

Interim Report for GreatSchools-Stanford Study: Findings from YouGov Surveys on How Americans Assess School Quality

* * * * *

August 11, 2013

Overview

This memo describes findings from a project examining how the American public consumes information about schools when assessing school quality. This project is part of a broader collaboration between researchers at Stanford University's Center for Education Policy Analysis and the staff at GreatSchools.

All findings reported here come from an online survey administered by the research and polling firm YouGov from June 24-July 5, 2013. The sample consists of 1000 respondents who were matched and weighted by YouGov to represent the general adult population of the United States (with respect to gender, age, race, education, party identification, ideology, and political interest). The memo is organized around our three primary topics of examination and subsequent findings.

- First, we show that Americans report trusting public school academic ratings from independent nonprofit organizations more than public school academic ratings from their state governments.
- Second, we present evidence that when ranking schools rated differently in “achievement,” “gains” (value-added), and “school climate,” Americans first prefer schools without “low” ratings in any category and then express preferences, in order, for achievement and then gains and then school climate.
- Third, we present evidence that seeing positive or negative parent comments can be remarkably influential on Americans' assessments of local schools, even when those comments come alongside numerical academic ratings.

In each section we report both on findings for our full, nationally representative sample and for important subgroups within that sample (e.g., parents of children under the age of 18).

Finding 1. Americans report trusting public school academic ratings from independent nonprofit organizations more than public school academic ratings from their state governments.

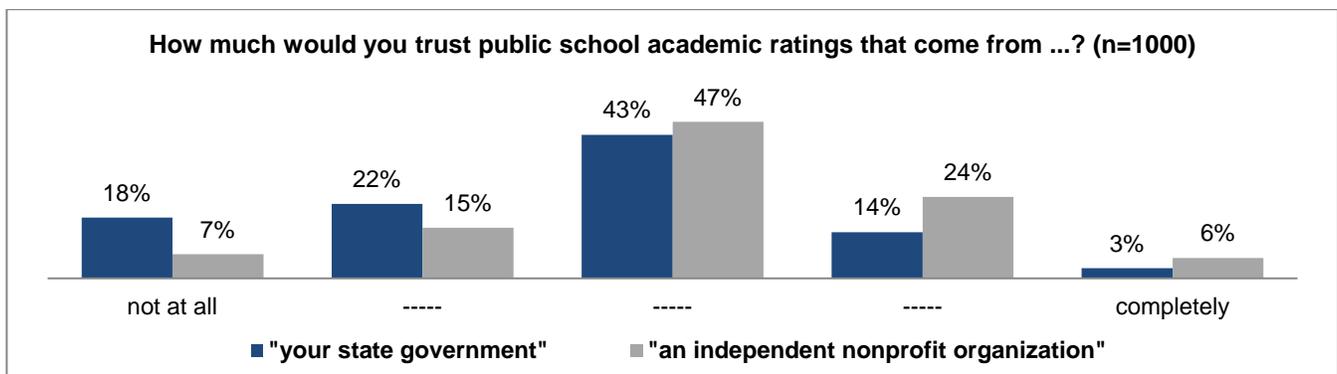
What we did:

We conducted two experiments, one direct and the other subtle. For the more direct experiment, respondents were asked, “How much would you trust public school academic ratings that come from [an independent nonprofit organization / your state government]?” Respondents were randomly assigned to either the nonprofit group or the government group. They reported their level of trust on a 1-5 scale that ranged from “not trust at all” to “completely trust.”

For the more subtle experiment, we embedded language about these sources into the broader “actual schools” items described below (under “Finding 3”). We presented respondents with information about two schools in their communities and then asked them to assess those schools from A+ to F. That information included academic ratings. Respondents were randomly assigned to see the source of these ratings as either “an independent nonprofit organization” or “your state government.” (A single respondent had the same assignment for both the “direct” experiment and this “subtle” experiment.) We tested whether academic ratings affected respondents’ school assessments differently depending on whether they saw the source described as “an independent nonprofit organization” or “your state government.” We call this “subtle” because we considered it highly possible that respondents would overlook our description of the academic rating’s source amid a much more elaborate and complicated series of items.

What we found:

Where we asked respondents directly about how much they would trust a particular source of public school academic ratings, we see large, statistically significant differences, with respondents reporting considerably more trust in ratings from “an independent nonprofit organization” (p-value <0.001). This trust gap appears in every tested subgroup related to gender (male or female), race/ethnicity (White, Black, or Hispanic), educational attainment (no college, < 4 years of college, 4+ years), and parent status (with children under 18 or not). The trust gap is perhaps slightly larger for males than females. On our 1-5 trust scale, those in the full “nonprofit” group provided a mean response of 3.07 while those in the full “state government” group provided a mean response of 2.61. The weighted response distributions for the full sample appear below.



For our more subtle experiment, we tested whether academic ratings were more influential upon respondents’ assessments of schools depending on whether the source of those ratings was described as “an independent nonprofit organization” or “your state government.” We did not detect statistically significant differences.

Finding 2. When ranking schools rated differently in “achievement,” “gains,” and “school climate,” Americans first prefer schools without “low” ratings in any category and then express preferences, in order, for achievement and then gains and then school climate.

What we did:

We asked respondents to imagine they were choosing a school for a child they know well. Some respondents were randomly assigned to choose a high school; others were randomly assigned to choose an elementary school. We then presented each respondent with profiles of four hypothetical schools. (We used six hypothetical schools in all, but each respondent only saw four of them in order to reduce the task’s complexity.) Each of these schools was rated “High,” “Medium,” or “Low” in each of three categories: “Achievement,” “Gains,” and “School Climate.” A school received either (A) two high ratings and one low rating or (B) one high rating and two medium ratings. Respondents then read definitions for each category and ranked the schools “from best to worst for this child.”

The category descriptions appear below. In order to test whether responses to the “Achievement” and “Gains” (value-added) ratings were sensitive to the types of descriptions provided, we randomly assigned some respondents to see *basic* descriptions and others to see *detailed* descriptions.

- [Basic] "**Achievement**" rates how well students understand the material and skills tested.
 [Detailed] "**Achievement**" rates how well students understand the material and skills tested. Differences between schools could be due to differences in how much students learn at school or other differences between students (e.g., how much they learn at home).
- [Basic] "**Gains**" rates how much students' understanding of the material and skills tested increases during a school year.
 [Detailed] "**Gains**" rates how much students' understanding of the material and skills tested increases during a school year. It compares the learning of students in one school to the learning of similar students in other schools (ruling out factors like how much they learn at home).
- "**School Climate**" shows how students, parents, and staff rate the quality of the relationships and environment in the school.

What we found:

The patterns were clear and highly consistent across different groups of respondents and treatment conditions. First, schools with a “Low” rating in any of the three categories generally sunk to the bottom of respondents’ lists. The three schools with a “Low” rating received the worst average rankings, despite having two “High” ratings to accompany their “Low” rating. A second pattern appears after this “Low” distinction. On average, respondents preferred, in order, higher ratings in “Achievement” and then “Gains” and then “School Climate.” This was true both for the full sample and across a wide range of respondent characteristics.

Below, the schools appear in order from the best-ranked to the worst-ranked, on average, for our full sample.

Rank	School Name	School Ratings		
		Achievement	Gains	School Climate
#1	Oak	High	Medium	Medium
#2	Redwood	Medium	High	Medium
#3	Pine	Medium	Medium	High
#4	Maple	High	High	Low
#5	Willow	High	Low	High
#6	Hickory	Low	High	High

Finding 3. Seeing positive or negative parent comments can be remarkably influential on Americans' assessments of local schools, even when those comments come alongside numerical academic ratings.

What we did:

In addition to providing profiles of hypothetical schools, as described above, we also asked respondents to assess actual schools in their communities. We provided each respondent with the names of two schools that shared their home's 5-digit zip code. After answering questions about their familiarity with each school, respondents saw two types of information for each school.

- First, they saw that school's most recent "GreatSchools Rating." These ratings show each school's performance on the most recent state assessments. Schools are rated 1-10, with 10 being best, based on how their students performed, on average, relative to students in other schools. (These are not value-added scores.)
- Second, they saw two comments for each school. We pulled actual parent comments from the GreatSchools website, cleaned them up (e.g., eliminated some spelling errors), and then categorized them by tone ("positive," "mostly positive," "mostly negative," or "negative") and content ("instruction/learning" or "school culture"). Respondents did not see these classifications. For one school a respondent saw one positive comment and one mostly positive comment; for the other school she saw one negative comment and one mostly negative comment. We randomly determined which school would have positive comments attached and which would have negative comments attached. We explained to respondents that they were seeing a random selection of comments from a website with parent comments on these and other schools.

After respondents read these academic ratings and comments, they were asked, "All in all, how would you grade each school?" They could assign a grade from A+ through F. It is important to note that many respondents had prior opinions of and experiences with these schools. We wanted to take an authentic setting – one in which people have preexisting attitudes toward schools – and see how attitudes might change upon seeing additional information.

The advantage in randomly determining which schools would have two positive comments and which schools would have two negative comments is that any differences that arise between the average ratings of positively reviewed schools and the average ratings of negatively reviewed schools would be plausibly caused by these comments. In other words, there is no reason to believe that schools showing positive comments are better, on average, than schools showing negative comments.

We also experimentally manipulated the described source of the academic ratings ("an independent nonprofit organization" or "your state government"). However, as described above, we did not observe any statistically significant differences associated with this experimentation.

What we found:

The gap between respondents' grades for schools with positive comments and schools with negative comments was striking. On average across our sample, seeing positive comments rather than negative comments led respondents to rate schools approximately 2/3 of a full grade better (e.g., the difference between an "A-" and a "B"). For context, a one-point difference in GreatSchools Rating – the numerical academic rating on a 1-10 scale – was associated with ratings of only about 1/10 of a grade better (e.g., only

a small fraction of the difference between a “B+” and a “B”). GreatSchools Ratings were not randomly assigned, which makes comparing these effects difficult.

There are some interesting, statistically significant differences across respondent groups in how they responded to the information provided. Their parent reviews were particularly influential with women and respondents less familiar with the schools reviewed. Among respondents with more formal education, grades were more strongly associated with academic ratings.

Still, while these differences are notable, so, too, is our finding that parental reviews were influential with every subset of our sample that we examined (for which we have sufficient data). Take, for example, parents with children under the age of 18 who report being “very familiar” with the schools they are evaluating. Even though many of these respondents likely have children in the schools presented, the difference between seeing a couple of positive comments and seeing a couple of negative comments from an unknown source – while looking at academic ratings and having their own preexisting opinions – was enough to shift their on-the-spot assessments by approximately 1/2 of a full grade (e.g., 1/2 of the way from a “B” to an “A”).

The effects of parent comments were very large regardless of whether respondents saw comments about instruction/learning or comments about school culture. This feature was also randomly assigned, such that a single respondent saw comments about only one of these characteristics or the other.



For additional information, please contact:

Susanna Loeb
Professor of Education, Stanford University
sloeb@stanford.edu

Jon Valant
Doctoral Candidate, Stanford University
jvalant@stanford.edu