

Making Sure All Children Matter: Getting School Accountability Signals Right

BY NATASHA USHOMIRSKY, DAVID WILLIAMS, AND DARIA HALL

School accountability systems have the potential to be a powerful tool to help close the long-standing gaps in achievement that separate low-income students and students of color from their peers. They can do this by

- ▶ Setting a clear expectation that schools have to serve all of their students — not just some — well;
- ▶ Drawing attention to how schools are performing for all student groups; and
- ▶ Prompting action when schools don't meet expectations for a group of students.

Closing gaps is critical to our communities and to our nation. Low-income students and students of color now make up the majority of the nation's public school students. Yet many of these students are not getting the quality education they need and deserve, leaving them unprepared for success in our economy and our democratic society. If we want to live up to the American ideal of equal opportunity for all, we have to turn these patterns around.

Even the strongest critics of No Child Left Behind will acknowledge that the law did one thing right — create an expectation that in order to be considered successful, a school had to be successfully educating all groups of students. Yet in issuing waivers from that law, U.S. Secretary of Education Arne Duncan allowed states to walk away from that framework. To obtain a waiver, he required states to set improvement goals for each group of students that would result in gap narrowing. But, in a move that had civil rights advocates shaking their heads, performance against those goals did not have to matter at all in school accountability ratings.

To be clear, new state accountability systems designed under the waivers represent an advance in many ways — including innovative uses of different student outcome indicators, inclusion of growth measures, and greater differentiation between schools. But we and many other equity-minded advocates have raised serious questions about whether these systems are built to play their critical gap-closing function.

We now have data to test these concerns, and as it turns out, they're well founded.

A key way that an accountability system signals its expectations is through the ratings it assigns to schools. A system that gives high marks only to schools that demonstrate high performance, growth, or improvement for all of their student groups — including low-income students and students of color — sends one signal. A system that gives out high marks to schools that are doing well on average, but have low results for their African American students, for example, sends another, very different message.

To better understand the signals that accountability systems are currently sending about group outcomes, we've analyzed student performance data from three states — Florida, Kentucky, and Minnesota. For each state, we asked, "How are schools that earn the highest accountability rating, as well as lower ratings, performing for all students? How about for low-income students and students of color?"

Our intent is not to pick on these three states, for each represents some innovative policy design and is a leader in making available the kind of data necessary to even do this analysis. Rather, our goal is to test the signals that school accountability systems are sending in a diverse group of states that represent policy choices made by a lot of other states.

What we've found is that school ratings are not powerful signals of the performance of every individual group of kids. In each state, schools are getting top ratings despite low performance for some groups. In fact, the differences are so large that top-rated schools often perform similarly for their students of color and low-income students as middle- to low-rated schools do for their white and higher income peers. Moreover, where we examined either improvement over time or growth, those data were no more encouraging.

These findings don't answer big questions about whether and how much achievement is increasing for different groups of

Natasha Ushomirsky is senior data and policy analyst, David Williams is K-12 data analyst, and Daria Hall is director of K-12 policy development at The Education Trust. © Copyright 2014 The Education Trust. All rights reserved.



What We Did

During the past year, states have started to release data on the ratings their schools receive, as well as the percentages of students meeting academic standards in each school. These data enabled us to look at how schools receiving a particular grade or rating are performing for students overall, as well as for each individual group of students. In some states, we were also able to look at whether that performance is improving and whether students in each group are making sufficient growth to be on track to meet standards.

For this analysis, we focused on elementary and middle school results only. Whenever possible, we excluded alternative schools and schools that predominantly serve special needs students from the calculations because these schools are often treated differently in accountability systems. High school accountability systems also tend to be structured differently and use more, and more varied, indicators, making it hard to look at high school outcomes at the same time. It would be worthwhile, however, to ask the same analytical questions of high school accountability systems.

To look at how students are performing in schools receiving each accountability designation, we averaged the percentages of students meeting standards in reading and math, by group, across all schools receiving a particular rating. To be as consistent as possible with each state's policies, we only included performance data in our analysis when a school tested enough students in a given group to meet the state's reliability criteria for accountability. Kentucky and Florida, for example, require 10 students in a group to be tested for the data to be included. In Minnesota, that number is 20.¹

children under new systems. It's just too early to say. But they do make clear that the signals that school ratings send about group performance are not strong enough to prompt wide-scale attention to and action around gap closing.

This should be a serious wake-up call both for the Secretary of Education whose guidance invited these changes and for the state education chiefs who took him up on that invitation. Despite advances in some areas of school accountability, when it comes to increasing equity in our education system, today's accountability systems still represent, at best, a missed opportunity.

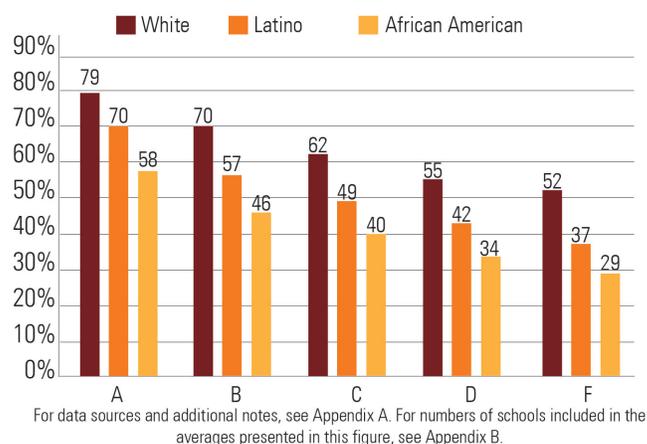
HOW ARE SCHOOLS THAT EARN EACH ACCOUNTABILITY RATING PERFORMING FOR ALL STUDENTS, AS WELL AS FOR LOW-INCOME STUDENTS AND STUDENTS OF COLOR?

In each of the three states we looked at, schools can get — and are getting — top ratings despite low performance for some groups. In fact, top-rated schools' proficiency rates for students of color and low-income students are, on average, about the same as those of white and higher income stu-

dents in middle- to low-rated schools. Moreover, in many states, whether or not a school is improving results for its lower performing groups doesn't affect its rating.

Consider Florida. Figure 1 shows the average reading proficiency rates of white, African American, and Latino students in schools that receive A's, B's, C's, D's, and F's under the state's accountability system. On average, results for all three groups of students are higher in "A" schools than in lower graded schools. This is a good thing. But the average proficiency rate for African American students in "A" schools is lower than that of white students in "C" schools. In other words, an "A" grade tells you something very different about the performance of African American students than that of white students.

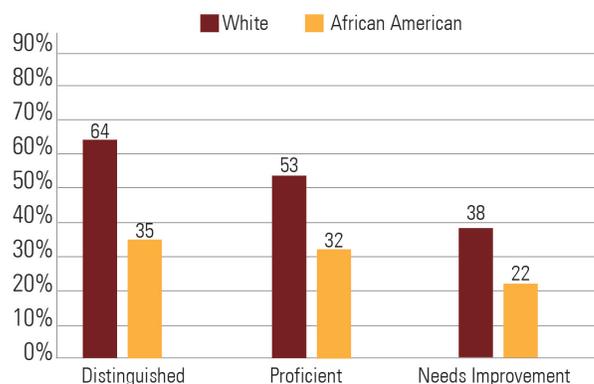
Figure 1: Florida Reading Proficiency Rates, 2014



Similarly, the average proficiency rate for Latino students in "B" schools is about the same as that for white students in "D" schools.

Next, let's look at Kentucky. In Kentucky we looked at three broad categories of schools — Distinguished, Proficient, and Needs Improvement. In Figure 2, we show average math proficiency rates for white and African American students in schools earning each designation. Once again, proficiency rates for both groups are higher in Distinguished

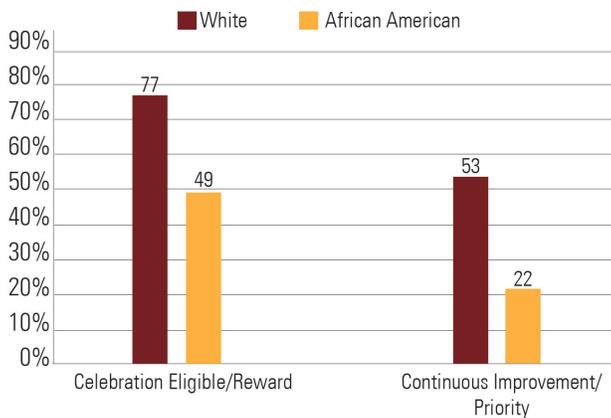
Figure 2: Kentucky Math Proficiency Rates, 2013



and Proficient schools than in Needs Improvement schools. But average math proficiency rates of African American students at schools earning a Distinguished rating are lower than average math proficiency rates of white students in Needs Improvement schools. Here, too, a Distinguished rating tells us something very different about the performance of African American students than it does about the performance of white students.

Lastly, let's look at *Minnesota*. Figure 3 shows math proficiency rates in the state. Here we see average results for white and African American students in the state's Celebration Eligible and Reward schools (schools singled out for recognition), and Continuous Improvement and Priority schools (schools identified for intervention).² Although proficiency rates for both white and African American students are higher in Celebration Eligible/Reward schools than in those identified for intervention, on average, Celebration Eligible/Reward schools demonstrate similar results for their African American students as Priority and Continuous Improvement schools do for their white students.

Figure 3: Minnesota Math Proficiency Rates, 2013



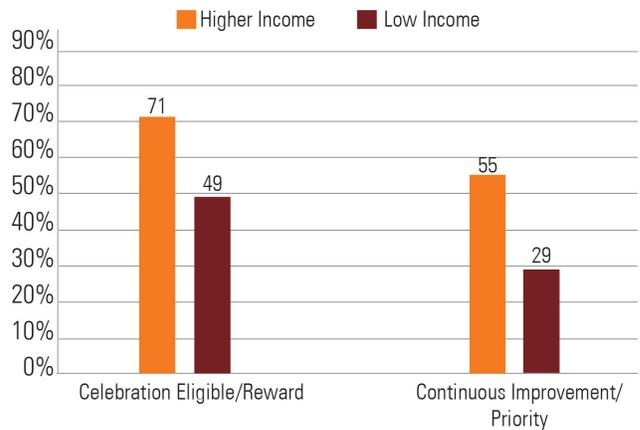
For data sources and additional notes, see Appendix A. For numbers of schools included in the averages presented in this figure, see Appendix B.

The patterns are similar in reading. Figure 4 presents reading proficiency rates for low-income and higher income students in Minnesota. Schools receiving recognition are doing better for both their low-income and higher income students than schools identified for intervention, but Celebration Eligible and Reward schools' proficiency rates for low-income students are similar to higher income students' results in Continuous Improvement and Priority schools.

ARE SCHOOLS EARNING HIGH RATINGS IMPROVING THE PERFORMANCE OF HISTORICALLY UNDERSERVED STUDENTS?

Of course, one-year proficiency rates are not the only important measure of achievement. If top-rated schools are improving a lot from year to year, especially if they are improving more for student groups that start out further behind, then the state accountability system may well be sending the

Figure 4: Minnesota Reading Proficiency Rates, 2013



For data sources and additional notes, see Appendix A. For numbers of schools included in the averages presented in this figure, see Appendix B.

right signals after all. If these schools are narrowing achievement gaps over time, they deserve the recognition they are getting.

The data indicate, however, that often enough this is not the case. In fact, 39 percent of Florida's "A" schools with data for African American students actually demonstrated lower reading proficiency rates for this group in 2014 than in 2013. Thirty percent showed declines in math. During the same time period, 45 percent of schools that earned B's and had data for Latino students lost ground for this group in reading, and a similar share declined in math.

Similarly, between 2012 and 2013 in Kentucky, reading proficiency rates for African American students declined at about 40 percent of Distinguished schools with data for this group. Math proficiency rates fell at about the same share of top-rated schools.³

ARE SCHOOLS EARNING HIGHER RATINGS GETTING STUDENTS ON TRACK ACADEMICALLY?

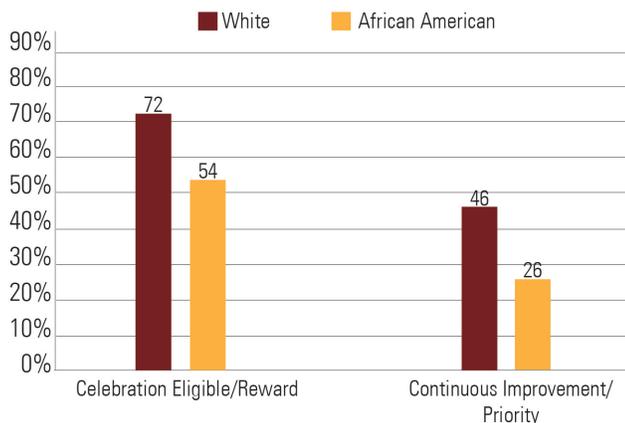
Schools also need to get credit for taking students who come in behind and helping them catch up academically. These students may not be reaching state standards yet, but they are showing sufficiently high growth to be considered academically on track.

Minnesota makes available data on the percent of on-track students by group. According to the state's definition, a student is on-track if she is not currently meeting standards, but is making high growth, or if he is meeting standards and is making medium or high growth. Minnesota identifies each student as making low, medium, or high growth depending on how that student's progress compares with the progress of other students with similar past performance.⁴

Our analysis indicates that while schools that earn Celebration Eligible or Reward status under Minnesota's accountability system demonstrate higher on-track rates for both

white and African American students than schools identified for intervention, the difference in on-track rates for white and African American students among these recognized schools are still vast (see Figure 5). In fact, the recognized schools' average on-track rate for African American students is close to that of Continuous Improvement and Priority Schools for white students.

Figure 5: Minnesota Average On-Track Rates: Math, 2013



For data sources and additional notes, see Appendix A. For numbers of schools included in the averages presented in this figure, see Appendix B.

In other words, while highly rated schools are producing higher on-track rates for African American students than are low-rated schools, these numbers do not paint a picture of gap closing that would eliminate concerns about gaps in proficiency rates.

WHY DOES THIS MATTER, ANYWAY?

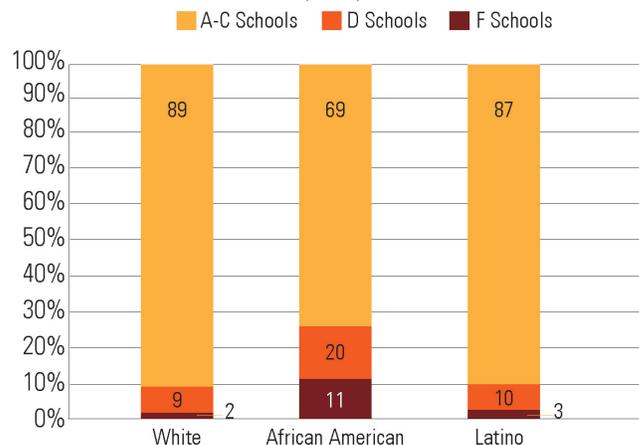
We cannot close achievement gaps on the backs of low-performing schools alone.

Recently, both federal and state accountability policy has focused heavily on improving performance in the lowest achieving schools. But while these schools likely need the most intensive support from their districts and states, improving their outcomes won't, on its own, close achievement gaps. Why? Because although low-income students and students of color are overrepresented in the lowest rated schools, in most states, most of them actually attend other, higher rated schools. Some of these higher rated schools are recognized or receive accolades under these systems. Others are simply left alone. Either way, the message is that what they're doing is okay, even if, as the data above show, they are not doing nearly as well for some groups as they are for others.

Take Florida, for example. Improving student achievement in Florida's "D" and "F" schools, which receive the most support from the state, will certainly make a difference for the students in those schools. But "D" and "F" schools only serve about 30 percent of Florida's African American elementary and middle schoolers, and only about 13 percent of the

state's Latino students (see Figure 6). In Florida, as in other states, closing achievement gaps will require improvement in more than just the lowest performing schools.

Figure 6: Percent of Students Attending Florida's A-C, D, and F Schools by Group, 2014



For data sources and additional notes, see Appendix A. For numbers of schools included in the averages presented in this figure, see Appendix B.

It's about transparency.

Accountability systems are meant to send powerful signals that drive action. Many parents make decisions about where to send their child to school at least in part based on that school's accountability rating. Unfortunately, current accountability systems aren't communicating to all parents the information they need to decide how well a school is likely to serve their children. On average, "A" (or "Distinguished," or "Reward") schools are performing far better for their white and higher income students than they are for their low-income students and students of color, meaning an "A" school for one child may not be an "A" for another. Moreover, the lack of transparency about what an "A" grade (or a "Celebration Eligible" rating) actually means for a school's low-income children and children of color limits parents' and community members' ability to advocate on behalf of these students.

Likewise, educators gauge progress in part based on these systems. A high rating despite low performance for some groups paints a false picture of success and allows schools to overlook some students.

WHAT CAN BE DONE?

States are undergoing seismic shifts as they transition to college- and career-ready standards and aligned assessments. And, they are rightly reconsidering their accountability systems in light of these shifts. As hard as this work is, it creates a real opportunity to put the focus squarely back on the performance of all groups of students.

To be clear, we are not advocating for a return to NCLB-style accountability. But as states do the work of designing their post-transition accountability systems, they should be careful to ensure that school ratings reflect how schools are serving all groups of students. One of the ways states can do that is by using data from new assessments to set gap-closing goals for every group of students and making performance against those goals a key factor in the ratings schools receive. For example, a state could drop schools by one rating (e.g., from a “4 star” to a “3 star”) if they miss goals for any group for two years in a row.

There is no one best way to do this. But by making these kinds of adjustments, states could ensure that gap closing and achievement for all groups of students matter in their rating systems, and, as a result, that the achievement of every individual group of students counts. Such changes would make group performance a central part of the main system of rewards, supports, and consequences tied to the rating system.

Accountability systems in and of themselves don’t close gaps and raise achievement. Only the hard work of educators and students can do that. But well-designed accountability systems can be a much-needed source of pressure and support in this work.

Our hope is that this analysis will prompt policymakers, advocates, and educators to put equity squarely back on the table in each and every conversation about accountability. The Secretary of Education can re-start that focus by making group performance matter in the upcoming waiver renewal process. ■

ENDNOTES

1. Because our analysis seeks to understand the signals that each state’s accountability system sends to its schools and the public, all schools are weighted the same in our calculations. In other words, schools with lots of students tested count as much as schools with few students tested. Results for schools with very small numbers of students tested are generally less reliable than those for larger schools. To see whether the inclusion of results for schools with small numbers of students tested biases our analysis, we repeated our calculations using various n-size decision rules, ranging from including any school with 10 or more students to only including those with about 40-plus students in a group. We found that these decision rules did not affect our analysis; the results we saw were very similar regardless of the n-size. That being the case, we defaulted to the state-selected n-size for the calculations presented in this brief. Information on state n-sizes was obtained from Minnesota and Kentucky’s ESEA Flexibility Requests, available at <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>; and Florida Department of Education’s Reporting Florida’s Annual Measurable Objectives (AMOs) in Compliance with ESEA Flexibility Requirements Guide to Calculations, available at: <http://schoolgrades.fldoe.org/pdf/1314/Amo.pdf>.
2. The Continuous Improvement/Priority group of schools includes two categories of schools — Priority and Continuous Improvement. Similarly, the Celebration Eligible/Reward group includes Celebration Eligible and Reward schools, respectively.
3. Schools were considered to have declined if proficiency rates fell by 1 or more percentage points.
4. For more on Minnesota’s growth model and definition of on-track rates, see Minnesota Department of Education, Minnesota Persistently Lowest Achieving Schools Definition, available at: <http://education.state.mn.us/mdeprod/groups/educ/documents/basic/003865.pdf>.

Appendix A:

Data Sources and Relevant Data Notes for Figures 1-6

This Appendix contains the data sources used to produce Figures 1-6 in *Making Sure All Children Matter: Getting School Accountability Signals Right*. The notes below each data source document the information obtained from each respective dataset, as well as any data-cleaning decisions applied.

FIGURES 1 AND 6

Florida Department of Education, 2014, Annual Measurable Objectives (AMOs) for Schools, Districts and the State, available at: <http://schoolgrades.fldoe.org/>.

This file contains 2014 school grades, as well as Florida's standardized assessment proficiency rates by student group for 2011-2014. It also includes information on school type and alternative school status.

Proficiency rates are not reported when fewer than 10 students are tested. Proficiency rates below 5 percent are reported as asterisks. Our analysis assumes a 2.5 percent proficiency rate whenever an asterisk is reported. Notably, there were very few instances of proficiency rates below 5 percent for the student groups and schools included in the analysis. That being the case, this assumption likely had no impact on the values reported in Figure 1.

High schools were excluded from the analysis, as were alternative schools and any additional schools that did not receive a school grade.

Florida Department of Education, 2014, Membership by School by Grade, 2013-14, available at: <http://www.fldoe.org/eias/eiaspubs/pubstudent.asp>.

This file contains 2014 enrollment data by grade for every school. These data were used to exclude a small number of schools that were not flagged as high schools in the Annual Measurable Objectives file, but that had high school enrollment.

Florida Department of Education, 2014, Membership by School by Race/ Gender, 2013-14, available at: <http://www.fldoe.org/eias/eiaspubs/pubstudent.asp>.

This file contains 2014 enrollment by race/ethnicity and gender for every school.

Enrollment numbers are suppressed whenever there are fewer than 10 students in a group enrolled in a particular school. Suppressed values are treated as 0's in our analysis. Given the relatively small number of suppression cases and the large number of schools included, this likely had no impact on the values reported in Figure 6.

FIGURE 2

Kentucky Department of Education, 2013, Delivery Targets Proficiency Gap, available at: <https://applications.education.ky.gov/src/DataSets.aspx>.

This file contains 2012 and 2013 Kentucky standardized assessment proficiency rates by student group for each school and grade level (elementary, middle, and high school).

Proficiency rates are suppressed whenever fewer than 10 students are tested.

High school level records were excluded from the analysis.

Kentucky Department of Education, 2013, Accountability Profile, available at: <https://applications.education.ky.gov/src/DataSets.aspx>.

This file contains accountability ratings by grade level for each school (elementary, middle, high school).

High school level records were excluded from the analysis.

Some schools with multiple grade levels (such as a K-8 school) received different accountability ratings for each level. For example, the elementary grades may receive a Distinguished rating, while the middle grades are rated as Proficient. In such cases, we included each grade level in the averages for its assigned designation. For example, the proficiency rate for the Distinguished elementary grades was included in the Distinguished average, and the proficiency rate for the Proficient middle grades was included in the Proficient average.

FIGURES 3, 4, AND 5

Minnesota Department of Education, 2013, 2013 MCA-III Results, available at: <http://w20.education.state.mn.us/MDE-Analytics/Data.jsp>.

These files for both math and reading contain Minnesota Comprehensive Assessment proficiency rates by school and student group. Data are reported by grade (grades 3-8 in math and grades 3-8, plus grade 10, in reading).

Data are suppressed whenever fewer than 10 students are tested.

We eliminated grade 10 results to exclude all high school assessment data, and aggregated the remaining grade-level results to the school building level to calculate school-wide proficiency rates. To be consistent with Minnesota's

accountability policy, we excluded all school-level records where fewer than 20 students in a given group were tested.

Minnesota Department of Education, 2013, 2013 Multiple Measurement System – Statewide Results, available at: <http://w20.education.state.mn.us/MDEAnalytics/Data.jsp>.

This file contains accountability ratings by school, as well as a school type indicator (elementary, middle, high school and other school).

All high school and “other school” records were excluded from the analysis. The “other school” category captures predominantly alternative/non-regular education programs. In addition, schools that did not receive a Multiple Measurement Rating were excluded from the analysis.

Minnesota Department of Education, 2013, 2013 Growth Summary – All Schools and Districts by Grade & Category, available at: <http://w20.education.state.mn.us/MDEAnalytics/Data.jsp>.

This file contains student growth information at the grade, school, and district levels, including the number of students classified as high/medium/low growth and the number of students classified as high/medium/low growth by proficiency status. All of the data are presented by subgroup and subject.

Data are suppressed when there are fewer than 10 students in a group with growth information. Our analysis relied on school-level records, and growth was only included for schools that also had proficiency rate information for the all student group.

Appendix B:

Numbers of schools included in Figures 1-5

Table 1: Number of Schools Included in Figure 1 (Florida Reading Proficiency Rates, 2014)

Number of schools with data for	School Grade				
	A	B	C	D	F
White students	867	369	619	263	90
African American students	735	366	671	313	144
Latino students	893	395	675	299	121

Table 2: Number of Schools Included in Figure 2 (Kentucky Math Proficiency Rates, 2013)

Number of schools with data for	School Rating		
	Distinguished	Proficient	Needs Improvement
White students	114	254	693
African American students	35	96	292

Table 3: Number of Schools Included in Figure 3 (Minnesota Math Proficiency Rates, 2013)

Number of schools with data for	School Rating	
	Celebration Eligible/Reward	Continuous Improvement/Priority
White students	304	66
African American students	67	33

Table 4: Number of Schools Included in Figure 4 (Minnesota Reading Proficiency Rates, 2013)

Number of schools with data for	School Rating	
	Celebration Eligible/Reward	Continuous Improvement/Priority
Higher Income students	292	59
Low-Income students	291	86

Table 5: Number of Schools Included in Figure 5 (Minnesota Math On-Track Rates, 2013)*

Number of schools with data for	School Rating	
	Celebration Eligible/Reward	Continuous Improvement/Priority
White students	307	68
African American students	62	32

*Note: The number of schools included in Figures 3 and 5 differ slightly because some schools that had proficiency data for 20 or more students did not have growth data for 20 students, and vice versa.

ABOUT THE EDUCATION TRUST

The Education Trust promotes high academic achievement for all students at all levels – pre-kindergarten through college. We work alongside parents, educators, and community and business leaders across the country in transforming schools and colleges into institutions that serve all students well. Lessons learned in these efforts, together with unflinching data analyses, shape our state and national policy agendas. Our goal is to close the gaps in opportunity and achievement that consign far too many young people – especially those who are black, Latino, American Indian, or from low-income families – to lives on the margins of the American mainstream.



The Education Trust

1250 H STREET, NW, SUITE 700, WASHINGTON, D.C. 20005

P 202-293-1217 F 202-293-2605 WWW.EDTRUST.ORG