Student Coaching: How Far Can Technology Go?

Philip Oreopoulos
philip.oreopoulos@utoronto.ca
University of Toronto
National Bureau of Economics Research
Canadian Institute for Advanced Research

Uros Petronijevic
upetroni@yorku.ca
York University

Abstract: Recent studies show that programs offering structured, one-on-one coaching and tutoring tend to have large effects on the academic outcomes of both high school and college students. These programs are often costly to implement and difficult to scale, however, calling into question whether making them available to large student populations is feasible. In contrast, interventions that rely on technology to maintain low-touch contact with students can be implemented at large scale and minimal cost but with the risk of not being as effective as one-on-one, in-person assistance. In this paper, we test whether the effects of coaching programs can be replicated at scale by using technology to reach a larger population of students. We work with a sample of over four thousand undergraduate students from a large Canadian university, randomly assigning students into one of the following three interventions: (i) a one-time online exercise designed to affirm students' values and goals; (ii) a text messaging campaign that provides students with academic advice, information, and motivation; and (iii) a personal coaching service, in which students are matched with upper-year undergraduate coaches. We find large positive effects from the coaching program, as coached students realize a 0.3 standard deviation increase in average grades and a 0.35 standard deviation increase in GPA. In contrast, we find no effects from either the online exercise or the text messaging campaign on any academic outcome, both in the general student population and across several student subgroups. A comparison of the key features of the text messaging campaign and the coaching service suggests that proactively and regularly initiating conversations with students and working to establish trust are important design features to incorporate in future interventions that use technology to reach large populations of students.

1. Introduction

Policymakers and academics share growing concerns about stagnating college completion rates and negative student experiences. Between 1970 and 1999, for example, while college enrollment rates of twenty-three-year-old students rose substantially, completion rates fell by 25 percent (Turner, 2004). More recent figures suggest that only 56 percent of students who pursue a bachelors' degree complete it within six years (Symonds et al., 2011) and recent research questions whether students who attain degrees acquire meaningful new skills along the way (Arum and Roska, 2011). Students are increasingly entering college underprepared, with those who procrastinate, do not study enough, and have superficial attitudes about success performing particularly poorly (Beattie, Laliberté, and Oreopoulos 2016).

Much existing research focuses on lacking financial resources among both students and the institutions they attend as explanations for low completion rates and negative experiences. The impediments of student resource constraints are highlighted by youth from high-income families being more likely to attend college, even after accounting for cognitive achievement, family composition, race, and residence (Belley and Lochner, 2007) and by student average work hours increasing during recent decades when college prices continued to rise but sources of financial aid did not follow suit (Scott-Clayton, 2012). Financially constrained students are often forced to under-invest in higher education or to take on part-time employment, thereby reducing the time available for schoolwork.[1] Resource constraints among less-selective public universities and community colleges, where there are fewer resources per student, also contribute toward low completion rates and student dissatisfaction. Completion times have increased most among

---

[1] Simply providing access to financial aid may not enough, as the application process can be prohibitively complex for some students. Bettinger et al. (2012) show that providing assistance with the Free Application for Federal Student Aid (FAFSA) increases both college entry and persistence, while Castleman and Page (2014) demonstrate that reminding students about the steps to renew FAFSA aid also leads to higher renewal and persistence.

students who start college at these institutions (Bound, Lovenheim, and Turner, 2012), where increases in student demand for higher education are not fully offset with increases in resources – something top-tier schools do by regulating enrollment size (Bound and Turner, 2007).

While the economics of education literature has devoted much attention to the role of resource constraints, comparatively less attention has been given to understanding the role that students themselves play in the production of higher education. Yet, at both the high school and college levels, an emerging recent literature demonstrates the benefits of helping students foster motivation, effort, good study habits, and time-management skills through structured tutoring and coaching. Cook et al. (2014) find that cognitive behavioral therapy and tutoring generate large improvements in math scores and high school graduation rates for troubled youth in Chicago, while Oreopoulos, Lavecchia, and Brown (forthcoming) show that coaching, tutoring, and group activities lead to large increases in high school graduation and college enrollment among youth living in a Toronto public housing project. Structured coaching has also recently been shown to improve outcomes among college students. Scrivener and Weiss (2013) find that the Accelerated Study in Associates Program (ASAP) – a bundle of coaching, tutoring, and student success workshops – in CUNY community colleges nearly doubled graduation rates and Bettinger and Baker (2014) show that telephone coaching by *Inside Track* professionals boosts two-year college retention by 15 percent.

While structured, one-on-one support services can have large effects on student outcomes, they are often costly to implement and difficult to scale up to the student population at large (Bloom, 1984). In this paper, we build on recent advances in social-psychology and behavioral economics, investigating whether technology – specifically, online exercises and text

and email messaging – can be used to generate comparable benefits to one-on-one coaching interventions but at lower costs among first-year university students.

Several recent studies in social-psychology find that one-time, short interventions occurring at an appropriate time can have lasting effects on student outcomes (Yeager and Walton, 2011; Cohen and Garcia, 2014; Walton 2014). Relatively large improvements on academic performance have been documented as a result of several types of intervention, including those that help students define their long-run goals or purpose for learning (Morisano et al., 2010; Yeager, Henderson et al., 2014), teach the "growth-mindset" idea that intelligence is malleable (Yeager et al., 2016), help students keep negative events in perspective by self-affirming their values (Cohen and Sherman, 2014), and help teachers change the tone of feedback to students in order to build trust (Yeager, Vaughns et al., 2014).[2] As a contrast to one-time interventions, other studies in education and behavioral economics attempt to maintain constant, low-touch contact with students or their parents at a low cost by using technology to provide consistent reminders aimed at improving outcomes. For example, several studies have shown that providing text, email, and phone call reminders to parents about their students' progress in school boosts both parental engagement and student performance (Kraft and Dougherty, 2013; Bergman, 2016; Kraft and Rogers, 2014; Mayer et al., 2015). Researchers have also used text messaging communication with college and university students directly, both to increase the likelihood of students renewing financial aid (Castleman and Page, 2014) and improve academic outcomes (Castleman and Meyer, 2016).

---

[2] While the studies cited above all carefully explain the settings and times in which the interventions are likely to be effective, there is growing skepticism about the generalizability of some interventions due to recent failed replication attempts (see, for example, Kost-Smith et al., 2012; Dee, 2015; Harackiewicz et al. (forthcoming))

We contribute to these literatures by examining whether benefits comparable to those obtained from one-on-one coaching can be achieved at lower costs by either a one-time, online intervention designed to affirm students' goals and purpose for attending university or a full-year text and email messaging campaign that provides weekly reminders of academic advice and motivation to students. We work with a sample of more than four thousand undergraduate students who are enrolled in introductory economics courses across all three campuses of the University of Toronto, randomly assigning students to one of three treatment groups or a control group. The treatment groups consist of (i) a one-time, online exercise completed during the first two weeks of class in the fall, (ii) the online intervention plus text and email messaging throughout the full academic year, and (iii) the online intervention plus one-on-one coaching in which students are assigned to upper-year undergraduate coaches. Students in the control group are given a personality test measuring the Big Five personality traits.

We find large positive effects from the one-year coaching service, amounting to approximately a 5 percentage-point increase in average course grades and a 0.35 standard deviation increase in GPA. In contrast, we find no effects on academic outcomes from either the online exercise or the text messaging campaign, even after investigating potentially heterogeneous treatment effects across several student characteristics, including gender, age, incoming high school average, international-student status, and whether students live on residence. Our results suggest that the benefits of personal coaching are not easily replicated by low-cost interventions using technology. As we describe below, coaches were instructed to be proactive by regularly initiating contact with their students and, whenever possible, to provide concrete actionable steps for solving a given problem. Our text messaging approach was not able to replicate this proactive approach. Students had to initiate contact and our team was unable to

"dig deep" into each problem to the same degree as the coaches. We discuss the key challenges – and potential solutions – to using technology to implement coaching-type support at large scale in our discussion of the results.

While our main contribution is assessing the scope for technological interventions to reproduce the benefits of one-on-one coaching, our paper also makes two more general contributions. To our knowledge, we provide the first causal analysis of the effects of a large-scale text messaging campaign on the academic outcomes of college students. The most closely related to work to ours in this respect is Castleman and Meyer (2016), who analyze the effects of text message reminders on academic outcomes such as GPA, the number of credits attempted, and persistence. The authors work with a sample of rural, low-income college students in West Virginia and find that text campaign participants attempted more credits that non-participants, although they are unable to make causal claims about the program's effectiveness because students were not randomly assigned to participation. In contrast, we randomly assign students into the messaging treatment, estimating no effects on academic outcomes. We also show that assigning students to upper-year undergraduate coaches can lead to potentially large academic improvements without the need for professionally trained coaches, as in the Bettinger and Baker (2014) study. Instead, a consistent characteristic across a variety of effective coaching studies appears to be proactive coaches or mentors who regularly contact students to provide support (Cook et al., 2014; Beattie et al. (2016)).[3]

The remainder of this paper is organized as follows: In the next section, we describe the intervention in greater detail, explaining how each treatment and control group exercise was

---

[3]Having our coaches be proactive is a key difference between our coaching program and that which resulted in negligible effects in Angrist, Lang, and Oreopoulos (2009).

implemented. Section 3 describes the data and our empirical strategy for estimating the effects of the intervention, while Section 4 presents the results. We discuss the results in Section 5 and provide concluding remarks in Section 6.

2. Description of the Intervention

We implemented our intervention across all three University of Toronto (U of T) campuses, working with a sample of all students registered for first-year economics classes in the fall of 2015. We cooperated with the instructors of each of these classes, having them agree to make completion of our online "warm-up" exercise worth 2 percent of students' final grade. Students had to complete the exercise in the first two weeks of the fall semester in order to receive credit.[4] The type of online exercise each student had to complete depended on whether the student was randomly assigned to one of the three treatment groups or the control group. Each student created an account and completed the same introductory survey, in which we asked several background questions, including the highest level of education obtained by students' parents, the amount of education they expect to obtain, whether they are first-year or international students, and their work and study time plans for the upcoming year.

Students in the first treatment group then worked through an online exercise, designed to get them thinking about the future they envision and the steps they could take in the upcoming year at U of T to help make that future a reality. The online module lasted approximately 60 to 90 minutes and led students through a series of writing exercises in which they wrote about their ideal futures, both at work and at home, what they would like to accomplish in the current year at

---

[4] We describe our sample, randomization strategy, and balancing tests in the next section.

U of T, how they intend on following certain study strategies to meet their goals, and whether they want to get involved with extracurricular activities at the university. Varying minimum word-count and time restrictions were placed on several pages of the online exercise to ensure that students gave due consideration to each of their answers before proceeding.[5] The exercise aimed to make students' distant goals salient in the present and to provide information on effective study strategies and how to deal with the inevitable setbacks that arise during the course of an academic year. At the conclusion of the exercise, students were emailed a copy of the answers they provided to store for future reference. A full description of the online exercise is available in Appendix A.

Students in the second treatment group completed the same online exercise but were additionally offered the opportunity to provide their phone numbers and participate in a text and email messaging campaign lasting throughout both the fall semester in 2015 and the winter semester in 2016. The messaging campaign was called *You@UofT* – a name we chose to emphasize that the program was geared to provide personalized assistance and help students reach their individual definitions of success. Students had the opportunity to choose the frequency with which they received text and email messages, with choices including once a week, two to three times per week, and three or more times per week. All students who were randomly sorted into this treatment received email messages, while only those students who provided their cell phone numbers received text messages throughout the year.[6] Students were

---

[5] Nearly all students took the exercise seriously, writing coherent statements that served as logical answers to the relevant questions. There were very few instances of students writing random words to hit the word-count minimum, and the students who did only did so on some questions, not throughout the entire exercise.

[6] A total of 2,024 students were randomly sorted into the messaging campaign treatment and 1,540 (76 percent) provided their phone numbers.

free to opt out of receiving email messages, text messages, or both at any time after the exercise, although few chose to do so.

A full documentation of all of the text and email messages we sent throughout the experiment is available in Appendix B. Our messages typically focused on three themes: academic and study preparation advice, information on the resources available at the university, and motivation and encouragement. Students always received both a text and email message. Text messages were typically three to four lines in length while emails were longer and provided more detailed information with which students could follow up. The *You@UofT* program offered personalized two-way communication, as both text and email messages regularly encouraged students to look further into the content and to reach out to us if they had specific or general questions. Approximately 25 percent of students engaged in two-way communication with our team via text messages, compared to only 3 percent of students responding via email.

There was wide variation in the types of response we received from students. For example, some students asked for locations of certain facilities on campus or how to stay on residence during the holiday break, while others said they need help with English skills or specific courses. Some students expressed relatively deep emotions, such as feeling anxious about family pressure to succeed in school or from doing poorly on an evaluation. We also received messages of thanks for our appropriately-timed advice or motivation and several students messaged us to tell us how well they were doing in their courses and how much they appreciated the communication. No matter the type of message received and when, we attempted to provide a personalized support service, typically responding to the inquiry within a few hours (and usually within less than one hour). The *You@UofT* program served as a virtual coach from whom students could expect a rapid response at any time. In this sense, the program leveraged

technology to provide a personalized coaching service at large-scale to all students while keeping the cost per-student lower than what is typically incurred with one-on-one coaching.

To test how the effects of the *You@UofT* program compare to those of in-person, one-on-one coaching, a third group of students also completed the online exercise and was offered the opportunity to participate in a pilot project in which they would be assigned to an upper-year undergraduate student acting as a personal coach. Coaches were available to meet with students to answer any questions via Skype, phone, or in person, and would send their students regular text and email messages of advice, encouragement, and motivation, much like the *You@UofT* program described above. In contrast to the messaging program, however, coaches were instructed to be proactive and regularly monitor their students' progress. Whereas the *You@UofT* program attempts to "nudge" students in the right direction with academic advice, coaches play a greater "support" role, sensitively guiding students through problems.

The coaching program was offered only to students at one of the university's satellite campuses, the University of Toronto at Mississauga campus. Our coaching treatment group was established by randomly drawing twenty-four students from the group of students that were randomly assigned into the text message campaign treatment. At the conclusion of the online exercise, instead of being invited to provide a phone number for the purpose of receiving text messages, these twenty-four students were given the opportunity to participate in a pilot coaching program. A total of seventeen students agreed to participate in the coaching program, while seven students declined. These seventeen students were assigned to a team of four upper-year undergraduate coaches, who participated in our program as part of a research opportunity program. Each coach originally agreed to coach six students throughout the academic year but

was eventually responsible for only four or five students as result of seven students declining participation.

Our coaches describe providing support to their students on a wide variety of issues, including questions about campus locations, booking appointments with counsellors, selecting majors, getting jobs on campus, specific questions about coursework, and feelings of nervousness, sadness, or anxiety. Coaches and students scheduled their own regular meetings, approximately half of which occurred face-to-face and half of which occurred via Skype or text messaging. Since each coach was responsible for only four or five students, they were able to remember the issues each student was dealing with, proactively reach out to do regular status checks, and provide specific advice for dealing with each unique problem. The extra time afforded to coaches with low student-to-coach ratios allowed them to befriend their students, communicate informally and with humor, and slowly prompt students about their issues through a series of gentle, open-ended questions until students felt comfortable to open up about the details of their particular problems. Once trust was established between coaches and students, students felt more comfortable discussing challenging problems, making it easier for coaches to provide clear advice.

Students assigned to the control group were given a personality test measuring the Big Five personality traits. The test could be completed in approximately 45 to 60 minutes and, at the conclusion of the exercise, students were emailed their scores in a report describing how they fair on each of the Big Five traits.[7]

---

[7] Beattie et al. (2016) use the data resulting from the personality test exercise to explore non-academic predictors of performance in university.

3. Data Description and Empirical Strategy

In this section, we describe the data we collected from the experiment and how we estimate the effects of the three treatments.

3.1. Data Description

Our experiment is registered with the American Economic Association's registry for randomized controlled trials. Prior to the experiment, we intended to sort 30 percent of students into the control group, 20 percent into the online-exercise-only treatment, and 30 percent into the treatment group that received the text messaging campaign in addition to the online exercise.[8] Students were sorted into one of the treatment groups or the control group according to the randomly-generated last digits of their student numbers, which they provided upon registering online for our experiment.[9] As mentioned, we established the personal coach treatment group by drawing twenty-four students at random from the group of students that was intended to be a part of the text messaging campaign. Table 1 shows some basic statistics about our randomization strategy among first-year students, which indicate that we successfully reached each of our randomization targets.[10] Furthermore, high fractions of students completed each exercise, with completion rates ranging from 95 to 99 percent.

---

[8] The remaining 20 percent of students were sorted into an online belonging exercise similar to the exercise that appeared in Walton et al. (2015). The effects of this treatment will be presented in a separate, standalone paper and we therefore do not discuss this treatment throughout the remainder of this paper. Students in this treatment are dropped from the analysis.

[9] Since completing the exercise was a course requirement worth 2 percent of the final grade in introductory economics, students had a high incentive to provide their real student numbers and complete the exercise.

[10] The table conditions on first-year students because the online belonging exercise mentioned above was only offered to students in their first year of study. Thus, students registered in second year or above are more likely (than

Table 2 shows summary statistics for baseline characteristics among students in the control group along with differences between each treatment group mean and control group mean for each baseline characteristic. The treatment indicators are never jointly significant in explaining variation in any student characteristic. The only individual exceptions are that students in the online-only group are slightly more likely to live on residence and to be fist-generation students while students in the personal coaching group report having a slightly less difficult time transitioning to university.

In terms of the descriptive statistics, approximately half of our sample is female and the average student 18.5 years old. Approximately half of the students are non-native English speakers and half are not Canadian citizens. Only 30 percent of students live on residence, but this fraction is pulled down by the two satellite campuses of U of T, the Mississauga and Scarborough campuses, which are both commuter campuses. At the main campus, St. George, 40 percent of students live on residence. The main campus also has students with higher incoming high school average grades: while the average is 87 percent across all students, it is 90 percent at the St. George campus. Approximately 24 percent of students are first generation and 43 percent have international status.

3.2. Empirical Strategy

Since we successfully randomized students into various treatment groups, we estimate the effects of each treatment by simply comparing mean outcomes in a regression framework. These estimates are 'Intent to Treat' effects, each representing the average impact from being invited to

---

the intended pre-randomization fractions) to be in one of the other three treatment groups. We account for this in our empirical strategy below by including first-year fixed effects in every regression.

complete the exercise, regardless of whether students actually completed or not. Given that almost all students finished, however, the estimates are likely close to 'Average Treatment Effects', measuring the average effect from completing the exercise for the entire sample. More formally, we estimate the following equation:

$$y_{ij} = \alpha + \beta_1 Online_i + \beta_2 Text_i + \beta_3 Coach_i + \delta_j + \mu FY_i + \epsilon_{ij},$$

where the outcome of student i who attends campus j is regressed on indicators for each of the three treatment exercises students were given, campus fixed effects, and a first-year student indicator. We include campus fixed effects because the coaching treatment was only offered at the Mississauga campus, which accepts students with lower high school averages who tend to perform worse in university than students who attend the main campus, St. George. We include the first-year indicator to account for students who are enrolled in second year and above being more likely to be in one of the three treatment groups than students in first year, as only first-year students were randomly assigned to the online belonging exercise that is not analyzed in this paper. The main parameters of interest are β_1, β_2, and β_3, which represent the effect of the online treatment, the online plus messaging treatment, and the online plus coaching treatment, respectively. As mentioned, we include all students in the analysis, irrespective of whether they completed the online exercise, provided a cell phone number, or agreed to participate in the coaching program, implying that our parameter estimates all represent intent to treat effects.

Our main outcomes of interest are course grades, grade point average (GPA), number of credits earned, and number of credits failed. When the outcome is course grades, we stack all of the reported course grades for a given student and run a regression at the course-student level in

13

which we cluster the standard errors by student. For all other outcomes, we run the regression at the student-level and report robust standard errors.

4. Results

4.1. Main Results

Table 3 presents the results from stacked regressions at the student-course level, with course grades from all courses (fall semester, winter semester, and full-year courses) as the dependent variable. Standard errors are clustered by student. The results in columns (1) show that neither the online exercise on its own nor the online exercise and texting messaging treatment had any effect on course grades. The insignificant effects are not due to statistical imprecision. We can rule out impacts above 6 percent of a standard deviation using a 95 percent confidence interval. In contrast, the personal coaching treatment had relatively large effects, boosting the average course grade by 4.92 percentage points, which amounts to 30 percent of the control group standard deviation. Reassuringly, including student age and gender as additional control variables in column (2) does not change the result. Columns (3) and (4) use a student's course-specific grade-point relative to the average course grade-point as the dependent variable. While the coaching effects are not statistically significant, they are substantially larger than the effects of the online exercise or the messaging campaign, each of which are zero.

Columns (5) to (14) regress indicators variables for whether students had a grade above the cutoff indicated in each column header. When all course grades are considered as the dependent variable, the coaching treatment has the effect of decreasing the likelihood of students earning extremely low grades. Coaching students are 8 percentage points less likely to earn a

grade below 60 percent. These students are also more likely to earn higher grades, although the effects are measured more imprecisely.[11]

In Table 4, we consider grades from courses taken only in the fall semester as the outcome of interest. The coaching effects on fall grades are slightly weaker than those on grades from all courses but coached students still earn higher grades, on average. In Table 5, we show treatment effects on grades from courses taken only in the winter semester. Here, the coaching treatment effects are even stronger, as coaching boosts the average grade by 6.7 percentage points (or 39 percent of the control group standard deviation). Students in the coaching treatment also tend to earn higher relative grades in their winter courses, scoring approximately 0.43 grade points higher than the average student in their courses. Columns (5) to (12) show that the coaching treatment shifted the performance distribution rightward, as coached students are more likely to earn a grade above 75 percent and are much less likely to earn a grade below 60 percent in their winter semester courses. It thus appears that the effects of the coaching treatment strengthened over time. It may be the case that students developed more trust with their coaches as the academic year progressed and that they learned how to use resources more effectively.

Figures 1 to 3 show graphically the effects of the coaching treatment strengthening over time. Each figure shows the treatment-group specific distributions of residual grades, after campus and first-year effects are removed. Figure 1 reports the residual-grade distributions for all courses (full-year, winter semester, and fall semester) and clearly shows that the coaching distribution is shifted rightward relative the control group distribution and the distributions for the online and texting treatments. Indeed, a Kolmogorov-Smirnov test rejects that the coaching

---

[11] Including as additional controls those variables that statistically differed between the control group and the online-only or coaching students (living in residence, first-generation student, and university transition difficulty) results in very similar point estimates as those reported here.

distribution is the same as the control and texting distributions at the 1 percent level and the online-only distribution at the 5 percent level. Contrasting Figures 2 and 3 shows that the strongest coaching effects emerge in the winter semester, as the coaching distribution's rightward shift relative to the other distributions is much more pronounced in the winter semester in Figure 3 than in the fall semester in Figure 2.[12]

Table 6 shows treatment effects on other academic outcomes with one observation per-student and student-level regressions. The dependent variables are constructed using outcomes from all courses. The coaching treatment causes a 0.35 grade-point increase in student GPA, equivalent to approximately 35 percent of the control group standard deviation. Coached students failed fewer credits and earned more credits, on average, than students in the control group. As with stacked grade outcomes, there are no detectable effects on GPA or the number of credits failed or earned from the online exercise treatment or the text messaging campaign. Although we do not report these results separately, the effects of the coaching treatment on these outcomes are again stronger in the winter semester than in the fall semester.

4.2. Heterogeneous Treatment Effects

In this section, we explore heterogeneous treatment effects across different student subgroups and across the three U of T campuses. As mentioned above, only twenty-four students are in the coaching treatment. We therefore investigate the heterogeneous effects of coaching along with the other treatments only for completeness; with a sample size of only twenty-four students, we

---

[12] Note that the coaching group density for fall grades in Figure 2 does not have overlapping mass with the other densities in the left tail of the grade distributions. Thus, while the coaching program does not cause a pronounced shift of the grade distribution in the fall, it does appear to prevent students from earning extremely low grades, as the grade distribution is truncated at a residual grade of -14.6.

lack the necessary power to meaningfully distinguish potential differences in the effects of coaching across different subgroups.

Table 7 shows treatment effects on all course grades across a variety of student subgroups. The effects of the online-only and text messaging treatments are not statistically significant for any type of student, with the lone exception being a small positive effect of the online-only treatment on students whose mother tongue is English. We find that coaching effects are stronger for men, students who are 20 years of age or older, first-generation students, and students who are not in first year. Given the small coaching treatment sample size, however, we are hesitant to push these results further without investigation on a bigger a sample of students.

We also investigate whether there are heterogeneous treatment effects across the three U of T campuses. Tables 8, 9, and 10 show the effects of the three treatments on grade outcomes from all courses at the St. George, Scarborough, and Mississauga campuses, respectively. The effects of the coaching treatment are only reported among students attending the Mississauga campus, as only these students were randomly offered the coaching service. Table 8 shows that neither the online exercise nor the text messaging campaign had any effect on student grade outcomes at the main campus, St. George. Table 9 shows similar patterns at the Scarborough campus, where there are also no significant effects from either the online exercise or the text messaging program.[13]

Table 10 shows the effects of the three treatments on grade outcomes from all courses at the Mississauga campus. The estimated effects of the online exercise and the texting campaign are larger on this campus than those found in the pooled sample and at the other two campuses

---

[13] Tables 8 and 9 show treatment effects on all course grades. While we do not report the results, there are also virtually no effects on grades from full-year courses, winter courses, and fall courses at both the St. George and Scarborough campuses.

separately. Columns (1) and (2) show that the online exercise boosts course grades by 1.93 percentage points, on average. The estimate is significant at the 10 percent level and implies that the online exercise increases grades by 11 percent of the control group standard deviation. This is a relatively small effect when compared to the coaching treatment, which increases grades by 5.95 percentage points or 35 percent of a standard deviation. The online exercise also boosts the likelihood that students earn a grade above 80 percent and decreases the likelihood of earning a grade below 60 percent, but the effects are again smaller than those from the coaching treatment.

In sum, there is robust evidence that the neither the online exercise nor the text messaging campaign were effective at improving students' academic outcomes, both in the general population and across various student subgroups. The lone exception may be the positive effects of online exercise among students at the Mississauga campus, although these effects are small relative to the coaching treatment. They are also larger in magnitude than the effects of the text messaging treatment, calling into question whether their statistical significance is due to real treatment effectiveness or random chance. In contrast to the one-time online intervention and the consistent-contact text messaging campaign, we find economically and statistically significant effects of the personal coaching treatment and a wide variety of academic outcomes. We discuss the potential reasons why the coaching treatment was more effective than the other two treatments and how the text messaging campaign can be adjusted to improve its effectiveness in the following section.

5. Discussion

We find that neither the one-time online intervention nor the text messaging campaign have significant effects on student outcomes while personal coaching boosts students' grades and GPA by approximately 35 percent of a standard deviation. The key disadvantage of our coaching programs – and others like it – is that it is costlier to implement and scale-up than one-time online interventions or interventions that rely heavily on technology for constant contact with students.

Although our upper-year coaches participated in the experiment as part of a research opportunity program (for course credit), such students would typically require at least $20 per-hour from the university to provide coaching services. With each of our coaches devoting approximately seven total hours per-week to coaching, this conservative wage rate implies that the coaching program would regularly cost over $13,000 to service seventeen student participants. In contrast, after the initial setup costs, the online intervention is done at no additional cost and the total cost of the messaging campaign that serviced more the 1500 student participants was approximately $1,200 for the entire academic year. Given the large differences in relative costs, it is worth discussing the key differences between the coaching treatment and the text messaging campaign, with the goal of learning how to modify the texting initiative to increase its effectiveness.[14]

A common characteristic across many successful coaching programs is regular student-coach interaction facilitated either by mandatory meetings between coaches and students or proactive coaches who regularly initiate contact (Scrivener and Weiss, 2013; Bettinger and

---

[14] An alternative way to reduce the costs of one-on-one coaching is to recruit upper-year undergraduates to volunteer their time as coaches, with the promise of gathering valuable experience to place on a resume. Universities can help make this type of volunteer work attractive by creating a system that official recognizes students volunteer investments. The U of T Co-Curricular Record, for example, is designed to give students explicit credit for their experiences outside of the classroom (https://ccr.utoronto.ca/home.htm).

Baker, 2014; Cook et al., 2014; Oreopoulos et al. (forthcoming)). Indeed, having proactive coaches is the key difference between our coaching service and that offered in Angrist, Lang, and Oreopoulos (2009), which was also conducted at the Mississauga campus of U of T but resulted in negligible effects. In that study, one treatment arm had student coaches email once a week and offer to meet at a student service office. In contrast, this study had student coaches aggressively initiate contact and build trust with students over time, in person and through text. Our coaches were able to clearly understand the problems students were facing through a series of open-ended gentle questions. Upon understanding the problem, the coaches could provide clear advice, resulting in most conversations ending with students knowing at least one specific action to take to help them solve their current problems.

Our text messaging campaign sent weekly messages of academic advice, resource information, and motivation, but we did not initiate contact with individual students to specifically ask how they were doing or whether they wanted to talk about something specific. The text messaging team often did invite students to reply to our texts and share their concerns, but we were unable to do this from the perspective of a coach – that is, we did not present ourselves as a real person (with a name) and we did not try to establish a rapport with the students. Our inability to reach out to all students and softly guide the conversation likely prevent us from learning the key details of their specific problems. Although we would provide answers and advice to the questions we received, we did not have as much information on the students' backgrounds as our coaches did, and thus could not tailor our responses to each student's specific circumstances.

Our coaches were also able to build trust with their students by often fulfilling a support role for them. Figure 4 provides an example of how the coaching service was far more effective

than the text messaging campaign in this respect. The text messages attempted to "nudge" students in the right direction, rather than provide tailored support. The left panel of Figure 4 shows three consecutive text messages, in which we provide a tip on stress management, an inspirational quote, and a time-management tip around the exam period. As shown in this example, it was very often the case that students would not respond to such messages. In contrast, the student-coach interaction in the right panel of Figure 4 shows our coaches offering more of a supportive role to students rather than trying to simply nudge them toward a certain path. The coach starts by asking an open-ended question, to which the student responds and guides the conversation forward. In this particular example, the coach assures the student that they will be available to help with a pending deadline and shows a genuine interest in the events in the student's life.

Coaches also kept a record of their evolving conversations with each student and were able to check in with students to ask how previously discussed issues were being resolved. Although we kept a record of all text message conversations we had with students, a lack of resources did not allow us to regularly check in with students to see how previous events had unfolded. A lack of these regular check-ups likely prevented us from helping students effectively with the problems they told us about and from establishing the trust required for students to share additional problems.

In sum, the two key features that distinguish the coaching service from the texting campaign are that coaches proactively initiated discussion with students about their problems and were able to establish relationships based on trust in which students felt comfortable to openly discuss their issues. Our coaches being able to slowly guide the conversations and inquire about

previously discussed events likely contributed in large part to establishing the required trust between students and coaches.

Future work that attempts to improve academic outcomes in higher education with interventions that use technology to maintain constant contact with students should keep in mind that simply nudging students in the right direction may not be enough. A more personalized approach may be required, in which coaches or mentors initially guide students through a series of gentle conversations and subsequently show a proactive interest in students' lives. These conversations need not necessarily occur during face-to-face meetings, but the available evidence suggests that they should occur frequently and be initiated by the coaches. While such an intervention is likely to be costlier than the text messaging campaign in this study, it is also likely to be more effective than the current messaging campaign but still less costly than the personalized coaching treatment analyzed here.

6. Conclusion

Building on recent insights from social psychology and behavioral economics, we estimated the effects of the following three treatments on students' academic outcomes: (i) an online exercise designed to affirm students' values and goals in university, (ii) a two-way text and email messaging campaign that provided students with academic advice and motivation, and (iii) a one-on-one coaching program in which students were matched with upper-year undergraduate students acting as coaches. We found no effect from either the online exercise or the messaging campaign, across the general student population and many student subgroups. In contrast, we found large and significant effects from the coaching program, which increased

average course grades by 0.3 standard deviations and GPA by 0.35 standard deviations. Coached students also failed fewer credits and earned more credits, on average.

Contrasting the designs of the text messaging campaign treatment and the coaching treatment, we argued that coaches being proactive in contacting students was a critical feature of the program's success. Coaches were also better able to keep an account of previously discussed issues, subsequently inquire about how those issues were being resolved, and build the required trust that made students feel comfortable enough to keep returning for help.

We intend to test these conjectures in ongoing work, which features another phase of the experiment where the text messaging treatment has been modified to more closely reflect the coaching program studied here. Our support team has expanded to include ten "virtual" coaches who are each assigned 80 to 100 students. Each virtual coach will regularly engage with their students via text and email, refer back to previously discussed events, and provide consistent support throughout the academic year. We also intend to expand the sample size of the personal coaching treatment and again test how far technology can go in replicating the effects of one-on-one coaching. These interventions are attractive because they are relatively inexpensive and scalable, and can be implemented across a wide range of settings. With continued research, there exists real potential for developing a consensus around the types of student services that are most effective for improving student progress and well-being.

References

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics.* 1(1): 136-163.

Beattie, Graham, Jean-William P. Laliberté and Philip Oreopoulos (2016). "Thrivers and Divers: Using Non-Academic Measures to Predict College Success and Failure," NBER Working Paper #XXXXX.

Belley, Philippe and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital.* 1(1): 37–89.

Bergman, Peter. 2016. "Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment." *Columbia University Working Paper*.

Bettinger, Eric and Rachel Baker. 2014. "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising." *Educational Evaluation and Policy Analysis*. 35(1): 3-19.

Bettinger, Eric, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics* 127(3): 1205-1242.

Bloom, Benjamin. 1984. "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring." *Educational Researcher*. 13(6): 4-16.

Bound, John and Sarah E. Turner. 2007. "Cohort Crowding: How Resources Affect Collegiate Attainment." *Journal of Public Economics*. 91 (5–6): 877–99.

Bound, John, Michael Lovenheim, and Sarah E. Turner. 2012. "Increasing Time to Baccalaureate Degree in the United States." Education Finance and Policy. 7(4): 375-424.

Castleman, Benjamin and Lindsay Page. 2014. Freshman "Year Financial Aid Nudges: An Experiment to Increase FAFSA Renewal and College Persistence." Center for Education Policy and Workforce Competitiveness Working Paper No. 28. Charlottesville, VA: University of Virginia.

Castleman, Benjamin and Katharine Meyer. 2016. "Freshman "Can Text Message Nudges Improve Academic Outcomes in College? Evidence from a West Virginia Initiative." Center for Education Policy and Workforce Competitiveness Working Paper No. 43. Charlottesville, VA: University of Virginia.

Cohen, Geoffrey and Julio Garcia. 2014. "Educational Theory, Practice, and Policy and the Wisdom of Social Psychology." *Policy Insights from the Behavioral and Brain Sciences*.1(1): 13-20.

Cohen, Geoffrey and David Sherman. 2014. "The Psychology of Change: Self-Affirmation and Social Psychological Intervention." *Annual Reviews of Psychology*. 65: 333-371.

Cook, Philip J., Kenneth Dodge, George Farkas, Roland G. Fryer, Jr, Jonathan Guryan, Jens Ludwig, Susan Mayer, Harold Pollack, Laurence Steinberg. 2014. "The (Surprising) Efficacy of Academic and Behavioral Intervention with Disadvantaged Youth: Results from a Randomized Experiment in Chicago." National Bureau of Economic Research Working Paper 19862. Cambridge, Mass.

Dee, Thomas. 2015. "Social Identity and Achievement Gaps: Evidence from and Affirmation Intervention." *Journal of Research on Educational Effectiveness*. 8(2): 149-168.

Harackiewicz, Judith M., Elizabeth A. Canning, Yoi Tibbetts, Stacy J. Priniski, and Janet S. Hyde. "Closing Achievement Gaps with a Utility-Value Intervention: Disentangling Race and Social Class." *Journal of Personality and Social Psychology, forthcoming*.

Kost-Smith, Lauren E., Steven J. Pollock, Noah D. Finkelstein, Geoffrey L. Cohen, Tiffany A. Ito and Akira Miyake. 2012. "Replicating a Self-Affirmation Intervention to Address Gender Differences: Successes and Challenges." *Physics Education Research Conference, AIP Conf. Proc*. 231-234.

Kraft, Matthew. A., and Dougherty, Shaun. M. 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, Matthew. A., and Todd Rogers. 2014. "The Underutilized Potential of Teacher-to-Parent Communication: Evidence from a Field Experiment." *Harvard Kennedy School Faculty Research Working Paper Series, RWP-14-049.*

Mayer, Susan E., Ariel Kalil, Philip Oreopoulos, and Sebastian Gallegos. 2015. "Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together(PACT) Intervention." National Bureau of Economic Research Working Paper 21602. Cambridge, Mass.

Morisano, Dominique, Jacob Hirsh, Jordan Peterson, Robert Pihl, and Bruce Shore. 2010. "Setting, Elaborating, and Reflecting on Personal Goals Improves Academic Performance." *Journal of Applied Psychology*. 95(2): 255–264.

Oreopoulos, Philip, Adam Lavecchia and Robert S. Brown. "Pathways to Education: An Integrated Approach to Helping At-Risk High School Students." *Journal of Political Economy*, forthcoming.

Richard Arum and Jospia Roksa. 2011. Academically Adrift: Limited Learning on College Campuses Chicago, IL: University of Chicago Press.

Scott-Clayton, Judith. 2012. "What Explains Trends in Labor Supply among U.S. Undergraduates, 1970–2009?" National Bureau of Economic Research Working Paper 17744. Cambridge, Mass.

Scrivener, Susan and Michael J. Weiss. 2013. "More Graduates: Two-Year Results from an Evaluation of Accelerated Study in Associate Programs (ASAP) for Developmental Education Students." Policy Brief, MDRC.

Symonds, William C., Robert Schwartz, and Ronald F. Ferguson. 2011. "Pathways to prosperity: Meeting the challenge of preparing young Americans for the 21st century." *Pathways to Prosperity Project*, Harvard University Graduate School of Education.

Turner, Sarah. 2004. Going to college and finishing college. Explaining different educational outcomes. In C. Hoxby (Ed.), *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 13–62). Chicago, IL: University of Chicago Press.

Walton, Gregory. 2014. "The New Science of Wise Psychological Interventions." *Current Directions in Psychological Science*. 23(1): 73-82.

Walton, Gregory, Christine Logel, Jennifer Peach, Steven Spencer, and Mark Zanna. 2015. "Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering." *Journal of Educational Psychology*. 107(2): 468-485.

Yeager, David and Gregory Walton. 2011. "Social-Psychological Interventions in Education: They're Not Magic." *Review of Educational Research*. 81(2): 267–301.

Yeager, David, Marlone D. Henderson, David Paunesku, Gregory M. Walton, Sidney D'Mello, Brian J. Spitzer, Angela Lee Duckworth. 2014. "Boring but Important: A Self-Transcendent Purpose for Learning Fosters Academic Self-Regulation." *Journal of Personality and Social Psychology*. 107(4): 559–580.

Yeager, David, Valerie Purdie-Vaughns, Julio Garcia, Nancy Apfel, Patti Brzustoski, Allison Master, William T. Hessert, Matthew E. Williams. 2014. "Breaking the Cycle of Mistrust: Wise Interventions to Provide Critical Feedback Across the Racial Divide" *Journal of Experimental Psychology*. 143(2): 804–824.

Yeager, David, Carissa Romero, Dave Paunesku, Christopher S. Hulleman, Barbara Schneider, Cintia Hinojosa, Hae Yeon Lee, Joseph O'Brien, Kate Flint, Alice Roberts, and Jill Trott. 2016. "Using Design Thinking to Improve Psychological Interventions: The Case of the Growth Mindset During the Transition to High School." *Journal of Educational Psychology*. 108(3): 374–391.

Table 1: Treatment randomization among First-year students

	Treatment Group			
	Control	Goal-setting alone	Goal-setting with reminders	Belonging
Number of students	1,455	968	1,505	972
(i) Fraction of total	29.61%	19.62%	30.9%	19.78%
(ii) Intended fraction	30%	20%	30%	20%
p-value of (i) = (ii)	0.55	0.50	0.14	0.70
Completed the exercise	1,440	922	1,449	960
Fraction selecting SMS frequency				
1 time a week	-	-	40%	-
2 times a week	-	-	19%	-
3+ per week	-	-	42%	-

Note: The intended fraction of students in the goal-setting with reminders group comprises the twenty-four students who were assigned to a coaching program.

Table 2: Summary Statistics and Balancing Tests

		Treatment Status			
Student Characteristics	Control Sample Mean [Standard Deviation]	Online Only Difference [Standard Error]	Text Messaging Difference [Standard Error]	Personal Coaching Difference [Standard Error]	P-Value from F-Test of No Difference
Female	0.524	0.011	0.013	0.098	0.700
	[0.500]	[0.018]	[0.017]	[0.105]	
Age in first year	18.511	0.025	0.064	-0.533	0.216
	[1.428]	[0.058]	[0.052]	[0.332]	
High school average grade	87.284	-0.093	-0.208	0.120	0.642
	[5.122]	[0.182]	[0.164]	[1.004]	
Non-English mother tongue	0.551	-0.002	-0.005	-0.064	0.932
	[0.498]	[0.018]	[0.016]	[0.105]	
Non-Canadian citizenship	0.541	0.002	-0.001	-0.069	0.926
	[0.498]	[0.018]	[0.017]	[0.105]	
Living in residence	0.300	0.035**	0.019	-0.041	0.143
	[0.458]	[0.016]	[0.014]	[0.090]	
Expects to get more than undergraduate degree	0.657	0.008	0.007	0.069	0.87
	[0.475]	[0.017]	[0.015]	[0.097]	
First generation student	0.237	0.027*	0.011	-0.036	0.375
	[0.425]	[0.016]	[0.014]	[0.092]	
Expect to receive A- average or more	0.659	0.008	0.022	-0.082	0.39
	[0.474]	[0.017]	[0.015]	[0.097]	
Hours expected to study	18.243	-0.323	-0.029	-0.175	0.855
	[10.890]	[0.398]	[0.355]	[2.257]	
Expect to work more than 8 hours/week	0.376	-0.002	-0.013	0.019	0.847
	[0.485]	[0.017]	[0.016]	[0.099]	
Economics is a required course	0.577	-0.001	-0.003	0.084	0.842
	[0.494]	[0.017]	[0.015]	[0.096]	
International Student	0.436	-0.024	-0.016	-0.008	0.571
	[0.496]	[0.018]	[0.016]	[0.102]	
Have a clear vision for future (1 to 7)	4.842	0.023	-0.028	-0.299	0.603
	[1.467]	[0.054]	[0.048]	[0.308]	
Think about the future (1 to 7)	5.559	0.041	-0.003	-0.199	0.605
	[1.226]	[0.045]	[0.040]	[0.256]	
Tend to cram for exams (1 to 7)	4.105	-0.035	-0.008	-0.002	0.936
	[1.521]	[0.055]	[0.049]	[0.312]	
Identify with university (1 to 7)	5.646	-0.028	0.000	0.145	0.893
	[1.427]	[0.053]	[0.047]	[0.300]	
Transition has been so far challenging (1 to 7)	4.774	0.001	0.015	-0.782**	0.146
	[1.670]	[0.061]	[0.054]	[0.344]	

Summary statistics and differences are calculated using the full sample of students. All differences are conditional on campus location and first-year fixed effects. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 3: Effects on All Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Grades		Grade Relative to Class Average		Grade Above 60		Grade Above 65		Grade Above 70		Grade above 75		Grade Above 80	
	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	0.469	0.462	0.063	0.063	0.016	0.016	0.020*	0.020*	0.020	0.020	0.015	0.015	0.003	0.004
	[0.471]	[0.472]	[0.041]	[0.041]	[0.010]	[0.010]	[0.012]	[0.012]	[0.013]	[0.013]	[0.013]	[0.013]	[0.012]	[0.012]
Text Messaging	-0.116	-0.132	-0.002	-0.003	-0.002	-0.002	-0.002	-0.003	-0.007	-0.007	-0.002	-0.003	-0.005	-0.005
	[0.424]	[0.423]	[0.038]	[0.038]	[0.009]	[0.009]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.010]	[0.010]
Coaching	4.920**	4.876**	0.270	0.273	0.086*	0.082	0.102	0.099	0.086	0.085	0.090	0.094	0.051	0.058
	[1.930]	[1.927]	[0.175]	[0.174]	[0.050]	[0.050]	[0.062]	[0.062]	[0.069]	[0.068]	[0.066]	[0.066]	[0.058]	[0.058]
Control Mean	68.905		0.077		0.784		0.671		0.570		0.401		0.275	
[std. dev]	[16.166]		[1.181]		[0.412]		[0.470]		[0.495]		[0.490]		[0.447]	
Observations	29,591	29,591	18,162	18,162	29,591	29,591	29,591	29,591	29,591	29,591	29,591	29,591	29,591	29,591
R-squared	0.021	0.022	0.013	0.014	0.009	0.011	0.011	0.011	0.015	0.015	0.016	0.017	0.016	0.017
F-stat	2.691	2.675	1.760	1.792	2.163	2.061	2.293	2.260	2.047	2.038	1.275	1.340	0.442	0.529
Prob > p	0.0447	0.0456	0.153	0.146	0.0902	0.103	0.0760	0.0794	0.105	0.106	0.281	0.260	0.723	0.663

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. All regressions control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 4: Effects on Fall Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Grades		Grade Relative to Class Average		Grade Above 60		Grade Above 65		Grade Above 70		Grade above 75		Grade Above 80	
	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	0.143	0.143	0.024	0.024	0.005	0.005	0.007	0.007	0.020	0.020	0.012	0.013	0.001	0.001
	[0.575]	[0.575]	[0.043]	[0.043]	[0.013]	[0.013]	[0.015]	[0.015]	[0.016]	[0.016]	[0.017]	[0.017]	[0.015]	[0.015]
Text Messaging	0.073	0.075	0.016	0.016	0.008	0.008	0.003	0.003	-0.001	-0.001	0.005	0.005	-0.002	-0.002
	[0.505]	[0.506]	[0.038]	[0.038]	[0.012]	[0.012]	[0.014]	[0.014]	[0.015]	[0.015]	[0.015]	[0.015]	[0.013]	[0.013]
Coaching	4.897***	4.935***	0.269	0.267	0.041	0.036	0.096	0.093	0.101	0.101	0.082	0.090	0.082	0.094
	[1.874]	[1.880]	[0.167]	[0.168]	[0.069]	[0.069]	[0.075]	[0.075]	[0.087]	[0.087]	[0.082]	[0.082]	[0.079]	[0.080]
Control Mean	69.428		0.066		0.785		0.677		0.573		0.412		0.287	
[std. dev]	[15.744]		[1.174]		[0.411]		[0.468]		[0.495]		[0.492]		[0.453]	
Observations	10,106	10,106	10,065	10,065	10,106	10,106	10,106	10,106	10,106	10,106	10,106	10,106	10,106	10,106
R-squared	0.029	0.029	0.017	0.017	0.011	0.011	0.013	0.014	0.024	0.024	0.030	0.031	0.034	0.038
F-stat	2.290	2.311	0.910	0.895	0.245	0.218	0.596	0.559	1.091	1.077	0.489	0.554	0.385	0.481
Prob > p	0.0763	0.0742	0.435	0.443	0.865	0.884	0.617	0.642	0.352	0.358	0.690	0.646	0.764	0.696

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. All regressions control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 5: Effects on Winter Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Grades		Grade Relative to Class Average		Grade Above 60		Grade Above 65		Grade Above 70		Grade above 75		Grade Above 80	
	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	0.596	0.581	0.111*	0.111*	0.017	0.017	0.030**	0.030*	0.018	0.018	0.004	0.003	0.003	0.003
	[0.607]	[0.607]	[0.057]	[0.057]	[0.013]	[0.013]	[0.015]	[0.015]	[0.017]	[0.017]	[0.017]	[0.017]	[0.015]	[0.015]
Text Messaging	-0.163	-0.197	-0.020	-0.021	-0.002	-0.002	-0.005	-0.005	-0.015	-0.016	-0.022	-0.022	-0.018	-0.018
	[0.549]	[0.549]	[0.053]	[0.053]	[0.012]	[0.012]	[0.014]	[0.014]	[0.015]	[0.015]	[0.015]	[0.015]	[0.014]	[0.014]
Coaching	6.659***	6.525***	0.431**	0.431**	0.121**	0.112**	0.145**	0.139*	0.161*	0.157*	0.167*	0.169*	0.040	0.047
	[2.107]	[2.094]	[0.207]	[0.207]	[0.048]	[0.048]	[0.073]	[0.072]	[0.086]	[0.086]	[0.096]	[0.096]	[0.074]	[0.074]
Control Mean	68.826		0.073		0.786		0.670		0.575		0.416		0.285	
[std. dev]	[16.820]		[1.176]		[0.410]		[0.470]		[0.494]		[0.493]		[0.452]	
Observations	10,554	10,554	5,864	5,864	10,554	10,554	10,554	10,554	10,554	10,554	10,554	10,554	10,554	10,554
R-squared	0.015	0.016	0.014	0.014	0.005	0.007	0.007	0.007	0.009	0.009	0.012	0.012	0.016	0.017
F-stat	3.918	3.865	3.352	3.360	2.800	2.504	3.202	3.111	2.552	2.513	2.163	2.216	0.976	1.073
Prob > p	0.00833	0.00897	0.0183	0.0181	0.0386	0.0574	0.0223	0.0253	0.0539	0.0567	0.0902	0.0842	0.403	0.359

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. All regressions control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 6: Effects on All Courses Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	GPA		Credits Failed		Credits Earned	
	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	0.022	0.021	-0.012	-0.012	-0.039	-0.037
	[0.037]	[0.037]	[0.027]	[0.026]	[0.050]	[0.050]
Text Messaging	-0.019	-0.021	0.008	0.009	-0.042	-0.040
	[0.033]	[0.033]	[0.024]	[0.024]	[0.045]	[0.044]
Coaching	0.358**	0.345*	-0.143	-0.136	0.501*	0.429*
	[0.182]	[0.183]	[0.152]	[0.117]	[0.283]	[0.247]
Control Mean	2.380		0.346		3.684	
[std. dev]	[0.990]		[0.717]		[1.360]	
Observations	4,840	4,840	4,840	4,840	4,840	4,840
R-squared	0.039	0.042	0.010	0.012	0.094	0.114
F-stat	1.750	1.660	0.496	0.670	1.520	1.461
Prob > p	0.155	0.173	0.685	0.570	0.207	0.223

The dependent variable in each regression is indicated by the column headings. All regressions are run at the student level and control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Robust standard errors are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 7: Effects on All Course Grades Across Student Subgroups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Gender		Age		High School Grade		Mother Tongue		International Status		First Generation		First-Year Student		Residence Status	
	Female	Male	20 or Older	19 or Younger	Grade Above Campus Median	Grade Below Campus Median	English	Not English	International	Domestic	First Gen.	Non-First Gen	First-Year	Non-First Year	Lives on Residence	Does not live on Residence
Online Only	-0.083	1.120	-1.518	0.804	-0.618	0.664	1.245*	-0.185	0.194	0.638	0.720	0.515	0.793	-1.203	0.158	0.444
	[0.587]	[0.757]	[1.357]	[0.499]	[0.659]	[0.796]	[0.660]	[0.665]	[0.770]	[0.594]	[1.045]	[0.560]	[0.504]	[1.201]	[0.726]	[0.601]
Text Messaging	-0.493	0.339	-0.366	-0.123	-0.518	-0.235	0.150	-0.364	-0.756	0.333	0.606	-0.380	-0.013	-0.762	-0.442	-0.064
	[0.539]	[0.668]	[1.177]	[0.451]	[0.614]	[0.687]	[0.613]	[0.584]	[0.688]	[0.534]	[0.867]	[0.512]	[0.462]	[1.032]	[0.693]	[0.525]
Coaching	3.176	6.728*	7.036	4.755**	2.242	4.829*	5.104	4.834***	5.665***	4.486	7.504***	3.273	3.061	8.579**	6.533***	4.726**
	[2.208]	[3.455]	[5.859]	[2.132]	[2.611]	[2.646]	[3.269]	[1.799]	[2.135]	[2.877]	[1.623]	[2.770]	[2.086]	[3.965]	[1.869]	[2.182]
Constant	69.804	68.070	67.208	69.186	72.492	63.942	68.642	69.121	68.998	68.998	68.038	69.470	69.506	66.192	71.748	67.650
[std. dev]	[14.958]	[17.264]	[17.952]	[15.836]	[14.564]	[16.617]	[16.119]	[16.203]	[16.524]	[15.766]	[15.977]	[16.170]	[15.733]	[17.744]	[14.730]	[16.608]
Observations	16,016	13,575	4,500	25,091	11,441	11,090	13,293	16,298	12,321	17,270	6,631	20,702	23,539	6,052	9,373	20,218
R-squared	0.017	0.027	0.018	0.024	0.031	0.011	0.027	0.019	0.021	0.026	0.020	0.023	0.013	0.021	0.020	0.014
F-stat	1.113	1.863	1.016	2.858	0.702	1.519	2.087	2.786	3.234	1.092	7.430	1.324	1.701	2.235	4.677	1.801
Prob > p	0.343	0.134	0.385	0.0357	0.551	0.208	0.0999	0.0394	0.0215	0.351	6.29e-05	0.265	0.165	0.0826	0.00295	0.145

The dependent variable in each regression is course grades. The unit of observation is a student-course. All regressions control for campus fixed effects, first-year status, age in first year, and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 8: Effects on All Grades at the St. George Campus of U of T

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Grades		Grade Relative to Class Average		Grade Above 60		Grade Above 65		Grade Above 70		Grade Above 75		Grade Above 80	
	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	-0.386	-0.387	-0.069	-0.069	-0.000	-0.001	0.007	0.007	0.002	0.002	0.004	0.004	-0.014	-0.013
	[0.645]	[0.645]	[0.066]	[0.065]	[0.013]	[0.013]	[0.016]	[0.016]	[0.017]	[0.017]	[0.018]	[0.018]	[0.017]	[0.017]
Text Messaging	-0.362	-0.362	-0.017	-0.019	-0.001	-0.001	0.000	0.000	-0.006	-0.005	-0.003	-0.003	-0.007	-0.007
	[0.554]	[0.554]	[0.054]	[0.054]	[0.011]	[0.011]	[0.014]	[0.014]	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]
Control Mean	70.694		0.081		0.811		0.702		0.613		0.447		0.325	
[std. dev]	[15.439]		[1.140]		[0.392]		[0.457]		[0.487]		[0.497]		[0.468]	
Observations	15,846	15,846	4,510	4,510	15,846	15,846	15,846	15,846	15,846	15,846	15,846	15,846	15,846	15,846
R-squared	0.009	0.009	0.009	0.011	0.008	0.009	0.007	0.008	0.007	0.007	0.004	0.005	0.002	0.003
F-stat	0.273	0.274	0.563	0.571	0.00402	0.00395	0.137	0.135	0.125	0.122	0.0924	0.101	0.328	0.313
Prob > p	0.761	0.761	0.569	0.565	0.996	0.996	0.872	0.874	0.882	0.885	0.912	0.904	0.720	0.732

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. All regressions control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 9: Effects on All Grades at the Scarborough Campus of U of T

|	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Grades		Grade Relative to Class Average		Grade Above 60		Grade Above 65		Grade Above 70		Grade Above 75		Grade Above 80	
	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	0.950	0.936	0.074	0.077	0.019	0.019	0.025	0.025	0.012	0.012	0.001	0.002	0.008	0.009
	[0.957]	[0.959]	[0.072]	[0.072]	[0.021]	[0.021]	[0.025]	[0.025]	[0.027]	[0.027]	[0.028]	[0.028]	[0.025]	[0.025]
Text Messaging	-0.957	-1.008	-0.065	-0.063	-0.023	-0.024	-0.029	-0.030	-0.043*	-0.043*	-0.038	-0.037	-0.029	-0.029
	[0.912]	[0.913]	[0.068]	[0.068]	[0.019]	[0.019]	[0.024]	[0.024]	[0.025]	[0.025]	[0.025]	[0.025]	[0.023]	[0.023]
Control Mean	68.894		0.162		0.785		0.665		0.571		0.406		0.267	
[std. dev]	[15.988]		[1.178]		[0.411]		[0.472]		[0.495]		[0.491]		[0.442]	
Observations	7,162	7,162	7,089	7,089	7,162	7,162	7,162	7,162	7,162	7,162	7,162	7,162	7,162	7,162
R-squared	0.005	0.006	0.009	0.010	0.002	0.004	0.004	0.005	0.005	0.006	0.003	0.004	0.003	0.003
F-stat	2.008	2.088	1.813	1.862	2.080	2.209	2.226	2.308	2.311	2.326	1.401	1.368	1.347	1.295
Prob > p	0.135	0.124	0.164	0.156	0.125	0.110	0.109	0.100	0.0997	0.0983	0.247	0.255	0.260	0.274

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. All regressions control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 10: Effects on All Grades at the Mississauga Campus of U of T

|	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Grades		Grade Relative to Class Average		Grade Above 60		Grade Above 65		Grade Above 70		Grade Above 75		Grade Above 80	
	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls	no controls	with controls
Online Only	1.905*	1.958**	0.143**	0.146**	0.048**	0.050**	0.044*	0.045*	0.071***	0.072***	0.055**	0.055**	0.040**	0.040**
	[0.989]	[0.984]	[0.068]	[0.068]	[0.022]	[0.022]	[0.025]	[0.025]	[0.025]	[0.025]	[0.024]	[0.024]	[0.020]	[0.020]
Text Messaging	1.453	1.486	0.079	0.081	0.022	0.023	0.020	0.022	0.033	0.034	0.040*	0.040*	0.029	0.029
	[0.931]	[0.924]	[0.065]	[0.064]	[0.021]	[0.021]	[0.023]	[0.023]	[0.024]	[0.023]	[0.022]	[0.022]	[0.018]	[0.018]
Coaching	5.974***	5.930***	0.331*	0.330*	0.104**	0.101**	0.117*	0.114*	0.115*	0.113	0.120*	0.123*	0.076	0.080
	[2.013]	[2.024]	[0.180]	[0.179]	[0.051]	[0.051]	[0.063]	[0.063]	[0.070]	[0.069]	[0.067]	[0.066]	[0.059]	[0.059]
Control Mean	64.762		-0.009		0.721		0.604		0.463		0.286		0.167	
[std. dev]	[17.069]		[1.202]		[0.448]		[0.489]		[0.499]		[0.452]		[0.373]	
Observations	6,583	6,583	6,563	6,563	6,583	6,583	6,583	6,583	6,583	6,583	6,583	6,583	6,583	6,583
R-squared	0.016	0.025	0.022	0.028	0.007	0.014	0.008	0.013	0.010	0.016	0.015	0.019	0.010	0.014
F-stat	3.507	3.512	2.230	2.296	2.490	2.495	1.896	1.918	3.059	3.150	2.639	2.732	1.870	1.923
Prob > p	0.0149	0.0148	0.0830	0.0761	0.0589	0.0584	0.128	0.125	0.0274	0.0242	0.0482	0.0426	0.133	0.124

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. All regressions control for campus and first-year fixed effects. Additional control variables include age in first year and gender. Standard errors clustered at the student level are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

**Figure 1: Grade Distributions Across All Courses by Treatment Status**
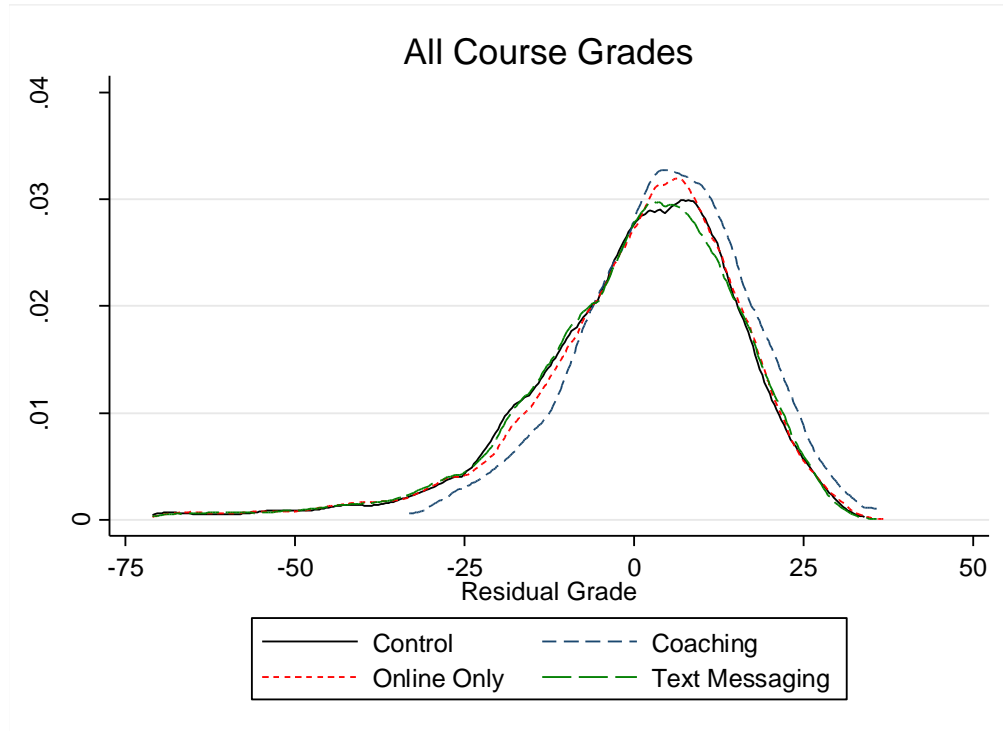
This figure presents residual grade distributions using grade outcomes from all (full-year, winter, and fall) courses. To construct the figure, we stack course grade outcomes for all students from all courses, regress course grades on campus and first-year fixed effects, and obtain the residuals from this regression. The figure shows the density of these residuals for each of the three treatment groups and the control group.

Figure 2: Grade Distributions Across Fall Courses by Treatment Status
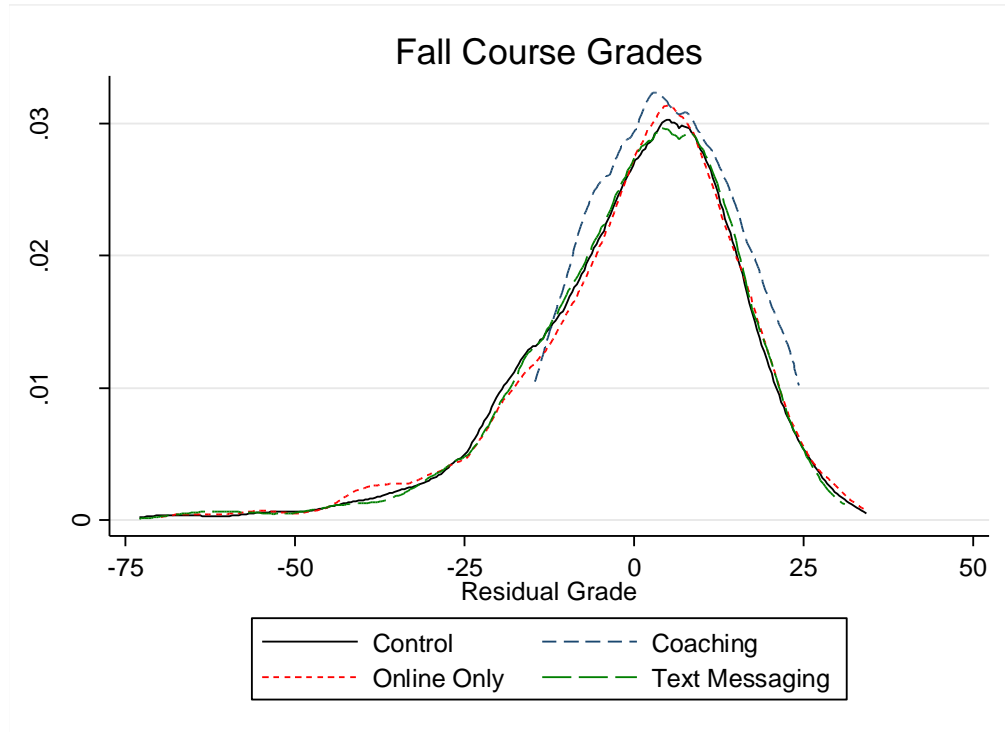
This figure presents residual grade distributions using grade outcomes from fall courses only. To construct the figure, we stack course grade outcomes for all students from all fall courses, regress course grades on campus and first-year fixed effects, and obtain the residuals from this regression. The figure shows the density of these residuals for each of the three treatment groups and the control group.

Figure 3: Grade Distributions Across Winter Courses by Treatment Status
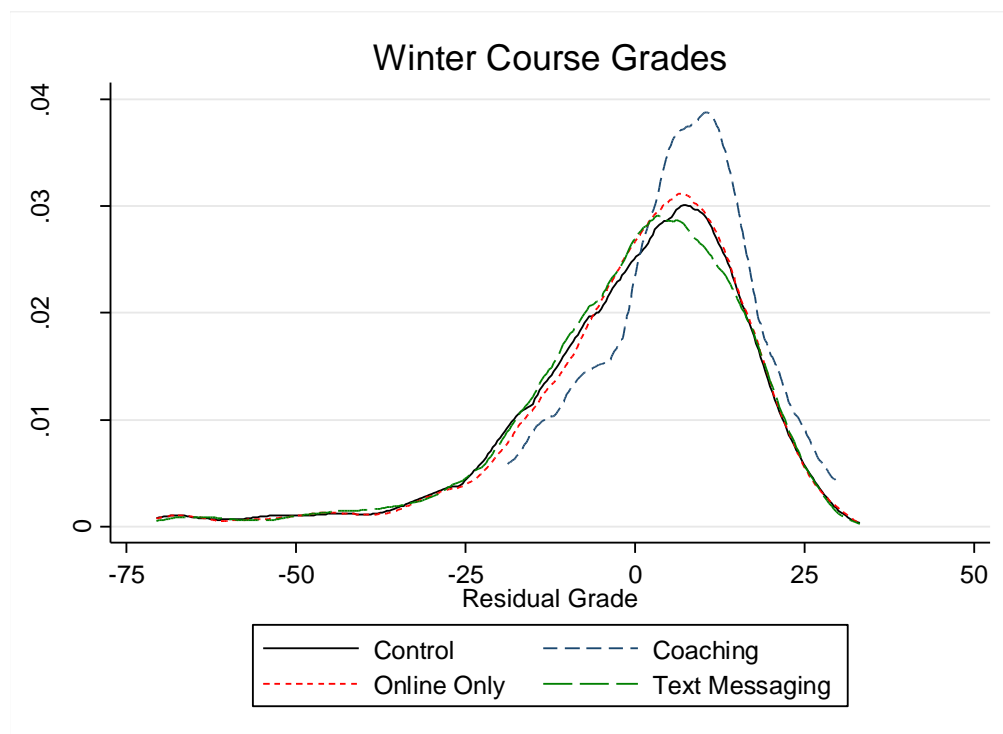
This figure presents residual grade distributions using grade outcomes from winter courses only. To construct the figure, we stack course grade outcomes for all students from all winter courses, regress course grades on campus and first-year fixed effects, and obtain the residuals from this regression. The figure shows the density of these residuals for each of the three treatment groups and the control group.

Figure 4: Contrasting Text Messaging and Coaching