

The Effects of Achievement and Growth Information on Perceptions of School Quality: Evidence from a Survey Experiment

Pre-Analysis Plan

David M. Houston, Harvard University
Michael Henderson, Louisiana State University
Paul E. Peterson, Harvard University
Martin R. West, Harvard University

May 30, 2019

Project Overview

We seek to identify the effects of providing district-level average student achievement information and/or average student growth information on participants' attitudes towards the public schools in their community.

By achievement, we refer to students' average academic performance at the end of one period of time. By growth, we refer to students' average rate of improvement in academic performance over the same period of time. Please see the "Other Data Sources" section for more information on how we measure achievement and growth.

This project consists of an online survey experiment embedded in the 2019 Education Next Poll, a nationally representative survey of Americans' attitudes towards education issues conducted by the polling firm Ipsos via its KnowledgePanel® platform. The experiment contains four parts:

1. Participants will estimate their district's percentile ranking in terms of average achievement and their district's percentile ranking in terms of average growth (compared to all districts nationwide).
2. Participants will be randomly assigned to receive either their district's average achievement percentile ranking, their district's average growth percentile ranking, both, or neither.
3. Participants will answer a question about how they would grade the public schools in their community on a scale from A – F.
4. Participants will answer a question about how much schools should focus on student academic performance versus student social and emotional wellbeing.

We are conducting this experiment with approximately 3,000 participants in May 2019. We expect to receive data in the first half of June.

Research Questions

Because of the large number of research questions, we are dividing them into three categories: primary, secondary, and tertiary. The results of the analyses associated with the secondary research questions should be viewed as exploratory. The results of the analyses associated with

the tertiary research questions should be viewed as speculative.

Below the primary and secondary research questions, we provide a short summary of the associated analyses. For more detail, please see the “Analyses” section.

Primary Research Questions

1. To what extent are individuals able to estimate average achievement and average growth in their school districts?
 - On average, how close are participants’ estimates of their districts’ achievement and growth percentiles to the districts’ actual achievement and growth percentiles?
2. Does the provision of achievement and/or growth information affect confidence in the local public schools?
 - What are the average effects of receiving achievement and/or growth information on the grades that individuals assign to the public schools in their community?
3. Does the provision of this information increase the salience of academic performance at the expense of other educational objectives?
 - What are the average effects of receiving achievement and/or growth information on the importance of academic performance relative to social and emotional wellbeing?
4. Do these effects vary by the academic performance of participants' districts?
 - Do the average effects in RQ2 and RQ3 vary by 1) the district's achievement percentile, 2) the district's growth percentile, or 3) the difference between the district's growth and achievement percentiles?

Secondary Research Questions

5. Do these effects vary by the demographic composition of participants' districts?
 - Do the average effects in RQ2 and RQ3 vary by 1) the percentage of white students in the district or 2) the percentage of FRPL-eligible students in the district?
6. Could these effects be the result of participants updating their prior beliefs about academic performance in their districts?
 - Do the average effects in RQ2 and RQ3 vary by 1) the difference between estimated and actual achievement percentiles or 2) the difference between estimated and actual growth percentiles?

Tertiary Research Questions

7. Does the provision of achievement and/or growth information affect other attitudes towards education issues?
8. Is there evidence of treatment effect heterogeneity by 1) participant characteristics or 2) other characteristics of participants' districts?

Participant Pool

The 2019 Education Next Poll is currently being administered to a nationally representative sample of the general population of American adults as well as to oversamples of teachers, African-Americans, and Hispanics. By early June 2019, we expect a total of approximately 3,000 participants to complete the survey.

The 2019 Education Next Poll also oversamples parents of high school students as well as their children for a separate study that compares the responses of these parent-child pairs. These individuals will not participate in the experiment described here.

Other Data Sources

For measures of district-level average achievement, average growth, free and reduced-price lunch (FRPL) eligibility, and racial composition, we use the Stanford Education Data Archive v2.1 (SEDA). SEDA contains data from state standardized tests in reading and math in grades 3-8 administered from 2009-2015 for almost every school district in the US. SEDA defines school districts in geographic terms: the dataset contains student performance data for all public schools located in the geographic boundaries of the district, including charter schools. For each district, SEDA contains average achievement and growth in reading and math as well as the average across both subjects (we employ these combined values in our experiment). The student test score data have been converted to a common scale that allows district-to-district comparisons across the country. We use the empirical Bayes grade cohort scale (GCS) estimates for the measures of achievement and growth. Achievement is measured such that a score of six represents a school district where the average student scores at about the same level as the average sixth grader in the national reference cohort (students who entered fourth grade in 2009 and eighth grade in 2013). Growth is measured such that a score of 1.2 represents a school district in which the average student's test scores improve about 1.2 grade level equivalents in one year. To aid in the interpretability of these values for participants, we provide achievement and growth scores in terms of national percentiles.

Random Assignment

Participants will be randomly assigned with equal probability to one of four experimental conditions:

1. Participants will receive their district's percentile ranking in terms of average achievement (the achievement group).
2. Participants will receive their district's percentile ranking in terms of average growth (the growth group).
3. Participants will receive both their district's percentile ranking in terms of average achievement and their district's percentile ranking in terms of average growth (the both group).
4. Participants will receive neither their district's percentile ranking in terms of average achievement nor their district's percentile ranking in terms of average growth (the control group).

Intervention

First, participants will estimate how their district performs in terms of average achievement:

The next few questions are about the **current level** of student academic performance and the **rate of growth** or improvement in student academic performance in your school district from one year to the next.

Enter any number from 0 to 100.

I think the **current level** of student academic performance in my school district is better than [number box, range 0-100] percent of other districts in the United States.

[NUMBER BOX, 0-100]

Next, participants will estimate how their district performs in terms of average growth:

Enter any number from 0 to 100.

I think the **rate of growth** in student academic performance in my school district is better than [number box, range 0-100] percent of other districts in the United States.

[NUMBER BOX, 0-100]

Participants assigned to the control group will receive the following:

Students are often given the grades A, B, C, and D, and Fail to denote the quality of their work. Suppose the public schools themselves were graded in the same way. What grade would you give the public schools in your community?

1. A
2. B
3. C
4. D
5. Fail

Participants assigned to the achievement group will receive the following:

According to the most recent information available, the **current level** of student academic performance in your school district is better than in [INSERT X_ach_pct] percent of districts and worse than in [INSERT _100MINUSX] percent of districts.

Students are often given the grades A, B, C, and D, and Fail to denote the quality

of their work. Suppose the public schools themselves were graded in the same way. What grade would you give the public schools in your community?

1. A
2. B
3. C
4. D
5. Fail

Participants assigned to the growth group will receive the following:

According to the most recent information available, the **rate of growth** in student academic performance in your school district is better than in [INSERT Y_gro_pct] percent of districts and worse than in [INSERT _100MINUSY] percent of districts.

Students are often given the grades A, B, C, and D, and Fail to denote the quality of their work. Suppose the public schools themselves were graded in the same way. What grade would you give the public schools in your community?

1. A
2. B
3. C
4. D
5. Fail

Participants assigned to the both group will receive the following:

According to the most recent information available, the **current level** of student academic performance in your school district is better than in [INSERT X_ach_pct] percent of districts and worse than in [INSERT _100MINUSX] percent of districts.

Additionally, the **rate of growth** in student academic performance in your school district is better than in [INSERT Y_gro_pct] percent of districts and worse than in [INSERT _100MINUSY] percent of districts.

Students are often given the grades A, B, C, and D, and Fail to denote the quality of their work. Suppose the public schools themselves were graded in the same way. What grade would you give the public schools in your community?

1. A
2. B
3. C
4. D
5. Fail

All participants will receive the following (the sequence of “student academic performance” and “student social and emotional wellbeing” will be randomly assigned to eliminate ordering effects):

How much should schools focus on student academic performance versus student social and emotional wellbeing?

Please give a percentage for each. Your answers should add to 100%.

1. Student academic performance	[NUMBER BOX, 0-100] %
2. Student social and emotional wellbeing	[NUMBER BOX, 0-100] %
Total	[SHOW SUM OF BOXES]

All participants will then complete the remainder of the survey.

Outcome Measures

There will be two primary outcome measures:

1. Local school grade (1 – 5 scale)
2. Relative importance of academic performance (0-100%)

For RQ7, we will consider participants’ answers to survey questions that occur after the experimental intervention (attitudes towards education spending, policy preferences, etc.).

Covariate Measures

Ipsos will provide us with demographic information for each participant: race/ethnicity, teacher status, parent status, Spanish language status, political party identification, political ideology, household income, US citizenship status, age, educational attainment, gender, head of household status, housing type, marital status, and employment status.

Missing Data

We expect a small number of participants will live in areas where we do not have data on district achievement and/or growth. In these cases, participants will receive the median value (the 50th percentile). When exploring treatment effect heterogeneity by district characteristics, we will also impute the median values of these characteristics for these participants.

If participants are missing data on a demographic variable, we will recode the missing value with “999” and control for an indicator of missingness in subsequent analyses.

If participants do not answer the local school grades question or the relative importance of academic performance question, they will be dropped from the analysis. We will report the

percentage of missing cases in each experimental group.

Weighting

Ipsos will compute post-stratification statistical weights for the total sample and by sample group (i.e., oversamples of teachers, African-Americans, and Hispanics) to ensure these samples are representative of their respective adult populations.

Ipsos describes their weighting procedure in the following way:

“Depending on the specific target population for a given study, geodemographic distributions for the corresponding population are obtained from the CPS, the American Community Survey (ACS), or in certain instances from the weighted KnowledgePanel® profile data. For this purpose an iterative proportional fitting (raking) procedure is used to produce the final weights. In the final step, calculated weights are examined to identify and, if necessary, trim outliers at the extreme upper and lower tails of the weight distribution. The resulting weights are then scaled to aggregate to the total sample size of all eligible respondents.”

In this process, Ipsos weights samples by age, gender, race/ethnicity, census region, metropolitan status, education, and household income.

All analyses will incorporate these weights.

Analyses

To check for balance between experimental groups, we will compare the demographic composition of the control group with the demographic compositions of each of the other randomly assigned groups. To accomplish this, we will use a series of ordinary least squares (OLS) regressions:

$$X_i = \beta_0 + \beta_1 \text{Achieve}_i + \beta_2 \text{Growth}_i + \beta_3 \text{Both}_i + \varepsilon_i,$$

where X_i is one of the available demographic covariates; Achieve_i , Growth_i , and Both_i are indicators of experimental group status; and ε_i is the error term for individual i .

To answer RQ1, we will calculate a range of descriptive statistics for estimated and actual achievement and growth:

- Mean/SD of estimated achievement
- Mean/SD of actual achievement
- Mean/SD of the difference between estimated and actual achievement
- Correlation between estimated and actual achievement
- R^2 of relationship between estimated and actual achievement
- Mean/SD of estimated growth
- Mean/SD of actual growth

- Mean/SD of the difference between estimated and actual growth
- Correlation between estimated and actual growth
- R^2 of relationship between estimated and actual growth

The treatment effect heterogeneity analyses associated with RQ5 are based on the premise that district-level student racial composition and district-level student socio-economic composition have different underlying relationships with average achievement and average growth. We will calculate the following correlations for the districts of the participants in our sample:

- Percent white and achievement percentile
- Percent white and growth percentile
- Percent FRPL and achievement percentile
- Percent FRPL and growth percentile

When calculating average treatment effects, we will rely on the following general model:

$$Y_i = \beta_0 + \beta_1 \text{Achieve}_i + \beta_2 \text{Growth}_i + \beta_3 \text{Both}_i + \delta' X_i + \varepsilon_i,$$

where Y_i is the outcome (local school grades or relative importance of academic performance) and X_i is a vector of demographic characteristics. The default comparison will be to the control group, but we will also be interested in calculating average treatment effects with each of the other three experimental conditions considered as the comparison group.

When calculating heterogeneous treatment effects, we will rely on the following general model:

$$Y_i = \beta_0 + \beta_1 \sum_j \text{Group}_i^j + \beta_2 Z_i + \beta_3 (Z_i \sum_j \text{Group}_i^j) + \delta' X_i + \varepsilon_i \quad \text{with } j \in \{a, g, b\},$$

where Group_i^a represents the achievement group, Group_i^g represents the growth group, Group_i^b represents the both group, and Z_i is the heterogeneous dimension (district achievement percentile, district growth percentile, the difference between the achievement and growth percentile, the difference between estimated and actual achievement, the difference between estimated and actual growth, percent white, or percent FRPL).

To test for statistical significance, we will rely on two-tailed hypothesis tests with $\alpha = 0.05$.

Institutional Approval

This project has been approved by the Harvard University Institutional Review Board (Protocol: IRB19-0507)

Pre-Registration

This study is registered on the American Economic Association RCT Registry.