



## SHANKER BLOG

# What Florida's School Grades Measure, And What They Don't

by Matthew Di Carlo -- July 19, 2012

A while back, I argued that Florida's school grading system, due mostly to its choice of measures, does a poor job of gauging school performance *per se*. The short version is that the ratings are, to a degree unsurpassed by most other states' systems, driven by absolute performance measures (how highly students score), rather than growth (whether students make progress). Since more advantaged students tend to score more highly on tests when they enter the school system, schools are largely being judged not on the quality of instruction they provide, but rather on the characteristics of the students they serve.

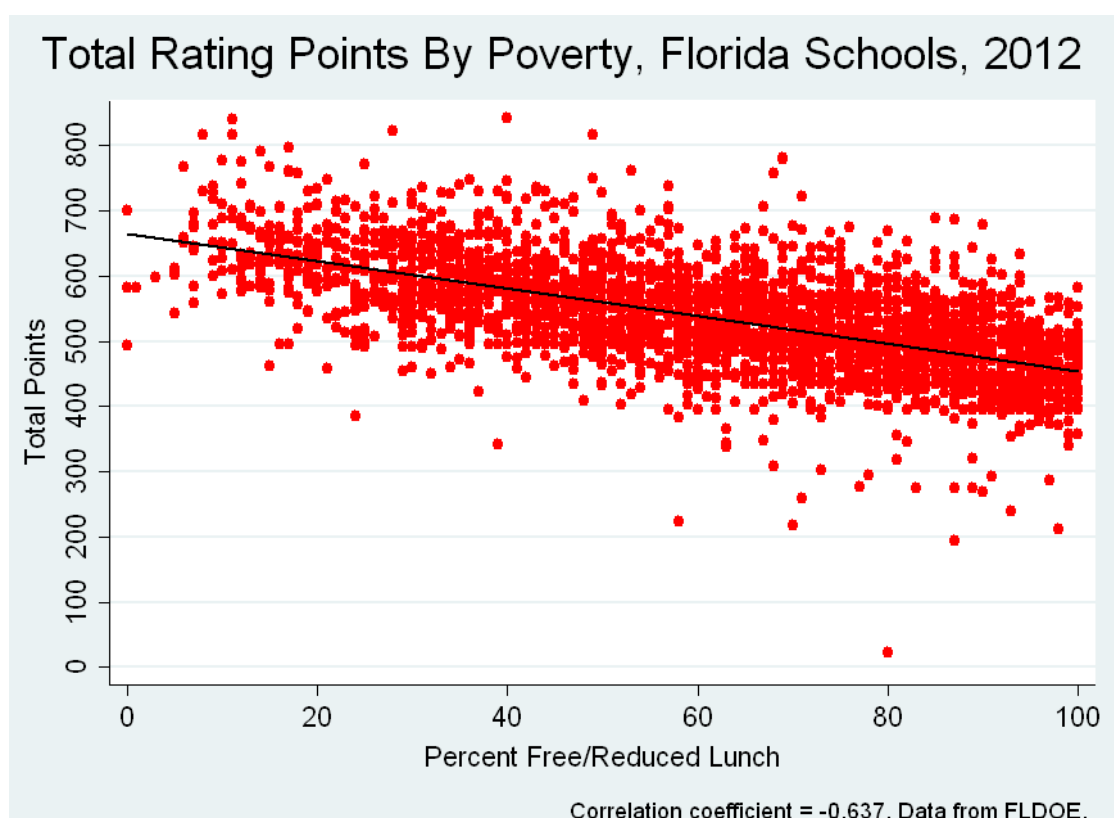
New results were released a couple of weeks ago. This was highly anticipated, as the state had made controversial changes to the system, most notably the inclusion of non-native English speakers and special education students, which officials claimed they did to increase standards and expectations. In a limited sense, that's true - grades were, on average, lower this year. The problem is that the

system uses the same measures as before (including a growth component that is largely redundant with proficiency). All that has changed is the students that are included in them. Thus, to whatever degree the system now reflects higher expectations, it is still for outcomes that schools mostly cannot control.

I fully acknowledge the political and methodological difficulties in designing these systems, and I do think Florida's grades, though exceedingly crude, might be useful for some purposes. But they should not, in my view, be used for *high-stakes* decisions such as closure, and the public should understand that they don't tell you much about the actual effectiveness of schools. Let's take a very quick look at the new round of ratings, this time using schools instead of districts (I looked at the latter in my previous post about last year's results).

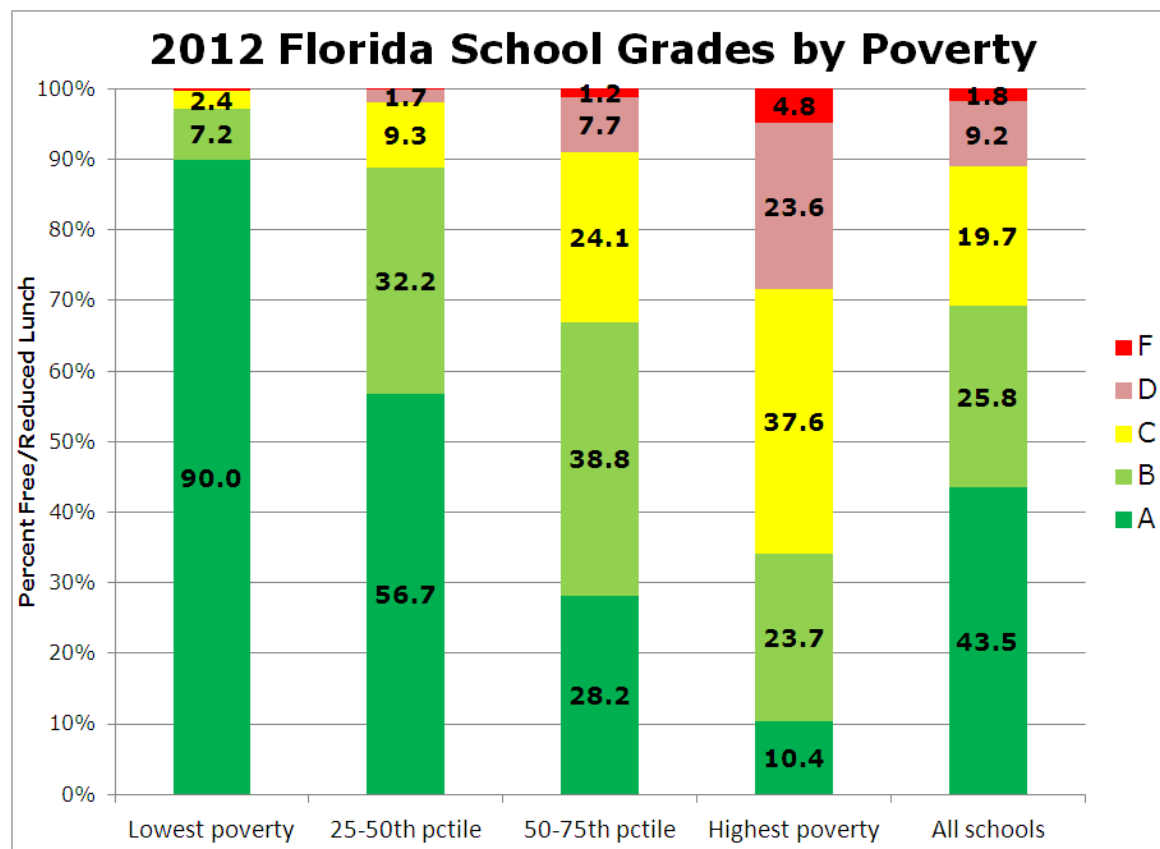
One rough, exploratory way to assess whether a high-stakes grading system may have problems is to see if the ratings are heavily associated with student characteristics such as poverty. It's important to note that the mere existence of a relationship does not *necessarily* mean that the latter are biased or inaccurate, since it may very well be the case that lower-performing schools are disproportionately located in poorer areas. But, if the association is unusually strong, and/or clearly due to the choice of measures rather than "real" variation in performance (as we already know is probably the case in Florida), this may be cause for concern, depending on how the ratings are used. I'll discuss this issue of use, which is the important one in the context of school ratings, after presenting some simple data.

In the scatterplot below, each dot is an elementary or middle school (almost 3,000 in total; ratings are pending for any schools that serve high school grades). The vertical axis is the number of points the school received (0-800; these points are sorted into A-F letter grades), while the horizontal axis is the percent of the schools' students receiving subsidized lunch, a rough proxy for poverty.



This is a fairly strong relationship. You can see that the number of points schools receive tends to decline as poverty increases. That's not at all surprising, of course - Florida's system relies heavily on absolute performance, which is correlated with the proportion of students receiving subsidized lunch.

Now let's recast these data using the letter grades, which are the basis for policy decisions. Just to keep it simple, I divided schools into quartiles based on their poverty rates – schools with the top 25 percent rates ("highest poverty"), the 25 percent with the lowest rates and the two quartiles in between. The graph below presents the grades for each group.



This graph shows that the schools with the lowest poverty (the bar all the way on the left) are virtually guaranteed to receive either an A (90 percent, dark green) or a B (another 7 percent, light green).

In contrast, the highest-poverty schools were likely to receive grades on the opposite end of the spectrum – around two-thirds received a C (yellow) or lower. Of the roughly 300 schools receiving a D or F, only 13 are in the two low-poverty quartiles (some of these percentages are not listed in the graph, due to space limitations).

So, according to Florida's system, almost every single low-performing school in the state is located in a higher-poverty area, whereas almost every single school serving low-poverty students is a high performer. This is not plausible. There is a very big difference between being a low-performing school and being a school that happens to serve lower-performing students. Confusing the two serves no one, especially not these students.

What do I mean? Accountability systems can be valuable to the extent that they identify both ineffective *schools* and well-functioning schools with lower-performing *students* in need of additional resources and support. But only if the system is capable of distinguishing between the two.

Either type of school might benefit from many types of interventions, such as additional resources or support, for which these ratings can be used. But the grades Florida assigns are also the basis for high-stakes, punitive actions, including closure and restructuring.\*

And, for these drastic decisions, it is crucial to avoid conflating student and school performance.

For example, a significant proportion of the schools receiving D or F grades this year are among the highest rated schools in the state in terms of "gains" among their lowest-performing students (which is, in my view, the only measure in the system that is even remotely defensible for gauging school effects *per se*). No doubt many of these D/F schools are also compelling large gains from their other students (though Florida's measure cannot really tell us which ones).\*\*

In other words, there are many Florida schools with lower-performing students that are actually very effective in accelerating student performance (at least insofar as tests can measure it). This particular ratings system, however, is so heavily driven by absolute performance - how highly students score, rather than how much progress they have made - that it really cannot detect much of this variation.

Closing or reconstituting these schools is misguided policy; their replacements are unlikely to do better and are very likely to do worse. Yet this is what will happen if such decisions are made based on the state's ratings.

Florida will have to do a lot more than make tweaks to truly improve the high-stakes utility of this system. In the meantime, one can only hope that state and district officials exercise discretion in how it is applied.

- Matt Di Carlo

\*\*\*\*\*

\* According to the state's waiver application, schools will be given one year, during which they receive state support, to improve their F grades. If they fail to do so, they must implement one of five drastic "turnaround strategies," such as closure, charter conversion or contracting with a private company to run the school.

\*\* As I discussed in my prior post, 25 percent of the ratings are based on the percent of all students "making gains." However, while the data are (to Florida's credit) longitudinal, it's not really a growth measure as traditionally understood. Students are coded as "making gains" if they remain proficient or better between years, without going down a category (i.e., from advanced to proficient). As a result, schools with high proficiency rates get much higher "gains" scores, which exacerbates the degree to which absolute performance drives the overall ratings (the correlation between "gains" and proficiency rates is 0.67 in math, 0.71 in reading). Another 25 percent is the measure discussed in the post's text - the percent of the lowest-scoring students "making gains." This indicator, while deeply flawed for several reasons (e.g., reliance on cutpoints), is at least not heavily dependent upon

absolute performance, and thus may provide some signal of "actual" school performance. In either case, it is a safe bet that a better growth measure would show that many of the D/F schools score highly.

## Blog Topics:

Education Education Reform Testing Data Standardized Tests Accountability florida

 Share

0

 Tweet

94

 +1

2

[Printer-friendly version](#)

## Authors

Matthew Di Carlo

## Comments

Submitted by Charlie Morris on August 3, 2012

You might note that Florida also awards teacher bonuses based upon FCAT scores. The teachers in the wealthiest districts almost always get annual bonuses (originally \$1000, now less); teachers in the poorest districts almost never get them.

Submitted by Catherine on July 22, 2012

Dear Mr. DiCarlo, Thank you so much for proving scientifically what those of us working in Title One schools sensed from anecdotal observation. It is no accident that none of the schools that service students from family above the 50% percentile in wealth are graded "F" while the lowest 25% have 2/3 of the F and D schools and only a about 11% of the A schools. It may also explain why there has been relatively little protest against the system. Those most adversely affected belong to the poor and minorities who are often politically marginalized while the middle and upper classes have even been favored (through school bonuses) by the system.

## RECENT BLOG POSTS

**The Smoke And The Fire From Evaluations Of Teach For America**

**How Not To Improve New Teacher Evaluation Systems**

**The Status Fallacy: New York State Edition**

**Turning Conflict Into Trust Improves Schools And Student Learning**

**Actual Growth Measures Make A Big Difference When Measuring Growth**

## UPCOMING EVENTS

### **A Diverse Teacher Force**

March 13, 2015

There is concern that, as the U.S. population and student body is growing more racially and ethnically diverse, the teacher workforce does not yet reflect this diversity. In fact, diversity should go beyond having more black and brown teachers in front of students. Diversity is also about equipping all teachers (regardless of race) to work with heterogeneous classrooms and diverse schools.

### **The Social Side of Teaching — A New Framework For Improving The Profession**

March 13, 2015

Teaching and learning are not primarily individual accomplishments but rather social endeavors that are best achieved and improved through trusting relationships and teamwork. Yet, most policies focus on improving the individual capacities of teachers. What is to be gained from approaches that strike a balance between human and social capital?

### **Fairness & Effectiveness in School Discipline**

March 14, 2015

How do we teach discipline and maintain order, while protecting against the effects of persistent, unconscious biases? How do we ensure that schools are warm, welcoming, fair, and effective in the treatment of all students?

## BLOG ARCHIVES

<div>- Choose -</div>	<div>Go</div>
-----------------------	---------------

## RECENT COMMENTS

Sneha on on Sample Size And Volatility In School Accountability Systems

Sal Corda on on Relationships Matter: Putting It All Together

Joe Miller on on Relationships Matter: Putting It All Together

Peter Hofman on on Fixing Our Broken System Of Testing And Accountability: The Reauthorization Of ESEA

john on on New Research On School Discipline

## OUR MISSION

The Albert Shanker Institute, endowed by the American Federation of Teachers and named in honor of its late president, is a nonprofit, nonpartisan organization dedicated to three themes - excellence in public education, unions as advocates for quality, and freedom of association in the public life of democracies. With an independent Board of Directors (composed of educators, business representatives, labor leaders, academics, and public policy analysts), its mission is to generate ideas, foster candid exchanges, and promote constructive policy proposals related to these issues.

This blog offers informal commentary on the research, news, and controversies related to the work of the Institute.

### DISCLAIMER

This web site and the information contained herein are provided as a service to those who are interested in the work of the Albert Shanker Institute (ASI). ASI makes no warranties, either express or implied, concerning the information contained on or linked from shankerblog.org. The visitor uses the information provided herein at his/her own risk. ASI, its officers, board members, agents, and employees specifically disclaim any and all liability from damages which may result from the utilization of the information provided herein. The content in the Shanker Blog may not necessarily reflect the views or official policy positions of ASI or any related entity or organization.

## SIGN UP FOR EMAIL UPDATES

## JOIN US

### EDUCATION

Career & Technical Education

Civic Education

Early Childhood

K-12 Education

School Culture

## **LABOR**

Healthcare

Public Sector Unions

Workforce Development/CTE

Worker Rights

## **DEMOCRACY**

Civic Education

Democracy Web

International Democracy

Worker Rights

## **RESOURCES**

Lesson Plans

Other Materials

Publications

Supported Work

Audio Visual

## **ABOUT US**

About Al Shanker

ASI Staff

Board of Directors

Mission Statement

Contact ASI

Albert Shanker Institute, 555 New Jersey Ave, NW, Washington, DC 20001, Tel: 202-879-4401