Rohen Shah
University of Chicago

Pre-Analysis Plan

March 15, 2024

Study Title:
When The Student Becomes the Master: Learning by Creating Math Tutoring Videos

AEA Registry: AEARCTR-0011884 **IRB Number**: IRB23-0041

Study Overview:
Teaching is often considered a good way to improve one's own knowledge. While there is lab evidence that preparing to tutor helps one's own understanding of content, the extent to which this generalizes to real world settings is unclear. In this study, I will partner with 22 middle and high schools in the midwestern United States (Illinois, Wisconsin, Iowa, Ohio, and Michigan) to test the impact of creating math tutoring videos on students' math skills. The study will take place during the Spring 2024 semester, and depending on the success of implementation, additional rounds might be conducted during the summer 2024 and fall 2024 semesters. Teachers who teach at least two periods of the same math class will be eligible to participate. These periods will randomly assigned to the "Creators" and "Watchers" conditions. In cases where teachers teach more than two periods, one period will also be assigned to a control condition. The class period assigned to the "Watchers" condition will be asked to solve a weekly math problem (in addition to their usual assignments), primarily taken from the PSAT and ACT, and will be provided with a "help" video that shows how to answer a similar problem. The class period assigned to the "Creators" condition will receive the same problem, but their task will be to create a video explaining the solution. Control periods will not receive any weekly tasks beyond their regular class assignments. 8 tasks will be given in total between March 2024 and May 2024. Students will be given a pre- and post- test on the PSAT/ACT questions that cover material related to the treatment tasks. Additionally, I will track how often the "help" video link is clicked in each class period, as well as the compliance and accuracy of the weekly tasks for both Creators and Watchers.

Sampling

Recruitment Phase 1: Districts

School districts were recruited via email to superintendents during December 2023 and January 2024. The recruitment email asked whether the district would be interested in participating in the study, and was sent to all districts with at least 5,000 students in Illinois, Wisconsin, Iowa, Michigan, and Ohio. I received 94 replies in total, of which 29 indicated interest in learning more. Of these, 13 agreed to participate and shared information about the study with middle and high school teachers. The most common reasons for not being able to participate was scheduling conflicts with other school improvement initiatives and extensive research review processes that would not fit in our timeline for the current school year.

Recruitment Phase 2: Teachers

Teachers at participating school districts were sent an interest form by their districts with information about the study. Teachers receive $500 for their effort in helping implement the study. Teachers who teach at least two periods of the same math class were eligible. Teacher recruitment took place on a rolling basis between February 5th and March 8th, 2024. In total, 47 teachers filled out the interest form, of which 41 teachers (representing 10 middle schools and 12 high schools) chose to proceed with the study after learning more.

Recruitment Phase 3: Parent Permission

The teachers distributed parent permission forms to all students in their classes. The permission form indicated whether the students were allowed to take a short survey with demographic information and their math background, as well as whether the teacher could share identifiable data with the researchers (including student-generated videos). Students whose parents did not give permission still participated in the tasks according to their class period, but they did not take the survey and only de-identified data was shared for these students.

Randomization

Level of Randomization

To improve compliance and minimize spillover concerns, students are randomized at the classroom level rather than individual level. The treatment tasks are assigned as homework assignments for completion credit.

Randomization Timeline

Students took a pre-test consisting of 15 multiple choice questions taken from either the PSAT, ACT, SAT, or grade-level state standardized math test. The questions were chosen based on the topics the teachers indicated they planned to cover between March and May 2024. The teachers administered the pre-test along with a survey that asked about students' age, gender, race, math confidence, growth mindset, as well as past experience with tutoring and video creation. Only students whose parents gave data sharing permission took the survey. This in-class pre-test was 25 minutes, and teachers administered these on a rolling basis between March 4th and March 29th. Each class period will be randomized to a treatment condition only after all participating periods for that teacher have completed the pre-test. Randomization will occur on a rolling basis beginning March 18th, 2024, stratified by teacher.

Randomization Protocol

Each period will be randomized to either "Creators", "Watchers", or "Control" conditions. Randomization will occur one teacher at a time. The following protocol will be used to determine the number of periods that will be assigned to each condition in order to minimize the likelihood of a lopsided randomization for a given teacher:

1) If there are two periods: one is "Creators", one is "Watchers".
2) If there are three periods:
    a. If each period has at least 15 students, then one will be randomly assigned to "Creator", one will be randomly assigned to "Watchers", and one will be randomly assigned to "Control".

b. If at least one period has fewer than 15 students, then there are two possibilities:
    i. If the sum of the number of students in the smaller two periods is *at least equal to* 1.25 times the number of students in the largest period, then the three periods will be randomly assigned to Creator, Watcher, and Control conditions.
    ii. If the sum of the number of students in the smaller two periods is *less than* 1.25 times the number of students in the largest period, then the smaller two periods will be treated as a single "unit" for the purpose of randomization, and then the two total "units" (consisting of the largest period as one unit, and the two smaller periods as the second unit) will be randomized to either the Creator or Watcher condition.

3) If there are four periods:
    a. The smallest two periods will be combined into a single unit for the purpose of randomization, and then the three units of randomization (which consists of 1) the smallest two classes, 2) the second-to-largest class, and 3) the largest class) will be randomly assigned to either Control, Watcher, or Creator conditions, provided that each unit of randomization has at least 15 students.
    b. If at least one unit of randomization (from 3a) has fewer than 15 students, then there are two possibilities:
        i. If the sum of the number of students in the smallest two units of randomization is *at least equal to* 1.25 times the number of students in the largest period, then the three units of randomization will be randomly assigned to Creator, Watcher, and Control conditions.
        ii. If the sum of the number of students in the smallest two units of randomization is *less than* 1.25 times the number of students in the largest period, then there will only be two units of randomization. These two units of randomization will either consist of the following two possible groupings: 1) The smallest 3 periods as one unit, and the largest period as the second unit, or 2) The largest and smallest periods combined as one unit, and the second largest and third

largest periods combined as the second unit. The grouping that minimizes the difference between the number of students in the larger and smaller unit of randomization will be used. The two units of randomization will be randomized to Watcher or Creator.

4) If there are five periods:
    a. If each period has at least 15 students, then two will randomly be assigned "Creator", two will be randomly assigned "Watcher", and one will be randomly assigned to "Control"
    b. If at least one period has fewer than 15 students, then the smallest two periods will be combined so that there are 4 total units of randomization. Then, the steps in item 3a) and 3b) above will be followed for assigning these four units of randomization to Creator, Watcher, or Control.

Randomization in Stata

Each unit of randomization (period or set of periods as described above) will be assigned a random number from a uniform distribution between 0 and 1. The period with the largest randomly generated number will be assigned Creator, the second-largest number will be Watcher, and the smallest will be Control. The seed in Stata will be set to 42, and thereafter each teacher's class periods will be randomized in the order in which they complete the pre-assessment.

Study Variables

Survey Variables

The following variables will be collected via survey for students with parent permission:

- Age
- Gender
- Race
- Math confidence
- Growth mindset for math
- Past experience receiving tutoring
- Past experience providing tutoring
- Past experience with video creation

Any of the 3 demographic variables above (Age, Gender, and Race) that are imbalanced at baseline will be controlled for. Math confidence is a secondary outcome of interest. The remaining variables will be used for exploratory heterogeneity analysis.

Primary Outcome Variables

The two primary outcomes will be 1) the researcher-provided post-test score, and 2) the final math class grade. The post-test will be a 15-question multiple choice question containing problems taken from the SAT, ACT, PSAT, or grade-level state tests. These problems will be similar to the types of problems that were given as treatment tasks.

Secondary Outcome Variables

The two secondary outcomes are 1) the students' self-reported math confidence, and 2) the students' likelihood of viewing the "help" video. The math confidence level will be measured on a scale of 1 to 10 during the student survey in May 2024. The students' likelihood of viewing the "help" video is a proxy for the amount of effort they put into the task, and will be measured as the number of total clicks on their period's link divided by the total number of students in that period.

Regression Analysis

Primary Treatment Effect Estimate:
The regression below will be the primary estimate of the treatment effect:

$$y_{it} = \beta_0 + \beta_1 T_i + \beta_2 C_i + \beta_3 y_{i,t-1} + \tau_j + \varepsilon_i$$

Here, y_{it} is the endline level of the outcome variable (either post-test, grade, or math confidence) and $y_{i,t-1}$ is the baseline level of the outcome variable. T_i is a binary variable with a value of 1 for students who are in either the Creator or Watcher group, and 0 otherwise. C_i is a binary variable with a value of 1 for students who are in the Creator group, and 0 otherwise. τ_j is a teacher fixed effect, and ε_i is the standard error clustered at the classroom level.

Three teachers indicated that they may choose to do the tasks as an in-class activity each week rather than as homework, because they did not have a way to enforce homework completion with grades. If this happens, then the main treatment effect regression will drop these teachers so that the treatment effect being estimated is that of a comparable treatment across all observations.

Video Views Estimate:
The following regression will estimate the impact of the Creators treatment on the likelihood of viewing videos. Here, each observation is a task-period dyad:

$$V_k = \alpha_0 + \alpha_1 C_k + u_k$$

Here, V_k is the "Viewing Likelihood", computed as the total number of views for a given task's "help" video divided by the number of students in that class period. C_k is a binary with a value of 1 for class periods that are assigned to Creators and 0 otherwise, and u_k is the standard error. Exploratory analysis here will include whether the viewing likelihood rate varies over time and whether it depends on the level of task difficulty.

Alternative Outcome Specification:

As a robustness check, I will additionally run the main regression but with an alternate specification for the outcome to measure nonlinearity in treatment effect. I will run 3 linear probability model regressions where the outcome is a binary variable indicating whether the student is in the 1) top quartile, 2) top half, or 3) bottom quartile of grades and test scores.

Local Average Treatment Effect:

In addition to the Intent-to-Treat analysis in the main regression, I will estimate the Local Average Treatment Effect (LATE) using task completion as the measure for compliance. I will display regression results for three alternative measures of compliance: 1) Percentage of all tasks completed, 2) Binary variable indicating completion of at least half of the treatment tasks, and 3) Binary variable indicating a task completion rate above the 75^{th} percentile.

Attrition Analysis:

Attrition might be an issue for analyses that include survey variables, as well as administrative variables where some students might have left the school district and don't have an endline score. I will test whether attrition leads to internal validity concerns by conducting a test for selective attrition to determine whether treatment status is predictive of attrition. I will additionally show whether any of the covariates are predictive of attrition, as well as bounding analyses using Lee bounds to show a range of possible treatment effects.

Exploratory Heterogeneity Analysis:

In addition to the main analyses above, I will also conduct exploratory analysis to see whether the treatment effect is moderated by race, gender, confidence, growth mindset, and prior experience with tutoring and video creation.