, unlike GreatSchools and other school rating systems profiled in this report, the Broad Prize does not create a quality rating score or performance classification. It uses the data to select a group of finalists from a pool of eligible school districts and CMOs for a more thorough review for award decisions. By not assigning weights to metrics, the Broad Prize methodology allows its district/CMO evaluation committees to independently review and interpret the full range of performance data.

Nationally Normed Assessments

Although many school districts, CMOs, and individual schools use nationally normed referenced grade/subject tests (e.g. MAP, SAT-10, ITBS) to evaluate student academic growth and proficiency, these assessments were not included in any of the school quality rating systems identified for this report. However, nationally normed assessments that focus on college-readiness indicators (e.g., ACT/SAT, AP/IB assessments) are being incorporated into many of the school rating systems profiled in this report. This includes many state accountability systems that use national assessment data as one component of their overall evaluation system, as well as several private organizations that publish school quality rating reports on high schools. *Newsweek*'s America's Best High Schools, *U.S. News & World Report*, and *The Washington Post* Challenge Index each use AP/IB test participation and success as a prominent metric for evaluating high school quality.

Adjustments for Relative Rigor of State Assessments

School proficiency rates in different states cannot be compared due to differences in assessment quality and proficiency cut points. The Charter School Growth Fund, which does not create a school ranking but does evaluate schools in different states with a uniform set of performance measures, has been working to use a regression model that accounts for the relative strength of state assessments. Similarly, the Broad Prize recognizes the challenge of evaluating district performance across states even though state tests cannot be directly compared. The Broad Prize does not use a quantitative approach to adjust for these differences, but gives its review panel state comparison data from NAEP and the Northwestern Evaluation Association (NWEA) to provide context for differences in proficiency levels across states.

Implications of Common Core Assessments on School Quality Rating Systems

States' transition to Common Core Standards and their adoption of Common Core-aligned assessments will present new opportunities to develop rating systems that provide meaningful cross-state school comparisons. As of March 2013, 43 states and the District of Columbia have become member states of either the Smarter Balanced or PARCC Consortia of Common Core assessments.[37] This creates the expectation that by school year 2014–15, nearly all states will use one of two common national assessments, making measures of student proficiency and growth comparable across states. Although common assessments will not necessarily lead to comparable state accountability systems, they will create new opportunities for organizations (and states) to create national school quality rating systems.

In interviews conducted for this report, rating systems experts expressed cautious optimism about the impact of Common Core assessments. While many acknowledged the upsides of national, common assessments, several indicated that it was too early to predict the impact on ratings and highlighted several challenges that will confront education policymakers. For many states, the transition

to Common Core assessments will precipitate a dramatic decrease in student proficiency rates. This will undoubtedly create pressure on education leaders as families, educators, and political leaders react to a new perception of school quality. Although several states have experienced the impact of dramatic reductions in student proficiency levels as a result of changes to state assessment rigor or cut scores, the public fallout from this transition may erode the will of state policymakers to fully embrace the adoption of national Common Core assessments.[38]

The transition to Common Core assessments will also create technical challenges for state accountability systems, including how to report reliable student growth measures during the transition from old to new assessments, and how to adapt their school quality rating system to account for differences in underlying assessment data. These generally represent short-term technical challenges, but may complicate or delay the development of national school quality rating systems. In addition, the adoption of Common Core assessments will not necessarily solve the limitations of existing state criterion-referenced assessments to accurately measure student academic growth—namely, that assessments evaluate student academic levels in a narrow range focused on grade-specific proficiency levels. On this dimension, the Smarter Balanced assessments will provide an opportunity to see how the use of computer adaptive technology can generate more accurate and meaningful measures of student academic growth.

CONCLUSION

In our review of state, district, and national systems, we found clear trends that should inform how state departments of education, school districts, charter authorizers, and the public think about the optimal design of school quality rating systems.

- **Inclusion of student growth.** When they have the flexibility and technical capability, rating systems have consistently incorporated student growth measures. This trend reinforces the intuitive appeal of evaluating school quality based on the academic progress of students. Incorporating student growth has become a new standard for school quality rating systems.

- **Expansion of college- and career-readiness measures.** The universal goal of public pre-K–12 education is to prepare students for success in college and careers. The measures for evaluating progress against this goal will continue to evolve, but effective school quality rating systems have clearly moved well beyond high school graduation rates to include more meaningful indicators for student attainment and readiness.

- **Exploration of new ways to focus attention on lowest-performing students.** Great schools are great for *all* students in the building. This value remains reflected in the newer generation of school quality rating systems, even as the metrics for evaluating achievement gaps have evolved and diversified.

- **Interest in valid measures of student engagement.** Though ratings systems continue to focus on student academic outcomes, some seek to add qualitative measures that capture school culture.

- **Simplified reporting formats to categorize school quality.** While the measures to evaluate school quality have expanded and become more complex, many systems have simplified the format for reporting school quality.

- **Increase in data transparency and public accessibility.** The emphasis on accountability and school choice has generated increased public demand for transparent and accessible school quality data. The value of a school quality rating system should be judged not just by its reliability and accuracy, but also by how effectively this data is made available to and actionable by educators, parents, and policymakers.

These trends represent progress in the movement toward more meaningful measures of school quality, but the challenge of comparing school quality across states remains. Proficiency rates and growth measures across states are not comparable, and the student-level data required to calculate student growth is not easily available to the national organizations interested in developing such measures.

Though the ability to rank schools nationally presents challenges, our study revealed useful approaches to creating valuable rating systems, including the use of statewide ranking based on proficiency rate, adjustments to state proficiency rates based on the relative strength of state assessments, use of nationally normed assessments, and expanded use of college- and career-readiness measures that can be applied equally across states. Most significantly, the adoption of Common Core-aligned assessments will provide new opportunities to develop a more robust and reliable system for evaluating school quality across states.

ENDNOTES

1 National Association of Charter School Authorizers. (2013, March). *Core performance framework and guidance: Academic, financial, and organizational frameworks for charter school accountability. Chicago: Author.* Retrieved from http://chartercommission.idaho.gov/faq/documents/NACSA%20Core%20Performance%20Framework%20and%20Guidance.pdf

2 National Consensus Panel on Charter School Academic Quality. (2008, June). *A report from the National Consensus Panel on Charter School Academic Quality: A framework for academic quality.* Retrieved from: http://www.charterschoolquality.org/media/1186/FrameworkForAcademicQuality.pdf

3 Assessment requirements for science are less frequent, requiring annual assessments at least once in grades 3-12. For more information, see: U.S. Department of Education. (2003, March 10). *Standards and assessments: Non-regulatory guidance* (pp. 2–3). Retrieved from http://www2.ed.gov/policy/elsec/guid/saaguidance03.doc

4 U.S. Department of Education. *Standards and assessments* (pp. 10, 16–23).

5 Schramm, J.B., & Zalesne, E.K. (2011, October). *Seizing the measurement moment: Why now is the time for states to help high schools get the postsecondary data they need and want.* College Summit. Retrieved from http://www.collegesummit.org/images/uploads/CollegeSummitWhitePaper2011.pdf

6 U.S. Department of Education. *Standards and assessments* (p. 25).

7 Phillips, G. (2010, October). *International benchmarking: State education performance.* AIR. Retrieved from http://www.air.org/files/AIR_Int_Benchmarking_State_Ed__Perf_Standards.pdf

8 Phillips. (2010). *International benchmarking.*

9 Data Quality Campaign. (2012, November). *Data for action 2012: Focus on people to change data culture* (p. 13). Retrieved from http://dataqualitycampaign.org/files/DFA2012%20Annual%20Report.pdf

10 Domaleski, C., & Perie, M. (2012, November). *Promoting equity in state education accountability systems.* Center for Assessment. Retrieved from http://www.nciea.org/publication_PDFs/Promoting%20Equity%20CSDMP110712.pdf

11 National Alliance for Public Charter Schools (June 27, 2013). "National Charter School Waitlist Numbers Approach One Million." Retrieved from http://www.publiccharters.org/pressreleasepublic/default.aspx?id=1068.

12 Bushaw, William J. and Lopez, Shane J. (September 2013). The 45th Annual *PDK/Gallup Poll of the Public's Attitudes Toward the Public Schools.* Arlington, VA: PDK International. Retrieved from http://pdkintl.org/noindex/2013_PDKGallup.pdf.

13 For more information, see: Center on Reinventing Public Education. (n.d.). *Portfolio strategy.* Retrieved from http://www.crpe.org/portfolio

14 For more information, see: The School District of Philadelphia. (n.d.). *School performance index.* Retrieved from http://webgui.phila.k12.pa.us/offices/a/accountability/school-performance-index-spi; Denver Public Schools. (n.d.). *School performance framework.* Retrieved from http://spf.dpsk12.org/

15 For more information, see: Chicago Public Schools. (n.d.). *2012 school progress report cards.* Retrieved from http://www.cps.edu/schools/pages/2012schoolprogressreportcards.aspx; NYC Department of Education. (n.d). *Progress reports: Methods and documentation.* Retrieved from http://schools.nyc.gov/Accountability/tools/report/default.htm#Methods

16 For examples, see: District of Columbia Public Charter School Board. (n.d.). *Performance management framework: Frequently asked questions.* Retrieved from http://www.dcpubliccharter.com/data/images/pmf%20frequently%20asked%20questions.pdf; Charter School Growth Fund. (n.d). *Portfolio performance: Academics.* Retrieved from http://chartergrowthfund.org/wp-content/uploads/2012/11/2011-Scorecard-Charts-Final-website.pdf; Texas Charter Schools Association. (n.d) *Quality framework.* Retrieved from http://www.txcharterschools.org/members/quality-framework.php#CommQualContImp

17 National Association of Charter School Authorizers. (2012). *Principles and standards for quality charter school authorizing.* Retrieved from http://www.qualitycharters.org/publications-resources/principles-standards

18 For examples, see the New Orleans Parents' Guide website, http://neworleansparentsguide.org/, and the School Choice for Kids website, http://www.schoolchoiceforkids.org/english.php

19 Table 3 includes only components that are included in school rating calculations, and does not include data presented to the public for informational purposes. For example, the new school report cards in South Carolina present the grade each school earned, supplemented by enrollment information and a comparison to the performance of similar schools. Because the similar schools comparison does not contribute to the school grade, it is not considered a component of the rating system and is not presented in Table 3. In-depth profiles of seven of the rating systems are presented in Appendix A.

20 Durant, S. & Dahlin, M., (2011, June). *The state of proficiency: How student proficiency rates vary across states, subjects, and grades between 2002 and 2010.* Kingsbury Center at NWEA. Retrieved from: http://www.kingsburycenter.org/sites/default/files/_State%20of%20Proficiency%202011_0628.pdf.

21 Cronin, J., Dahlin, M., Adkins, D., & Kingsbury, G., (2007). *The proficiency illusion.* The Fordham Institute. http://www.edexcellence.net/publications/theproficiencyillusion.html

22 Mathematica Policy Research. (n.d.). *Using value-added growth models to track teacher and school performance.* Retrieved from http://www.mathematica-mpr.com/education/value_added.asp

23 Betebenner, D.W. (2009). *A primer on student growth percentiles.* The Center for Assessment. Retrieved from http://www.azed.gov/research-evaluation/files/2011/07/growth_percentile_primer_030809.pdf

24 Domaleski, C., & Perie, M. (2012, November). *Promoting equity in state education accountability systems.* Center for Assessment. Retrieved from http://www.nciea.org/publication_PDFs/Promoting%20Equity%20CSDMP110712.pdf; U.S. Department of Education. (n.d). *ESEA flexibility.* Retrieved from: http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html

25 Castellano, K., & Ho, A. (2013, February). *A practitioner's guide to growth models.* The Council of Chief State School Officers. Retrieved from Retrieved from: http://scholar.harvard.edu/files/andrewho/files/a_pracitioners_guide_to_growth_models.pdf

26 Ernst, J., & Wenning, R. (2009, July). *Leave no charter behind: An authorizer guide to the use of growth data.* National Association of Charter School Authorizers. Retrieved from http://www.qualitycharters.org/images/stories/publications/Issue_Briefs/July09_Issue_Brief_Growth_Data.pdf

27 Castellano, K.E., & Ho, A.D. (2013). *A practitioner's guide to growth models.* Council of Chief State School Officers. Retrieved from http://scholar.harvard.edu/files/andrewho/files/a_pracitioners_guide_to_growth_models.pdf. Domaleski & Perie. *Promoting equity;* Auty, W. (2008). *Implementer's guide to growth models.* Washington, DC: The Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Documents/2008/Implementers_Guide_to_Growth_2008.pdf

28 May, H., Perez-Johnson, I., Haimson, J., Sattar, S. & Gleason, P. (2009, November). *Using State Tests in Education Experiments: A Discussion of the Issues* (NCEE 2009-013). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from: http://ies.ed.gov/ncee/pdf/2009013.pdf.

29 Examples of widely utilized norm-referenced tests in public schools include the Stanford Achievement Test series (SAT-10), the Iowa Test of Basic Skills (ITBS), and the Measures of Academic Progress (MAP) assessment developed by the Northwest Evaluation Association (NWEA).

30 Davis, M.R. (2012, October 17). Adaptive testing evolves to assess common-core skills. *Education Week.* 6(1),12–16. Retrieved from http://www.edweek.org/dd/articles/2012/10/17/01adaptive.h06.html

31 Schramm, J.B. & Zalesne, (2011). E.K. *Seizing the measurement moment.* College Summit. Retrieved from: http://www.collegesummit.org.

32 Data Quality Campaign. *Data for action 2012* (p. 10).

33 Data Quality Campaign. *Data for action 2012* (p. 9).

34 For more information, see: The Colorado Department of Education. (n.d.) *Schoolview.* Retrieved from http://www.schoolview.org/

35 Que, Nancy. (2013, January 16). Telephone interview by Tim Field and Lyria Boast.

36 The Broad Prize for Urban Education. (n.d.). *Process and criteria.* Retrieved from http://www.broadprize.org/about/process.html

37 Smarter Balanced. (n.d.). *Member states.* Retrieved from http://www.smarterbalanced.org/about/member-states/; Partnership for Assessment of Readiness for College and Careers. (n.d.). *About PARCC.* Retrieved from http://www.parcconline.org/about-parcc

38 Alapo, L. (2010, September 20). TCAP scores might be a jolt. Knoxville News Sentinel. Retrieved from: http://www.knoxnews.com/news/2010/sep/20/tcap-scores-might-be-jolt/

APPENDIX A:

California Charter Schools Association: Accountability Framework

Overview

The Member Council of the California Charter Schools Association (CCSA) developed the Accountability Framework to set minimum expectations of charter school academic performance. The system helps California charter schools examine their performance and compare it with other schools in the state serving similar student populations. It also assists CCSA in determining, based on school quality, what support to give schools during renewal periods.

CCSA staff and its member council developed the framework over three years with support from technical experts. CCSA developed the model using the following criteria:
1. Simplicity and transparency
2. Reliance on publicly available data
3. Applicability to both charter and traditional schools
4. Ability to highlight both high-performing and low-performing schools
5. Fair assessment of schools serving traditionally low-performing students

CCSA staff piloted the model in 2011, using data from the 2009–10 school year and conducted an in-depth study of the 58 lowest performing charter schools in California to test the model's validity. Results of the final framework were published for the 2010–11 and 2011–12 school years.

Methodology 2011–12

COMPONENTS OF THE FRAMEWORK

The framework includes three components:

1. Academic Performance Index (API) score
The California Department of Education (CDE) calculates an API score for all California public schools annually, based on proficiency rates in all subject assessments the CDE administers.

2. Change in API over time
CCSA uses change in API for the past three years to evaluate improvement over time for each school.

3. Similar-School Measure (SSM)
Using API scores and school-level student characteristics for each school, CCSA uses regression analysis to assess whether the school has performed above or below its expected performance in a given year, taking into consideration its student population. The SSM is calculated for the current and prior two years.

2nd Look Process: CCSA allows schools that fall below criteria 1-3 the opportunity to provide supplemental student-level data to show they are adding value.

DATA SOURCES

All of the data used to complete the framework are publicly available from CDE.

California's standard state exam is the California Standards Test (CST). Students take CSTs in the following subjects, based on their grade level:

Grades 2–8:
- ELA
- Math (including algebra 1 if applicable)
- Science (grades 5 and 8)
- History (grade 8)

High school
- ELA (all students, grades 9–11)
- Math
  - General mathematics
  - Algebra 1, 2
  - Geometry
  - Integrated mathematics 1–3
  - Summative high school mathematics
- Science (at least one exam in addition to general science)
  - General science (grade 10)
  - Biology
  - Chemistry
  - Earth science
  - Physics
  - Integrated coordinated science 1–4
- History
  - U.S. history (grade 11)
  - World history

APPLYING THE FRAMEWORK

Based on the scores of the three components, CCSA determines whether each school is above or below its minimum recommended criteria for renewal, described below. Though schools receive scores on each component, the final result is either "above minimum criteria" or "below minimum criteria." For schools that fall "below minimum criteria," CCSA provides schools with an opportunity to provide additional student-level data to show they are adding value in a "2nd look" process.

The framework only evaluates schools that have been open for at least four years, have at least 20 valid test scores in their API,[1] and do not have a CDE Alternative Schools Accountability Model (ASAM) designation.

Calculating the components:

1. Academic Performance Index (API)
To calculate the API, CDE awards the following points based on individual student proficiency levels on CST results on all subjects assessed:
- Advanced = 1,000 points
- Proficient = 875 points
- Basic = 700 points
- Below basic = 500 points
- Far below basic = 200 points

CDE calculates an average point score for each subject assessed; the API is the weighted average for all subjects and grade levels assessed.[2]

Though not part of the CCSA framework, CDE also calculates an API for each eligible subgroup, which includes ethnic and racial minority groups, socioeconomically disadvantaged students, English language learners, and students with disabilities. To be included, the student subgroup must contain at least 100 students, or at least 50 students who make up 15 percent or more of the school's total population.[3]

2. Change in API over time
CCSA evaluates four years of API data to assess the change in API over three time periods. For example, the change in API calculated for 2012 would include changes in:
- Period 1: 2008–09 to 2009–10
- Period 2: 2009–10 to 2010–11
- Period 3: 2010–11 to 2011–12

The change in API over time equals the sum of change across the three periods (Period 1 growth + Period 2 growth + Period 3 growth).

CDE has made numerous changes to the subjects assessed annually, affecting the number of subjects included in API calculations. To account for the frequent changes to the API methodology, CDE publishes "growth" and "base" API results for all schools annually. CDE calculated the 2012 Growth API with 2011–12 CST results and the 2011 API methodology, and calculated the 2012 Base API using 2011–12 CST results and the new methodology adopted for 2012. To calculate growth, or change, in API, the first year's base API would be compared with the second year's growth API.

3. Similar-School Measure (SSM)

CCSA uses regression analysis to evaluate each school's annual performance, taking into consideration the following school-level student characteristics, all available from CDE:

- Percentage enrollment of students qualifying for free or reduced-price lunch
- Percentage enrollment of English language learners
- Percentage enrollment of students with disabilities
- Percentage enrollment of students by ethnicity
- Education level of parents
- Student retention rates
- Grades enrolled

The framework compares the school's actual API score to the expected API calculated by the SSM regression and assigns one of five ratings for each of the past three years:

- Far above
- Above
- Within range of
- Below
- Far below

The framework includes three years of SSM results to assess whether schools are consistently out-performing or under-performing their expected API.

Assigning a rating:

The framework results in a rating of "above minimum criteria" or "below minimum criteria" for each school. To earn the "above" rating, schools must meet at least one of the following targets:

1. API score of 700 or above[4]
2. API growth over 3 years of at least 50 points
3. SSM results better than "below" for two of the past three years.

Results are shared with schools and schools have the opportunity to review results before they are publically released and provide additional student-level data for the 2nd look process.

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

CCSA publishes report cards annually for every California charter school. They include school results on each of the three framework components and indicate whether the school meets CCSA's requirements for support at renewal.

CCSA also provides a series of snapshot reports that permits comparison of charter schools of various types and non-charter schools across the state, county, and district. Using API scores and SSM data, the website allows the user to compare schools by a variety of measures, including enrollment, subgroup performance, and performance over time. Individual schools can compare their academic performance to geographically proximate non-charter schools, higher performing charter schools with similar demographics, or charter schools serving similar high-risk populations.

PUBLIC REACTION

Advocates for increased charter school accountability praised the CCSA framework and its recommendations for renewal standards. CCSA participated in the National Association of Charter School Authorizers' launch of the "Million Lives Campaign," which advocates for the closure of the lowest-performing charter schools. A recent editorial in the LA Times lauded CCSA's efforts to increase expectations for charter school performance.[5]

Critics of the CCSA framework complain that it relies too heavily on state assessment data. Gov. Jerry Brown opposed a bill to strengthen assessment-based performance standards for charter renewal, and has suggested that qualitative measures of school quality should play a larger part in renewal decisions.[6]

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

32

Appendix A — Endnotes

1 Beginning in 2013–14, the framework will only evaluate schools with at least 50 valid test scores in their API.

2 English language arts and math tests receive higher weights. For more information, see: California Department of Education. (2012, May). *2011–12 Academic Performance Index reports: Information guide.* Retrieved from http://www.cde.ca.gov/ta/ac/ap/documents/infoguide12.pdf. The final score is adjusted with the Scale Calibration Factors (SCF), which are applied to maintain consistency in the API across years, given frequent changes to the API methodology.

3 California Department of Education. (2012, May). *Parent and guardian guide to California's 2011–12 Accountability Progress Reporting System.* Retrieved from http://www.cde.ca.gov/ta/ac/ay/documents/parentguide12.pdf

4 In 2013–14, Criteria #1 will be changed from an API score of 700 or above to an API score above the 25th percentile of the statewide distribution. This threshold will gradually increase to 33rd percentile over five years.

5 Charter schools – a report card. [Editorial]. (2013, July 15). *Los Angeles Times.* Retrieved from http://www.latimes.com/news/opinion/editorials/la-ed-charter-schools-stanford-study-20130713,0,1522096.story

6 Fensterwald, J. (2012, February 27). *Charter movement's U-shape: Clumps of high and low performers.* Silicon Valley Education Foundation. Retrieved from http://toped.svefoundation.org/2012/02/27/charter-movements-u-shaped-dilemma/

APPENDIX B:

Colorado: School Performance Framework

Overview

Colorado released its School Performance Framework (SPF) for the 2010–11 school year, but components of the framework had been in use and in development for years. As early as 2003, the Colorado Association of School Executives (CASE) and the Donnell-Kay Foundation supported the creation of the Colorado Accountability Project, which recommended improvements to the state accountability system. One of the cornerstones of the resulting accountability system is the Colorado Growth Model, referred to generally as the Student Growth Percentile model. The Colorado Growth Model was developed in partnership with the Colorado Charter School Institute and is used by the Colorado League of Charter Schools to report on charter school quality. The Colorado Growth Model was adopted by the state in 2008 and since then more than 20 additional states have begun to use the model as an important component of school rating systems.

In 2009, state legislation required a system of aligned state, district, and individual school performance framework reports that use a common set of metrics to report on academic quality. Improvements to the state assessment system followed, including expanded postsecondary and workforce readiness standards. Colorado was one of the first states to apply for an ESEA flexibility waiver in 2012. The system approved under the waiver is a more unified version of the old model, with improved public reports and online interactive tools for viewing student, school, and district performance.

Methodology 2011–12

COMPONENTS OF THE FRAMEWORK

Colorado's SPF includes four indicators and eight measures:

1. **Academic Achievement**—Applies to elementary, middle, and high schools.
- Proficiency rate in:
   - Reading
   - Math
   - Writing
   - Science

2. **Academic Growth**—Applies to elementary, middle, and high schools. Growth is measured in each subject using the Transitional Colorado Assessment Program (TCAP) and the Colorado English Language Assessment (CELA).[1] The Colorado Growth Model evaluates both peer-referenced growth (using median growth percentile) and criterion-referenced growth (using adequate growth percentile).
- Median Student Growth Percentile (MGP)
   - Reading
   - Math
   - Writing
- Adequate Student Growth Percentile (AGP)
   - Reading
   - Math
   - Writing

3. **Academic Growth Gaps**—Applies to elementary, middle, and high schools.
- Difference between MGP and AGP (calculated separately for reading and math) for:
   - Student eligible for free or reduced-price lunch
   - Minority students (non-white or Hispanic)
   - Students with disabilities
   - English language learners
   - Students needing to catch up (below proficient in prior year)

4. **Postsecondary and Workforce Readiness**—Applies to high schools.
- Graduation rate (highest of the 4-, 5-, 6-, or 7-year adjusted cohort rate is included in the final score)
- Graduation rate disaggregated by subgroup(highest of the 4-, 5-, 6-, or 7-year adjusted cohort rate is included in the final score):

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

34

- ■ Students eligible for free or reduced-price lunch
- ■ Minority students (non-white or Hispanic)
- ■ English language learners
- ■ Students with disabilities
■ Dropout Rate
■ Colorado ACT composite score (administered to all 11th-graders statewide)

DATA SOURCES

All data used in calculating the SPF results are collected by the Colorado Department of Education, with the exception of ACT results, which CDE receives from the College Board.

Colorado measures student performance using the Transitional Colorado Assessment Program (TCAP), first used in 2011–12. The TCAP tests students in reading, writing, and math in grades 3 through 10, and in science in grades 5, 8, and 10. A Spanish version accommodates students in reading and writing in grades 3 and 4.[2]

SCORING THE FRAMEWORK

The state takes three steps to assign each school a final rating:

1. Assign points for each measure

Schools receive points for each measure based on whether the school or subgroup exceeded, met, was approaching, or did not meet the standard. The cut-off points for each category are higher for schools and subgroups that did not meet the median adequate student growth percentile (see table below).

Measure	Targets	Points
Academic Achievement (Proficiency) (points are assessed separately for reading, math, writing, and science) Percentile rates are based on 2009–10 proficiency rates statewide.	Proficiency rate meets or exceeds the 90th percentile of all schools statewide.	4
	Proficiency rate is between the 50th and 89th percentiles of all schools statewide.	3
	Proficiency rate is between the 15th and 49th percentiles of all schools statewide.	2
	Proficiency rate is below the 15th percentile of all schools statewide.	1
Academic Growth (points are assessed separately for reading, math, and writing)	AGP meets the state median and SGP is at or above 60 or AGP falls below the state median and SGP is at or above 70.	4
	AGP meets the state median and SGP is below 60 but at or above 45 or AGP falls below the state median and SGP is below 70 but at or above 55.	3
	AGP meets the state median and SGP is below 45 but at or above 30 orAGP falls below the state median and SGP is below 55 but at or above 40	2
	AGP meets the state median and SGP is below 30 or AGP falls below the state median and is below 40.	1
Academic Growth Gaps (points are assessed for reading and math for each subgroup)	AGP meets the state median and SGP is at or above 60 or AGP falls below the state median and SGP is at or above 70.	4
	AGP meets the state median and SGP is below 60 but at or above 45 or AGP falls below the state median and SGP is below 70 but at or above 55.	3
	AGP meets the state median and SGP is below 45 but at or above 30 or AGP falls below the state median and SGP is below 55 but at or above 40	2
	AGP meets the state median and SGP is below 30 or AGP falls below the state median and is below 40.	1

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

35

Graduation Rate	At or above 90%	4
	Above 80% but below 90%	3
	At or above 65% but below 80%	2
	Below 65%	1
Dropout Rate	At or below 1%	4
	At or below the state average but above 1%	3
	At or below 10% but above the state average	2
	At or above 10%	1
Average Colorado ACT Composite Score	At or above 22	4
	At or above the state average but below 22	3
	At or above 17 but below the state average	2
	At or below 17	1

2. Calculate weighted score

Next, the state adds the points the school earns for each measure within the indicators and weights each indicator total to calculate a final score (see table below). The SPF places a high weight on growth; the combined weight of the Student Growth and Growth Gaps indicators counts toward 75 percent of the final elementary and middle school scores, and toward 50 percent of the final high school score.

Indicator	Elementary/Middle School Weight	High School Weight
Academic Achievement (Proficiency)	25%	15%
Academic Growth	50%	35%
Academic Growth Gaps	25%	15%
Postsecondary and Workforce Readiness	N/A	35%

3. Assign rating

Based on the final weighted score, schools are placed into one of four performance categories:
Accredited, Improvement, Priority Improvement, or Turnaround (see table below).

Rating	Score Range
Accredited	At least 60
Improvement	47 to 59
Priority Improvement	22 to 46
Turnaround	Less than 22

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

36

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

Colorado provides detailed reports for each school online on the SchoolView website. For each school, it shows an overview of the evaluation metrics, a detailed breakdown of the school's performance on each metric, and the overall school rating. The website also includes state and district data on accountability, performance, demographics, and finance. The Schoolview tool presents detailed graphs and tables with results of the Colorado Growth Model and allows users to compare achievement and growth across schools. Using password-protected login access, parents and students can view individual student proficiency and growth results.

PUBLIC REACTION

The SPF and SchoolView reporting format have received praise for focusing on closing achievement gaps and for presenting comprehensive data on school performance.[3] Critics say that the school results are difficult to decipher and that the SPF designations do not effectively differentiate levels of school quality.[4] (When results of the current SPF system were first released in 2010, 62 percent of schools received the highest rating.) In response, Denver Public Schools created an additional category, Distinguished, and a coalition of community organizations came together to create Colorado School Grades, which publishes school report cards that assign schools an A–F grade using the SPF score.[5]

Parts of SPF have received national attention. More than 20 states have adopted the Colorado Growth Model, making it the most commonly used growth model in state accountability systems across the country. In addition, a number of states are in the process of contracting to use the SchoolView web tool for displaying school and student results.

LINKS TO PUBLIC REPORTS

- Schoolview Data Center—accountability results, academic performance, student demographics: http://www.schoolview.org/
- School Performance Framework Results—interactive map: http://www.cde.state.co.us/Accountability/PerformanceFrameworks.asp
- Colorado Growth Model School Comparison Tool: http://www.schoolview.org/ColoradoGrowthModel.asp
- Colorado Performance Framework Online Tutorial: http://www.cde.state.co.us/media/training/SPF_Online_Tutorial/player.html

Appendix B — Endnotes

1 The Colorado Department of Education (CDE) will begin using the ACCESS assessment, developed by the World-Class Instructional Design and Assessment (WIDA) consortium, to assess ELL proficiency starting in 2012-13.

2 Colorado Department of Education. (n.d.). *About CSAP/TCAP*. Retrieved from http://www.cde.state.co.us/assessment/CoAssess-About.asp

3 McGraw, C., & Iodice, K. (2010, December 11). School district rankings point out strengths, weaknesses. *Colorado Springs Gazette*. Retrieved from http://www.gazette.com/articles/school-109424-received-education.html#ixzz2ODSNXxDi

4 Carroll, V. (2013, Februrary 6). Colorado School Grades tell the truth about our schools. *The Denver Post*. Retrieved from http://www.denverpost.com/carroll/ci_22526096/carroll-telling-truth-about-our-schools

5 Mitchell, N. (2010, November 3). State releases new school ratings. *EdNews Colorado.* Retrieved from http://www.ednewscolorado.org/news/education-news/state-releases-new-school-ratings; Colorado School Grades. (n.d.). *About us*. Retrieved from http://coloradoschoolgrades.com/AboutUs.aspx

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

37

APPENDIX C:

District of Columbia Public Charter School Board: Performance Management Framework

Overview

The District of Columbia Public Charter School Board (PCSB) developed the Performance Management Framework (PMF) to guide its oversight of Washington, D.C., charter schools. PCSB previously used individualized academic evaluations that varied from one charter school to another. In contrast, the PMF provides a set of common academic measures that PCSB applies consistently to all charter schools.

PCSB developed the PMF over three years, working with stakeholders, charter schools, technical consultants, and staff to develop, test, and revise the system. The result is a framework that includes 16 measures under four indicators—student growth, student achievement, gateway measures (postsecondary readiness and success), and leading indicators (student engagement).

PCSB piloted the PMF using academic data for the 2008–09 school year, and launched it in the 2009–10 school year.

 "The board took its time to make sure we got it right," former PCSB Chairman Brian Jones said.[1] "We spent the past seven months talking to stakeholders, national experts, and especially schools who played a huge part in the development and validation of the PMF."

Methodology 2011–12

COMPONENTS OF THE FRAMEWORK

The PMF includes four indicators and 16 measures:

1. **Student Growth** ("Progress")—Applies to elementary, middle, and high schools.
■ School median growth percentile (MGP)
   ■ Math
   ■ Reading

2. **Student Achievement/Proficiency** ("Achievement")—Applies to elementary (tested grades 3-5), middle, and high schools (10th grade)
■ Percentage of students achieving proficiency in:
   ■ Math
   ■ Reading
■ Percentage of students achieving advanced proficiency in:
   ■ Math
   ■ Reading
■ AP/IB performance (high schools only)

3. **Postsecondary Readiness and Success** ("Gateway measures")—Applies to elementary, middle, and high schools.
■ 3rd-grade reading proficiency rate (elementary schools only)
■ 8th-grade math proficiency rate (middle schools only)
■ Graduation rate (four-year adjusted cohort rate)
■ PSAT performance (11th grade)
■ SAT/ACT performance (12th grade)
■ College acceptance rate

4. **Student Engagement** ("Leading indicators")—Applies to elementary, middle, and high schools.
■ Attendance rate
■ Re-enrollment rate
■ 9th-grade credits on track (high schools only)

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

38

DATA SOURCES

PCSB collects the data necessary to complete the PMF from charter schools and the Office of the State Superintendent of Education (OSSE).

OSSE overseas the DC Comprehensive Assessment System (DC CAS), which assesses student proficiency and mastery of the DC Content Standards. Tested subjects include reading (grades 3–8 10), composition (grades 4, 7, and 10), mathematics (grades 3–8 and 10), science (grades 5 and 8), and health (grades 5 and 8). High school students are tested in biology and health.

DC CAS proficiency rates and growth results (median growth percentiles) are calculated annually by OSSE and provided to PCSB. PCSB also receives SAT, PSAT, and graduation rates from OSSE and re-enrollment using the annual school audits.

Charter schools submit 9th grade credit accumulation reports, college admission information, and attendance data directly to PCSB. PCSB calculates the rates and verifies these submissions before they are included in the PMF calculations.

SCORING THE FRAMEWORK

1. Assign points for each measure

Each measure of the framework receives a score from 0 to 100. PCSB established "floors" and "targets" for each measure which are reviewed annually. The floor represents the minimum performance the school must demonstrate before receiving any points for the measure; schools can receive the maximum 100 points by meeting or exceeding a measure's target. Schools receive a prorated value for scores that fall between the floor and the target. The floors and targets for elementary school (ES), middle school (MS), and high school (HS) for school year 2011-2012 were:

Component	Measure	ES Floor	MS Floor	ES/MS Target	HS Floor	HS Target
Student Growth	Median Growth Percentile—Reading	30	30	70	30	65
	Median Growth Percentile—Math	30	30	70	30	65
Student Achievement/ Proficiency	Reading—% of proficient and advanced students	24	28	100	26	100
	Math—% of proficient and advanced students	15	24	100	20	100
	Reading—% of advanced students	0	0	25	0	25
	Math—% of advanced students	0	0	25	0	25
	AP and IB Performance (12th grade)	N/A	N/A	N/A	0	15
Postsecondary Readiness and Success	3rd-Grade Reading—% of proficient and advanced students (used as an early indicator of success)	17	N/A	100	N/A	N/A
	8th-Grade Math—% of proficient and advanced students (used as an early indicator of success)	N/A	22	100	N/A	N/A
	Graduation Rate	N/A	N/A	N/A	57	100
	PSAT[2] (11th grade)	N/A	N/A	N/A	3	50
	SAT/ACT[3] (12th grade)	N/A	N/A	N/A	10	75
	College Acceptance Rate	N/A	N/A	N/A	63	100
Student Engagement	Attendance Rate	85	85	95	85	95
	Re-enrollment Rate	57	57	90	64	90
	% of 9th-graders on track to graduate	N/A	N/A	N/A	44	100

Note: PMF floor and target values are reviewed, and if necessary, revised, annually.
Source: DC Public Charter School Board. (2011, November). PMF guidelines and technical guide. Retrieved from http://www.dcpcsb.org/data/images/2010-2011%20pmf%20guidelines%2011_1_11.pdf

2. Calculate weighted score

The PMF also includes an overall school score comprising a weighted average of the individual component scores, which emphasizes the importance of some components over others (see table below).

Component	Elementary/Middle School Weight	High School Weight
Student Growth	40%	15%
Student Achievement/Proficiency	25%	30%
Postsecondary Readiness and Success	15%	30%
Student Engagement	20%	25%

3. Assign rating

The PMF then places each school within a tier ranging from Tier 1 (high-performing) to Tier 3 (low-performing) based on the overall scores (see table below).

Rating	Final Overall Score	Percentage of Schools in Rating Category 2011–12
Tier 1—High-performing	65% or higher	31%
Tier 2—Mid-performing	35% to 64.9%	55%
Tier 3—Low-performing	Less than 35%	14%

Source: 2012 DC public charter school performance reports. Retrieved from
http://www.dcpubliccharter.com/data/images/dc%20pcsb%20pmf%20book_nov2.pdf

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

PCSB produces annual report cards for each charter school, published in its annual PCSB School Performance Report. The report cards present the overall school rating, together with results in a format that is easy to interpret for each component of the framework. The reports also provide demographic information about the school's students and the school mission.

In 2012, PCSB published the Parent Guide to Public Charter School Performance summarizing the PMF methodology and providing guidance on using the report cards to make school selection decisions. The guide is available online, at libraries, and at other government and community organizations (see link under "Links to public reports," below).

PUBLIC REACTION

Public reaction to the school report cards has generally been favorable; the PCSB credits this to stakeholder participation in developing the framework. Schools' concerns were addressed early on as they participated in PMF development and revisions, so the PMF results did not surprise them. Media reports portrayed the PMF as a useful tool for parents and the public. "While some of the information in the assessments is already available in annual performance reports, the new system creates a more detailed and easily accessible snapshot for parents and families" The Washington Post reported.[4]

LINKS TO PUBLIC REPORTS

- School performance reports: http://www.dcpcsb.org/Oversight/PMF-Results.aspx
- Parent guide: http://www.dcpcsb.org/data/images/pcsb%20mini%20guide-nov2-spreads.pdf
- System overview: http://www.dcpubliccharter.com/PCSB-Accountability/Performance-Management-Framework.aspx

Appendix C — Endnotes

1 DC Public Charter School Board. (2011, December 6). *Charter School Board announces list of high performing charter schools; report card shows schools' annual performance (press release).* Retrieved from http://www.dcpubliccharter.com/News-Room.aspx?id=232

2 Percentage of students with combined math and verbal PSAT scores of 80 or higher.

3 Percentage of students achieving math and verbal SAT scores of 800 or higher or an ACT score of 16 or higher.

4 Turque, B. (2011, December 6). District unveils first ranking of public charter schools. *The Washington Post.* Retrieved from http://articles. washingtonpost.com/2011-12-06/local/35286978_1_charter-schools-school-report-cards-charter-campuses

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

41

APPENDIX D:

Florida: School Grades

Overview

In 1998, Jeb Bush won election as the governor of Florida on a platform of accountability and school choice. In response, Florida state legislation, passed in 1999, established the mandate for Florida's School Grades system. The system assigns an A–F grade to every public school to identify schools in need of intervention and to publicly acknowledge the state's highest-performing public schools. Florida was the first state to implement and report an A–F grade for schools each year.

Since the implementation of No Child Left Behind (NCLB), Florida has maintained two accountability systems—the A–F grades and NCLB school report cards. The A–F system has been revised several times since its adoption, adding and expanding some metrics and, most recently, increasing the cut points. Each adjustment resulted in a year or two of transition for the distribution of school grades. The state education department made concerted communication efforts at each transition to address school, parent, and community concerns about the changes and the results.

Starting with the 2012–13 school year, the state will report only A–F grades in compliance with Florida's ESEA Flexibility Waiver Application, approved in July 2012. The waiver development process took into account feedback from students, schools, teachers, and community stakeholders.[1]

Methodology 2011–12

COMPONENTS OF THE FRAMEWORK

The A–F grading system includes three indicators and eight measures:

1. **Student Growth** ("Learning gains")—Applies to elementary, middle, and high schools. Percentage of students making learning gains[2] from spring to spring in math and English:
   - All students
     - Math
     - English
   - Lowest-performing 25% of students
     - Math
     - English

2. **Student Achievement/Proficiency** ("Performance")—Applies to elementary, middle, and high schools. Percentage of students meeting proficiency on the state assessments in:
   - Math
   - Reading
   - Writing
   - Science

3. **Postsecondary Readiness and Success** ("Other components")—Applies to high schools, although middle schools may earn bonus points for the percentage of students participating in and performing well on high school-level courses/tests). Measures include:
   - Participation in advanced curricula
   - Performance in accelerated curricula
   - Graduation rate (federal four-year rate and federal five-year modified rate)
   - Graduation rate for at-risk students
   - College readiness—Percentage of on-time graduates scoring college-ready reading and math scores based on ACT, SAT, CPT, or PERT results

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

42

DATA SOURCES

The Florida Department of Education uses its longitudinal data system to collect data for all of the measures evaluated.

The FCAT or Florida Comprehensive Assessment Test was introduced in 1998 for grades 3–11. It consisted of criterion-based assessments in math, reading, science, and writing and was introduced in an effort to raise achievement standards. Transition to the FCAT 2.0 began in 2010–11 and was completed in the 2011–12 school year. (FCAT writing and FCAT reading and math retakes were still administered in 2011–12).

SCORING THE FRAMEWORK

The state takes three steps to assign each school a final rating:

1. Assign points for each measure
Schools earn points for each measure based on their performance on each. The maximum amount of points a school can earn differs by measure, from 50 to 150 points.

Indicator	Measure	Elementary/Middle school points equal:	High school points equal:	Max Points
Student Growth	Learning Gains Math, all students	Percentage of students making learning gains*	Percentage of students making learning gains*	100
	Learning Gains English, all students			100
	Learning Gains Math, lowest-performing 25% of students	Percentage of lowest-performing 25% of students making learning gains*	Percentage of lowest-performing 25% of students making learning gains*	100
	Learning Gains English, lowest-performing 25% of students			100
Student Achievement	Proficiency Rate Math	Percentage of students achieving proficiency	Percentage of students achieving proficiency	100
	Proficiency Rate Reading	Percentage of students achieving proficiency	Percentage of students achieving proficiency	100
	Proficiency Rate Writing	Percentage of students achieving proficiency	Percentage of students achieving proficiency	100
	Proficiency Rate Science	Percentage of students achieving proficiency	Percentage of students achieving proficiency	100

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

43

Postsecondary Readiness and Success	Participation in Advanced Curricula AP, IB, AICE, industry certifications or dual enrollment	N/A	Percentage of participating 11th- and 12th-graders times 1.5	150
			Percentage of participants eligible to earn college credit times 1.5	150
	Performance in Accelerated Curricula	MS only—Percentage of students taking high school-level EOCs divided by 2	N/A	50
		MS only—Percentage of students scoring 3 or higher divided by 2	N/A	50
	Graduation Rate Federal four-year rate	N/A	4-year grad rate	100
	Graduation Rate Federal five-year modified rate	N/A	5-year grad rate	100
	At-risk Graduation Rate Federal four-year rate	N/A	At-risk 4-year grad rate divided by 2	50
	At-risk Graduation Rate Federal five-year modified rate	N/A	At-risk 5-year grad rate divided by 2	50
	College Readiness Percentage of on-time graduates scoring college-ready reading and math scores based on ACT, SAT, CPT, or PERT results	N/A	Percentage of on-time graduates scoring college-ready—reading	100
		N/A	Percentage of on-time graduates scoring college-ready—math	100

* Students who move from non-proficiency to proficiency or who show higher than expected growth are counted 1.1 times towards the numerator in the calculation of percentage of students showing learning gains.

2. Calculate weighted score

The A–F model places an equal value on student growth and proficiency. For high schools, post-secondary readiness and success measures account for half of the points awarded in the model. The points awarded to each measure result in the following weight for each indicator:

Indicator	Elementary/Middle School Weight	High School Weight
Student Growth	50%	25%
Student Achievement/ Proficiency	50%	25%
Postsecondary Readiness and Success	N/A	50%

3. Assign rating

The final grade for each school is assigned using the point ranges shown below. In the 2011–12 school year, nearly half of all schools received an "A" grade.

Rating	Final Overall Score— Elementary (800 possible points)	Final Overall Score— Middle School (900 possible points)	Final Overall Score— High School (1500 possible points)	Percentage of Schools in Rating Category 2011–12
A	At least 525	At least 590	At least 1,050	47%
B	495–524	560–589	990–1,049	24%
C	435–494	490–559	870–989	19%
D	395–434	445–489	790–869	7%
F	Less than 395	Less than 445	Less than 790	2%

Note: Though not included in the overall score, assessment participation rates of 90 percent must be met. Schools that fail to meet participation requirements may receive a grade penalty.

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

School report cards with A–F grades and school information are published annually on the DOE website. Additional spreadsheets that present each of the components of the A–F model are also available.

PUBLIC REACTION

The Florida A–F grades have elicited both strong praise and harsh criticism. Supporters point to an increase in transparency and accountability, and underline improvements in student performance and school grades.

Opponents disagree with the emphasis placed on student testing. They question whether it is appropriate to rely on student assessment results when making high-stakes accountability decisions. They voice concern that labeling schools as low-performing also unfairly labels students, and argue that the state's system of awarding high-performing schools with financial incentives is unfair. Opponents were particularly alarmed when the state education department found an error in its learning gains calculation in 2012, which resulted in incorrect grades for 213 of the state's 2,586 schools.[3]

The 2012 changes to the state grading system resulted in a 90 percent increase in the number of schools receiving failing grades.[4] Proponents say that it is essential to periodically "raise the bar" and that even though this results in a temporary decrease in grades, schools have ultimately improved after every increase in rigor to the A–F system. In response, some parents and school leaders have continued to protest against the changes, especially where they affect schools that have worked to increase their grades from failing to "passing" over the past decade. They fear that with lower grades their schools will face the loss of financial incentives.[5]

LINKS TO PUBLIC REPORTS

- Report cards, 2011–12: http://schoolgrades.fldoe.org/
- Report cards, 1999–2011: http://schoolgrades.fldoe.org/default.asp?report=RC
- A–F guide: http://schoolgrades.fldoe.org/pdf/1112/Guidesheet2012SchoolGrades.pdf
- Technical guide to calculations in the grading system: http://schoolgrades.fldoe.org/pdf/1112/SchoolGradesTAP2012.pdf

Appendix D — Endnotes

1 U.S. Department of Education. (n.d.) *Florida ESEA flexibility request.* Retrieved from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/map/fl.html

2 Students are considered to show learning gains who improve one or more performance levels on the state assessment, maintain proficiency, or show more than a year's worth of growth based on scale score growth.

3 Strauss, V. (2012, July 23). Florida gives incorrect grades to hundreds of its public schools. *The Washington Post.* Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/florida-gives-incorrect-grades-to-hundreds-of-its-public-schools/2012/07/23/gJQA7uqg4W_blog.html

4 O'Donnell, C. (2012, July 11). Sarasota bucks state trend on school grades. *Sarasota Herald-Tribune.* Retrieved from http://www.heraldtribune.com/article/20120711/ARTICLE/120719943

5 Postal, L. (2012, February 25). Florida schools brace for tougher new grading system. *The Orlando Sentinel.* Retrieved from http://articles.orlandosentinel.com/2012-02-25/features/os-florida-school-grade-changes-20120225_1_grade-level-school-grades-face-state-oversight

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

45

APPENDIX E:

GreatSchools

Overview

A nonprofit, San Francisco-based organization, GreatSchools was founded in 1998 by Bill Jackson, a Silicon Valley entrepreneur, to support parents in making informed choices about their children's schooling and inspire parental involvement in schools.

The GreatSchools system includes ratings for more than 200,000 public, private, and charter schools across the country.[1] GreatSchools assigns every school a rating from 1 to 10, based on school proficiency rates by grade and subject. According to Alexa Internet, which tracks web traffic, GreatSchools is the most used school ratings website, attracting more than 40 million visitors annually.[2] In 2011 the U.S. Department of Housing and Urban Development partnered with GreatSchools to offer resources to parents in public housing on understanding school options.[3]

GreatSchools attracts funding from philanthropic organizations that support school choice. In 2011, these included the Bill & Melinda Gates Foundation, the Robertson Foundation, the Walton Family Foundation, and the Goldman Sachs Group.[4]

Methodology 2011–12

GreatSchools ratings are based solely on student proficiency rates.[5] For each grade and subject combination, GreatSchools compares proficiency rates at each school to all other schools across the state and assigns the school a decile ranking.[6] A school in the 10th decile ranks in the top 10 percent of schools in the state, while a school in the 1st decile is in the bottom 10 percent.

The decile rankings for all grade and subject combinations are averaged to arrive at an overall score for each school. For instance, in the example below, for a grades 6–8 middle school in a state that administers assessments to students in reading, math, and science, the GreatSchools rating would be 6, the average of all of the statewide decile rankings across grades and subjects at the school.

Sample Middle School Overall Rating Calculation

Subject/Grade Combination	Statewide Decile Ranking
6th-grade reading	6
6th-grade math	5
6th-grade science	8
7th-grade reading	7
7th-grade math	4
7th-grade science	6
8th-grade reading	7
8th-grade math	6
8th-grade science	9
Overall Rating (average of all combinations)	6

The site also reports proficiency rates by student subgroup, including ethnicity, gender, students qualifying for free or reduced-price lunch, English language learners, and students with disabilities, as reported by each state. These proficiency rates, however, do not factor into a schools' overall rating.

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

46

DATA SOURCES

GreatSchools uses publicly available state assessment results available from each state department of education.

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

The GreatSchools website allows users to search for individual school reports. The GreatSchools rating is presented along with grade-level state assessment results. Additional information about each school, including student demographics, per-pupil funding, and teacher credentials is provided. Schools may provide curriculum or programming information, though participation is voluntary. Parents and community members may enter comments or school reviews, compiled on the site.

PUBLIC REACTION

Although users generally consider Greatschools a useful resource, some critics see limited value in its ratings because they consider only standardized test results. These critics encourage users to also consider a variety of school rating tools,[7] such as school district report cards and other published school data. Others complain that that the 1–10 rating scale is too narrow to provide a useful comparison between schools.[8] Some have found it difficult to understand a school's 1–10 rating, and argue that there is little correlation between GreatSchools ratings and published school test results.[9]

In an effort to expand the ratings to include more metrics, GreatSchools is developing a more complex system that includes trend and school climate components. The system, piloted in Indianapolis and Milwaukee in October 2012, has five criteria, including robust teacher support, active family engagement, and high expectations within the school.[10] GreatSchools' long-term aim is to roll out the new system across the U.S.

LINKS TO PUBLIC REPORTS

- School reports: http://www.greatschools.org
- FAQ: http://www.greatschools.org/find-a-school/defining-your-ideal/2423-ratings.gs

Appendix E — Endnotes

1 Chaltain, S. (2012, July 5). What makes a great school? *Forbes*. Retrieved from http://www.forbes.com/sites/ashoka/2012/07/05/what-makes-a-great-school/. Private schools are included only if they report state standardized test results.

2 Samuels, C.A. (2012, April 3). GreatSchools finds a niche in school ratings: School ratings service has designs on parent market. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2012/04/04/27greatschools_ep.h31.html

3 For more information, refer to the HUD website: http://portal.hud.gov/hudportal/HUD?src=/program_offices/housing/mfh/nnw/nnwaboutnn/GreatSchools

4 Samuels. (2012). GreatSchools finds a niche.

5 Results are not reported for categories that contain fewer than 10 students.

6 Decile rankings are not calculated if fewer than 50 schools report proficiency results for a subject/grade combination.

7 Home by School. (2012, March 1). Are you only researching on GreatSchools when choosing a school? Retrieved from http://fairfax.homebyschool.com/are-you-only-using-greatschools-when-choosing-a-school/

8 Samuels. (2012). GreatSchools finds a niche.

9 Allegheny Institute for Public Policy. (2010, April 14). *GreatSchools ranking of schools flawed and misleadin*g. Retrieved from http://alleghenyinstitute.org/education/pittsburghpublicschools/374-greatschools-ranking-of-schools-flawed-and-misleading.html

10 Olivieri, S. B. (2013, January 23). *Evolution of the GreatSchools rating*. Retrieved from http://blogs.greatschools.org/greatschoolsblog/2013/01/evolution-of-the-greatschools-rating.html

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

47

APPENDIX F:

New York City
Performance Management Framework

Overview

In 2002, the New York State Assembly established mayoral control over New York City (NYC) schools, and Mayor Michael Bloomberg appointed Joel Klein as the schools' chancellor. Klein implemented extensive reforms, including the NYC Department of Education's publication of annual progress reports for all NYC schools, beginning with fall 2007 reports on the 2006–07 school year.

The progress reports provide the public with information on school quality and guide the district's accountability efforts. Higher-performing schools are eligible for financial rewards, while poorer results may trigger restructuring or school closure.

Methodology 2011–12

COMPONENTS OF THE FRAMEWORK

The NYC Progress Reports include five indicators composed of more than 30 individual measures:[1]

1.   Student Progress (Growth)—Applies to elementary, middle, and high schools
Elementary and Middle Schools
- Adjusted[2] school median growth percentile (MGP) for the school overall:
  - Math
  - English Language Arts (ELA)
- Adjusted school median growth percentile (MGP) for the school's lowest-performing third of students:
  - Math
  - ELA
- Early-grade progress—Third-grade proficiency of high-need students
  - Math
  - ELA

High Schools
- Credit accumulation
- Credit accumulation of the school's lowest-performing third of students
- End-of-course Regents[3] completion rates
- End-of-course Regents proficiency rates

2.   Student Performance—Applies to elementary, middle, and high schools
Elementary and Middle Schools
- Proficiency rate in:
  - Math
  - ELA
- Average proficiency level[4] in:
  - Math
  - ELA

Middle and K–8 Schools
- Course pass rates in:
  - English
  - Math
  - Science
  - Social studies

High Schools
- Graduation rate (4-year and 6-year cohort rates)
- Diploma quality (type of diploma earned—4-year and 6-year cohort rate)

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

48

3. **School Environment**—Applies to elementary, middle, and high schools.
- Student attendance
- Results of NYC Learning Environment Survey (annual survey of middle and high school students and parents)

4. **College and Career Readiness**—Applies to high schools only
- College and Career Preparatory Index
- 4-Year and 6-Year College Readiness Index
- Enrollment in post-secondary institutions

5. **Closing the Achievement Gap**—Bonus points available for elementary, middle, and high schools.
Elementary and Middle Schools
- Percentage of high-need students showing high student growth in math or ELA
- Percentage of high-need students achieving proficiency in math or ELA
- Movement of students with disabilities to less restrictive environments
- Movement of English language learners toward English language proficiency
- Percentage of 8th-grade students earning high school credit

High Schools
- Graduation rate for high-need students
- College and Career Preparatory Index for students in the lowest-performing third citywide
- College Readiness Index for students in the lowest-performing third citywide
- Post-secondary enrollment of students in the lowest-performing third citywide
- Movement of students with disabilities to less restrictive environments

---

DATA SOURCES

All of the components of the NYC Progress Reports are collected or calculated by the NYC Department of Education (DOE).

The New York State Testing Program (NYSTP) assesses student achievement in grades three through eight in ELA and math and assesses student performance in science in fourth and eighth grades.[5] NYSTP discontinued testing for social studies in fifth and eighth grades in 2010. The New York State Education Department uses Regents exams to assess student achievement in high school. To graduate, students who entered high school in 2008 or later qualify to receive a Regents Diploma by passing five Regents exams. The exams are administered in global history and geography; U.S. history and government; comprehensive English; integrated algebra; geometry; algebra 2 and trigonometry; earth science; biology; chemistry; and physics. Foreign-language Regents exams were discontinued in the 2010–11 school year.

SCORING THE FRAMEWORK

1. Assign points for each measure
Schools receive points for each measure of the framework based on the school's performance in comparison with two groups:
A. All schools in the city
B. A group of 40 similar schools identified based on student characteristics
   (The NYC Peer Index is used to identify schools with similar enrollment of high-need students)
   i. **Elementary peer schools** are identified by enrollment rates of students with economic need,[6] students with disabilities, black and Hispanic students, and English language learners.
   ii. **Middle school peer schools** are identified by fourth-grade proficiency rates and enrollment rates of students with disabilities.
   iii. **High school peer schools** are identified by eighth-grade proficiency rates and enrollment rates of students with disabilities, students with self-contained placements, and over-age students.

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

49

NYC DOE establishes a range for each measure and comparison group by calculating a *minimum value* two standard deviations below the group average and a *maximum value* two standard deviations above the group average. Schools are evaluated against each of the two comparison groups:

$$\frac{\textbf{\textit{School Result}} \text{ minus } \textbf{\textit{Minimum Value}}}{\textbf{\textit{Maximum Value}} \text{ minus } \textbf{\textit{Minimum Value}}} = \textbf{Percent of Range}$$

Schools earn points for each metric based on an average of the comparison to the peer schools and all city schools based on the following formula:

$$\left[\left[\ \textbf{Percent of \textit{Peer} Range} \ \text{times } \textbf{0.75}\ \right]\ \text{plus}\ \left[\ \textbf{Percent of \textit{Peer} City}\ \text{times } \textbf{0.25}\ \right]\right]\ \text{times}\ \textbf{Maximum Points Possible}$$

Source: Educator Guide http://schools.nyc.gov

The maximum points available for each measure are presented in the table below.

	Maximum Points Possible			
	Elementary	**Middle**	**K–8**	**High School**
1. Student Progress	60	60	60	55
ELA—Median Adjusted Growth Percentile	10	15	12.5	--
ELA—Median Adjusted Growth Percentile for students in the school's lowest-performing third*	10	15	12.5	--
ELA—Early-Grade Progress (proficiency level of third-grade high-need** students)	10	--	5	--
Math—Median Adjusted Growth Percentile	10	15	12.5	--
Math—Median Adjusted Growth Percentile for students in the school's lowest-performing third*	10	15	12.5	--
Math—Early-Grade Progress (proficiency level of third-grade high-need** students)	10	--	5	--
Credit Accumulation (percentage of students earning 10 or more credits in the current year)	--	--	--	Year 1 – 4.17 Year 2 – 4.17 Year 3 – 4.17
Credit Accumulation (percentage of the lowest-performing third* of students earning 10 or more credits in the current year)	--	--	--	Year 1 – 4.17 Year 2 – 4.17 Year 3 – 4.17
Regents Completion Rate	--	--	--	5
Weighted Regents Proficiency Rate[7]—English, math, science, U.S. history, global history	--	--	--	Up to 25 (5 points for each subject)

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

50

2. Student Performance	25	25	25	20
ELA—Percentage of Students at Proficiency	6.25	5	5	--
ELA—Average Student Proficiency Level	6.25	5	5	--
Math—Percentage of Students at Proficiency	6.25	5	5	--
Math—Average Student Proficiency Level	6.25	5	5	--
English Core Course Passing Rate	--	1.25	1.25	--
Mathematics Core Course Passing Rate	--	1.25	1.25	--
Science Core Course Passing Rate	--	1.25	1.25	--
Social Studies Core Course Passing Rate	--	1.25	1.25	--
4-Year Cohort Graduation Rate	--	--	--	5
6-Year Cohort Graduation Rate	--	--	--	5
4-Year Weighted Diploma Rate	--	--	--	5
6-Year Weighted Diploma Rate	--	--	--	5
3. School Environment	15	15	15	15
Academic Expectations***	2.5	2.5	2.5	2.5
Communication***	2.5	2.5	2.5	2.5
Engagement***	2.5	2.5	2.5	2.5
Safety and Respect***	2.5	2.5	2.5	2.5
Attendance	5	5	5	5
4. College and Career Readiness (HS only)	NA	NA	NA	10
College and Career Preparatory Course Index—successful completion of Regents, AP, IB, or dual credit coursework or CTE or arts diploma or industry certification	--	--	--	3.33
4-Year College Readiness Index—Regents diploma or achievement of CUNY standards	--	--	--	1.67
6-Year College Readiness Index—Regents diploma or achievement of CUNY standards	--	--	--	1.67
College Enrollment—within 6 months of high school graduation	--	--	--	1.67
College Enrollment—within 18 months of high school graduation	--	--	--	1.67

Source: 2011–12 Educator Guides. Retrieved from http://schools.nyc.gov/Accountability/tools/report/
*For elementary and middle schools, the lowest-performing third includes student who earned the lowest scores on state exams in math or ELA in the previous year (Spring 2011). For high schools, the lowest-performing third includes students who earned the lowest average math and ELA scores on state exams in the eighth grade.
**Black or Hispanic students, students with economic need, students IEPs with SETSS, ICT, or self-contained recommendations, or English language learners.
***NYC Learning Environment Survey

2. Award Additional Credit

Schools may earn additional points for performance on achievement gap measures. Schools receive additional points for each metric, up to the maximum value in the table below, based on the percentage of students who meet the conditions set by each metric. For example, if a middle school had 50 percent of ELL students show high growth, the school would receive .5 additional points (.5 times the maximum points awarded – 1).

	Maximum Points Possible			
	Elementary	**Middle**	**K–8**	**High School**
5. Closing the Achievement Gap	14	15	15	16
Percentage of eighth-graders earning high school credit[8]	--	1	1	--
ELA—Percentage of high-needs students achieving high growth (75th percentile): – Students with disabilities (self-contained, ICT,* SETTS**) – English language learners – Students in the lowest-performing third citywide*** – Black and Hispanic males in the lowest-performing third citywide***	4 (1 point for each group)	4 (1 point for each group)	4 (1 point for each group)	--
Math—Percentage of high-needs students achieving high growth (75th percentile): – Students with disabilities (self-contained, ICT,* SETTS**) – English language learners – Students in the lowest-performing third citywide*** – Black and Hispanic males in the lowest-performing third citywide***	4 (1 point for each group)	4 (1 point for each group)	4 (1 point for each group)	--
ELA—Percentage of students with disabilities achieving proficiency: – Self-contained – ICT* – SETTS**	3 (1 point for each group)	3 (1 point for each group)	3 (1 point for each group)	--
Math—Percentage of students with disabilities achieving proficiency: – Self-contained – ICT* – SETTS**	3 (1 point for each group)	3 (1 point for each group)	3 (1 point for each group)	--
4-Year Graduation Rate for high needs groups: – Students with disabilities (self-contained, ICT,* SETTS**) – English language learners – Students in the lowest-performing third citywide*** – Black and Hispanic males in the lowest-performing third citywide***	--	--	--	8 (2 points for each high-needs group)
College and Career Preparatory Index for students in the lowest-performing third citywide***	--	--	--	2
College Readiness Index for students in the lowest-performing third citywide***	--	--	--	2
Postsecondary enrollment by students in the lowest-performing third citywide***	--	--	--	2
Students with disabilities' movement to less restrictive environment	1	1	1	2
English language learner progress	1	1	1	--

Source: 2011–12 Educator Guides. Retrieved from http://schools.nyc.gov/Accountability/tools/report/
*Integrated Co-Teaching
**Special Education Teacher Support Services
***For elementary and middle schools, the lowest-performing third includes student who earned the lowest scores on state exams in math or ELA in the previous year (Spring 2011). For high schools, the lowest-performing third includes students who earned the lowest average math and ELA scores on state exams in the eighth grade.

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

52

3. Assign school grade

Score ranges for elementary and middle schools are set each year in order to assign approximately 25 percent A scores, 35 percent B scores, 30 percent C scores, 7 percent D scores, and 3 percent F scores. High school score ranges are presented in the table below and are used annually by DOE.

Rating	Elementary and Middle Schools	High Schools	
	Percentage of Schools in Rating Category 2011–12	Final Overall Score Range for High Schools	Percentage of Schools in Rating Category 2011–12
A	25%	70 or greater	35%
B	35%	58.0 to 69.9	37%
C	30%	47.0 to 57.9	20%
D	7%	40.0 to 46.9	5%
F	3%	39.9 or less	3%

NYC DOE does not calculate school grades for schools in the first year of operation, schools with reported growth results for fewer than 25 students, schools designated for "phase-out," or high schools without a graduating class.

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

NYC DOE publishes the NYC Progress Reports each year. Detailed reports for each school present the results for each of the metrics, the overall score, the school's relative ranking, and assigned grade. Excel data files with performance data are available for download. Guidance documents are published in English and nine other languages.

PUBLIC REACTION

The NYC Progress Reports have elicited both support and criticism.[9] Proponents say the school grades provide the public with comprehensive school information, bring attention to school quality, and allow the New York DOE to hold schools accountable. Critics argue that the system relies too heavily on testing data and misses important measures of school quality.

Changes in methodology across years have created confusion for parents and schools. While grades have remained consistent for most schools, the large changes in some schools' grades raised skepticism about the grading system's validity[10]—several schools targeted for closure by the NYC DOE in recent years moved from an F to a B in two years.

LINKS TO PUBLIC REPORTS

- School Progress Reports: http://schools.nyc.gov/ProgressReport
- Educator Guides: http://schools.nyc.gov/Accountability/tools/report/default.htm

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

53

Appendix F — Endnotes

1 NYC Department of Education publishes reports for alternative schools and early-childhood (PK-2) schools; we profile the system applied to non-alternative schools serving grades 3 through 12.

2 Adjustments are made to growth results for special education students and students with economic need.

3 Regents exams are New York State high school end-of-course exams administered in English, math, science, U.S. history, global history.

4 Students are assigned one of four levels on state examinations in reading and math: Below Standard, Meets Basic Standard, Meets Proficiency Standard, or Exceeds Proficiency Standard

5 For more information, see http://www.p12.nysed.gov/assessment/about-osa.html

6 Students with economic need include students in temporary housing, students eligible for assistance from the Human Resources Administration (HRA), or students eligible for free lunch.

7 Additional weight is given to the students with lower eighth-grade proficiency levels.

8 The eighth-grade credit metric is the only additional credit metric that is scored by ranking the school result against both citywide and similar school performance.

9 Making the grade in New York City. (2012, October 9). *The New York Times*. Retrieved from http://www.nytimes.com/roomfordebate/2012/10/09/making-the-grade-in-new-york-city

10 Chapman, B., Samuels, T., & Monahan, R. (2012, October 1). 217 city elementary and middle schools could face closure after earning bad grades on Department of Education report cards. *The Daily News*. Retrieved from http://www.nydailynews.com/new-york/education/269-city-elementary-amp-middle-schools-bad-report-cards-article-1.1172051#ixzz2b3CxQyXn

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

54

APPENDIX G:

U.S. News & World Report
"Best High School" Rankings

Overview

U.S. News & World Report first published its list of "Best High Schools" in 2007, and has done so three times since, including in 2012. U.S. News aims for its high school rankings to provide a "clear, unbiased picture of how well public schools serve all of their students" to help parents make education choices for their children.[1]

To measure high school success, the rankings attempt to answer two primary questions:
- Are schools successfully serving all students, including disadvantaged students, based on state assessments in reading and math?
- Are schools preparing students for college, based on participation and performance in Advanced Placement (AP) or International Baccalaureate (IB) exams?

For its 2012 report, U.S. News evaluated data from nearly 22,000 public schools in 49 states and Washington, D.C.[2] For the first time, it also collaborated with the American Institutes for Research (AIR), a behavioral and social science research organization, while still using largely the same methodology.

Methodology 2011–12

COMPONENTS OF THE FRAMEWORK

U.S. News considers three components in its rating system:
- Proficiency rates on state assessments in reading and math for:
  - All students
  - Black/African-American students
  - Hispanic/Latino students
  - Economically disadvantaged students
- Percentage of economically disadvantaged students enrolling at the schools
- Participation and performance on AP and IB exams

DATA SOURCES

The U.S. News rankings use the following data sources:
- School and student descriptive information from the Common Core of Data by the National Center for Education Statistics[3]
- College Board AP test data for high school seniors, provided by the College Board
- International Baccalaureate test data for high school seniors, provided by the International Baccalaureate Organization
- Publicly available state assessment results, collected from state departments of education[4]

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

55

SCORING THE FRAMEWORK

To calculate the final rankings, all schools go through a three-step screening process. A school must pass the first two steps to be included in the final published rankings.

Step 1. Does the high school exceed expected performance on state assessments?
Each school receives a performance index score—a composite score based on the number of students at different proficiency levels in reading and math.

Proficiency level	Points awarded
Two or more levels below proficiency	0
One level below proficiency	0.5
Proficient	1.0
One level above proficiency	0.5
Two levels above proficiency	2.0

Schools receive full credit (1 point) for each proficient score, extra credit (1.5 or 2 points) for each advanced score (scores one or two levels above proficiency), partial credit for each non-proficient score (0.5 points) and no points for scores two or more levels below proficiency. The points for all students are averaged and multiplied by 100 to arrive at a performance index score between 0 and 200.

U.S. News then uses regression analysis to evaluate the relationship between the performance index score and socioeconomic status, and to calculate each school's expected performance index score based on the school's percentage of economically disadvantaged students. Each school's actual performance index is compared to its expected performance index. Only high schools that meet or exceed the expected performance index level pass on to step 2.

Step 2. Is the performance of the least advantaged students higher than average statewide performance?
For each school, the model compares the proficiency rates in math and reading for three student subgroups—black/African-American students, Hispanic/Latino students, and economically disadvantaged students—to the state average proficiency rate for the same subgroup. If each of the proficiency rates for each of the student subgroups at the school is equal to or greater than the statewide average for that student subgroup, the high school passes on to step 3. If a school did not enroll students in these subgroups, the school also moves on to step 3.

Step 3. Do students in the high school have access to college-level coursework, and are they prepared for college?
For the final step, the model calculates a college readiness index based on AP and IB participation and performance rates. The participation rate equals the percentage of 12th-grade students who take an AP or IB exam while in high school. The performance rate equals the percentage of 12th-grade students who scored 3 or higher on an AP test or 4 or higher on an IB exam while in high school.

To calculate the college readiness index, U.S. News averages the participation and performance rates, with the participation rate counting for 25 percent of the final score and the performance rate counting for 75 percent. The resulting college readiness index ranges from 0 to 100, with 100 indicating that every 12th-grader took and passed one or more AP or IB test during high school.

Calculating a school's ranking
All schools that successfully pass steps 1 and 2 are awarded gold, silver, or bronze medals based on their college readiness index for the current year.

- Gold-medal schools have a college readiness index equal to or greater than the median college readiness index value and are in the top 500 schools based on the college readiness index.
- Silver-medal schools have a college readiness index that is equal to or greater than the median index value but are not in the top 500 schools based on the college readiness index
- Bronze-medal schools have a college readiness index that is less than the median index value or did not offer AP or IB courses.

Schools earning gold and silver medals are ranked both nationally and within their states. Charter and magnet schools that earn gold or silver medals are ranked separately.

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

56

Public Presentation

HOW ARE RESULTS PRESENTED TO THE PUBLIC?

Schools are listed by rank and state in the print edition of U.S. News & World Report, and may be searched by name online.[5]

PUBLIC REACTION

The U.S. News rankings received some negative press in 2012, when Green Valley High School in Henderson, Nev., was ranked 13th due to a federal and state clerical error that led U.S. News to use faulty data.[6]

States that do not offer IB or AP tests have also criticized the rating system because of the weight it places on those exams, putting states that do not offer the exams at a disadvantage. These states argue that an evaluation of curriculum rigor should be included in the methodology.[7]

LINKS TO PUBLIC REPORTS

- Rankings FAQs: http://www.usnews.com/education/high-schools/articles/2012/05/07/frequently-asked-questions-best-high-schools-rankings
- Analytical Methodology and Technical Appendix: http://www.usnews.com/pubfiles/Identifying_Top_Performing_High_Schools_May2012.pdf

Appendix G — Endnotes

1 See the rankings by state at http://www.usnews.com/education/best-high-schools

2 Nebraska was not included in the rankings due to insufficient data reporting.

3 U.S. Department of Education National Center for Education Statistics. (n.d.). Common core of data. Retrieved from http://nces.ed.gov/ccd/

4 The 2012 rankings used 2009-2010 results for all states except Wyoming. 2008-09 results were used for Wyoming because 2009-2010 results were not available.

5 See the rankings at http://www.usnews.com/education/best-high-schools

6 Rindels, M. (2012, May 10). Green Valley High School ranking as 13th in U.S. caused by state, federal errors. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/2012/05/11/state-fed-errors-rank-nev_n_1509147.html

7 Bowie, L. (2012, May 8). U.S. News and World Report high school rankings are out. *The Baltimore Sun*. Retrieved from http://articles.baltimoresun.com/2012-05-08/news/bal-us-news-and-world-report-high-school-rankings-20120508_1_ib-tests-rankings-baltimore-school

National Alliance for Public Charter Schools
QUALITY SCHOOL RATINGS: TRENDS IN EVALUATING SCHOOL ACADEMIC QUALITY — APPENDIX

57

ACKNOWLEDGEMENTS

National
Alliance
for Public
Charter
Schools