nature research

Corresponding author(s): Todd Rogers, Avi Feller

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

Describe how sample size was determined.

We used administrative records from the School District of Philadelphia in the summer preceding the 2014-2015 School Year (SY) to identify all students enrolled in a public school in the district. This yielded an initial study universe of 161,922 students in 103,408 households. After making the exclusions listed below, we were left with a final analytical sample of 28,080 households with the same number of focal students.

2. Data exclusions

Describe any data exclusions.

Prior to randomization, we excluded opt-outs and households whose consent forms bounced back; students who were enrolled in "non-regular status" schools; students who had graduated, withdrawn, or were not enrolled as of June 2014; students who were enrolled in the schools used for the pilot study; students who were homeless; students who were disabled; students whose home language was not one of the consent form languages; students who had perfect attendance in the prior school year; students with extremely high absenteeism (more than 2 standard deviations above the mean); students who had prior absences within three days (or fewer) of the modal number of absences for their specific school-grade; and students in overly large households. Finally, we selected one focal student from all households with more than one eligible student. All exclusions are described in greater detail in the manuscript and SOM.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

This study was a replication of a pilot study that is reported in the SOM.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

We randomly assigned focal students to one of four treatment conditions (including a control group) with randomization stratified by school, grade, and prior attendance. We divided each school-grade based on prior year absences. Due to differing sample sizes, we divided school-grades at the high school level into four prior attendance groups (i.e., by quartiles) and divided school-grades at the elementary and middle school level into two prior attendance groups (i.e., below/above the median). If any resulting strata had fewer than four qualifying students, we instead used student-grade as the stratum for all students in that student-grade.

We generated a standard uniform variate for each student and then ranked the students within each stratum. We first randomly selected a starting treatment condition for assignment and then assigned students to each treatment arm in the following order: (1) Control; (2) Reminder; (3) Total Absences; (4) Relative Absences. This sequence repeated until all students in the stratum were assigned to a treatment arm.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

The School District of Philadelphia and its teachers and staff were blinded to treatment group allocation for the entirety of the study. The principal investigators were aware of group allocation during analysis, but no data were analyzed before the analysis plan was pre-registered at socialscienceregistry.org.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☒	☐ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.*
☒	☐ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	StataSE 2013; R

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.	No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used.
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used.
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No commonly misidentified cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Our consent universe consisted of 161,922 students in 103,408 households. After pre-randomziation exclusions, we were left with 32,437 students in 40,326 households. We randomly assigned each household to one of four treatment regimes, and ensured all key covariates were balanced across conditions (see SOM for details).

The population of students is diverse: about 53% are Black/African American, 20% are Hispanic/Latino, and 17% are White/Caucasian. Almost three-quarters of students in our sample qualify for free or Reduced Price Lunch, and about 7% have limited English proficiency. Among the 110,229 students in the consent universe for whom we have baseline data, the average number of absences in 2013-2014 was 15.3 days (SD=17.7 days) with a minimum of zero and a maximum of 171 days absent. The average number of baseline absences in our final experimental universe was 16.3 days (SD=10.4 days), with a minimum of three and a maximum of 97. The tighter distribution of absences in our experimental universe is by design, as explained in the manuscript and SOM.