Working Paper:

The Effects of Nudges on Students' Effort and Performance: Lessons from a MOOC

Ignacio Martinez[1]

This paper uses a Massive Open Online Course (MOOC) as a laboratory to identify the effect of informational nudges on students' effort and performance. I find that: (i) very low-cost informational interventions nudge students into exerting more effort; (ii) more effort is, in some cases, translated into higher achievement; (iii) framing plays an important role. Positive effects on achievement are observed for new quizzes that are given after the intervention. Meanwhile, The negative treatment tends to change outcomes for those who were doing relatively poorly, and the positive treatment tends to work for those who were doing relatively well. As an example of the magnitude of the effects, students assigned to the negatively framed treatment (telling them the proportion of students doing better than they) attempt a quiz in the week of the intervention more times and are ranked, on average, 8.43 percentage points better than students in the control group. While these experiments are conducted using a MOOC, low-cost informational interventions could also improve achievement in traditional classrooms.

[1]University of Virginia

Updated February 2014

EdPolicyWorks
University of Virginia
PO Box 400879
Charlottesville, VA 22904

Acknowledgements: I am grateful to Sarah Turner, Leora Friedberg, Kristin Palmer, Eli Bildner, Chuong Do, Jacalyn Huband, and Katherine Holcomb. Also, thanks to the Jefferson Trust for their financial support. I am solely responsible for any errors. The most recent version of this paper is available at http://goo.gl/QPJpDo.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

The Effects Of Nudges On Students' Effort And Performance
Ignacio Martinez

1 Introduction

The online education environment provides a powerful space for improving the understanding of the incentives and informational environment that students face and how they affect student actions and, ultimately, learning. Massive Open Online Courses (MOOCs) have the potential to facilitate low-cost implementation of randomized control trials; the sample sizes are large (and growing); and while many student behaviors such as time spent on homework and class participation in bricks-and-mortar classrooms have historically been costly to measure except through self and/or teacher reports, the majority of online learning activities can be observed by the researcher. While many of these lessons may be translated to bricks-and-mortar classrooms, distance learning may itself become an important component of human capital formation.

In the classroom, teachers may provide informal information about average student performance, in addition to informing students individually about their own outcomes. However, we know little about whether providing information, much less the form in which it is provided, has any effect on student effort and probably most teachers have given little thought to how this may motivate future effort. This paper assesses the extent to which very low-cost informational interventions affect how students exert effort and, ultimately, their learning success. First, I study whether modest "informational nudges," a concept introduced by Sunstein and Thaler [2008], can affect students' behavior. Second, I look at the role that "framing," an idea introduced by Tversky and Kahneman [1985], has on how effective these nudges are. Finally, I examine how the informational nudges affect learning outcomes.

This study uses data from Coursera, a social entrepreneurship company partnering with 108 top universities in the world to offer MOOCs.[1] As of December 2013 Coursera had 5.8 million users. Coursera users watch video lectures to learn class material, are evaluated via online quizzes, and use forums to communicate with fellow students and the instructor. I use data from the second edition of *Foundations of Business Strategy (FBS)* by Michael J. Lenox of the University of Virginia. This Coursera course enrolled 64,415 students and ran from September 9, 2013, through October 11, 2013. For this study, only students who completed the first quiz before September 15 at 3:00 p.m. Eastern standard time were included. It

[1] Martinez and Diver [2014] discuss the challenges and potential of the data generated by Coursera.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

has been widely reported that most MOOC students do not complete the course nor even do any of the assignments, and this criterion reduces the number of observations to 7,924, which is still a very large number considering that the third largest college freshman class in the U.S. has 7,740 students.[2]

I partnered with Coursera, which allows me to observe many of the inputs (e.g., lecture watching, forum participation, number of attempts for each quiz) and outputs (e.g., grades for each attempt at each quiz, forums reputation), this is not sufficient data to understand how achievement is produced. This is because student inputs (e.g., effort) are endogenous choices. Therefore, to understand the causal effect of effort on achievement, I need an exogenous source of variation for effort. My research design generates that exogenous variation by randomly selecting a group of students to receive an email with information about their performance relative to their peers, with the goal of inducing them to alter their effort. Thus, I study the effect of both information on effort and effort on achievement.

Drawing from the economics of education and psychology literature, I designed two emails intended to nudge students into exerting more effort. These emails were sent to 65% of the 7,924 students who took Quiz 1. The remaining 35% of students form the control group. The emails contained information about the relative performance of the student on Quiz 1. The two sets of emails differ only in whether they informed the students of the percentage doing better or worse than themselves.

I provide evidence that this simple email that informs students of their relative performance on the previous quiz nudges them into exerting more effort and that this effort translates, in some cases, into higher achievement. Positive effects on achievement are observed for new quizzes that are given after the intervention. This suggests that students are better at adjusting so that they can learn new material rather than relearn stuff from the past. For example, students assigned to the negatively framed treatment (telling them the proportion of students doing better than they) attempt a quiz in the week of the intervention more times and were ranked, on average, 8.43 percentage points better than students in the control group. Finally, I show that the negatively framed nudge had persistent effects. In the last week students assigned to the negatively framed nudged attempted that week quiz more times and were ranked, on average, 5.77 percentage points better than students in the control group.

I also show that framing plays an important role. The negative treatment tends to

[2] In 2008, Arizona State University had the third largest freshman class size. The largest is Miami Dade College with 8,993. See Forbes: http://goo.gl/zllze8

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

change outcomes for those who were doing relatively poorly, and the positive treatment tends to work for those who were doing relatively well.

The remainder of this paper is organized as follows. In the next section, I discuss related research and the conceptual framework for this work. In section 3, I describe the data and demonstrate that the randomization worked. In section 4, I explain the intervention and show the results. In section 5, I conclude.

2 Related Research and Conceptual Framework

The conceptual framework for this analysis brings together ideas from economics, education, and psychology about how student time investments (effort) affect achievement, insights from psychology about the extent to which relative performance targets affect behaviors, and behavioral economics ideas about how framing motivates behavior. The objective of this inquiry is to understand whether very low cost informational interventions affect how students invest time in a course environment and, ultimately, their learning outcomes.

Fryer [2013] conducted a randomized field experiment in Oklahoma City Public Schools that provided information to students on the link between human capital and future outcomes such as unemployment, incarceration, and wages. The essential element of the experiment was a cellphone that was provided to 1,470 students in the treatment group. Students received one text message per day containing this information. Three facts emerge from this study: (1) students update their beliefs about the returns to education in response to the text messages (2) students report that they are putting more effort into their work, and (3) there are no detectable changes in academic achievement. Fryer argues that the explanation for this is that students do not fully understand the education production function. Earlier Fryer work shows that paying young people to finish reading books (that is, inducing them to invest in inputs) has a bigger effect than incentives to do well on exams, see Fryer [2011]. Therefore, to the extent that informational interventions affect productivity-enhancing behavior, such interventions might improve learning outcomes.

The term nudge was first used by Sunstein and Thaler [2008] to describe “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any option or significantly changing their economic incentives.” For example, some firms offer employees the option of joining a program in which their saving rates are automatically increased whenever the employee gets a raise. This plan tripled saving rates in those firms. In this paper, I nudge Coursera students into changing their behavior by sending them

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

information.

Although MOOCs are relatively new, research interest in them is rapidly growing. Liyanagunawardena et al. [2013] found in a systematic study of published literature that through 2012, there had only been 45 distinct articles where MOOCs or their use are the primary focus. However, the pattern of publication shows a quickly increasing trend (from one in 2008 to 26 in 2012). Of these, only 15 of those were classified under the broad heading of "Educational Theory." In 2013, research using MOOCS has continued, but it is evidently still in an early state. The nascent research is still trying to understand who is enrolling in MOOCs and why (see, for example, Christensen et al. [2013]) and overall trends in completion and engagement (see, for example, Kizilcec et al. [2013]). Those researchers that have made use of the MOOCs data have typically limited their focus to classification or broad associative observations. Kizilcec et al. [2013], for example, present a simple classification method that identifies a small number of longitudinal engagement trajectories in MOOCs. A limitation of descriptive studies is that they can detect association among variables, but they cannot rule out the possibility that the association was caused by an omitted factor that is correlated with outcomes. For example, it is possible that some unobserved characteristic, such us being a hard worker, explains both forum participation and course completion, rather than forum participation increasing course completion. To overcome this selection bias, I use a randomized controlled trial. This is a rigorous way of determining whether a causal relation exists between effort and achievement outcomes.

Tversky and Kahneman [1985] present evidence showing how seemingly inconsequential changes in the formulation of choice problems caused significant shifts of preferences. Zhang and Buda [1999] study positive and negative framing on the advertising industry. They find that framing has a significant influence on consumer responses to advertisements. They also find that this effect is more salient for people with low need for cognition than a high need for cognition.[3] Framing effects can be viewed as heuristic errors. That is, if people are boundedly rational, then the presentation of a choice may draw attention to new aspects of a problem, leading to people to make mistakes in pursuing their true underlying preferences [Rabin, 1998]. If this is the case, students may react differently to the information that they are performing better than 20% of their classmates than to the information that they are doing worse than 80% of their classmates, even though this could be interpreted as violating

[3]The need for cognition (NFC), in psychology, is a personality variable reflecting the extent to which individuals are inclined towards effortful cognitive activities. Need for cognition has been variously defined as "a need to structure relevant situations in meaningful, integrated ways" and "a need to understand and make reasonable the experiential world", see Cohen et al. [1955] and Cacioppo and Petty [1982].

the assumption of rationality. Levitt et al. [2012] conducted a series of field experiments involving thousands of primary and secondary school students to explore this. They find that incentives framed as losses have more robust effects than comparable incentives framed as gains.

Numerous studies find that students' effort is positively correlated with academic achievement; see for example Carbonaro [2005], Johnson et al. [2001]. However, it is very difficult to disentangle the roles of pedagogical methods, students' effort, and individual characteristics (including prior preparation and innate ability) in the measurement of achievement. One reason is that these studies rely on student reported or teacher reported effort. Yet, self-assessment may be subject to substantial misreporting which is, in turn correlated with underlying student characteristics, resulting in biased estimates. Another problem is that effort is an endogenous choice. I designed a scalable very low cost intervention that generates a source of exogenous variation in effort to enable me to estimate causal effects of effort.

3 Data

Data from Massive Open Online Courses (MOOCs) provide an unprecedented opportunity to get inside the black box of student learning. These data include time-stamped logs of student activities such as viewing lectures, submission of assignments; participating in forums; clickstream logs (which track user activity on the course website); page views and lecture video interaction (e.g., video seek events); geolocation information from Internet Protocol addresses (IP addresses); all the courses a student is currently and has previously taken; student background surveys. Additionally, survey data provide some demographic characteristics such as age (figure 2), and level of education (figure 3).[4]

I use data from the second edition of Foundations of Business Strategy (FBS) by Michael J. Lenox of the University of Virginia. This Coursera course had 64,415 students enrolled and ran from 09/02/2013 to 10/11/2013. This course explores the underlying theory and frameworks that provide the foundations of a successful business strategy. The class is divided into weekly modules. Each weekly module consists of an introductory video, a reading from the strategist's toolkit, a series of video lectures, a quiz, and a case study to illustrate points in the lectures. Students wishing to receive a Statement of Accomplishment

[4] Currently Coursera does not have a very good survey system. Each course has its own survey question, which are in many cases the same basic demographic questions, and data is not shared across courses from. Martinez and Diver [2014] discuss in more detail the opportunities and challenges of working with these survey data.

must satisfy the following criteria:

1. Complete 6 quizzes. Students can take each quiz as many times as they want. The score of record is the best score on each quiz. Quizzes have 10 questions with 4 choices each.

2. Submit a final project (a strategic analysis for an organization of their choosing).

3. Assess five peers' strategic analysis using the peer assignment function.

Final grades are out of 100 points. Out of the 100 points of the final grade, 50 points are for the final project, 42 points for the quizzes (spread evenly over the 6 quizzes), and 8 points for post and comment up-votes. Those who received 70 points or more, and assessed five peers' strategic analysis, received a Statement of Accomplishment.

In this paper I use the 7,924 students who completed Quiz 1 before September 15 at 3:00pm EST. Of these students only 1,539 received a Statement of Accomplishment.[5] Table 1 shows when each of the quizzes was published.

Table 1: Quiz open times, 2013

Quiz #	Date
I	September 1
II	September 8
III	September 15
Treatment	September 16
IV	September 22
V	September 29
VI	October 6
Deadline	October 15

All quizzes share the same deadline on October 15^{th}; students can take each quiz as many times as they want; and the score of record is the best score on each quiz. Therefore, although the intervention happened at the beginning of the third week, students can retake Quizzes 1 or 2 if they wish.

[5] 1,997 of these students came from the U.S., followed by 878 from India. Figure 1 shows the students' geolocation. An interactive version of this map can be found at http://goo.gl/pw9oHz.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

3.1 Treatment / Control balance

In this section, I show evidence that the randomization worked as intended. In the next section, I will describe the randomized treatments, in which 35% of the sample were in the control group and half of the treatment group received either the positive or negative treatment. In a randomized controlled trial such as this, the econometrics entail fairly simple comparisons between the treatment and control group, so long as the groups were randomly selected.

Given the use of a random number generator to assign students to treatment and control status, the large number of students and the Law of Large Numbers, then observable and unobservable characteristics should be the same across groups. Table 2 summarizes the following variables for treatments and control groups. "First Grade" is the grade out of 10 points on the first attempt a student took Quiz 1; "Play 1" is the number of times the student presses the play button on the video player to watch lectures; "Threads Views 1" is the number of threads views on the course forums before attempting the quiz for the first time.

Table 2: Treatment / Control balance

	Treatment 1	Control	90% Confidence Interval	p-value
First Grade	6.927	6.8945	(-0.0495, 0.1144)	0.5142
Play 1	19.8791	20.1758	(-1.5971, 1.0036)	0.7074
Threads Views 1	2.1116	2.2385	(-0.5107, 0.2571)	0.5868
n	2589	2730		
	Treatment 2	Control	90% Confidence Interval	p-value
First Grade	6.9512	6.8945	(-0.0256, 0.1391)	0.2571
Play 1	19.3808	20.1758	(-2.0982, 0.5082)	0.3156
Threads Views 1	2.2672	2.2385	(-0.505, 0.5624)	0.9295
n	2605	2730		

Note: "First Grade:" Grade on the first attempt a student made for Quiz 1; "Play 1:" number of time the student presses the play button; "Threads Views 1:" number of threads views before attempting the quiz for the first time.

The results on Table 2 are consistent with the randomization working as intended, as the means are quite close and the t-tests always fail to reject equality of the means. Additionally, Figures 2 to 4 show that the distributions of age, education, and expected hours of course work are also consistent with the randomization having worked as intended.

4 Experimental and Results

4.1 Experimental Design

I designed the following experiment to disentangle the role of information and framing on students' choices, and ultimately the role of these choices on achievement. The experiment consists of nudging students with an email containing information about their relative class performance. The nudge is framed in two different ways to allow me to test the hypothesis that students react differently to positive and negative framing.

A day after the third quiz of FBS was published, on September 16^{th} 2013, I sent emails to 65% of the $7,924$ students who had taken Quiz 1. The remaining 35% of students form the control group. These emails differ only in their framing. Treatment 1 (T1) informs the student of the percentage of the class doing better than the student; treatment 2 (T2) informs the student of the percentage of the class doing worse. These emails read as follows:

> Subject: [Foundations of Business Strategy] Quiz 1
>
> *Dear [name],*
>
> *This information about your performance may benefit you. You obtained a [maxGrade] on the first quiz. That means that you are doing [better] / [worse] than [%worse] / [%better] of the class.*
>
> *Best,*
>
> *University of Virginia MOOC Research Team*
>
> *PS: This email was generated with data from Sunday September 15 at 3:00pm EST.*

I sent these emails via sendgrid.com, which provides additional analytics on whether the emails are read. The confirmed unique open rate for T1 was 38.33% with 6 bounces and 1 spam report. For T2, the confirmed unique open rate was 39.39% with 8 bounces and 0 spam reports, so the open rates are not statistically different from each other. To get the unique open rate, sendgrid inserts a white pixel in the body of the email. Because images are blocked by default on most email clients, these numbers are lower bounds for the open rates. The open rate is crucial to calculating the effect of the treatment on the treated.

Next, I discuss how the treatments affected the following outcomes and choices that reflect both effort and achievement: the decision to retake Quiz 1 after September 16, "Go

Back;" the number of attempts a student makes for a given quiz, "Attempts;" the maximum grade the student received, "Max Grade;" the percentile ranking for the student, "Ranking;" whether the student takes the next quiz, "Drop-out;" and the maximum grade for a student that decides not to take the next quiz, "Drop-out grade."[6][7]

People may respond by changing their effort on any or all of the quizzes. I expect the effects of the experiment to be different across quizzes for two reasons. First, the intervention only gives information about relative performance for Quiz 1. Since the informational nudge makes Quiz 1 salient, the effects on Quiz 1 outcomes and choices should be greater than the effects on other quizzes. Second, the emails were sent the day after Quiz 3 was published. Thus, the effects on Quiz 3 may also be greater.

4.2 Experimental Effects

I find that a very-low cost intervention can nudge students into exerting more effort and that this effort translates, in some cases, into higher achievement. For example, students assigned to the treatment groups are more likely to go back and retake Quiz 1 and 2 after the intervention. For Quiz 1, the negatively framed nudge was more effective in leading to Quiz 1 retakes, and the results were driven by students who did not have a perfect Quiz 1 score before the intervention. For Quiz 2, the positively framed nudge proved to be more effective in generating retakes, specifically for students with a perfect Quiz 1 score. Yet, the increase in effort on both quizzes did not translate into a significant increase in scores. Quiz 3 is, as expected, the one that shows the most important changes in behavior and outcomes. Not only did the treatment increased student effort, but this additional effort significant raised student achievement. Additionally, I show that the intervention is more effective for students "on track," that is, students who took Quiz 3 before Quiz 4 was published, and that the negative treatment, in general, is more effective than the positive treatment, although the positive treatment significantly reduced the dropout rate between while T2 did not. Finally, Quizzes 4 to 6 reveal that the effect of the negatively framed nudge are persistent.

[6] GoBack is only relevant for quizzes 1 and 2 because these were published before the intervention. For the rest of the quizzes, I use Attempts as a measure of effort.

[7] Although it would be possible to compare the effects of the nudge on variables such as number of clicks on the play button or number of threads read, there is not obvious interpretation for these variables. Clicking fewer time the play button, or reading fewer threads, that not imply less time doing those activities. Alas, Coursera does not collect data on time use at the moment. Martinez and Diver [2014] discuss this in more detail.

4.2.1 Effects on Quiz 1

In this section, I describe the effects of the intervention on Quiz 1, which students can retake. Treatment 2 (the negative informational treatment) has a statistically significant effect on GoBack, the proportion of students who went back to take Quiz 1. 8.68% of these students went back to Quiz 1 after the email was sent while only 7.11% of students in the control group did. None of the other measure of effort were affected.

Table 3: Effect of Treatment Participation on Quiz 1, Full Sample

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	8.652	8.6018	(-0.0302, 0.1305)	0.3043
Go Back	0.0745	0.0711	(-0.0082, 0.0152)	0.6252
Drop-out	0.448	0.4626	(-0.0371, 0.0079)	0.2857
Drop-out grade	8.0388	7.9731	(-0.0705, 0.2019)	0.4274
n	2589	2730		
	Treatment 2	**Control**	**90% Confidence Interval**	**p-value**
Max Grade	8.643	8.6018	(-0.0392, 0.1216)	0.3996
Go Back	0.0868	0.0711	(0.0035, 0.0279)	0.0337
Drop-out	0.4541	0.4626	(-0.031, 0.0139)	0.533
Drop-out grade	7.9569	7.9731	(-0.1529, 0.1206)	0.8456
n	2605	2730		

Note: "Max Grade:" maximum grade the student received; "Raking:" percentile ranking for the student; "Go Back:" decision of retaking the quiz after September 16; "Drop-out:" choice to not take the next quiz; "Drop-out grade:" maximum grade for drop-out.

The distribution of grades before the intervention, Figure 5, shows bunching at 10. These students already have a perfect score and therefore no room for improvement. Next, I exclude these students from the sample. Table 4 summarizes the results for this population that had not gotten a perfect score on Quiz 1. Compared to an estimated effect of 1.57 percentage points on retaking Quiz 1 for the entire sample, T2 had a statistically significant estimated effect of 4.00 percentage points, or 40.12%, which is quite large. Meanwhile, Treatment 1, the positive treatment, had an effect of 1.65 percentage points, which fell a little short of statistical significance at the 90% level. These results indicate that informational nudges affect effort, and that negative information has a more powerful effect. T2 has an estimated effect that is of 2.29 percentage points greater on retaking Quiz 1 than does T1.

By choosing to go back and retake the Quiz, students are exerting more effort, and the

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

MaxGrade on Quiz 1 rose a little, but this increase fell just short of statistical significance for Treatment 2. MaxGrade for the sample that did not already have a grade of 10 rose by 0.10 (or a tenth of a point) for Treatment 2 (with a p-value of 0.1044) and by 0.05 for Treatment 1. This may support the hypothesis that students do not fully understand the education production function, and motivates future research in which the nudge guides students about how to exert effort (i.e. spend more time in forums, or watching videos).

4.2.2 Effects on Quiz 2

Of note, Quiz 2 was released before the treatment, though scores on Quiz 2 were not mentioned in the email. Following the same logic as before, a student could decide to go back to Quiz 2 in order to improve his ranking. This might be less costly than going back to Quiz 1 because less time had passed since the material had been covered. However, students have less information about their ranking for this Quiz.

Tables 5 summarizes the effects of treatment participation on Quiz 2. In order to deal with attrition from the course after the delivery of the treatment, students who took Quiz 1 but not Quiz 2 were included with a grade of zero.

Treatment 1 had a statistically significant estimated effect of 2.92 percentage points on retaking Quiz 2. Table 6 shows, interestingly, that this positive treatment was effective on students who had a perfect score on Quiz 1. Students with a perfect score on Quiz 1 and who were assigned to T1 are 4.06 percentage points more likely to retake Quiz 2. This perhaps suggests that the nudge revealed that perfect score is not something extraordinary but the norm.

4.2.3 Effects on Quiz 3

Performance on Quiz 3 provides a strong indication of whether the intervention affected the forward looking behavior of participants. If a student wants to improve his ranking, it might be ideal to exert more effort on Quiz 3. This quiz was published just a day before they received the email, and thus the material was fresh in their mind. Table 7 shows the effects of treatment participation on Quiz 3.

When looking at the population of all the students who took Quiz 1, the only statistically significant effect is on the number of times a student took Quiz 3 when assigned to the negative treatment. If we restrict the sample to those students who took Quiz 3 at all, we observe additional statistically significant effects on choices and outcomes for both

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

treatments. Students in the positive treatment T1 are ranked 2.02 percentage points better on Quiz 3 than students in the control group. The ranking for students in the negative T2 treatment rose on average by 3.39 percentage points. Their grade rose on average by 0.1951 points.[8] Finally, even students who dropped out after Quiz 3 performed better on this quiz if assigned to T2, by 0.4286 points. The improvements in outcomes for students in T1 and T2 is at least partially explained by an increase in effort. Students in T1 attempted the quiz, on average, 0.1193 more times than students in the control group, and students in T2 attempted the quiz, on average, 0.1541 times more than students in the control group.

I next focus on a narrower group, those students who were "on track" and took Quiz 3 before Quiz 4 was published.As seen in Table 8 on page 26, on-track students on track assigned to T1 attempted Quiz 3, on average, 0.2231 times more than students on the control group. Those assigned to T2 attempted the quiz 0.204 times more, and were ranked 2.6 percentage points higher than students in the control group. Further restricting the sample to only those who did not have a perfect score in Quiz 1 shows that students assigned to T1 took this quiz, on average, 0.2686 times more than students in the control group. Additionally, these students are more likely to take Quiz 4, with a dropout rate 8.26 percentage points lower. Students assigned to T2 attempted the quiz, on average, 0.5123 times more than students in the control group, and were ranked 8.43 percentage points higher. Once again, the fact that students assigned to T2 scored, on average, 0.43 points higher than students assigned to T1 indicates that framing matters.[9]

4.2.4 Effects on Quizzes 4 to 6

Quiz 4 was published a week after the intervention. Changes in choices and outcomes for this quiz provide information about the persistence of the nudge, depending at least partly on the extent to which learning earlier material helps students grasp later material and also, perhaps, on longer lasting effects of the informational nudges, although these may diminish over time. On one hand, if the material for each quiz is self-contained the additional knowledge from increased effort on the previous quiz should not affect choices nor behaviors. On the other hand, the nudge could have a persistent effect in motivating students to work harder. Table 9 on page 27 summarizes the effects for this quiz. Students on track who scored 9 or less on Quiz 1 attempted, on average, 0.3968 times more than students in the

[8] The standard deviation for this quiz is 1.61, therefore the benefit of treatment participation is 0.12 standard deviations.

[9] The standard deviation is 1.76, therefore being assigned to T2 instead of T1 implies and improvement of 0.25 standard deviations.

control group. Their ranking was 7.15 percentage points higher than students in the control group.

Table 10 on page 28 summarizes the results for Quiz 5. Students on track who did not have a perfect score on Quiz 1 before the intervention were ranked 7.26 percentage points higher than students in the control group on Quiz 5.

Three weeks after the nudge, the effects of the negatively framed nudge persist. Table 11 shows that students assigned to T2 worked harder and performed better on Quiz 6. These students attempted the quiz, on average, 0.2352 more times, and are ranked 5.77 percentage points higher than students than students in the control group. Moreover, students assigned to T2 scored 0.3747 points more than students assigned to T1.

5 Conclusions

This paper examines whether we can nudge students into exerting more effort and learning more. I use a MOOC as a laboratory by engaging in random assignment of positive and negative informational nudges sent to thousands of students. The nudges raise achievement on new quizzes that are given after the intervention. Some students also retake old quizzes, but without improvements in scores, suggesting that students are better at making adjustments to learn new material rather than relearning older material. For instance, I find that students who were on track but who did not obtain a perfect score on Quiz 1 and who were assigned to receive the negatively-framed email were ranked, on average, 8.43 percentage points higher than students in the control group on Quiz 3. Additionally, students assigned to the positively-framed email were more likely to take Quiz 4 than students in the control group, by 8.26 percentage points. This represents a big impact per dollar spent. This also shows that students want to improve their educational achievement. While these experiments are conducted using a MOOC, low-cost informational interventions could also improve achievement in traditional classrooms.

The fact that an increase in effort does not always translate into higher achievement, as I find here especially for increases in effort on previously published quizzes, supports the hypothesis that students do not fully understand the education production function. In future research, I will use more directive interventions to further test this hypothesis. For example, the emails will recommend specific behavior (e.g., taking notes, reading the forums, etc) or will point out the rewards for acting early and avoiding procrastination.

As important, I show that framing plays an important role when nudging students. The

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

negative treatment tends to change outcomes for those who were doing relatively poorly, and the positive treatment tends to work for those who were doing relatively well.

Future research can further exploit the “Big Data” nature of MOOCs to study more personalized nudges. To get there, further collaboration with a MOOCs provider is needed. For example, it would be desirable to combine data from courses delivered by different universities. Finally, it will be useful for Coursera and other MOOC providers to start collecting data on online time use in order to allow researchers to address additional important questions.

References

William G. Bowen. Academia online: Mussings. *Stafford Little Lecture, Princeton University*, pages 1–28, 2013.

William G. Bowen, Matthew M. Chingos, Kelly A. Lack, and Thomas I Nygren. *Interactive learning online at public universities: Evidence from randomized trials.* Ithaka S+ R, May 2012.

Leo Breiman. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3):199–231, 2001.

John T Cacioppo and Richard E Petty. The need for cognition. *Journal of personality and social psychology*, 42(1):116, 1982.

William Carbonaro. Tracking, students' effort, and academic achievement. *Sociology of Education*, 78(1):27–49, 2005.

Benjamin L Castleman. Prompts, personalization, and pay-offs: Strategies to improve the design and delivery of college and financial aid information. 254, 2013.

Benjamin L Castleman and Lindsay C Page. Personalized text messages and peer mentor outreach increase college going among low-income high school graduates?

Gayle Christensen, Andrew Steinmetz, Brandon Alcorn, Amy Bennett, Deirdre Woods, and Ezekiel J. Emanuel. The mooc phenomenon: Who takes massive open online courses and why? Submitted for publication, November 2013. URL http://ssrn.com/abstract=2350964.

Arthur R Cohen, Ezra Stotland, and Donald M Wolfe. An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology*, 51(2):291, 1955.

Céline Darnon, Benoît Dompnier, Ophélie Gilliéron, and Fabrizio Butera. The interplay of mastery and performance goals in social comparison: A multiple-goal perspective. *Journal of Educational Psychology*, 102(1):212, 2010.

Liran Einav and Jonathan D Levin. The Data Revolution and Economic Analysis. Technical report, National Bureau of Economic Research, 2013.

Andrew J Elliot. A conceptual history of the achievement goal construct. *Handbook of competence and motivation*, 16:52–72, 2005.

NS Fagley and Paul M Miller. The effect of framing on choice interactions with risk-taking propensity, cognitive style, and sex. *Personality and social psychology bulletin*, 16(3): 496–510, 1990.

Roland G Fryer. Financial incentives and student achievement: Evidence from randomized trials. *The Quarterly Journal of Economics*, 126(4):1755–1798, 2011.

Roland G Fryer. Information and student achievement: Evidence from a cellular phone experiment. Technical report, National Bureau of Economic Research, 2013.

Kirk Hallahan. Seven models of framing: Implications for public relations. *Journal of Public Relations Research*, 11(3):205–242, 1999.

Caroline Hoxby and Sarah Turner. Expanding college opportunities for high-achieving, low income students. Technical report, 2013.

Peter John, Sarah Cotterill, A Moseley, L Richardson, Graham Smith, Gerry Stoker, and Corinne Wales. *Nudge, nudge, think, think: experimenting with ways to change civic behaviour*. Bloomsbury Academic, 2011.

Monica Kirkpatrick Johnson, Robert Crosnoe, and Glen H Elder Jr. Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education*, pages 318–340, 2001.

René F Kizilcec, Chris Piech, and Emily Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 170–179. ACM, 2013.

Steven D Levitt, John A List, Susanne Neckermann, and Sally Sadoff. The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. Technical report, National Bureau of Economic Research, 2012.

Tharindu Rekha Liyanagunawardena, Andrew Alexandar Adams, and Shirley Ann Williams. Moocs: A systematic study of the published literature 2008-2012. *International Review of Research in Open & Distance Learning*, 14(3), 2013.

Helen M Marks. Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American educational research journal*, 37(1):153–184, 2000.

Ignacio Martinez and Paul Diver. MOOCs as a massive laboratory: Opportunities and challenges. 2014.

Cynthia McCarty, Gene Padgham, and Doris Bennett. Determinants of student achievement in principles of economics. *Journal for Economics Educators*, 6(2):1–9, 2006.

Matthew Rabin. Psychology and economics. *Journal of economic literature*, 36(1):11–46, 1998.

Bonnie Sibbald and Martin Roland. Why are randomised controlled trials important. *Bmj*, 316:201, 1998.

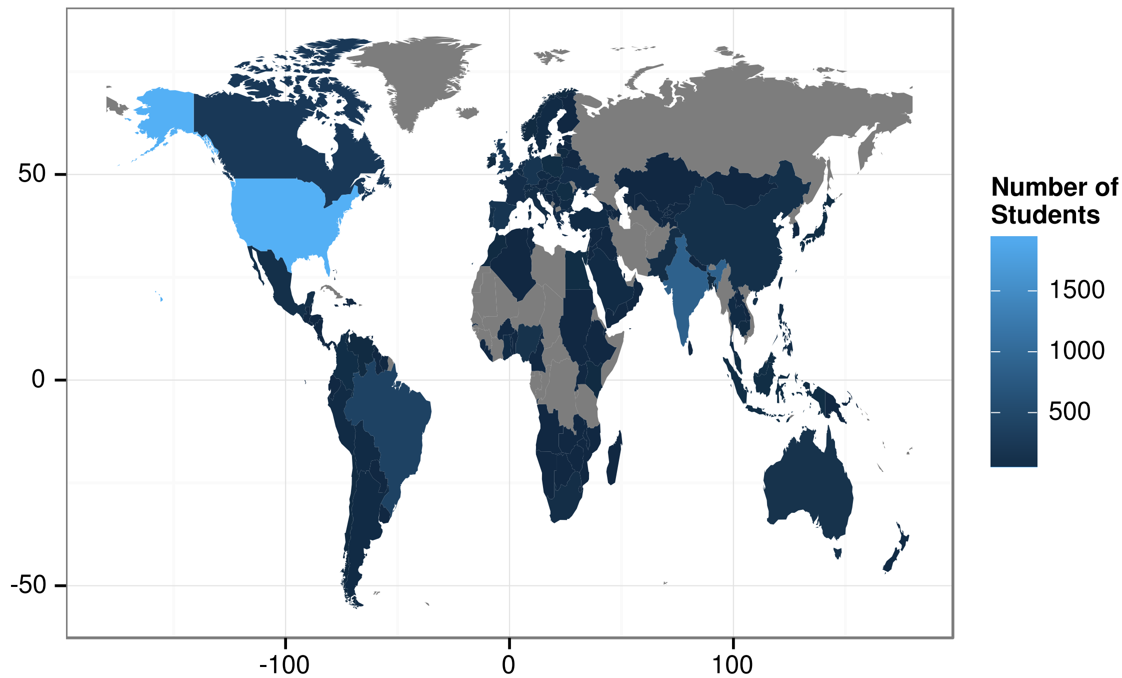
Cass R Sunstein and Richard Thaler. *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press New Haven, 2008.

Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. In *Behavioral decision making*, pages 25–41. Springer, 1985.

Bernard Weiner. Ability versus effort revisited: The moral determinants of achievement evaluation and achievement as a moral system. *Educational Psychologist*, 29(3):163–172, 1994.

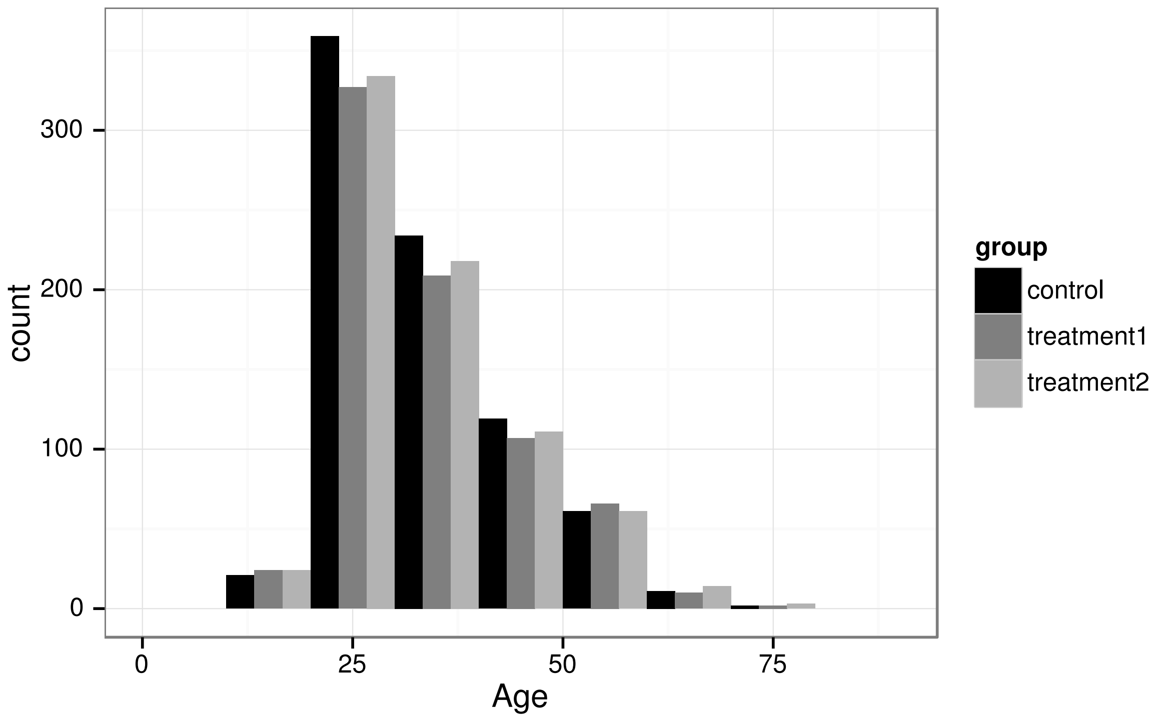
Yong Zhang and Richard Buda. Moderating effects of need for cognition on responses to positively versus negatively framed advertising messages. *Journal of Advertising*, 28(2): 1–15, 1999.

Figure 1: Students' geolocation



EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Figure 2: Please indicate your age (survey question)



EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Figure 3: Please indicate your highest level of education (survey question)

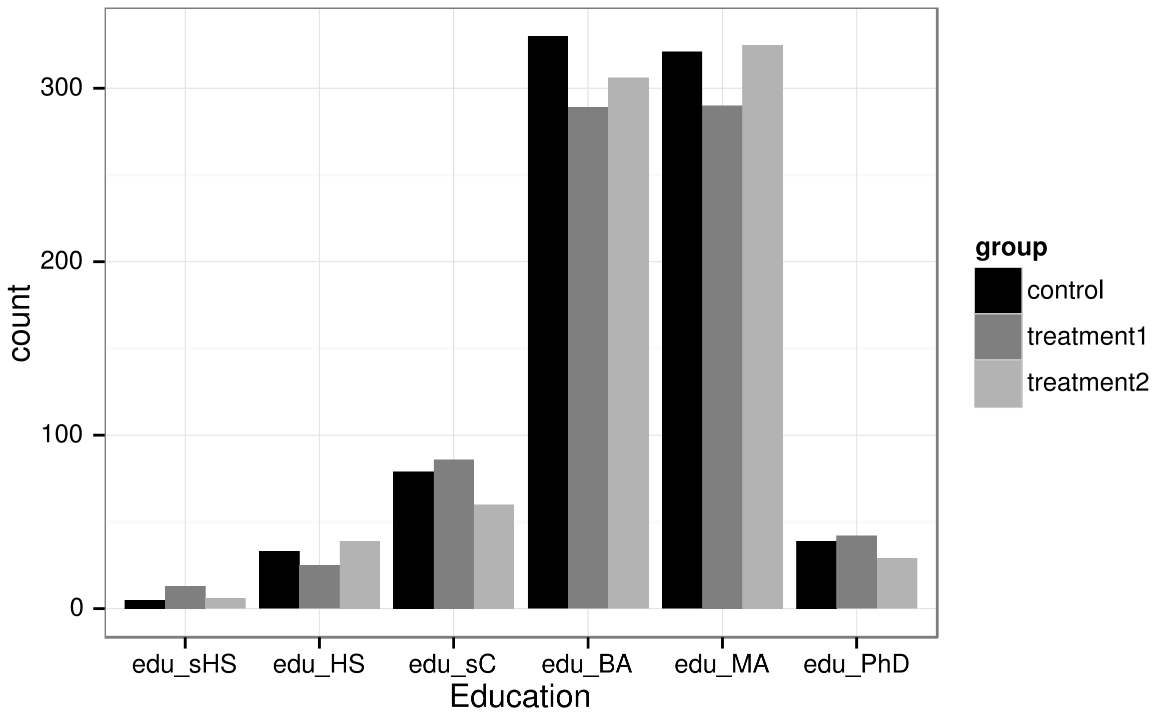
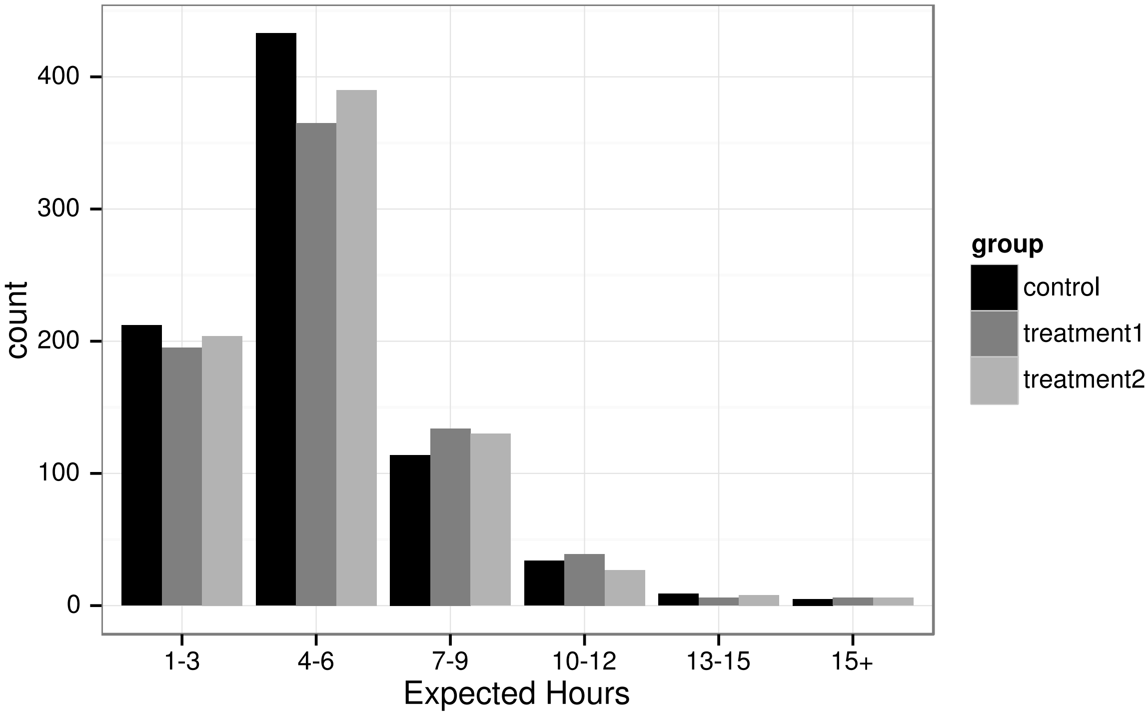


Figure 4: How many hours per week will you be spending on this course? (survey question)



EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Figure 5: Grade distribution before the intervention of Quiz 1 by group

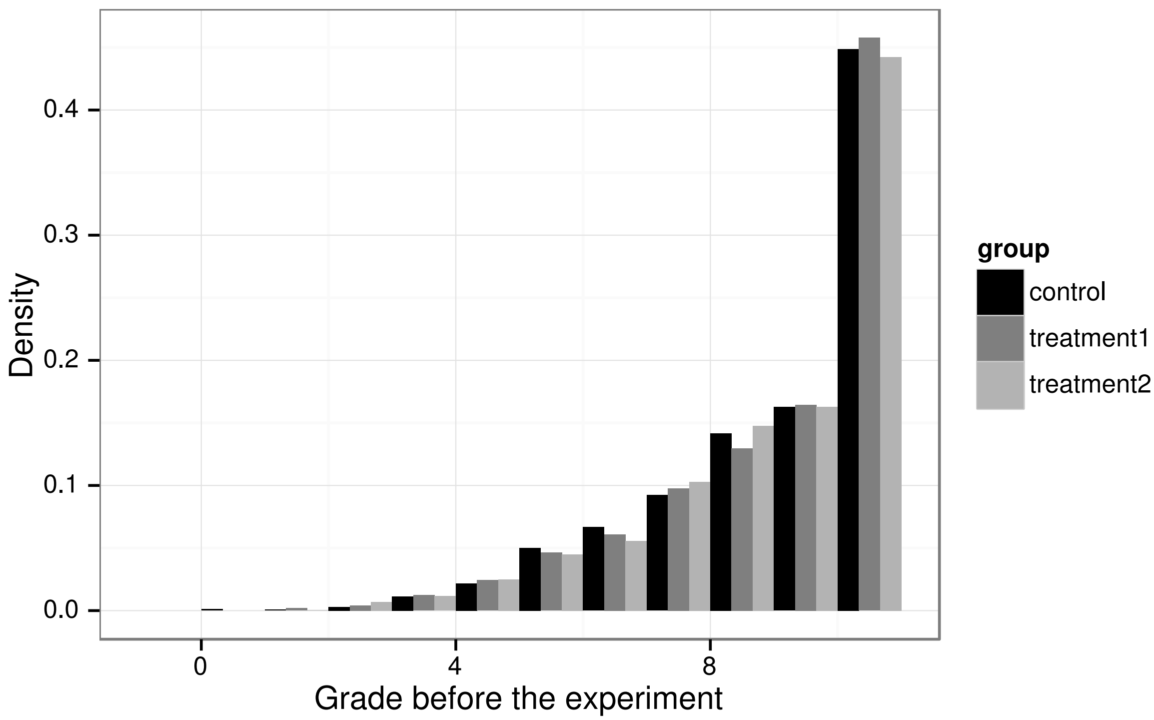


Table 4: Effect of Treatment Participation on Quiz 1, Restricted to Score<10 on Quiz 1

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	7.5142	7.4638	(-0.0542, 0.1551)	0.4276
Go Back	0.1168	0.0997	(-0.0018, 0.0361)	0.1376
Drop-out	0.5527	0.5608	(-0.0384, 0.0222)	0.6608
Drop-out grade	7.0683	6.9668	(-0.0461, 0.2491)	0.258
n	1404	1505		
	Treatment 2	**Control**	**90% Confidence Interval**	**p-value**
Max Grade	7.5671	7.4638	(-0.0013, 0.208)	0.1044
Go Back	0.1397	0.0997	(0.0204, 0.0597)	8e-04
Drop-out	0.5533	0.5608	(-0.0375, 0.0226)	0.6832
Drop-out grade	6.9938	6.9668	(-0.1201, 0.174)	0.7629
n	1453	1505		
	Treatment 2	**Treatment 1**	**90% Confidence Interval**	**p-value**
Go Back	0.1397	0.1168	(0.0023, 0.0435)	0.0671

Note: "Max Grade:" maximum grade the student received; "Raking:" percentile ranking for the student; "Go Back:" decision of retaking the quiz after September 16; "Drop-out:" choice to not take the next quiz; "Drop-out grade:" maximum grade for drop-out.

Table 5: Effect of Treatment Participation on Quiz 2, Full Sample

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	4.8297	4.6978	(-0.0753, 0.339)	0.2951
Go Back	0.3314	0.3022	(0.0082, 0.0502)	0.0222
Drop-out	0.6006	0.6026	(-0.024, 0.0202)	0.8848
Drop-out grade	1.9723	1.8012	(-0.0349, 0.3772)	0.1719
n	2589	2730		
	Treatment 2	**Control**	**90% Confidence Interval**	**p-value**
Max Grade	4.8307	4.6978	(-0.0748, 0.3406)	0.2925
Go Back	0.314	0.3022	(-0.009, 0.0326)	0.3504
Drop-out	0.5946	0.6026	(-0.03, 0.0142)	0.5544
Drop-out grade	1.856	1.8012	(-0.1501, 0.2597)	0.6598
n	2605	2730		

Note: "Max Grade:" maximum grade the student received; "Raking:" percentile ranking for the student; "Go Back:" decision of retaking the quiz after September 16; "Drop-out:" choice to not take the next quiz; "Drop-out grade:" maximum grade for drop-out.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Table 6: Effect of Treatment Participation on Quiz 2

	Restricted to Score<10 on Quiz 1			
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	3.4786	3.4452	(-0.2195, 0.2864)	0.8278
Go Back	0.3006	0.2817	(-0.0089, 0.0466)	0.264
Drop-out	0.6944	0.6884	(-0.0221, 0.0343)	0.7232
Drop-out grade	1.3856	1.2712	(-0.0976, 0.3264)	0.3745
n	1404	1505		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	3.565	3.4452	(-0.1341, 0.3738)	0.4374
Go Back	0.3001	0.2817	(-0.0092, 0.0458)	0.2724
Drop-out	0.6882	0.6884	(-0.0282, 0.0279)	0.9934
Drop-out grade	1.342	1.2712	(-0.1411, 0.2826)	0.5826
n	1453	1505		
	Restricted to Score=10 on Quiz 1			
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	6.4304	6.2367	(-0.1156, 0.5029)	0.3029
Go Back	0.3679	0.3273	(0.0087, 0.0725)	0.0365
Drop-out	0.4895	0.4971	(-0.0412, 0.0258)	0.7059
Drop-out grade	2.9586	2.7028	(-0.152, 0.6637)	0.302
n	1185	1225		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	6.4271	6.2367	(-0.1217, 0.5024)	0.3156
Go Back	0.3316	0.3273	(-0.0275, 0.036)	0.8257
Drop-out	0.4766	0.4971	(-0.0543, 0.0132)	0.3159
Drop-out grade	2.7923	2.7028	(-0.3217, 0.5009)	0.7201
n	1152	1225		

Note: "Max Grade:" maximum grade the student received; "Raking:" percentile ranking for the student; "Go Back:" decision of retaking the quiz after September 16; "Drop-out:" choice to not take the next quiz; "Drop-out grade:" maximum grade for drop-out.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Table 7: Effect of Treatment Participation on Quiz 3

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	3.6037	3.5498	(-0.1501, 0.2578)	0.6638
Ranking	0.4997	0.4955	(-0.0072, 0.0156)	0.5437
Attempts	1.1468	1.0938	(-0.0233, 0.1293)	0.2533
Drop-out	0.6667	0.6744	(-0.0289, 0.0135)	0.5509
Drop-out grade	0.8192	0.8647	(-0.1868, 0.0957)	0.596
n	2589	2730		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	3.6998	3.5498	(-0.0546, 0.3546)	0.2279
Ranking	0.5051	0.4955	(-0.0019, 0.0209)	0.169
Attempts	1.1781	1.0938	(0.0087, 0.16)	0.0666
Drop-out	0.3305	0.3256	(-0.0163, 0.026)	0.7046
Drop-out grade	0.9472	0.8647	(-0.0643, 0.2293)	0.3551
n	2605	2730		
		Attempts > 0		
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	9.0232	8.9318	(-0.0271, 0.2099)	0.2044
Ranking	0.5024	0.4822	(0.002, 0.0385)	0.0675
Attempts	2.8714	2.7521	(0.0082, 0.2304)	0.0775
Drop-out	0.8317	0.8166	(-0.0121, 0.0424)	0.3604
Drop-out grade	8.1264	8	(-0.2446, 0.4975)	0.5745
n	1034	1085		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	9.1269	8.9318	(0.0819, 0.3083)	0.0046
Ranking	0.5161	0.4822	(0.0161, 0.0519)	0.0018
Attempts	2.9062	2.7521	(0.0481, 0.2602)	0.0168
Drop-out	0.8144	0.8166	(-0.0298, 0.0254)	0.8959
Drop-out grade	8.4286	8	(0.0813, 0.7758)	0.0426
n	1056	1085		

Note: “Max Grade:” maximum grade the student received; “Raking:” percentile ranking for the student; “Attempts:” number of times the student attempted the quiz; “Drop-out:” choice to not take the next quiz; “Drop-out grade:” maximum grade for drop-out.

Table 8: Effect of Treatment Participation on Quiz 3 for students on track

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	9.1864	9.1265	(-0.096, 0.2159)	0.5272
Ranking	0.5058	0.4843	(-0.0037, 0.0468)	0.1612
Attempts	2.9138	2.6907	(0.0755, 0.3708)	0.013
Drop-out	0.8317	0.8074	(-0.0155, 0.0641)	0.3156
Drop-out grade	8.5595	8.2525	(-0.193, 0.807)	0.3114
n	499	514		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	9.2554	9.1265	(-0.0195, 0.2773)	0.153
Ranking	0.5103	0.4843	(0.0012, 0.0508)	0.0847
Attempts	2.8947	2.6907	(0.0622, 0.3459)	0.0181
Drop-out	0.8031	0.8074	(-0.045, 0.0365)	0.8629
Drop-out grade	8.6337	8.2525	(-0.0861, 0.8484)	0.1792
n	513	514		
9 or less in their best attempt to Quiz 1 before the intervention				
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	8.1638	8.0968	(-0.2492, 0.3833)	0.7267
Ranking	0.4805	0.4678	(-0.0355, 0.061)	0.663
Attempts	2.3277	2.0591	(0.0364, 0.5007)	0.0572
Drop-out	0.1808	0.2634	(-0.1544, -0.0109)	0.0581
Drop-out grade	7.375	7.3469	(-0.7326, 0.7887)	0.9511
n	177	186		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	8.5979	8.0968	(0.2076, 0.7947)	0.0051
Ranking	0.5521	0.4678	(0.0374, 0.1313)	0.0032
Attempts	2.5714	2.0591	(0.2772, 0.7474)	4e-04
Drop-out	0.2116	0.2634	(-0.1244, 0.0208)	0.2399
Drop-out grade	7.75	7.3469	(-0.3305, 1.1367)	0.3634
n	189	186		
	Treatment 2	Treatment 1	90% Confidence Interval	p-value
Max Grade	8.5979	8.1638	(0.1377, 0.7304)	0.0162

Note: “Max Grade:” maximum grade the student received; “Raking:” percentile ranking for the student; “Attempts:” number of times the student attempted the quiz; “Drop-out:” choice to not take the next quiz; “Drop-out grade:” maximum grade for drop-out.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Table 9: Effect of Treatment Participation on Quiz 4

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	3.0421	2.9505	(-0.1058, 0.2889)	0.4454
Ranking	0.5023	0.4969	(-0.0054, 0.0162)	0.4071
Attempts	0.9189	0.8656	(-0.0145, 0.1211)	0.1956
Drop-out	0.7107	0.7143	(-0.024, 0.0168)	0.7727
Drop-out grade	0.5364	0.4595	(-0.0322, 0.186)	0.246
n	2589	2730		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	3.0288	2.9505	(-0.1188, 0.2753)	0.5136
Ranking	0.5011	0.4969	(-0.0065, 0.015)	0.5151
Attempts	0.9017	0.8656	(-0.0304, 0.1028)	0.3718
Drop-out	0.7098	0.7143	(-0.0249, 0.0159)	0.717
Drop-out grade	0.5187	0.4595	(-0.0488, 0.1671)	0.3672
n	2605	2730		
Students on track				
Scored 9 or less in their best attempt to Quiz 1 before the intervention				
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	8.3203	8.4631	(-0.4684, 0.1828)	0.4698
Ranking	0.4769	0.4767	(-0.0528, 0.0532)	0.9942
Attempts	2.2484	2.0403	(-0.0313, 0.4475)	0.1525
Drop-out	0.1699	0.1678	(-0.0692, 0.0735)	0.9604
Drop-out grade	8	7.8	(-0.7686, 1.1686)	0.7306
n	153	149		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	8.8079	8.4631	(0.0515, 0.6382)	0.0534
Ranking	0.5482	0.4767	(0.0205, 0.1226)	0.0214
Attempts	2.4371	2.0403	(0.1641, 0.6295)	0.0052
Drop-out	0.1854	0.1678	(-0.0552, 0.0905)	0.6898
Drop-out grade	8.1786	7.8	(-0.4635, 1.2206)	0.4547
n	151	149		
	Treatment 2	Treatment 1	90% Confidence Interval	p-value
Max Grade	8.8079	8.3203	(0.1638, 0.8116)	0.0135

Note: "Max Grade:" maximum grade the student received; "Raking:" percentile ranking for the student; "Attempts:" number of times the student attempted the quiz; "Drop-out:" choice to not take the next quiz; "Drop-out grade:" maximum grade for drop-out.

Table 10: Effect of Treatment Participation on Quiz 5

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	2.6902	2.6278	(-0.1297, 0.2545)	0.5931
Ranking	0.5008	0.4979	(-0.0075, 0.0132)	0.6489
Attempts	0.735	0.7256	(-0.0525, 0.0712)	0.8027
Drop-out	0.7435	0.7473	(-0.0234, 0.0159)	0.7555
Drop-out grade	0.3995	0.3936	(-0.0909, 0.1026)	0.9207
n	2589	2730		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	2.7113	2.6278	(-0.1087, 0.2756)	0.4748
Ranking	0.5016	0.4979	(-0.0066, 0.014)	0.5565
Attempts	0.7363	0.7256	(-0.051, 0.0723)	0.7767
Drop-out	0.7455	0.7473	(-0.0214, 0.0178)	0.8824
Drop-out grade	0.4439	0.3936	(-0.0488, 0.1493)	0.4041
n	2605	2730		
Students on track				
Scored 9 or less in their best attempt to Quiz 1 before the intervention				
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	8.5305	8.6496	(-0.4304, 0.1921)	0.5281
Ranking	0.4612	0.4842	(-0.075, 0.0288)	0.4633
Attempts	2.0671	2.2774	(-0.5324, 0.1118)	0.282
Drop-out	0.1524	0.1752	(-0.0938, 0.0483)	0.5979
Drop-out grade	8	7.6667	(-0.5579, 1.2246)	0.5333
n	164	137		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	8.9739	8.6496	(0.019, 0.6295)	0.0807
Ranking	0.5568	0.4842	(0.0203, 0.1248)	0.0228
Attempts	2.2288	2.2774	(-0.3736, 0.2764)	0.8051
Drop-out	0.2092	0.1752	(-0.0426, 0.1105)	0.4645
Drop-out grade	8.1562	7.6667	(-0.366, 1.3452)	0.3425
n	153	137		
	Treatment 2	Treatment 1	90% Confidence Interval	p-value
Max Grade	8.9739	8.6496	(0.019, 0.6295)	0.0807

Note: "Max Grade:" maximum grade the student received; "Raking:" percentile ranking for the student; "Attempts:" number of times the student attempted the quiz; "Drop-out:" choice to not take the next quiz; "Drop-out grade:" maximum grade for drop-out.

EdPolicyWorks Working Paper Series No. 19. December 2013.
Available at http://curry.virginia.edu/edpolicyworks/wp.
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

Table 11: Effect of Treatment Participation on Quiz 6

	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	2.2897	2.2451	(-0.1359, 0.2251)	0.6842
Ranking	0.5011	0.4988	(-0.0076, 0.0122)	0.7039
Attempts	0.7992	0.7963	(-0.0679, 0.0735)	0.9478
Certificate	0.1379	0.1447	(-0.0225, 0.0089)	0.4768
No Certificate Grade	1.1447	1.027	(-0.0237, 0.2592)	0.1709
n	2589	2730		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	2.2856	2.2451	(-0.1399, 0.221)	0.7117
Ranking	0.5004	0.4988	(-0.0083, 0.0115)	0.7902
Attempts	0.795	0.7963	(-0.0711, 0.0685)	0.975
Certificate	0.1397	0.1447	(-0.0207, 0.0108)	0.6043
No Certificate Grade	1.1281	1.027	(-0.0401, 0.2423)	0.2388
n	2605	2730		
Scored 9 or less in their best attempt to Quiz 1 before the intervention				
	Treatment 1	Control	90% Confidence Interval	p-value
Max Grade	7.9198	8.0397	(-0.4385, 0.1988)	0.5356
Ranking	0.4704	0.4864	(-0.0555, 0.0235)	0.5049
Attempts	2.6412	2.704	(-0.2776, 0.1521)	0.6306
Certificate	0.4046	0.4585	(-0.1242, 0.0164)	0.2073
No Certificate Grade	7.3397	7.26	(-0.3845, 0.544)	0.7771
n	262	277		
	Treatment 2	Control	90% Confidence Interval	p-value
Max Grade	8.4144	8.0397	(0.0626, 0.6869)	0.0484
Ranking	0.5441	0.4864	(0.0187, 0.0967)	0.0151
Attempts	2.9392	2.704	(0.0273, 0.443)	0.0628
Certificate	0.4829	0.4585	(-0.0465, 0.0953)	0.5709
No Certificate Grade	7.7279	7.26	(-0.0269, 0.9628)	0.1197
n	263	277		
	Treatment 2	Treatment 1	90% Confidence Interval	p-value
Max Grade	8.4144	8.0397	(0.0626, 0.6869)	0.0484

Note: “Max Grade:” maximum grade the student received; “Raking:” percentile ranking for the student; “Attempts:” number of times the student attempted the quiz; “Certificate:” proportion of students who obtained a statement of accomplishment; “No certificate grade:” maximum grade for students who did not obtain the certificate.