Toward A Next-Generation School Accountability System

Design Priorities

Next-generation accountability systems must inspire schools and their communities to lift the achievement of all graduates to college- and career-ready levels. School rating systems must pinpoint challenges, spur more innovation and inspire much broader local support if these ratings are to provoke lasting change.

We've learned from the current accountability system that we've got a long way to go. Our system is designed to protect the state's role in monitoring school quality, to foster more educator and community engagement, and to avoid taking a short-sighted view based on today's backlash-driven malaise. Our design builds upon the belief that the academic achievement and growth of every child is most important—especially so for children of color and those from low income communities—and that non-test academic indicators should round out the picture of student and school success. We have imagined a system to guide resource allocation and rally school teams, school boards, communities and funders to the cause.

Our two design priorities:

1) We must give all students our attention, but focus on closing achievement gaps.

A high-quality education is the best way to improve the odds for all young people. No gaps are more pernicious or at odds with the promise of America than gaps among whites and people of color and gaps among youth at opposite rungs of the socioeconomic ladder. Given our nation's shifting demographics—in which people of color are on a trajectory to surpass whites as the majority (some estimate within 25 years)—there are few goals that should be more important to our nation than closing these persistent gaps.[1] We place our democracy, economy and standing in the world at risk if we do not. With more than 51% of American students on free and reduced lunch, we must also attend to the needs of low-income students.

2) Local communities should have real decisionmaking in accountability system design.

Past accountability systems were the darlings of policymakers, think tanks, foundations, editorial boards and advocates; they rarely had the support of educators, school communities and the public writ large. They were too often equated with excessive testing that many parents picking up their children after school tell each other "takes time away from learning." Our design provides school communities the opportunity to select additional indicators and measures in every component through discussion of what matters most to them, to share that publicly and to commit to work that addresses goals the community develops. Our design also acknowledges that one size does not fit all. It allows communities to apply for waivers to experiment with new assessments and processes, such as measuring growth more frequently or measuring growth to competency.

---

[1] Brookings reported that, as of the 2010 Census, 14 states already had majority-minority toddler populations, including Florida, Georgia, Mississippi, New Jersey, New Mexico and New York. Populations in these states' schools will be majority-minority quite soon.

Below we detail how the system works for an elementary school with grades PreK-5 in a mid-sized exurban school district with real socioeconomic diversity. In Appendix 2, we offer options for high school indicators.

System Overview, Including Summative Scoring and School Ratings

Our system relies on two scales: state-determined indicators ("base points") and locally-determined indicators ("local points"). Both the state and local indicators are in the four categories required by ESSA: academic achievement, growth, progress toward English language proficiency and student success/school quality.

On the first scale, schools acquire up to 100 "base points" through growth and proficiency on state assessments (80 points), reductions in chronic absenteeism and increases in social-emotional learning (SEL) indicators (20 points). On the second scale, school districts select the indicators for schools to acquire up to 20 "local points" by performing against goals set for each of the same four components. Districts decide, in collaboration with their school communities, what indicators and targets to set. Districts submit plans to the state (or a regional entity like the BOCES in New York or Education Service Centers in Texas) for approval.

Acquiring local points is not an opportunity to recover lost base points. Rather, schools are assigned a performance tier (green, yellow or red) **and** a level (1-10 stars) using **both** their base score **and** their local score. A green, 8-star school can score between 90 and 92 base points and at least 15 local points. Another school might score 15 on the locally selected measures but have its rating dragged down by a score of 60 on its base indicators – making it a red, 3-star school. Schools that have any racial/ethnic subgroup of students that is consistently below the state average for that subgroup, across all indicators with a three-year rolling average, cannot be rated above red/3-stars.

Schools are assigned colors and stars/levels with the table below. Red schools will be subject to interventions determined by the state; yellow schools are the responsibility of local school boards.

Table 1. Summative scoring & categories[2]

Tier	Levels	Base Points	Local Points
Green	10 stars	>= 96	20
	9 stars	93 – 95	15 – 19
	8 stars	90 – 92	15 – 19
Yellow	7 stars	80 – 89	15 – 20
	6 stars	70 – 79	10 – 14
	5 stars	60 – 69	5 – 9
	4 stars	60 – 69	0 – 4
Red	3 stars	<60	15 – 20
	2 stars	<60	10 – 14
	1 star	<60	0 – 9

We envision a clickable "dashboard" reporting results by each component, with clear delineation between performance on state indicators and local indicators, and enabling users to click through for more detail. We've included examples of the "dashboard" reporting below.

---

[2] The Appendix charts 2-4 detail all points and calculations for base points and local points.

Components and Indicators: Base Points

Component 1: Academic Achievement

This component accounts for 30 of 100 base points and relies on rigorous state-developed assessments in English language arts and mathematics in grades 3, 4 and 5, and a state-developed assessment in science in grade 4 or 5. There are four sub-indicators: all students, targeted students, gap-closing and preK-2 literacy. In the 1st three sub-indicators, points are weighted (or "pro-rated"), with more points for students achieving above standard and fewer for below standard. The school receives points for the performance of "all students," and it receives additional points for targeted subgroups (defaulted as racial subgroups unless the "n" size is too low or the district can justify to the state other subgroups) and for gap-closing between historically high-performing and low-performing students. Districts select a locally-administered measure for PK-2 literacy from a state-approved list.

Component 2: Academic Growth

This component accounts for 30 of 100 base points. While measuring growth to proficiency might be ideal, our imagined state cannot do this at scale; our default design uses student growth percentiles and awards points equally to mathematics and English language arts for grades 3, 4 and 5, and focuses on subgroups and gap-closing. The school receives points for the performance of "all students," and it receives additional points for the growth of targeted subgroups (defaulted as racial subgroups unless the "n" size is too low or the district can justify to the state other subgroups) and for gap closing between historically high-performing and lower performing groups. Our state allows a waiver process (discussed below) that allows school districts to experiment with growth to proficiency.

Component 3: Progress toward English Language Proficiency

This component accounts for 20 of 100 base points for schools with significant English language learner (ELL) populations, defined as having at least 5 percent of all students designated as ELLs. Schools with ELL populations below 5 percent will count this component as 10 base points (5 for each indicator). Schools with no reportable ELLs will not include this component. For schools below 5 percent or with no ELLs, the points—either 10 or 20—will be distributed equally across components one and two. This component awards schools points for two equally-weighted indicators: ELL proficiency rates, as measured by scores on a state-specific or consortium test (e.g., WIDA's ACCESS), and English learner re-designation rates, measured by the percent of ELLs who are reclassified as English learners within five years of being designated as ELLs. Like the proficiency component, the ELL proficiency indicator is weighted based on proficiency (i.e., 1 for at proficiency, .06 for approaching proficiency, etc.).

Component 4: Student Success/School Quality

This component accounts for 20 out of 100 base points. Students can't achieve if they aren't showing up, nor if their teachers are absent,[3] so our system emphasizes reducing teacher absenteeism (5 points) and reducing chronic student absenteeism (10 points).

The school also earns up to an additional 5 points by meeting a *district-determined* SEL goal measured by a district-selected assessment, such as DESSA.

---

[3] In a recent study of the 40 largest metropolitan areas, 16 percent of teachers are chronically absent, missing at least 18 days. A study by the National Bureau of Economic Research estimates that each 10 days of teacher absences reduce students' mathematics achievement. A growing body of evidence shows that student chronic absenteeism reduces literacy.

Dashboard example: Sample base point scale for all schools, statewide[4]

In this example dashboard, each indicator is shown as a "fuel gauge," with the school's performance (in orange) also compared against the statewide average for "demographically similar schools" (in black).

Achievement Base
School X: 27
Similar Schools: 17

Growth Base
School X: 22
Similar Schools: 20

ELL Base
School X: 11
Similar Schools: 12

School Quality Base
School X: 18
Similar Schools: 10

Base Scale	Achievement	Growth	ELL	School Quality
100 total	30	30	20	20

Total: 78/100

Components and Indicators: Local Points

For each of the four components, school districts select one indicator and one measure. Districts are responsible for ensuring these local measures meet state-determined criteria for quality. An indicator for component one might be 5th grade performance in fine arts and the measure a performance task scored using a district rubric. In component two, the district might measure growth in science. Component three could include bi-literacy (language proficiency in English and in another language), or reducing ELL designation rates in three years. In component four, school climate could be measured by questions on the TELL Survey and/or district-led surveys administered to families. Or schools might measure additional SEL goals. For each indicator, the district sets a goal and up to five points are awarded based on performance against the target. Ideally, the district plan lays out school-specific indicators, targets and goals.

Dashboard example: Local point scale with district-selected indicators[5]

Achievement Local
School X: 3
Similar Schools: 3.75

Growth Local
School X: 1
Similar Schools: 2.5

ELL Local
School X: 2
Similar Schools: 3.75

School Quality Local
School X: 5
Similar Schools: 3

Local Scale	Achievement	Growth	ELL	School Quality
20 total	5	5	5	5

Total: 11/20

---

[4] The Appendix charts 2-4 detail all points and calculations for base points and local points.
[5] The Appendix charts 2-4 detail all points and calculations for base points and local points.

Waivers for Innovation

Our waiver process is for districts that have the capacity and commitment to innovate. It invites districts to go beyond the growth and achievement components of the default state plan, while still using some of the state's year-end assessments as a public check.

The waiver process requires districts to:
- Select one state assessment in 3-5 for English language arts, one in 3-5 for mathematics for component one (plus science) and component two.
- Identify measures in one or more subject areas beyond the state assessments. The district could use portfolios and/or competency-based assessments in PK-5. It might experiment with assessments administered more frequently than annually or measure growth against competency, not age.
- Propose new calculations and weights for each measure.
- Submit a sampling of tasks, rubrics and processes the school/district will use to score measures, such as procedures for quality assurance, multiple scorers, etc.
- Submit its quality control plan for scorers and to assure validity and reliability of the measures.

State responsibilities include:
- Approving the waiver plan and publicly reporting on the district's tools and processes.
- Inspecting the results to make sure the assessments are being scored accurately.
- Providing technical assistance on assessment development/selection and scoring.

Conclusion

We began by imagining that next generation accountability systems can be an inspiration to school communities, engaging them in efforts to improve achievement for all while making sure that we are delivering on the promise of America. Here is how we think the system we've presented does that:

Gap Closing
- Both achievement and growth components weight gap closing. Nineteen of 100 base points are earned by gap closing. This means that low-income students and youth of color count more than middle- or upper-class white students. We don't apologize for that. Lifting people out of poverty and making sure that all students regardless of race achieve are among our greatest priorities.
- The system requires each school to address racial and socioeconomic gaps as subgroups as well, unless there are none or there is too little racial or socioeconomic diversity in the school.

Local Choice and Engagement
- Districts decide the details of its plan and submit the plan to the state.
- Districts and school communities choose a PK-2 literacy measure and goal.
- Districts and school communities select additional indicators and goals for academic achievement, growth, English language proficiency and student success/school quality. Within the district, some schools might use common measures for discipline/suspension rates; other schools might share common measures on working conditions.
- Districts can apply for a waiver to use additional measures for growth and achievement.

Appendix 1: Accountability System Component and Calculation Details

Table 2. Base components, indicators & points breakdown

Base Components	Indicators	Student Groups	Points
			Total Base Points: 100
Component 1: Academic achievement (as measured by proficiency) → **30 total base points**	Reading/ELA proficiency in grades 3-5	→ All students[6]	→ 3
		→ Targeted subgroups[7]	→ 3
		→ Gap closing[8]	→ 3
		→ *PK-2 literacy (district-selected)*	→ 3
	Math proficiency in grades 3-5	→ All students	→ 3
		→ Targeted subgroups	→ 3
		→ Gap closing	→ 3
	Science proficiency in at least one grade (3, 4 or 5)	→ All students	→ 3
		→ Targeted subgroups	→ 3
		→ Gap closing	→ 3
Component 2: Student growth (as measured by student growth percentiles) → **30 total base points**	Reading/ELA proficiency in grades 3-5	→ All students	→ 5
		→ Targeted subgroups	→ 5
		→ Gap closing	→ 5
	Math proficiency in grades 3-5	→ All students	→ 5
		→ Targeted subgroups	→ 5
		→ Gap closing	→ 5
Component 3: English language proficiency → **20 total base points**	ELL proficiency rates	→ ELL students	→ 10
	English learner designation rates[9]	→ ELL students (re-designated and non-redesignated)	→ 10
Component 4: Student success/school quality → **20 total base points**	Student absenteeism rate[10]	→ All students	→ 10
	Teacher absenteeism rate	→ All teachers	→ 5
	Non-cognitive or SEL competencies (district-selected measure)	→ All students	→ 5

---

[6] Each "all students" and "targeted subgroups" indicator for Component 1 is weighted based on Table 4 below. Schools can earn more than 3 points per indicator if many students score at the advanced (i.e., highest) achievement level.

[7] The state default system selects race/ethnic subgroups against white students and low-income against all students.

[8] As measured by the change in the difference in percentage of students at or above proficiency for the subgroup with the highest proficiency and the subgroup with the lowest proficiency. This indicator uses a 5-year average gap between each subgroup as the base for comparison.

[9] As measured by: The number of all ELLs who re-designate at the school in the current year (no matter how long they've been English learners), divided by the number of all ELLs who re-designate at the school in the current year + all 5-year-plus non-re-designated ELLs at the school.

[10] Emphasis is reducing chronic absenteeism. Measured by the percentage of student population that has been absent 10 percent or more of the school year (excused and unexcused combined). This indicator will use an index that awards points (up to 10) based on the school's rate (e.g., 0-1% = 10 points, 1.1-2% - 9 points) and based on state historical data for absence rates. Same calculation for teachers.

Table 3. Local components, indicators & points breakdown

Local Components	Indicators	Student Groups	Points
			Total Local Points: 20
Achievement	District selected	→ District selected	→ 5
Growth	District selected	→ District selected	→ 5
English Language Learners	District selected	→ District selected	→ 5
Student success/school quality	District selected	→ District selected	→ 5

Table 4. Example of weighting proficiency by achievement level

Achievement Level	# Students	(multiplied by)	Weight	(equals)	Points
Four (advanced)	20	x	1.2	=	24
Three (proficient)	50	x	1.0	=	50
Two (approaching)	20	x	0.6	=	12
One (beginning)	10	x	0	=	0
				Total	86

To calculate weighted points for each "All students" and "Targeted subgroups" indicator in Component 1:

(total points from Table 4 above)   86

divided by

(total possible points if all students scored at level three)   100

= 0.86

(indicator's weight)   multiplied by

3

= **2.58 points in indicator** (out of 3)

Appendix 2: High school indicators that promote college and career readiness

Under ESSA, states must include the four-year cohort graduation rate as one of the academic indicators, and could use this graduation rate in lieu of other "student success/school quality" indicators.

Education First believes that states should promote college and career readiness as the goal for high school accountability. Important indicators in the "school quality" category might include:

- Whether 9th-graders are "on track" for graduation
- The "college- and career-ready graduation rate," which is a measure of how many students graduate in four years not just with a diploma, but ready for postsecondary enrollment without the need for remediation
- The percentage of students enrolled in a college- and career-ready course of study
- The percentage of students enrolled in college and career ready dual credit programs (e.g., rigorous CTE pathways that award postsecondary credit, Advanced Placement, International Baccalaureate)
- The percentage of students who enroll in some form of postsecondary institution within two years of high school graduation
- The percentage of students who enroll in postsecondary and need remediation in reading, writing and/or mathematics

States also may decide to *report* a robust set of indicators such as these, while only holding schools accountable for a subset of such indicators. Achieve and NASDCTEc's 2014 *Making Career Readiness Count* report offers an approach to indicators that considers how well students are moving toward college and career readiness.

Table 5. Possible indicators for high schools[11]

	Toward College and Career Readiness	Meeting College and Career Readiness	Exceeding College and Career Readiness
Course completion/ success	• Timely credit accumulation along a college- and career-ready course of study	• Students in a graduating cohort who complete a college- and career-ready course of study	• Graduates who have completed AP, IB and/or dual enrollment courses and earned college credit
Achievement	• Students with "on track to college and career readiness" performance on aligned assessments in middle and early high school	• Students in a graduating cohort with a college- and career-ready level of performance on a high-quality assessment aligned to college- and career-ready standards	• Graduates with a college-level performance on an AP and/or IB exam
Attainment	• 9th grade students with "on track" to graduation status based on attendance and grades in core courses • High school graduation	• Students in a graduating cohort who earn a college- and career-ready diploma	• Graduates who enroll in postsecondary education with no need for remediation • Graduates who successfully complete at least one year of postsecondary education or a workforce/military training program

---

[11] Achieve and National Association of State Directors of Career and Technical Education Consortium (NASDCTEc), *Making Career Readiness Count*, 2014, http://achieve.org/publications/making-career-readiness-count