

11-28-2017

Improved grade outcomes with an e-mailed “grade nudge”

Ben O. Smith

Dustin R. White

Patricia C. Kuzyk

James E. Tierney

Follow this and additional works at: <https://digitalcommons.unomaha.edu/econrealestatefacpub>

 Part of the [Economics Commons](#)

IMPROVED GRADE OUTCOMES WITH AN E-MAILED “GRADE NUDGE”

Ben O. Smith^a, Dustin R. White^b, Patricia C. Kuzyk^c, and James E. Tierney^d

^{a,b}Department of Economics, University of Nebraska at Omaha, Omaha, NE, USA; ^cSchool of Economic Sciences, Washington State University, Pullman, WA, USA; ^dDepartment of Economics, The Pennsylvania State University, University Park, PA, USA.

Abstract: Information provided at the moment a person makes a decision can influence behavior in predictable ways. The United Kingdom’s Behavioural Insights Team have used this idea to help improve the insulation of lofts, collect taxes, and even reduce litter. The authors of this article developed software that appends a personalized message to each assignment in the class regarding the student’s current grade. This “grade nudge” explains precisely how the assignment will impact the student’s final grade given their current standing in the class. Through a randomized trial, the authors show that the nudge improves student homework performance by about four percentage points.

Keywords: bounded rationality, grading outcomes, nudge, selective transparency

JEL codes: A20, A22, A23

Contact: Ben O. Smith; bosmith@unomaha.edu; Assistant Professor, Department of Economics, University of Nebraska at Omaha, 6708 Pine Street, Omaha, NE 68182, USA.

The authors thank Jadrian Wooten, Jamie Wagner and Georg Schaur for their comments on earlier versions of this article. Additionally, they thank the two anonymous referees for their insightful comments.

Despite the proliferation of learning management systems (LMS), it is common for college instructors to hear questions about grade status. Usually, these well-meaning students have not concerned themselves with their grade up to this point in the semester, or they may not realize that their grade is available online. This concerns the authors for two reasons. First, in large classes this results in a negative externality endured by students who have content questions. Second, by the time this question is asked it is often too late in the semester for the student to substantially improve their grade.

Students often think they can “make up” for poor performance early in the semester through additional work near the end of the semester. This is problematic because it is often mathematically impossible to overcome the grade impact of missed assignments or previous poor scores. Additionally, in many classes the content builds from one topic to another. Poor performance early on could lead to worse performance later even with additional effort by the student.

We address the above problems by providing grade outcomes with each assignment. In principle, our solution should reduce the likelihood that the student is uninformed of their grade and “nudge” the students into better behavior early in the semester. In this article, we will show that our nudge improves student performance and is anecdotally liked by the students.

We have implemented our nudge with a Google App Script written by the authors. Any instructor can use our software for free to send the following message with each assignment in a course:

Hi [Name],

As of now, you have a(n) [Grade] in the class. This assignment is worth [Points] points. If you get more than [X] on this assignment, your class grade will increase to a(n) [Higher Grade]. If you get less than [Y] on this assignment, your grade will drop at least one grade. Not doing the assignment will result in a(n) [Lower Grade].

Here, [Name] is the student's name, [Grade] is the student's current grade, [Points] is the number of points the assignment is worth, [X] is the minimum number of points necessary to raise a student's grade, [Higher Grade] is a grade higher than the student's current grade and [Y] is the maximum points the student could receive and still drop a grade. [Lower Grade] is the updated grade of the student assuming they choose not to do the assignment.¹

The message can be delivered to the student in two different ways. If the homework assignment is a document, the message will be appended to the bottom of the assignment. The instructor can choose to share the document within the Google Drive environment or have a PDF sent by e-mail. Alternatively, if the instructor uses an online homework system, the grade nudge can be e-mailed separately from the assignment to each student. The software and instructions can be found online at <https://bensresearch.com/nudge/>.

SOME USES AND RESEARCH ON NUDGES

The concept of a nudge, popularized by Thaler and Sunstein (2008), often centers around providing information at the exact moment that an agent makes a decision. Past successes include informing tax evaders that most people pay their taxes (Lawton 2013), as well as increasing exercise through marketing techniques (Cohen et al. 2013). Perhaps most famously, the United Kingdom set up the Behavioural Insights Team in 2010 (BIT—often referred to as the “Nudge Unit”). The United States followed suit in 2014 with the Social and Behavioral Sciences Team (SBST).

This selective transparency of information is an application of bounded rationality. Due to the low cost of heuristics, Kahneman (2003) shows that relatively small barriers to information result in decisions that are based on intuitive thinking. The costlier the information is to acquire, the more likely it becomes that a heuristic will be used. This tendency to account for only some of the available information can result in socially suboptimal behavior. For instance, Bettinger et al. (2012) found that the complexity of the financial aid application process could lead some otherwise qualified students to forgo college; Bettman, Payne, and Staelin (1986) found that warning label design could change the consumer's perception of risk.

Using this basic insight from behavioral economics, the BIT and SBST have altered behavior by providing information in less costly ways. As an example, the BIT provides potential job market applicants information on recruiting events via text message (BIT 2015, 9–10). This has resulted in an improvement in recruitment event attendance. A similar program addresses the problem of missed hospital appointments by sending a reminder text message to the patient (Hallsworth et al. 2015). To improve university attendance, the BIT has encouraged more students to go to college by discussing the lifestyle benefits of higher education (BIT 2015, 26–27).

While the SBST has not existed for as long as the BIT, their first year was marked with a number of successes sending information nudges by short message service (SMS) or e-mail. For instance, sending text messages to low-income students has resulted in an increase in college matriculation (SBST 2015, 33). Another project improved student loan borrower behavior by e-mail notification (SBST 2015, 34). Finally, a single well-crafted and well-timed e-mail to military service members almost doubled the enrollment rate of a savings program (SBST 2015, 31).

Previous research has shown that information nudges can be an effective tool to improving education outcomes. Kraft and Dougherty (2013) and Kraft and Rogers (2015) show that frequent

communication from K–12 teachers to parents results in a reduction in undesirable student outcomes. Further, Rogers et al. (2017) found that a single postcard home improved K–12 attendance. At the collegiate level, Castleman and Page (2016) suggest that text messages encouraging students to file their financial aid forms results in higher rates of continuous enrollment. This article builds on existing literature by showing that timely grade information can result in higher homework scores.

RANDOMIZED TRIAL

To test the grade nudge treatment, we ran a randomized trial on a total of two sections of “Economics of Sports in America” (hereinafter “Sports Econ”) and one section of “Fundamentals of Microeconomics” (hereinafter “Microeconomics”). Both courses are offered at a large geographically isolated research 1 institution: Washington State University (WSU). WSU is in a traditional college town where the student body accounts for 70 percent of the total population. The nearest major city is 75 miles away. In addition to programs offered on its main campus, WSU offers extensive online programs.

Other than being online, both Microeconomics and Sports Econ are very similar to traditional on-campus courses, and involve a mixture of exams and electronically graded homework. Further, both classes exhibit a traditional grade distribution and have normal expected grades. Over the past three semesters, the median Microeconomics student has received a B+; the average Sports Econ student can expect a C+. Nonetheless, both classes exhibit the full spectrum of grades: many students will receive As and some students will fail.

Both Microeconomics and Sports Econ are offered as an online undergraduate class every fall semester. Additionally, the two sections of Sports Econ were taught by the same instructor using similar online homework assignments graded on the computer using identical rubrics. Due

to the similarity across terms and consistent grading, this is a particularly attractive class in which to run this experiment. Similarly, Microeconomics used computer-graded homework assignments throughout the term. However, this class was taught by a different instructor than the two Sports Econ sections.

The ideal way to implement a grade nudge is to append the treatment to the bottom of each assignment. With the treatment appended to the homework, the instructor is guaranteed that the student will see the grade outcome information if they look at the assignment at all. Unfortunately, assignments in all three sections were assigned through an LMS, and cannot be appended to by the Google App. Thus, our trial makes use of the e-mail notification method.

An additional requirement of a proper test of the treatment is the ability to form two randomly created groups of students on each assignment: those who received the grade nudge and those who did not. Ideally, the students in the two groups would not communicate with each other and would have a high likelihood of seeing the treatment (sent over e-mail). Because online students rarely live near the main campus, they do not form intimate study groups and by necessity communicate with the instructor and classmates through e-mail. Online sections are therefore perfect for our experiment.

For the fall 2014 (Sports Econ), fall 2015 (Sports Econ) and fall 2016 (Microeconomics) semesters, the instructors used the nudge software in a randomized setting. Homework in each class was deployed within the learning management system (LMS).

In each of the three sections, the instructors used the nudge software to perform a randomized trial once there was at least one score in the gradebook to ensure that each student had an established grade in the class. With each assignment that followed, students would receive an e-mail indicating that the homework would soon be due. With 50 percent probability, the student

also would receive the grade nudge message (copied at the beginning of this article) at the bottom of the e-mail (regardless of whether or not the student had received a nudge in the past). This setup allowed for an experiment that was both ethically and statistically valid.

We could have initially randomly assigned each student to the control or treatment group and used the same group assignments throughout the semester. This was ethically concerning to us and could possibly raise concerns with the Institutional Review Board. Based on the past successes of public policy information nudges, we suspected the treatment would have some positive impact on student behavior. Therefore, if the groups were consistently assigned, we could be potentially responsible for the control group receiving a marginally lower grade. Further, given that the control and treatment groups are randomly formed on every assignment, it is less likely that our results are spurious in the unlikely event that the treatment and control groups have different characteristics.

Using the randomly assigned treatments and final homework scores, we constructed a dataset containing four elements: did the student receive a grade nudge on this assignment (0 or 1), percent homework score ($[0-1]$), student identifier and homework identifier. Therefore, we have five to seven observations per student in each of the three sections.

TABLE 1: Summary Statistics

	Nudge (Treatment)	Assignment Score
Online Sports Econ – Fall 2014 (204 Obs.)		
Mean	0.446	0.789
Standard deviation	0.498	0.289
Online Sports Econ – Fall 2015 (110 Obs.)		
Mean	0.400	0.744
Standard deviation	0.492	0.371
Online Microeconomics – Fall 2016 (462 Obs.)		
Mean	0.502	0.710
Standard deviation	0.501	0.292

Table 1 shows the summary statistics for each of the three sections. Not surprisingly, a little less than half of the assignment/student pairs received a nudge and the assignment scores are relatively close to each other. We would note two anomalies with our data. First, the nudge procedure started earlier in the fall 2015 semester than in the fall 2014 sequences—some of the differences in effect size can probably be attributed to this difference. Second, due to logistical issues with late enrollments (and associated makeup assignments) in the first two weeks of class along with a misapplication of the treatment, the first nudge in our fall 2016 dataset is not the first treatment received by the students—however, it is unclear who was treated during the two weeks prior. Nonetheless, we have seven captured assignments from the class with which to run our analysis.

Utilizing the collected data, we propose the following econometric model for our analysis:

$$s_{ij} = \beta_0 + \beta_N \text{Nudge}_{ij} + \mu_i + \gamma_j + \varepsilon_{ij} \quad (1)$$

where s_{ij} is the score individual i received on homework j , μ_i represents the individual fixed effects and γ_j the assignment fixed effects. Nudge_{ij} takes the value one if student i received a nudge message on homework j , zero otherwise. For our purposes, when models contain multiple sections, we create assignment fixed effects for each section/assignment combination. While some assignments are identical in the Sports Econ sections, the material may have been taught differently. Therefore, we do not think we can rely on a single fixed effect for each assignment. Just like when we analyze a single class (or section), we drop only a single fixed effect of each type. Therefore, the control variables capture the differences between the classes. This is functionally equivalent to having separate baseline terms (β_0) for each class and dropping variables for each class—both methods produce identical results for β_N .

We have applied a fixed effects approach as a random effects approach is inconsistent. One of the primary characteristics of a good student is that they have good metacognition (Coutinho 2008). This results in reduced homework assignment performance variance. As there is less of an opportunity for the best students to be influenced by outside factors, the individual specific effects are negatively correlated with the independent variables. Our identification strategy assumes the treatment is strictly exogenous. A fixed effects model assumes the regressor is not correlated to the disturbance. If the effect of a nudge persists beyond the treatment period, this assumption may not hold. Partially for this reason, we examine the first nudge in a separate specification.

TABLE 2: Estimated Impact of Nudge Treatment – Individual and Assignment Fixed Effects

	Sports				
	Fall 2014	Sports Fall 2015	All Sports	Micro Fall 2016	All
β_N	0.044	0.052	0.047	0.034	0.038
Uncorrected p -value	0.101	0.390	0.083	0.047	0.008
Corrected p -value	0.086	0.222	0.036	0.026	0.002
Cluster-robust p -value	0.130	0.359	0.079	0.077	0.013
Cluster-bootstrap p -value	0.072	0.295	0.045	0.048	0.005
Observations	204	110	314	462	776
Students	34	22	56	66	122
Assignments	6	5	11	7	18

Results of our fixed effects regression model are included in table 2. As there is only one coefficient of interest (β_N), we present a total of five regression results in a single table: one for each section, one containing all sections and one grouping the two Sports Econ sections. Using White’s LM test (White 1980) and a Shapiro-Wilk test (Shapiro and Wilk 1965), we can reject the nulls of homoskedastic and normally distributed errors. This is true of all regressions throughout the manuscript. As a result, we report four p -values: using uncorrected errors, using heteroskedasticity-consistent errors (White 1980), using cluster-robust errors (Williams 2000) and using a cluster-bootstrap procedure that controls for both heteroskedasticity and normality issues.

The advantage of our bootstrap procedure is that it does not assume any underlying distribution. However, in small enough samples it may not be reasonable to assume that the sample is representative of the underlying population, in which case simply using a p -value that relies on asymptotic theory may be preferable.

The cluster-bootstrapping procedure mentioned above randomly selects individual student IDs in the panel with replacement—not rows. Using the randomly selected student IDs, we select the corresponding set of rows. In some cases, these rows may be selected multiple times if the student was randomly selected more than once. This approach is appropriate for panel data where the errors may be correlated to the individuals (Künsch 1989). In all cases, we use 10,000 replications in our cluster-bootstrap procedure.

While the β_N point estimates for the individual Sports Econ classes are insignificant at the 5 percent level, using the cluster-bootstrap p -value, the pooled Sports Econ classes' β_N is significant at 5 percent. Using any set of assumptions, the result is significant at the 10 percent level. Further, the Microeconomics class is significant at 5 percent using the cluster-bootstrap p -value. Again, using any set of assumptions, the result is significant at the 10 percent level. Pooling the data, β_N is significant at the 1.5 percent level using any set of assumptions, and is significant at the 1 percent level under the assumptions most accurately representing our data (evaluated by the cluster-bootstrap procedure).

While table 2 represents the average impact of the nudge throughout a 16-week semester, an instructor adopting the procedure should be careful to send grade nudges early in the semester: there is some evidence that these treatments are more impactful than those sent late in the semester. Suppose a student receives a nudge early in the semester but then was assigned to the control group on the later assignments. The first nudge could have “primed” the student to be more aware of

their grade throughout the semester (violating the assumptions of the fixed effects model). While our data contains only two first nudge assignments (as the first treatment in Microeconomics is unknown), we still wish to explore this idea. Consider the following alteration to our model:

$$s_{ij} = \beta_0 + \beta_{NF}\text{Nudge First}_{ij} + \beta_N\text{Nudge}_{ij} + \mu_i + \gamma_j + \varepsilon_{ij} \quad (2)$$

In this specification, First Nudge_{ij} takes the value one only if the nudge message is on the first treated assignment. Nudge_{ij} takes the value one on any treatment, including the first assignment. Running this specification on data from both Sports Econ classes and pooling all data reveals the results in table 3. This result shows limited statistical evidence that the first nudge might have a larger effect—we can reject the null with 90 percent confidence that the additional effect of the first nudge is zero using the Sports Econ data. When analyzing all data, we can reject the null with 95 percent confidence.

TABLE 3: Estimated Impact of Nudge Treatment – Individual and Assignment Fixed Effects with Separate Nudge First Coefficient

	Sports Classes	All Classes
β_{NF}	0.107	0.106
Uncorrected p -value	0.144	0.088
Corrected p -value	0.062	0.048
Cluster-robust p -value	0.126	0.093
Cluster-bootstrap p -value	0.057	0.043
β_N	0.031	0.033
Uncorrected p -value	0.277	0.029
Corrected p -value	0.199	0.011
Cluster-robust p -value	0.281	0.039
Cluster-bootstrap p -value	0.214	0.021
Observations	314	776
Students	56	122
Assignments	11	18

The larger impact of the first nudge could be due in part to the priming discussed above. However, there is an alternative explanation. The first nudge occurs early in the semester when students are the least concerned with their grade. Because a nudge is providing new information to students who are uninformed regarding the importance of homework assignments, the impact could be larger. Regardless of the reason, our results indicate that the first nudge sent early in the semester might be particularly important.

An information nudge often produces a small behavioral change for a very low cost. Past projects by the Behavioural Insights Team have also produced a small effect size. For instance, using lotteries increased electoral registration rates by 3.3 percent (Service et al. 2014, 26) and using a direct Web link improved tax collection by 4 percent (Service et al. 2014, 13). Nonetheless, those projects still have value as the benefit outweighs the cost.

If our point-estimates of the effect are valid, then implementing grade nudges in a course can improve homework scores by several percentage points. Presumably, this increases the knowledge accumulation of the student for a minimal cost: Once an instructor has become familiar with the application, it takes about three minutes to send subsequent messages to the entire class.

CONCLUSION

In most university classes, the student must visit the learning management system (LMS) to calculate the assignment's impact on the overall grade. This is not *easy*, one of the common characteristics of a good nudge. By adding messaging to the bottom of each assignment, we have lowered the workload for the students, and made students' grade information readily available. Through our randomized trial we have shown this grade nudge improves homework performance. It is notable that these gains are accomplished with a technique that requires no class time and is of minimal cost to the instructor when using the provided software. Our hope is that this article

will encourage instructors to provide grade information to their students in timelier and less costly ways than the traditional LMS.

NOTE

¹ The treatment message is simplified when achieving a higher grade is impossible or when a zero on the assignment will not result in a lower grade. For instance, a student with the highest possible grade in the course will not see the third sentence of the treatment message. Similarly, if a zero on the assignment will not result in a grade reduction, the fourth and fifth sentence will be omitted from the treatment. This selective transparency is designed to induce students to work harder in the course without causing unintended consequences for low-value assignments.

REFERENCES

- The Behavioural Insights Team (BIT). 2015. *The Behavioural Insights Team update report 2013–2015*. London, UK: Behavioural Insights Ltd., in partnership with the UK Cabinet Office. <https://goo.gl/QCmVwW> (accessed June 2017).
- Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. 2012. The role of application assistance and information in college decisions: Results from the H&R Block FAFSA Experiment. *The Quarterly Journal of Economics* 127 (3): 1205–42.
- Bettman, J. R., J. W. Payne, and R. Staelin. 1986. Cognitive considerations in designing effective labels for presenting risk information. *Journal of Public Policy & Marketing* 5 (1): 1–28.
- Castleman, B., and L. Page. 2016. Freshman year financial aid nudges: An experiment to increase FAFSA renewal and college persistence. *Journal of Human Resources* 51 (2): 389–415.
- Cohen, D. A., B. Han, K. P. Derose, S. Williamson, T. Marsh, and T. L. McKenzie. 2013. Physical activity in parks: A randomized controlled trial using community engagement. *American Journal of Preventive Medicine* 45 (5): 590–97.
- Coutinho, S. 2008. Self-efficacy, metacognition, and performance. *North American Journal of Psychology* 10 (1): 165–72.
- Hallsworth, M., D. Berry, M. Sanders, A. Sallis, D. King, I. Vlaev, and A. Darzi. 2015. Stating appointment costs in SMS reminders reduces missed hospital appointments: Findings from two randomised controlled trials. *PloS one* 10 (9): 1–14.
- Kahneman, D. 2003. Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review* 93 (5): 1449–75.

- Kraft, M., and S. Dougherty. 2013. The effect of teacher–family communication on student engagement: Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness* 6 (3): 199–222.
- Kraft, M., and T. Rogers. 2015. The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review* 47 (1): 49–63.
- Künsch, H. R. 1989. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* 17 (3): 1217–41.
- Lawton, G. 2013. Nudge in the right direction. *New Scientist* 218 (2922): 32–36.
- Rogers, T., T. Duncan, T. Wolford, J. Ternovski, S. Subramanyam, and A. Reitano. 2017. *A randomized experiment using absenteeism information to “nudge” attendance*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. <https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=440> (accessed June 2017).
- Service, O., M. Hallsworth, D. Halpern, F. Algate, R. Gallagher, S. Nguyen, S. Ruda et al. 2014. *EAST: Four simple ways to apply behavioural insights*. London, UK: Behavioural Insights Ltd., in partnership with the UK Cabinet Office. <https://goo.gl/PvNmuy> (accessed June 2017).
- Shapiro, S., and M. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52 (3/4): 591–611.
- Social and Behavioral Sciences Team (SBST). 2015. *Social and Behavioral Sciences Team annual report*. Washington, DC: National Science and Technology Council, Executive Office of the President. <https://goo.gl/rvIxbv> (accessed June 2017).
- Thaler, R. H., and C. R. Sunstein. 2008. *Nudge*. New Haven, CT: Yale University Press.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society* 48 (4): 817–38.

Williams, R. L. 2000. A note on robust variance estimation for cluster-correlated data. *Biometrics* 56 (2): 645–46.