Journal of Economic Behavior and Organization 171 (2020) 342–360

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Improving non-academic student outcomes using online and text-message coaching

Philip Oreopoulos[a,*], Uros Petronijevic[b], Christine Logel[c], Graham Beattie[d]

[a] Department of Economics, University of Toronto, Canada
[b] Department of Economics, York University, Canada
[c] Social Development Studies, University of Waterloo, Canada
[d] Loyola Marymount University, Canada

ARTICLE INFO

Article history:
Received 25 August 2018
Revised 20 December 2019
Accepted 10 January 2020
Available online 18 February 2020

Keywords:
Behavioral economics of education
Non-cognitive
Non-academic outcomes
Text message nudge
College coaching

ABSTRACT

We design and experimentally evaluate two low-cost, scalable interventions – an online preparatory module to help students reflect on and overcome barriers and a text-message coaching program – in a sample of over 3000 undergraduate students at a large Canadian university. Supplementing administrative data on academic outcomes with a unique follow-up survey on student well-being, we estimate positive program effects on non-academic outcomes such as feelings of satisfaction and belonging, despite estimating null effects on course grades and credit accumulation. Given the low costs associated with administering these programs, our results suggest that the positive impacts on student well-being may warrant program expansion even in the absence of impacts on academic outcomes.

1. Introduction

Individuals care deeply about how their actions are perceived by others, whether around home, at work, or at school (Akerlof and Kranton, 2000, 2005). The sense of belonging students experience at school is especially important, as it helps determine the effort they exert, the programs they select, the decision to pursue further education, and their overall human capital development (Coleman, 1961; Akerlof and Kranton, 2002). Yet in the current higher education landscape of limited skill development (Arum and Roksa, 2011), little study effort (Babcock and Marks, 2011), and low completion rates (Shapiro et al., 2019), much of the public policy agenda is based on designing student support programs that focus exclusively on fostering academic outcomes, such as grades and retention (Scrivener and Weiss, 2013; Bettinger and Baker, 2014; Castleman and Meyer, 2016; Evans et al., 2017; Clotfelter et al., 2018; Oreopoulos and Petronijevic, 2018; Oreopoulos et al., 2020). Comparatively less attention is devoted to how colleges can support and improve important non-academic outcomes, such as life satisfaction, engagement, and the extent to which students feel like they belong at and are supported by their institution.

In this paper, we experimentally evaluate two light-touch behavioral interventions designed to help college students improve both academic and non-academic outcomes, including overall satisfaction, a sense of belonging, confidence, and depression. A focus on well-being outcomes is particularly important in the current setting in higher education, as colleges

* Corresponding author.
E-mail addresses: philip.oreopoulos@utoronto.ca (P. Oreopoulos), upetroni@yorku.ca (U. Petronijevic), clogel@uwaterloo.ca (C. Logel), graham.beattie@lmu.edu (G. Beattie).

https://doi.org/10.1016/j.jebo.2020.01.009
0167-2681/

nation-wide face what some commentators in the press refer to as the 'mental health crisis' (David, 2019). Despite the importance of feeling like one belongs, many students today feel out of place: a recent national survey of American college students found that 43% have felt very lonely in the previous 30 days and 42% have felt overwhelming anxiety (American College Health Association, 2018). Students experiencing social isolation and a lack of belonging are more likely to experience mental health problems (Hefner and Eisenberg, 2009), which contribute to negative academic outcomes, including low grades and higher dropout rates (Eisenberg et al., 2009). Even among students with similar grades and persistence, those who endure greater levels of stress and depression may suffer lower levels of immediate (and long-term) utility. Unsurprisingly, as mental health awareness and focus on student experience spreads, student well-being is becoming an increasingly important concern for college administrators (Staglin, 2019).

There are many reasons students may feel discouraged and dissociated from their institution. Some students may find the pace of course progression overwhelming (Hofmann and Muhlenweg, 2018) and would benefit from learning better study or time-management skills, while others have effective study skills and would benefit more from initiatives that address low motivation. Other students have both skills and motivation but are hampered by stress in their personal lives, while students from historically underrepresented groups may underperform because of the pressure from contending with negative stereotypes (see Steele, 1997). All students may benefit from interventions that address these underlying issues, even if such interventions improve student mental health and well-being without affecting academic performance.

We address this underlying heterogeneity among students through two light-touch behavioral interventions, both designed to improve students' academic and non-academic outcomes. The first intervention incorporates a novel design that lets students personalize their experience according to their academic and social needs through a psychologically informed online module we named 'Choose Your Own Challenge,' or CYOC.[1] The module teaches students helpful academic behaviors while guiding them to reflect on, and then overcome, behavioral and psychological barriers to implementing those behaviors. The first part of the module presents students with six broad factors critical to academic success, with subsequent sections elaborating on each factor and taking students through tasks that draw on psychological research on attitude and behavior change. The second part of the module presents students with eight institutional barriers to success related to academic success factors, the implications of being part of a negatively stereotyped group, and the consequences of experiencing significant life challenges. Students are invited to choose the two barriers most relevant to future students like them, identify and write about a reason why students might struggle with this problem, and identify and write about a potential solution.[2]

Students in the second intervention complete both parts of the CYOC module and are additionally assigned to a text-message coaching program (Castleman and Meyer, 2016; Oreopoulos and Petronijevic, 2018) in which they are mentored by an upper-year undergraduate student coach who offers advice and consultation about students' specific challenges via text message throughout the academic year. Students in the control group are given a personality test measuring their relative ranking on each of the Big Five personality traits. We experimentally evaluate both treatments in a sample of over 3000 undergraduate students at the University of Toronto (UofT) during the 2016–2017 academic year and supplement administrative grades data with a unique (and mandatory) follow-up survey designed to capture important non-academic outcomes, such as life satisfaction, feelings of belonging, confidence, and depression.

Our results indicate that, while neither treatment is effective at improving student grades or credit accumulation, students' satisfaction with university life and sense of belonging in university improves. Specifically, we aggregate student responses to the follow-up survey into two main indices. The 'core well-being' index includes measures of life satisfaction, belonging, confidence, and depression and the 'success strategies' index includes measures of study strategies and help-seeking, such as time management and frequency of meeting with instructors. We estimate that, on average, the enhanced text-message coaching treatment improves the core well-being index by 4% of a standard deviation and the success strategies index by 5% of a standard deviation. In a more disaggregated analysis, we show that the effect on well-being is mainly driven by treated students feeling a greater sense of belonging at UofT and a greater sense of support from the institution.[3] Using student responses to direct questions about the coaching program, we also show that most treated students report feeling supported by their coaches, appreciating the messages they receive from their coaches, and having a better experience at UofT because of their coaches.[4]

[1] Interactive online modules are an increasingly common mode of delivering intervention materials, used, for example, in research on adaptive mindsets (Yeager and Walton, 2011; Walton, 2014; Bettinger et al., 2017), goal-setting interventions (Dobronyi et al., 2019; Clark et al., 2017), and timely information provision to graduating high school students (Oreopoulos and Ford, 2016; Hastings et al., 2018).

[2] To underscore the heterogeneity underlying students' challenges to academic success and overall well-being in college, Table 1 (which we discuss in greater detail in Section 2) summarizes the percentage of students who selected each of the potential barriers to success during the CYOC module. Poor time management and bad study strategies are the most popular barriers, accumulating 16.4% and 15.4% of votes, respectively. A lack of belonging and dissociation of one's identity from the university environment are also pervasive problems, however, as more than 30% of student votes identified as key challenges feeling socially isolated, unmotivated, or wondering whether other students are smarter than them and doubting whether they have what it takes to do well.

[3] In Section 4, we show that treatment effects on core well-being outcomes replicate at a satellite campus of UofT, where treated students participated in a one-way text-message coaching program. Treatment effects on the success strategies outcomes do not replicate, however, so we do not emphasize this result throughout the paper.

[4] The effects of the online CYOC module on its own trend in a positive direction on both indices, at approximately 3% of a standard deviation, but are not statistically differentiable from zero or from the effects of the enhanced text-message coaching treatment. Given that our design does not include a

To our knowledge, our paper is the first to provide evidence of a brief behavioral intervention influencing non-academic outcomes in an education-based context.[5] Our findings suggest that such interventions can improve student experiences in college, even without causing discernable improvement in course grades. Bolstering students' sense of belonging and the extent to which they feel supported by their institutions is important for several reasons. First, academic outcomes such as grades contribute only partially to students' immediate utility. Akerlof and Kranton (2002) formalize ideas from vast literatures in sociology and education into an economic model in which a student utility depends on both the return to accumulated human capital (proxied by grades) and the extent to which their sense of identity differs from the identity of their peers or the goals and values of their institutions. Conditional on grades and the expected future return to human capital, students who feel a lower sense of belonging and support will endure lower utility while in college.

Second, and related, students derive consumption value from attending college and do not choose majors or even schools based solely on the expected financial returns (Arcidiacono, 2004; Alstadsaeter, 2011). The experiences students have while in school contribute to their overall well-being and influence their educational decisions independently of human capital opportunities.[6] Institutions appear aware of students' desires for non-academic experiences, as Jacob et al. (2018) show that market pressure causes colleges to increase non-academic spending in response to student preferences for consumption amenities. At the same time, some studies suggest that half of all university students struggle with some form of mental illness but that fewer than half of those who struggle receive treatment (Zivin et al., 2009). Students who wish to seek treatment are often unaware of on-campus resources or face long wait times before receiving help (Reilly, 2018). Our results suggest that offering light-touch, large-scale student support programs could potentially allow schools to raise awareness of on-campus resources, improve student experience and, as result, better compete for student enrollment.[7]

Our interventions likely improve students' non-academic outcomes through two related mechanisms. First, treated students learn that they are not alone in facing doubts about belonging in college, as both the online module and subsequent text-message conversations with coaches clarify that many students initially face challenges to doing well in college and fitting in. Providing students with a more accurate perspective on the common challenges faced by all students and how these challenges are usually overcome with positive action is one of the key channels through which psychology interventions designed to improve social belonging operate (Walton and Cohen, 2011; Yeager et al., 2016a; Broda et al., 2018).

Second, on-campus services for helping students with belonging concerns, such as academic counseling or mental health centers, require students to initiate and maintain contact but many students are either not aware of these services and do not follow through with plans to use them. Our text message coaching program does not require students to initiate contact, instead allowing them to receive proactive support on a regular basis without having to ask for it. In this regard, the intervention operates similarly to other light-touch interventions that automatically provide students with timely information that causes improvement in outcomes that require immediate, concrete action, such as enrolling in college after being admitted, renewing financial aid applications, or choosing a selective college (Castleman and Page, 2015, 2016; Dynarski et al., 2018).

Our intervention demonstrates the potential for applying behavioral insights to help students feel more engaged and supported. In our case, the personalized nature of the online module and subsequent text-message conversations is likely important, as both interventions are flexibly designed to address the many underlying reasons students potentially face belonging concerns. More generally, new technologies like text messaging offer new means to reach students conveniently and quickly, and more detailed data can be used to predict which students are most likely struggle, allowing colleges to efficiently target the help they offer.

The remainder of this paper is organized as follows. The next section provides a detailed description of our interventions and their implementation. Section 3 describes the experimental data and outlines our empirical strategy for estimating the treatment effects. Section 4 presents the results, while Section 5 provides concluding remarks.

2. Description of the intervention

Throughout the 2016–2017 academic year, we ran an experiment at the main campus of the University of Toronto (St. George). We partnered with all first-year introductory economics instructors to make completion of our online 'warm-up' exercise worth 2% of students' final course grades. Students had to complete the exercise within the first two weeks of the fall semester to receive course credit. The type of exercise each student completed depended on whether he or she was randomly sorted to one of two treatment groups or the control group. We then administered a follow-up survey to all students in the final two weeks of the fall semester, approximately 12 weeks after the intervention exercise. The survey solicited students' feelings about non-grades outcomes, such as life satisfaction, feelings of support and belonging at the

condition testing the text-message coaching treatment on its own, we cannot determine whether treatment effects are additive, complementary, or entirely driven by the coaching treatment.

[5] For reviews on the literature of behavioral interventions in education, see Lavecchia et al. (2016) and Damgaard and Nielsen (2018).

[6] Pope and Pope (2009), for example, find that football and basketball team success increases the number of applications colleges receive and Alter and Reback (2014) find similar increases when an institution's quality-of-life reputation improves in college guidebooks.

[7] Improving well-being and sense of belonging may also improve long-term career outcomes, even without impacting short-term academic outcomes. Improvements to self-esteem and college satisfaction may lead to greater optimism and sense of control which could contribute to longer-term improvements to career. Further, friends and family members who observe a student have a positive university experience may be more likely to pursue higher education.

university, and self-reported study habits. Completion of the follow-up survey was worth 1% of students' final grade in their economics courses.

All students began the online warm-up exercise by creating an account and completing the same short introductory survey, in which they responded to background questions about their parents' education, their own expected educational attainment, first-year and international status, work and study plans, and tendencies to procrastinate or become distracted. After completion of the initial survey, students were randomly sorted to either a treatment group or the control group. Students sorted to the CYOC treatment group completed an online module. Students sorted to the text-message coaching treatment group also completed the online module but were additionally offered the opportunity to provide a cell phone number and participate in a text-message coaching program. Students assigned to the control group were given a personality test. The following sections describe the treatment and control modules, as well as the follow-up survey, in more detail.

2.1. Treatment one: choose your own challenge (CYOC) online module

We conceptualized the CYOC module based on the premise that students succeed and maintain well-being at least in part because of effective academic behaviors and adaptive perspectives. Teaching students effective academic behaviors is necessary but not sufficient for their success, as people can know effective behaviors but fail to follow through with them in nearly every domain. Students also face a diverse array of barriers to implementing and following through with effective behaviors. These barriers include low motivation, personal life stress, and common identity threat concerns (i.e., concerns that one's social identities, such as race, gender, or socioeconomic status, could be devalued in a given context). Moreover, taking adaptive perspectives requires not just learning what those perspectives might be, but having the opportunity to reflect on them and incorporate them into one's way of thinking about the world.

As such, the CYOC intervention aimed to teach students effective behaviors and perspectives, increase their likelihood of following through on the behaviors and taking on the perspectives as their own, and address the diverse array of barriers to success, all while being cost-effective and implementable to large numbers of students. Consisting of two parts, the entire CYOC online module was designed to take 60 min to complete. An information page emphasized that the purpose of the exercise was to allow UofT and the economics department to learn more about students' perceptions of the transition to university, with the intent of later using this information to create helpful resources for future students. Stipulating that the information students provided would be used to help future cohorts follows the format of most studies on the belonging mindset (see, for example, Walton et al., 2015). To underscore this framing, most segments included items asking students about the degree to which most students already understand each concept, or which one of a set of related statements they think is most relevant to most students.

In part one, students were asked to think about their own future and education, which we explained would help UofT better understand how students form strategies for achieving their goals.[8] An initial page listed six broad strategies critical to academic success: studying enough, studying effectively, getting help when you do not understand, keeping up and going to class, staying motivated, and being patient and taking a long-term perspective. Later subsections elaborated on each strategy, using elements designed to educate participants about the factor, change their attitude towards it, and then improve the intention-behavior link. For example, the subsection on 'studying enough' started with evidence of its efficacy (e.g., noting the strong relationship between studying and doing well) and provided detailed examples (e.g., finding at least 20 h a week; treating it as a full-time job). Next, it used an 'implementation intentions' writing task, shown to increase the likelihood of behavior change (e.g., Gollwitzer and Schaal, 1998), in which students created a plan for studying enough and committed to it. Part one of the CYOC module also incorporated visualization techniques (e.g., "Let your mind imagine details of your environment" and "Where are you sitting? What does the desk look like?").

Part two of the CYOC module was designed to address either barriers to carrying out the behaviors and perspective changes from part one or barriers that can remain even when those behaviors and perspectives are adopted (e.g., low motivation, personal life stress, and psychological threats experienced by students whose social identities are underrepresented or negatively stereotyped academically). It used a 'self-persuasion' design, in which students meaningfully engage to develop the material rather than passively reading through pre-determined content (see Canning and Harackiewicz, 2015). To ensure that due consideration was given to each component of the module, we placed minimum word-count and time restrictions on several pages of the exercise.

The activity was framed in the context of attributing academic struggles to changeable situation factors rather than to unchangeable personal factors. To connect it to students' identities while keeping it focused on helping future students, the introductory page described how UofT only accepts students whose records show that they have the motivation, background knowledge, and skills to succeed, so each student is capable of doing well academically. It then presented a set of situations that could interfere with academic success, ostensibly for the purpose of getting students' advice about which problems are most common and how to solve them.

To prompt participants to connect this task to their own social identities, the instructions described how UofT accepts more than 15,000 students each year, so it is highly likely that the following year there will be at least some students with

[8] There is a growing literature devoted to understanding whether encouraging students to focus on their goals can improve outcomes. See, for example, Dobronyi et al. (2019) and Clark et al. (2017).

Table 1
Students' barriers to success.

Barrier to success	Percentage of student votes
Trying to memorize course materials or focusing mostly on answers to practice questions and past tests instead of using effective study strategies.	15.4
Feeling out of place, lonely, or socially isolated.	10.3
Not devoting enough time to keeping up.	16.4
Waiting too long to seek out help.	11.1
Feeling unmotivated to devote time and energy toward doing well in university.	10.1
Wondering if they just don't have what it takes to do well academically. Feeling like other people are smarter than them and they can't compete.	12.5
Dealing with a great deal of personal stress and challenges.	12.1
Needing to build up more skills to get the most out of university-level material, such as becoming more fluent in English.	12.1

Notes: This table shows the percentage of student votes received by each of the eight barriers to success that were presented during the choose your own challenge treatment module.

the same ethnic, religious, national background, age, and many of the same strengths and struggles as the participant. The module then presented a series of situational barriers to success, including those that could be related to social identity threat (e.g., "feeling that maybe 'people like them' are not especially welcome at UofT" and "Wondering if they just don't have what it takes to do well academically at UofT"), significant life challenges (e.g., "dealing with a great deal of personal stress and challenges along with classes"), and the academic success factors identified in part one (e.g., "Feeling unmotivated to devote time and energy to doing well in university" and "waiting too long to seek out help when class concepts are unclear"). Participants were asked to choose the two most important barriers for 'people like them.'

Table 1 shows the most and least common barriers to success as voted by students in our sample. Note that no one barrier stands out—rather, students are heterogeneous in their concerns about the challenges to academic success and overall well-being. Poor time management and bad study strategies are the most common barriers, accumulating 16.4% and 15.4% of votes, respectively. A lack of belonging and dissociation of one's identity from the university environment are also pervasive problems, however, as more than 30% of student votes identified feeling socially isolated, unmotivated, or wondering whether other students are smarter than them as key challenges.

The survey program routed each student to content specific to the two barriers they chose. For each of those two barriers, students were provided with four potential reasons for that barrier and asked to choose one and write about why students might struggle with this problem. The four reasons focused on changeable situational factors and some included a subtle message to change students' attributions. For example, for the barrier of "feeling that maybe people like them were not especially welcome at UofT", the four reasons were designed to convey a message that belonging concerns are common (e.g., "Thinking they are the only ones wondering if they belong" and "Seeing other students and thinking that they seem to be completely comfortable at UofT"). Participants were then given a list of four solutions and asked to choose one and elaborate on it. In the above example, solutions include "giving it time – realizing that, in time, most students come to feel that they do belong at UofT." Psychological interventions that convey to students experiencing identity threat that belonging concerns are common and pass with time have been shown to reduce belonging uncertainty and improve academic retention and success (see Walton and Cohen, 2011).

Upon completion of the entire online module, students were emailed a printable poster of the tips for success presented in part one. Full documentation for the online exercise is available in Appendix A.

2.2. Treatment two: online exercise with follow-up text-message coaching

A random subset of the students who received the CYOC treatment was offered the opportunity to participate in a text-message coaching program. These students also completed the CYOC module described above but, upon completion, were asked to provide their phone numbers to participate in a text-message coaching program. Branded *You@UofT*,[9] the coaching program was active throughout both the fall semester of 2016 and the winter semester of 2017. The experiment featured a total of 10 coaches, each being assigned between 70 and 185 students. The coaching team consisted of 8 senior undergraduate students and two of the paper's coauthors, Oreopoulos and Petronijevic, making a team of 10 coaches. In the description of the data below, we provide details on the precise number of students assigned to each coach.

By assigning students to individual coaches, we attempted to combine the most promising features of the text-messaging and in-person coaching treatments that are evaluated in Oreopoulos and Petronijevic (2018). In that paper, we evaluated a mass two-way text-messaging campaign that facilitated communication with a large sample of students at low cost but was ineffective at improving outcomes. Students who participated in that text-messaging campaign were not assigned to individual coaches; instead, they were sent mass texts every week in which we invited them to share concerns and ask for help. In contrast, an in-person coaching treatment did improve academic outcomes, likely because coaches could proactively

[9] We chose the name to emphasize that the program would help coach students toward their individual definitions of success.

initiate discussions with students about challenges and establish relationships based on trust. The text-message coaching treatment in this paper is designed to continue to reach a large sample of students at low cost while integrating some of the features that made the in-person coaching program successful.

To that end, eight undergraduate students with previous student support experience were recruited to act as coaches.[10] Based on our results in Oreopoulos and Petronijevic (2018), coaches were trained to message their students regularly and to gently encourage those students to discuss the challenges they faced navigating through university. A web-based coaching platform we designed provided coaches with the ability to make notes about each individual student, allowing them to easily recall recently discussed topics with each student and follow-up regularly about specific issues. Coaches could also program batch messages – messages sent to multiple recipients in separate messaging chains – to be sent at specific times of the day, further specifying subgroups of students to which each batch would be sent. Students could be differentiated based on international or domestic status, first-year or non-first year status, and incoming high-school grades. Coaches were also able to categorize students into three distinct categories (red, yellow, and green), which indicated the degree of help or attention the coach deemed each student required.

Students who were randomly sorted to the text-message coaching treatment were offered the opportunity to enroll and had to make an active opt-in decision if they wanted to participate. Approximately 90% of students chose to opt-in, with less than 3% later choosing to opt out. As mentioned, coaches were instructed to initiate communication with each of their students at least once a week (often twice a week), which they typically did using pre-programed batch messages designed to stimulate conversation. Coaches were also encouraged to follow-up with individual students on the specific issues they had recently discussed to make sure that students were effectively progressing.[11] Once contact was established, conversations evolved organically, with coaches usually trying to determine how students were progressing throughout university, both academically and emotionally.[12] Appendix B provides categorized summaries of the different types of message that coaches sent.

2.3. Control group: personality test

As in Oreopoulos and Petronijevic (2018), students who were assigned to the control group at both campuses were given a personality test measuring the Big Five personality traits. The test could be completed in 45–60 min, and students were emailed a report describing their scores on each trait upon completion of the exercise. Beattie et al. (2018) use the data resulting from the personality test to explore non-academic predictors of performance in university. The appendix of that paper provides a full description of the personality test.

2.4. Follow-up survey

During the last two weeks of the fall semester, approximately 12 weeks after the initial treatments, all students were again required to log into their accounts and complete a 15-min follow-up survey for an additional 1% of their course grade.[13] We used the follow-up survey to solicit students' answers to questions about non-academic outcomes, such as life satisfaction, confidence, feelings of belonging, and study habits, such as study strategies and the frequency with which they sought help from instructors, tutors, and advisors. Most questions could be answered using dropdown menus of pre-populated choices, but we also did ask all students to freely write responses to open-ended questions about their biggest challenges to academic success, how the university could better help them, and how they could better help themselves. Students who were in the text message coaching treatment were also asked about their satisfaction with the coaching program and open-ended questions about how helpfulness of their coaches.

Beattie et al. (2019) use the data from this follow-up survey to descriptively explore behavioral and mental health differences between high- and low-achieving university students. The follow-up survey is documented in full in the appendix of that paper. In Section 3.2 below, we discuss how we categorize the variables from this survey to use them in our analysis of treatment effects on non-grade outcomes.

3. Data description and empirical strategy for estimating treatment effects

In this section, we describe the experimental data and sample, the follow-up survey, and the interactions that took place between students and coaches in the text-message treatment. Having described the environment, we then outline our empirical strategy for estimating treatment effects on both academic and non-academic outcomes.

[10] Most coaches participated through enrolling in a research opportunity course and received course credit rather than payment. Two coaches were not eligible to take the course and received a stipend of about $2000.

[11] Topics of conversation were also sometimes dictated by the events currently unfolding at the university. During midterm and exam periods, for example, coaches tended to guide conversations toward making study schedules and effective study strategies.

[12] Coaches did not act as tutors for course material; instead, they mentored students on effective study strategies, how to learn from past mistakes, and how to seek out campus resources when extra help was required.

[13] Both the initial survey and follow-up survey were graded based only on completion.

Table 2
Treatment randomization.

	Control	Online only	Text messaging
Number of students	1119	1154	1122
(i) Fraction of total	0.329	0.339	0.330
(ii) Intended fraction	0.33	0.33	0.33
p-Value of (i)–(ii)	0.96	0.22	0.95
Completed exercise	1110	1143	1113

3.1. Experimental randomization and sample description

We begin our data description by reporting the fractions of students sorted to treatment and control groups. Prior to the experiment, we intended to randomly sort one-third of the students to each group, based on the randomly-generated digits of their student numbers.[14] Table 2 shows that we successfully reached our randomization targets: the *p*-values for the tests of the difference between the realized and intended fraction are all well above conventional significance levels. Furthermore, the completion rates for both the personality test and the treatment modules are very high, each at 99%.

The full experimental sample consists of 3395 students, from which 1119 were assigned to the control group, 1154 were assigned to only complete the online module, and 1122 were assigned to the text-message coaching group. We could match 91% of the 3395 students to the university's administrative data on grades on background characteristics, leaving us with a final analysis sample of 3088 students for estimating treatment effects on course grades.[15]

Table 3 presents summary statistics and balancing tests. Across a rich set of student background variables – obtained from both our survey and the university's administrative data – almost no variables are statistically different, on average, between the control group and the two treatment groups. The only exceptions are that students assigned to complete only the online module are approximately 3.6 percentage points less likely to expect to earn an average grade of at least A- and to self-report checking their cell phones often, while students assigned to the text-message coaching treatment are 3.3 percentage points more likely to report expecting to work (for pay) more than 8 h per week in the upcoming academic year. We show below that our main results are robust to controlling for these and many other background variables.

In terms of sample composition, slightly more than half of our sample (55%) is female, and the average student starts university at 18.45 years of age and has an incoming high school average of 90.45%. Approximately 72% of our sample consists of first-year students, 45% speak English as their mother tongue, and 48% are international students. Approximately 43% of students live in residence and 29% of students are first generation (neither mother nor father attended university). Most students (68%) plan to earn more than a bachelor's degree and at least an A- average throughout their undergraduate studies, but the average student plans to study only 18.26 h per week, less than the amount of time one would typically devote to a part-time job.

3.2. Follow-up survey: sample description and outcomes

We use twenty-four questions from the follow-up survey in our main analysis of treatment effects.[16] To address the issue of multiple hypothesis testing directly and to draw general conclusions about treatment effects on non-grades outcomes, we group student answers to these questions into two main indices.[17] We construct each index using the method of Kling et al. (2007), standardizing students' answers to each question relative to the control group mean and standard deviation and then taking a simple (within-student) mean of the resulting standardized variables to construct the index. As shown in Table 4, we broadly define the two main indices as the 'core well-being' index and the 'success strategies' index, with the former being an aggregation of 14 variables and latter being an aggregation of 10 variables. The last column of Table 4 lists the questions that contribute to the construction of each index.

The questions contributing to the well-being index are motivated by extensive literatures in economics on happiness, life satisfaction, and well-being and in psychology on social belonging. The types of question we ask about satisfaction are standard and widespread in the well-being literature, with past research demonstrating, for example, relationships between such subjective measures of individuals' life satisfaction and unemployment rates (Di Tella et al., 2001), moments of the national income distribution (Senik, 2004), and one's relative rank in the earnings distribution in their local neighborhood

[14] Students provided their student numbers upon registering online for the experiment and had a strong incentive to provide the correct student number, as completion of the online exercise accounted for 2% of their final course grade.

[15] The matching rate is not significantly different across treatment and control groups. Regressing an indicator variable for whether a student cannot be matched to the grades data on treatment indicators results in coefficient estimates of 0.005 (se=0.012) and 0.019 (se=0.012) for the online only and the online with text message treatment groups, respectively.

[16] Among the students with available grades data, 465 (15%) did not complete the follow-up survey at the end of the fall semester. The attrition rate, however, is not significantly different across treatment and control groups. Regressing an indicator variable for whether a student completed the follow-up survey on treatment indicators results in coefficient estimates of 0.016 (se=0.015) and 0.024 (se=0.016) for the online only and the online with text message treatment groups, respectively.

[17] We perform explicit inference adjustments for multiple hypothesis testing in Section 4.

Table 3
Summary statistics and balancing tests.

Student characteristics	Control sample mean [standard deviation]	Online only difference [standard error]	Text message difference [standard error]	p-Value from F-test of no difference
Female	0.548	0.000	−0.003	0.987
	[0.498]	[0.022]	[0.022]	
Age in first year	18.453	0.057	0.015	0.624
	[1.388]	[0.061]	[0.062]	
High school average grade	90.448	−0.233	−0.292	0.340
	[4.032]	[0.207]	[0.213]	
First-year student	0.715	0.003	0.012	0.797
	[0.452]	[0.019]	[0.019]	
English mother tongue	0.446	−0.012	0.001	0.816
	[0.497]	[0.022]	[0.022]	
Lives on residence	0.431	−0.001	0.017	0.650
	[0.495]	[0.022]	[0.022]	
International student	0.478	0.002	0.012	0.828
	[0.500]	[0.021]	[0.021]	
First-generation student	0.239	0.006	0.006	0.925
	[0.426]	[0.018]	[0.018]	
Expects to earn more than BA	0.711	−0.015	0.009	0.463
	[0.453]	[0.019]	[0.019]	
Expects at least A- average	0.682	−0.036*	−0.025	0.172
	[0.466]	[0.020]	[0.020]	
Checks cell phone often	0.471	−0.037*	−0.034	0.147
	[0.499]	[0.021]	[0.021]	
Expected study hours/week	18.265	−0.139	−0.317	0.820
	[12.095]	[0.500]	[0.505]	
Expected work hours/week > 8	0.298	−0.002	0.033*	0.133
	[0.457]	[0.019]	[0.020]	
Not discouraged by setbacks (5-point scale)	3.440	−0.048	−0.047	0.407
	[0.964]	[0.040]	[0.041]	
Finish what I start (5-point scale)	3.814	−0.021	0.007	0.695
	[0.837]	[0.035]	[0.035]	
Think about future (7-point scale)	2.393	−0.025	−0.040	0.722
	[1.214]	[0.051]	[0.050]	
Tend to cram for exams (7-point scale)	3.941	0.001	0.055	0.602
	[1.521]	[0.064]	[0.063]	
Transition has been challenging (7-point scale)	4.685	0.036	−0.029	0.628
	[1.615]	[0.068]	[0.068]	

Summary statistics and differences are calculated using the full sample of students. Robust standard errors are reported in brackets. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.

(Luttmer, 2005). In addition, ample evidence exists that such satisfaction measures are predictive of tangible (and easily measurable) economic outcomes such as local quality of life indices (constructed from local income distributions and amenities), adult earnings up to a decade later, and productivity in the workplace (Oswald and Wu, 2010; De Neve and Oswald, 2012; Oswald et al., 2015).

Likewise, the questions we ask about students' sense of belonging in university are motivated by similar questions used in social psychology studies on using light-touch interventions to improve students' perspectives on adversity associated with social belonging in school. Several studies have used such questions to measure students' sense of baseline and post-intervention belonging in high school and university settings (Walton and Cohen, 2011; Yeager et al., 2016b; Broda et al., 2018).

As mentioned, we use these questions in the follow-up survey to define two main indices—the core well-being index and the success strategies index—as our main outcome variables from the follow-up survey. We provide a further disaggregation of the core well-being index by breaking it apart into four sub-indices, reflecting (i) overall satisfaction with life and the university, (ii) feelings of belonging at and support by UofT, (iii) confidence to succeed at UofT, and (iv) overall depression or stress. The success strategies index is also disaggregated into two sub-indices, reflecting (i) study strategies and (ii) help-seeking behavior. In our empirical analysis below, we report treatment effects on the two main indices and each of the sub-indices.[18]

3.3. Text message coaching program and data

Throughout the duration of the text-message coaching program, every text message sent by a coach or student was stored, allowing us to assemble a large dataset of the text dialog between each student and his or her coach. To provide a

[18] Appendix C offers a completely disaggregated analysis of treatment effects on each of the twenty-four variables in Tables C1 and C2.

Table 4
Outcome variables from follow-up survey.

Main index	Sub index	Survey question
Core well-being	Satisfaction (1–7 scale)	All things considered, how satisfied are you with your life as a whole these days?
		All things considered, how satisfied are you with your experience at University of Toronto so far?
	Belonging and support (1–6 scale)	I feel like I belong here at UofT
		Being a student at UofT is an important part of how I see myself
		UofT wants me to be successful here
		I know where to go if I need academic advice
		UofT does its best to help support me
		Other students understand more than I do about how things work here at UofT[a]
		I often remind myself of my goals and motivations for being here at UofT
		The transition to the University of Toronto has, so far, been challenging[a]
	Confidence (1–7 scale)	How important is it to you that you do well at UofT?
		How confident do you feel that you have the ability to do well at University of Toronto?
	Depression (1–4 scale)	Since the beginning of the academic year, I have felt sad or depressed[a]
		Since the beginning of the academic year, I have felt very stressed[a]
Success strategies	Study strategies (1–6 scale)	I manage my time well
		I try to learn from my mistakes on past tests and assignments
		I write thoughts and ideas down when I study to test my understanding
		I get feedback from my writing assignments before handing them in
	Help seeking (numeric)	Last term, how often did you meet with an instructor outside of class?
		Last term, how often did you meet with an academic advisor?
		Last term, how often did you use a free academic tutor?
		Last term, how often did you meet with a paid tutor?
		Last term, how often did you attend a workshop?
		Last term, how often did you participate in an informal study group?

[a] Responses to these variables are re-coded so that more beneficial outcomes have higher scores.

Table 5
Text message summary statistics.

	(1) Mean	(2) Standard deviation	(3) Minimum	(4) Maximum	(5) Number of observations
Coaches					
Number of students per coach[a]	100.8	45.75	70	185	10
number of messages sent by coaches					
Fall semester non-batch messages	1520.5	816.02	479	2792	10
Fall semester batch messages	3121.7	2729.71	834	9355	10
Fall semester unique batch messages[b]	46.5	24.80	21	87	10
Winter semester non-batch messages	481.60	225.67	171	854	10
Winter semester batch messages	1733.7	1908.55	488	6859	10
Winter semester unique batch messages[b]	18.7	10.28	9	40	10
Students					
Fraction of treated students sending at least one message	0.62	0.49	0	1	1122
Number of messages sent by students					
Fall semester messages	17.29	28.33	0	385	983
Winter semester messages	6.21	12.72	0	140	983
Number of messages received by students					
Fall semester total messages	47.22	25.62	0	287	983
Fall semester batch messages	31.76	14.33	0	61	983
Winter semester total messages	22.53	14.02	0	94	983
Winter semester batch messages	17.63	10.93	0	39	983

[a] Multiplying the average number of students per coach by the number of coaches gives only 1008 students. Of the 1122 students who were assigned to the text-messaging treatment, 114 declined to participate and were thus not assigned a coach. An additional 25 were asked to be removed from the program, bringing the number of students who appear in the text-message data down to 983.
[b] We avoid double counting by counting each pre-programmed batch message that coaches sent to their students only once, despite such messages being received by many students of a given coach.

fuller description of the text-message coaching program, we present summary statistics from this dataset in Table 5, focusing our discussion on the assignment of students to coaches and the resulting interactions that took place.

The average coach was assigned 100 students; however, there was wide variation in the number of students assigned to each coach: senior undergraduate coaches received between 70 and 95 students, while Oreopoulos and Petronijevic were assigned 185 students each. The variation in the number of students per coach occurred because we achieved a higher opt-in rate (90%) into the text-messaging program than we expected prior to launching the experiment. To avoid breaking the workload agreements that we made with senior undergraduate coaches, we allowed each coach to choose whether he or she would accept more than 70 students. Five coaches stopped at approximately 70 students, three coaches accepted 95 students, and the remaining students were assigned to Oreopoulos and Petronijevic.

Prior to the time we stopped sorting students to all coaches (when approximately 700 students had already completed the online exercise), students who were assigned to the text-message coaching treatment were also randomly assigned to a coach. After we stopped directing students to the coaches who did not wish to accept more than 70 students (and later, to coaches who did not wish to accept more than 95 students), the sorting of students to coaches was no longer random. Coaches who continued accepting students were less averse to a larger workload while the students who were being sorted to these coaches completed the warmup exercise relatively late. Using similar data from a prior experiment, Beattie et al. (2018) show that these students tend to have a strong tendency to procrastinate and do not perform well in their courses. We address concerns over non-random sorting of students to coaches below by estimating specifications that control for three ‘coach group’ binary variables, each of which captures the sorting rule of students to coaches that applied during the (calendar) time when a student completed the online module.[19] Controlling for these binary variables creates conditional random assignment of students to coaches.

Among the 1122 students who were randomly offered the coaching treatment, 114 opted not to provide a phone number and an additional 25 students asked to be removed from the program during of the academic year, leaving us with a total of 983 students in our dataset of text message exchanges. Most students who were offered the coaching treatment (62%) sent a text message to their coach at least once. On average, students sent 17 text messages to their coaches in the fall semester and received a total 47 messages, with 32 of those messages being a batch message the coach sent to many students at once. Engagement fell in the winter semester, with students sending only 6 messages, on average, and receiving 23 from their coach, 18 of which were batch messages.

The average coach sent 1521 non-batch messages and 3122 batch messages to their students in the fall semester. If we do not double count batch messages that go out to multiple students at once, the average coach sent 47 unique batch messages in the fall semester. The decline in engagement in the winter semester is again reflected in the number of messages sent by coaches, as the average coach sent 482 non-batch messages, 1734 batch messages, and 19 unique batch messages.

In Appendix B, we provide more detail about the batch text messages students received from their coaches, presenting summary statistics on text characteristics and categorized examples of select batch messages.

3.4. Empirical strategy for estimating treatment effects

Having successfully randomized students across treatment and control groups, we estimate the effects of the online-only and text-messaging treatments with a comparison of mean outcomes in a simple regression framework. The main specification we estimate is given by

$$y_i = \beta_0 + \beta_1 Online_i + \beta_2 Text_i + \rho' X_i + u_i, \tag{1}$$

where the outcome of student *i* is regressed on an indicator for the student being assigned to the online (CYOC) treatment only, an indicator for the student being assigned to complete the online module and receive text-message coaching and, in some specifications, additional student-level control variables.

The main parameters of interest are β_1 and β_2, the estimated average effects of the online module alone and online module enhanced with the text-message coaching. These parameters represent intent-to-treat effects, as students are included in the treatment groups if they are offered the *opportunity* to work through the online module or provide a cell phone number (based on the randomly-generated digits of their student numbers) without necessarily having to complete the module or opt-in to texting messaging. Given that our completion and opt-in rates are quite high, these estimates are likely close to the average treatment effect.

Our main academic outcomes of interest are course grades, overall grade point average (GPA), the number of credits earned, and the number of credits failed. In terms of non-academic outcomes, we explore treatment effects on our core well-being and success strategies indices as well as the effects on each of their sub-indices (described above). When the outcome of interest is course grades, we stack all course grades and run a regression at the student-course level. In these cases, we cluster standard errors at the student level. The effects on all other outcomes are estimated with regressions at the student level and robust standard errors are reported.

4. Experimental results

In this section, we present the estimated effects of the CYOC condition and the enhanced text-messaging condition, followed by a series of robustness checks and an exploration of heterogeneous treatment effects.

4.1. Grades outcomes

In Table 6, we estimate treatment effects on course grades using student-course level regressions and clustering standard errors at the student level. In columns (1) and (2), we report the estimated treatment effects from regressions without and

[19] Students who completed during the initial time period were sorted to all coaches; students who completed during the second time period were sorted to Oreopoulos and Petronijevic and senior undergraduate coaches who were willing to accept between 70 and 95 students; and students who completed during the third time period were sorted only to Oreopoulos or Petronijevic.

Table 6
Treatment effects on stacked grades.

	(1) All course grades	(2)	(3) Fall course grades	(4)	(6) Winter course grades	(7)	(8) Economics grade	(9)
Online only	−0.012	0.261	−0.687	−0.531	−0.163	0.085	0.452	0.823
	[0.575]	[0.532]	[0.624]	[0.574]	[0.761]	[0.707]	[0.761]	[0.730]
Text messaging	0.169	0.240	−0.434	−0.410	−0.015	0.012	0.618	0.717
	[0.554]	[0.521]	[0.603]	[0.567]	[0.720]	[0.679]	[0.751]	[0.729]
Control mean	70.851		72.999		70.993		67.957	
[Standard deviation]	[15.881]		[14.120]		[16.526]		[15.995]	
Controls?	No	Yes	No	Yes	No	Yes	No	Yes
Observations	20,786	20,786	6687	6687	7367	7367	2584	2584

The dependent variable in each regression is indicated by the column headings. The unit of observation is a student-course. Additional control variables include student age, self-reported expected study hours per week, expected paid-work hours per week, expected average grade, tendency to finish what he or she starts, tendency to get discouraged by setbacks, tendency to study at the last minute, and indicator variables for first-year status, international student status, first-generation status, gender, English mother-tongue status, whether the student lives on residence, a self-reported desire to earn more than an undergraduate degree, and a self-reported tendency to check his or her cell phone often. Standard errors clustered at the student level are reported in brackets in columns (1) to (7). In columns (8) and (9), we report robust standard errors in brackets. *** indicates significance at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level.

Table 7
Treatment effects on GPA and credit accumulation.

	(1) All courses	(2)	(3)	(4) Fall courses	(5)	(6)	(7) Winter courses	(8)	(9)
	GPA	Credits failed	Credits earned	GPA	Credits failed	Credits earned	GPA	Credits failed	Credits earned
Online only	0.041	−0.007	−0.036	0.023	0.013	−0.015	0.043	−0.002	−0.011
	[0.040]	[0.037]	[0.064]	[0.043]	[0.011]	[0.025]	[0.046]	[0.015]	[0.028]
Text messaging	0.034	−0.039	−0.009	0.012	0.005	−0.006	0.007	−0.020	−0.025
	[0.041]	[0.035]	[0.063]	[0.044]	[0.010]	[0.025]	[0.047]	[0.014]	[0.028]
Control mean	2.546	0.376	4.142	2.779	0.060	1.151	2.664	0.110	1.228
[Standard deviation]	[0.979]	[0.862]	[1.461]	[0.995]	[0.222]	[0.562]	[1.067]	[0.359]	[0.631]
Observations	3075	3075	3075	2783	2783	2783	2807	2807	2807

The dependent variable in each regression is indicated by the column headings. All regressions are run at the student level and control for student age, self-reported expected study hours per week, expected paid-work hours per week, expected average grade, tendency to finish what he or she starts, tendency to get discouraged by setbacks, tendency to study at the last minute, and indicator variables for first-year status, international student status, first-generation status, gender, English mother-tongue status, whether the student lives on residence, a self-reported desire to earn more than an undergraduate degree, and a self-reported tendency to check his or her cell phone often. Robust standard errors are reported in brackets. *** indicates significance at the 1% level; ** indicates significance at the 5% level; and * indicates significance at the 10% level.

with additional control variables. Importantly, the set of control variables includes 'coach group' binary variables for the date the student completed the online module, which preserves random assignment of students to coaches.[20] Treatment effects for the online module and text-messaging program across all course grades are small and statistically insignificant in both cases. Columns (3) and (4) show estimated effects on fall semester (September 2016–December 2016) courses, while columns (6) and (7) show effects on winter semester (January 2017–April 2017) courses.[21] Treatment effects are again statistically and economically insignificant in both cases. In the last columns, (8) and (9), the dependent variable is students' first-year economics course grades, the class in which they completed the online module. Here, the treatment effects are larger than those estimated across all courses, but they remain small and insignificant.

In Table 7, we further present estimated treatment effects on student GPA and credit accumulation. Neither the online module nor the text-messaging treatment affected student GPA or credit accumulation, as we estimated null effects across all courses taken throughout the academic year, fall-semester courses, and winter-semester courses. Overall, the evidence demonstrates that the online-module and text-message coaching treatment are ineffective at improving student grades and credit accumulation, as we can rule out effect sizes of 1.3 percentage-points on overall grades (0.08 standard deviations), 0.12 points on overall GPA (0.12 standard deviations), and 0.12 credits on overall credits earned (0.08 standard deviations). These are fairly precise null effects, although given the low cost of the types of light-touch program evaluated in this paper,

[20] Additional control variables include student age, self-reported expected study hours per week, expected paid-work hours per week, expected average grade, tendency to finish what he or she starts, tendency to get discouraged by setbacks, tendency to study at the last minute, and indicator variables for first-year status, international student status, first-generation status, gender, English mother-tongue status, whether the student lives on residence, a self-reported desire to earn more than an undergraduate degree, and a self-reported tendency to check his or her cell phone often.

[21] Treatment effects in columns (1) and (2) were estimated using all available courses, which include those from the fall semester, those from the winter semester, and those that run across both semesters and conclude in the winter.

Table 8
Treatment effects on core well-being index.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
			Sub-indices							
	Core well-being index		Satisfaction index		Belonging and support index		Confidence index		Depression index	
Online only	0.018	0.028	0.047	0.061	0.012	0.018	0.047	0.066*	−0.016	−0.005
	[0.025]	[0.024]	[0.043]	[0.041]	[0.026]	[0.025]	[0.036]	[0.034]	[0.041]	[0.039]
Text messaging	0.046*	0.041*	0.039	0.034	0.058**	0.053**	0.036	0.034	0.012	0.007
	[0.025]	[0.024]	[0.043]	[0.042]	[0.026]	[0.025]	[0.037]	[0.035]	[0.042]	[0.040]
p-Value for test of treatment differences	0.275	0.577	0.859	0.517	0.0771	0.163	0.763	0.350	0.503	0.763
Controls?	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2623	2623	2623	2623	2623	2623	2623	2623	2623	2623

The dependent variable in each regression is indicated by the column headings. Control variables include student age, self-reported expected study hours per week, expected paid-work hours per week, expected average grade, tendency to finish what he or she starts, tendency to get discouraged by setbacks, tendency to study at the last minute, and indicator variables for first-year status, international student status, first-generation status, gender, English mother-tongue status, whether the student lives on residence, a self-reported desire to earn more than an undergraduate degree, and a self-reported tendency to check his or her cell phone often. Robust standard errors are reported in brackets. * indicates is significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

even the potential for these relatively small effects may warrant further investigation into whether such programs can be refined to improve academic outcomes.[22]

4.2. Non-grades outcomes

Having shown that neither treatment affected student grades, we now explore treatment effects on the indices created from our follow-up survey. Recall that the questions used to construct each index are reported in Table 4.

In Table 8, we report treatment effects on the core well-being index and its four sub-indices. The estimates in columns (1) and (2) show that students in the coaching treatment scored 4% of a standard deviation higher on the aggregate core well-being index. Breaking the index apart into four sub-indices, the treatment effect is driven by the effect on students' sense of belonging and support, as treated students score 5.5% of a standard deviation higher on the belonging and support index. Treatment effects on the three other sub-indices are also positive, although they are not statistically distinguishable from zero.

Although the estimated treatment effect of the online module on the core well-being index is not statistically different from zero, it also not statistically different from the effect of the text-message coaching treatment, as indicted by the *p*-values pertaining to the tests for treatment differences. Comparing the effects of the online module and the text-message coaching treatments on the sub-indices of the core well-being index indicates that, despite it not having an overall effect on the core well-being index, the point estimates for the effects of the online module are sometimes larger than the point estimates for the text-messaging treatment. Students who only completed the online module reported higher overall satisfaction with life and the university and indicated greater confidence to succeed in university. Taken together, we cannot rule out that the significant effect for the text-message coaching treatment is driven, at least in part, by students having completed the CYOC online module.

For completeness and to further investigate mechanisms, we report treatment effects on each of the 14 variables used to construct the core well-being index in Appendix Table C1. The estimates imply that the effects of the text-message coaching program operate by causing students to feel more like being a student at UofT is a large part of their identity, like UofT wants them to succeed, and like they know where to get advice at UofT. Treatment effects on these variables range between 9.4% and 12.5% of a standard deviation and are statistically significant. Students in the text-message coaching treatment also felt more supported by UofT and were less likely to report having a tough transition to the university, although these effects are not statistically significant at conventional levels. Overall, the evidence is consistent with the text-message coaching program causing students to form stronger ties with the university and feel like they are better equipped to navigate the challenges in their environment.[23]

We now turn to treatment effects on the success strategies index, which are reported in Table 9. Students who were in the text-message coaching treatment scored 5% of a standard deviation higher on the aggregate success strategies index.

[22] We note, however, that we have explored other one-way and two-way text-messaging programs in other work focused on helping students keep their long-run goals salient (Dobronyi et al., 2019), providing them with generic academic advice and motivational messages (Oreopoulos and Petronijevic, 2018), and helping them stay committed to their study time goals (Oreopoulos and Petronijevic, 2019). In each case, we found null effects on course grades, GPA, and credit accumulation, with levels of statistical precision similar to those reported here. None of our prior work, however, explores whether such light-touch support programs can be designed to improve student well-being by affecting non-academic outcomes, such as life satisfaction and feelings of belonging.

[23] As is the case for the main indices, despite the treatment effects for the online module not being statistically different from zero, they are also not statistically different from the effects of the coaching treatment, suggesting that the students in the coaching treatment may have benefited from completing the online module.

Table 9
Treatment effects on success strategies index.

	(1)	(2)	(3) Sub-indices	(4)	(5)	(6)
	Success strategies index		Strategies index		Help-seeking index	
Online only	0.029	0.034	0.032	0.041	0.027	0.030
	[0.023]	[0.022]	[0.033]	[0.031]	[0.027]	[0.026]
Text messaging	0.055**	0.051**	0.050	0.043	0.057**	0.057**
	[0.023]	[0.022]	[0.032]	[0.030]	[0.027]	[0.027]
p-Value for test of treatment differences	0.271	0.431	0.655	0.966	0.263	0.306
Controls?	No	Yes	No	Yes	No	Yes
Observations	2623	2623	2623	2623	2623	2623

The dependent variable in each regression is indicated by the column headings. Control variables include student age, self-reported expected study hours per week, expected paid-work hours per week, expected average grade, tendency to finish what he or she starts, tendency to get discouraged by setbacks, tendency to study at the last minute, and indicator variables for first-year status, international student status, first-generation status, gender, English mother-tongue status, whether the student lives on residence, a self-reported desire to earn more than an undergraduate degree, and a self-reported tendency to check his or her cell phone often. Robust standard errors are reported in brackets. * indicates is significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

Breaking the index apart into its two sub-indices in columns (3)–(6) shows that the effect is mainly driven by the help-seeking index, although the point estimates of the effects on the strategies index are approximately the same magnitude but estimated less precisely. Once again, we find that the effects of the text-message treatment and the online module are not statistically different from each other, despite the latter not having a statistically significant effect on student outcomes. The evidence is therefore consistent with at least part of the text-message coaching treatment effect being driven by students having completion of the CYOC module, although we cannot determine whether treatment effects are additive, complementary, or driven entirely by the text-message coaching.

In Appendix Table C2, we break apart the treatment effects on the success strategies index and its two sub-indices by estimating treatment effects on each of the 10 variables used to construct these indices. The effect of the online module combined with the text-message coaching program on the aggregate indices is driven mainly by students meeting with instructors and free tutors more often. Relative to the control group mean, the text-message coaching treatment increased the number of times students met with instructors by 16.5% and the number of times they met with free tutors by 17%.

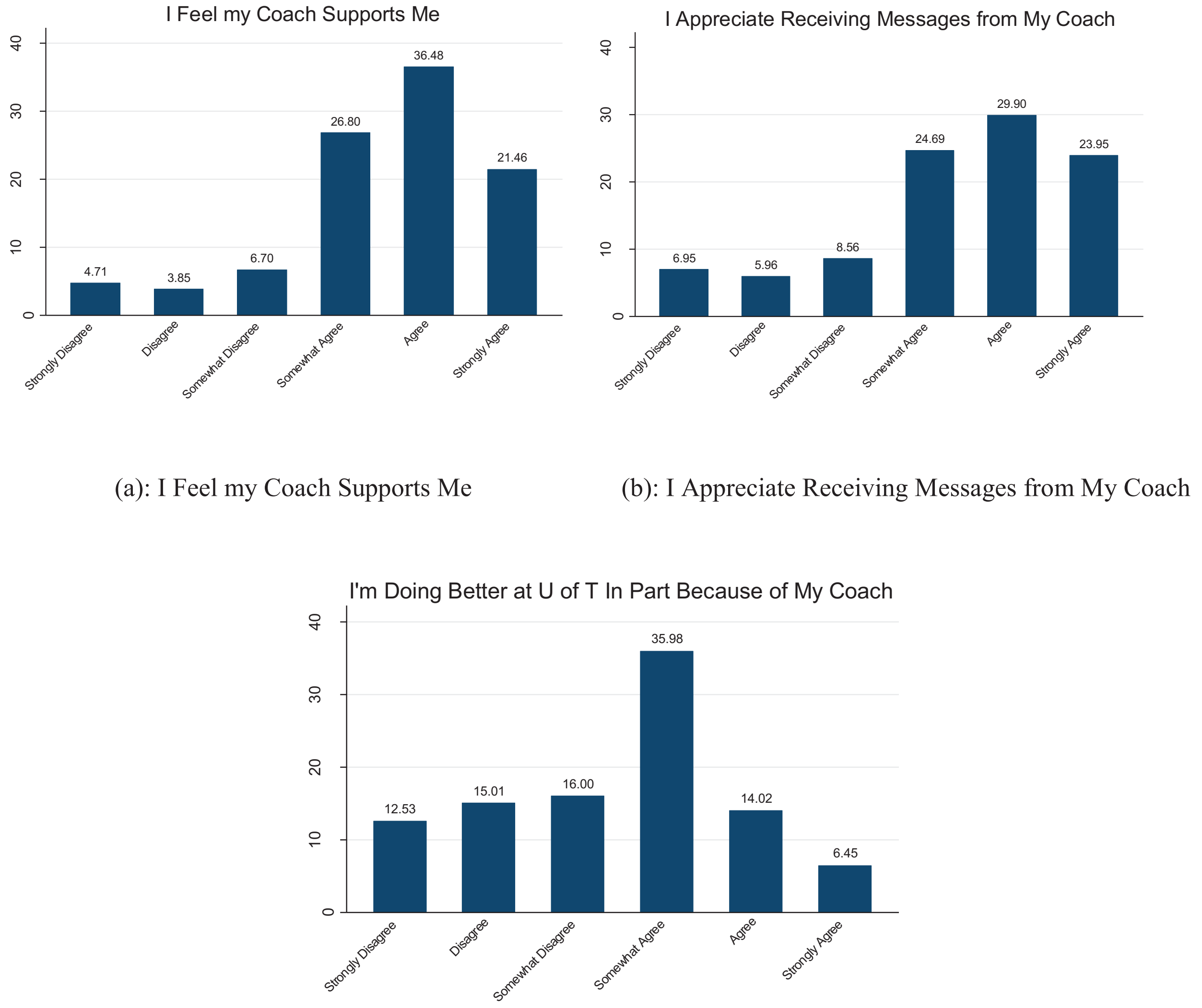
Given the evidence that the text-message program resulted in students meeting more often with tutors and instructors, it may appear somewhat surprising that we estimate no treatment effect on grades. There are two ways to reconcile these results. First, although the treatment increased the frequency of student meetings with instructors by 16.5%, students in the control group only met with an instructor 1.2 times, on average, during the fall semester. Therefore, despite visiting instructors more frequently, treated students still did not meet with instructors often. Second, although we do not report these estimates here, we used the follow-up survey to estimate treatment effects on students' (self-reported) independent study hours outside of the classroom. The text-message coaching treatment did not cause students to spend more time devoted to independent weekly study throughout the fall semester.[24] Because students in the text-messaging treatment did not study more on their own and visited instructors rarely, it is perhaps not surprising that we find no treatment effect on course grades.

The evidence presented thus far shows that the text-messaging treatment caused students to feel a greater sense of belonging and support at university. Students' answers to follow-up questions about the text-message coaching program demonstrate that most students enjoyed the program and felt that their experience in first semester was better at least in part because of the program. During the follow-up survey, treated students were asked to express the extent to which they agree with the following statements: "I feel my coach supports me;" "I appreciate receiving messages from my coach;" and "I am doing better at UofT in part because of my coach." Fig. 1 shows the percentage of respondents who selected each of the possible categories for each statement.

Nearly 60% of students agree or strongly agree that their coach supports them, while 85% of students at least somewhat agree with that statement. Similarly, 54% of students agree or strongly agree with the statement that they appreciate receiving message from their coaches, while 79% of students at least somewhat agree. Relative to the first two statements, student support for the statement that they are doing better at UofT because of their coaches is a little lower, as only 20% of students agree or strongly agree with this statement, but 56% of students at least somewhat agree. The more tepid responses to this question are understandable given that the coaching program had no effect on student grades and that many students likely interpreted the question as referring to doing better with respect to grades. It is clear, however, that nearly all students (80%) felt supported by their coaches and appreciated receiving text messages from them.[25]

[24] These results are available upon request.

[25] Coaches often interacted with students about serious anxiety or depression. In such cases, coaches promptly responded advice similar to that offered by the University Wellness Centre and referred some cases to professional support. Even students facing these more serious concerns expressed gratitude and appreciation for the support that coaches provided.



(a): I Feel my Coach Supports Me

(b): I Appreciate Receiving Messages from My Coach

(c): I am Doing Better at UofT In Part Because of My Coach

Fig. 1. Student feelings about the text-message coaching program. *Notes*: This figure shows the percentages of students in the text-message coaching program who strongly disagree, disagree, somewhat disagree, somewhat agree, agree, and strongly agree with the statement that appears as the title of each panel.

Overall, the analysis of treatment effects on non-grades outcomes suggests that the text-message coaching treatment caused students to feel a greater sense of belonging and support at UofT than students in the control group. Descriptive evidence from the follow-up survey also demonstrates that nearly all students felt supported by their coaches and appreciated receiving text messages from them. Despite the online module not having a statistically significant impact on non-grades outcomes, the evidence is consistent with at least part of the text-message coaching treatment effects being driven by students completing the online exercise.

4.3. Robustness checks

In this subsection, we argue that our results are robust to coach heterogeneity and concerns about multiple hypothesis testing.

4.3.1. Coach heterogeneity

Apart from ensuring the treatment and control groups are balanced on pre-determined variables, one may also be concerned about the importance of coach heterogeneity. Specifically, with only 10 coaches, if coaches vary considerably in their quality, one may worry that the estimated treatment effects are quite sensitive to the presence of particular coaches. We address this concern by re-estimating many treatment effects for several outcomes of interest, with each estimated effect being obtained from a specification in which we drop the students of a given coach from the regression. The results from this exercise are presented in Appendix Table C3 for student grades (across all courses), the core well-being index, the belonging and support sub-index, the success strategies index, and the help-seeking sub-index. Each cell contains a treatment effect estimated from a separate regression and the column numbers indicate the number corresponding to the coach whose students we drop from the regression.

Regardless of which coach is dropped, and consistent with our main results above, the estimates in Appendix Table C3 indicate that the treatment was ineffective at improving student grades but was effective at improving students' non-grade outcomes. Further, all ten point estimates for a given outcome are very similar in magnitude and are not statistically differentiable from the main point estimates in Tables 6, 8, and 9. We therefore conclude that the estimated treatment effects are not driven by the effectiveness of any one coach.

4.3.2. Multiple hypothesis testing

Because we estimate the effects of two treatments on three main outcomes – grades, the core well-being index, and the success strategies index – one may be concerned about statistically significant effects on non-academic outcomes arising by chance because of the multiple-hypothesis testing problem. In general, one can account for multiple hypothesis testing by performing readily available inference adjustments that directly account for the number of tests being conducted or by replicating the results in other settings, thereby showing the original results are not driven by noise (see, for example, Alan et al., 2019).[26] In this section, we discuss both approaches.

Table C4 in Appendix C presents treatment effects from student-level regressions of the three main outcomes of interest – mean first-year grade, the core well-being index, and the success strategies index – on treatment indicators. To explore multiple hypothesis testing concerns, in addition to conventional *p*-values, we report List–Shaikh–Xu *p*-values (see List et al., 2019), which are calculated by directly considering the number of hypotheses (and relationships between variables) being tested. When performing inference adjustments, we assume six hypotheses are tested, one for each outcome-treatment pair. Although the statistical significance of the treatment effect on non-academic outcomes does not survive multiple testing adjustments, we argue that this is due to limited statistical power, and show that we replicate our effects on core well-being at one of the University of Toronto's satellite campuses, the University of Toronto at Mississauga (UTM).

UTM is a less-selective, suburban, satellite campus of the University of Toronto, located in a suburb west of Toronto called Mississauga. In the same year we ran the main experiment at the St. George campus (the 2016–2017 school year), we partnered at the UTM campus with an existing for-profit company in the business of sending one-way text messages to college students with the goal of improving academic achievement and persistence. Students in the treatment group at UTM also completed the CYOC module and were all additionally offered the one-way text messaging program. Although students received texts from the commercial organization (and not our research team), they were not told that the messages were coming from an outside organization and instead received messages appearing to come from *You@UofT* support. Students in the control group at UTM were assigned to the same Big 5 personality test as control students at St. George. Appendix D provides more information about the UTM campus and the text messages received by treated students.[27]

Table C5 shows treatment effects on the three main outcomes at UTM along with *p*-value adjustments for multiple testing, where we also report Romano–Wolf *p*-values.[28] The text-message treatment effect on the core well-being index is not statistically significant after performing *p*-value adjustments but the point estimate is very similar to the text-message treatment effect on core well-being in main sample at St. George. Somewhat surprisingly, treatment had a negative effect on the success strategies index at UTM.[29] Because the effects of text-message coaching on core well-being are so similar across campuses, we combine the St. George and UTM samples and estimate treatment effects in the pooled sample in Table C6. The increase in statistical power that comes from the larger sample results in a statistically significant effect of text-message coaching on the core well-being index even after adjusting *p*-values for multiple testing. The point estimate implies that students who received text-message coaching scored 5% of a standard deviation higher on this index, a magnitude nearly identical to our main estimate in column (1) of Table 8. Treatment effects on the success strategies index are no longer economically or statistically significant, as the negative effects at UTM cancel the positive effects at St. George.

[26] A complementary approach is to register the experiment prior to unblinding (revealing treatment status in) the experimental data. We registered the current experiment with the AEA Registry, under the ID AEARCTR-0000810. This registration ID covers many years of experiments, and the update done on December 17, 2016 corresponds to this experiment. In that update, the well-being outcomes analyzed here are reported as secondary outcomes of interest.

[27] At UTM, 25.6% of students did not complete the follow-up survey but attrition from the survey is not differential by treatment status. Regressing an indicator variable for whether a student completed the follow-up survey on a treatment indicator results in a coefficient estimate of 0.017 (se=0.024).

[28] We do not report Romano–Wolf *p*-values in Table C4 because that procedure is not designed to account for multiple treatments.

[29] The negative effect at UTM is driven by treated students making less frequent use of study groups, workshops, and both paid and free tutors (results available upon request). There are two potential explanations for this change in behavior. First, looking ahead to treatment effects at UTM on the sub-indices of the well-being index in Table C7, they are generally larger in magnitude than those at St. George. Feeling even more supported by the institution than treated students at St. George, it is possible that treated students at UTM chose to access study groups and tutors less because they did not require these services to boost their well-being. Second, many of the conversations between coaches and students at St. George inevitably touched on study strategies and seeking help from available resources. The messages from the for-profit company, in contrast, usually consisted of general support or one-time suggestions about on-campus resources, without the opportunity for subsequent discussion (see Appendix Table D2). This discrepancy in messaging style may have led treated students at UTM to feel less of need to seek help.

Table 10
Treatment effects across student subgroups.

	Gender		First year status		International student		Transition difficulty (TD)	
	(1) Female	(2) Male	(3) First year	(4) Not first year	(5) International	(6) Domestic	(7) TD > Median	(8) TD < Median
Panel (a): Grades								
Online only	−0.435	1.121	0.439	−0.436	−0.221	0.848	0.272	0.129
	[0.657]	[0.863]	[0.540]	[1.405]	[0.816]	[0.683]	[0.779]	[0.712]
Text messaging	−0.362	0.924	−0.395	2.686**	0.378	0.205	−0.112	0.287
	[0.662]	[0.826]	[0.543]	[1.336]	[0.783]	[0.693]	[0.744]	[0.721]
Control mean	70.976	70.696	72.305	66.407	69.645	71.933	72.017	69.896
[Standard deviation]	[15.388]	[16.472]	[14.756]	[18.206]	[16.603]	[15.124]	[15.125]	[16.414]
Panel (b): Core well-being index								
Online only	0.022	0.029	0.027	0.024	−0.002	0.047	0.052	0.023
	[0.031]	[0.037]	[0.026]	[0.058]	[0.033]	[0.035]	[0.034]	[0.033]
Text messaging	0.057*	0.016	0.030	0.101	−0.003	0.084**	0.042	0.052*
	[0.031]	[0.036]	[0.025]	[0.066]	[0.033]	[0.034]	[0.035]	[0.031]
Panel (c): Belonging and support index								
Online only	0.004	0.029	0.020	0.004	−0.008	0.034	0.037	0.016
	[0.032]	[0.039]	[0.027]	[0.062]	[0.034]	[0.036]	[0.036]	[0.034]
Text messaging	0.075**	0.019	0.044*	0.107	0.015	0.089**	0.046	0.070**
	[0.032]	[0.038]	[0.026]	[0.072]	[0.033]	[0.036]	[0.036]	[0.033]

The dependent variable in each regression is indicated by the panel headings. Control variables include student age, self-reported expected study hours per week, expected paid-work hours per week, expected average grade, tendency to finish what he or she starts, tendency to get discouraged by setbacks, tendency to study at the last minute, and indicator variables for first-year status, international student status, first-generation status, gender, English mother-tongue status, whether the student lives on residence, a self-reported desire to earn more than an undergraduate degree, and a self-reported tendency to check his or her cell phone often. Standard errors clustered at the student level are reported in brackets in panel (a). Robust standard errors are reported in brackets in panels (b) and (c). * indicates is significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

Our main results in Table 8 imply that the overall effect of the text-message treatment on the core well-being index is driven mainly by students feeling a greater sense of belonging at and support from the university, although treatment effects on all other sub-indices are positive. Table C7 shows that the text-message treatment at UTM also had uniformly positive effects on all the core well-being sub-indices, with the magnitude of the effect on the belonging and support index being very similar to the magnitude at the St. George campus. In Table C8, we again pool all observations and estimate treatment effects on each sub-index of core well-being. As before, our main result is that the effects on student well-being are predominately driven by students experiencing a greater sense of belonging and support, as treated students score 5% of a standard deviation higher on this index, an effect significant at the 5% level and robust to inference adjustments for multiple hypothesis testing.

Given the similarity of the point estimates across campuses and robustness of the results to inference adjustments in the pooled sample, we conclude that our main finding that support programs of this type increase student well-being is not an artifact of multiple-hypothesis testing concerns. In sum, there is robust evidence that the text-message coaching intervention did not improve student grades but it did modestly improve students' sense of belonging and support.

4.4. Heterogeneous treatment effects

Having established that the text-message coaching intervention caused modest improvements in students' non-grades outcomes, we now explore potentially heterogeneous treatment effects across student subgroups. We focus attention on subgroups defined by gender, first-year student status, international student status, and whether the student reports (in the pre-randomization survey) experiencing below or above median difficulty in transitioning to university (on a 7-point scale). Potentially differential effects by gender are often of interest in education interventions. We further chose to explore the other three student subgroups because they each partition students into groups that differ in their familiarity or comfort with university and life in Toronto and both of our treatments are designed to help students form study strategies and adjust to university.

The panels of Table 10 report treatment effects across student subgroups on grades and non-grades outcomes. Turning first to heterogeneous treatment effects on course grades in panel (a), neither treatment was effective at improving student grades in any subgroup, except for the text-message coaching treatment having a positive effect among non-first-year students. Non-first-year students comprise only 25% of our sample. Given the lack of an effect on grades in the overall sample and in any other student subgroup, we believe it is likely that this result is due to chance.

Focusing next on non-grades outcomes in panels (b) and (c), text-message coaching treatment effects are larger for female students than for male students, but the point estimates are not statistically differentiable. Treatment effects on the core well-being index and its belonging and support sub-index are larger among non-first year students than first-year students but are never statistically significant or statistically differentiable from effects on first-year students. A similar theme emerges for the international-domestic student comparison, as treatment effects on well-being are larger for domestic students but only the effect on the main core well-being index is statistically differentiable across the two groups. Splitting students by self-reported difficulty in transitioning to university reveals slightly larger treatment effects on the core well-being index and its belonging and support sub-index for students with less difficulty transitioning.

Although treatment effect estimates are rarely statistically differentiable across subgroups, the evidence is potentially suggestive of two conclusions. First, the text-message coaching treatment may have been more effective for women than for men. Second, treatment may have been more effective at improving well-being outcomes among students who are more familiar with university and life in Toronto (non-first year students, domestic students, and those with less difficulty transitioning to university). Ultimately, however, treatment effects in the overall sample and across all subgroups are modest, making it difficult to draw definitive conclusions about heterogeneous treatment effects.

5. Conclusion

In this paper, we designed and evaluated two large-scale behavioral interventions aiming to improve student experiences in college. Students assigned to the first intervention completed a novel online preparatory module while students assigned to the second intervention were also offered the opportunity to participate in a coaching program, in which they were matched with senior undergraduate coaches who mentored them via text message. We found that neither intervention improved student grades or credit accumulation but that the online module enhanced with the text-messaging program increased students' sense of belonging and support at the university. Although the treatment effects we estimate are modest, we note that most students in the text-message coaching intervention report feeling supported by their coaches, appreciating their coach's messages, and feeling like they are doing better in university partly because of their coaches.

Given the pervasive sense in today's higher-education landscape that the many colleges are facing a 'mental health crisis' among their students, our results are important despite their small relative magnitude. Colleges can potentially use light-touch support programs to promote a greater sense of belonging among their students and to make students feel like their schools are invested in their success. Such interventions may be worth implementing even when they do not affect student grades, as they enhance the college experience for students and can do so at scale for relatively little cost. The interventions we evaluated cost approximately $12 per student enrolled when we account for the setup costs of building the online platform. Conditional on having the platform, the costs of the text messages we sent are approximately $2 per student.[30] Our results from UofT's satellite campus show that even a one-way text-messaging service featuring automated replies can positively affect student well-being, implying that colleges need not spend great resources on support programs for them to be beneficial and appreciated by students.

We also note that our treatment effects of student well-being potentially represent a lower bound for the true effect, as studies show there exists peer spillovers in mental health services usage (Golberstein et al., 2016), suggesting potential benefits even for students who do not directly receive the messages.[31] Considering these possible spillover effects, text-message coaching services may represent an efficient way to reach a large number of students very quickly and can be particularly cost-effective when aiming to raise awareness of on-campus mental health resources.

Taken together, we view our results as the start of a conversation about how colleges can use light-touch behavioral interventions to combat mental health issues among their students and to raise awareness of available services on campus. Future programs should to be designed to explicitly consider treatment effects on non-academic outcomes, such as belonging, satisfaction, confidence, and depression, in addition to commonly explored academic outcomes like grades, credit accumulation, and persistence. As we demonstrate in this paper, an intervention can improve student experiences in college and be perceived as beneficial by students while not necessarily affecting grades and credit accumulation.

Further research is needed to better understand how the effects we found generalize across other institutions and student populations. It is also worth exploring whether such interventions can be made even more targeted and cost-effective by tailoring intervention materials to explicitly focus on providing mental health support and using survey data to predict which students are likely to experience the greatest barriers to their well-being while in college. We are currently exploring both avenues in ongoing work.

Declaration of Competing Interest

I declare no conflict of interest for this manuscript.

[30] This paper is part of a broader research agenda on low-cost, scalable interventions in education. To encourage further research and help reduce the costs of building an online platform, we have made all interventions available online and made them customizable for other researchers who are interested in implementation at their institutions. For details, see https://studentachievementlab.org/.

[31] It is also possible that the estimated treatment effects on well-being are driven in part by treated students comparing their emotional states to control students, who did not receive support, and feeling better as a result of the comparison. This mechanism hinges, however, on control students being aware of their treatment status and being disappointed because of not receiving treatment materials. Here, we note that in both this experiment and in several others we have conducted at UofT over six years (for a total experimental sample approximately 25,000 students), we have never recorded an instance of a student in the control group complaining to either our research team or their instructor about not receiving treatment materials.

Acknowledgments

We are indebted to the first-year economics instructors at the University of Toronto for their willingness to incorporate an experiment into their courses for a third consecutive year. We especially thank Aaron de Mello, our web developer, for his tireless commitment to designing and perfecting the experiment's website, as well as for his help with organizing and extracting the experimental data. Dana Britton, Holly Fenton, Ayush Gupta, Aamir Husain, Anthony Koundourakis-Soares, Julianna Lu, Sukhi Singh, and Kadachi Ye showed great enthusiasm and professionalism in their role as coaches. Seminar participants at the Canadian Institute for Advanced Research provided useful feedback. We also thank participants at the SOLE 2018 conference in Toronto for helpful comments and suggestions. All remaining errors are our own.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2020.01.009.

References

Akerlof, G., Kranton, R.E., 2000. Economics and identity. Q. J. Econ. 3, 715–753 CXV.
Akerlof, G., Kranton, R.E., 2002. Identity and schooling: some lessons for the economics of education. J. Econ. Lit. XL, 1167–1201.
Akerlof, G., Kranton, R.E., 2005. Identity and the economics of organizations. J. Econ. Perspect. 19 (1), 9–32.
Alan, S., Boneva, T., Ertac, S., 2019. Ever failed, try again, succeed better: results from a randomized educational intervention on grit. Q. J. Econ. 134 (3), 1121–1162.
Alstadsaeter, A., 2011. Measuring the consumption value of higher education. CESifo Econ. Stud. 57 (3), 458–479.
Alter, M., Reback, R., 2014. True for your school? How changing reputations alter demand for selective U.S. colleges. Educ. Eval. Policy Anal. 36 (3), 346–370.
American College Health Association, 2018. American College Health Association-National College Health Assessment II: Reference Group Executive Summary Fall 2017. American College Health Association, Hanover, MD.
Arcidiacono, P., 2004. Ability sorting and the returns to college major. J. Econom. 121, 343–375.
Arum, R., Roksa, J., 2011. Academically Adrift: Limited Learning on College Campuses. University of Chicago Press, Chicago, IL.
Babcock, P., Marks, M., 2011. The falling time cost of college: evidence form half a century of time use data. Rev. Econ. Stat. 93 (2), 468–478.
Beattie, G., Laliberté, J.-W.P., Oreopoulos, P., 2018. Thrivers and divers: using non-academic measures to predict college success and failure. Econ. Educ. Rev. 62, 170–182.
Beattie, G., Laliberté, J.-W.P., Michaud-Leclerc, C., Oreopoulos, P., 2019. What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health. Econ. Lett. 178, 50–53.
Bettinger, E., Baker, R., 2014. The effects of student coaching: an evaluation of a randomized experiment in student advising. Educ. Eval. Policy Anal. 35 (1), 3–19.
Bettinger, E., Ludvigsen, S., Rege, M., Solli, I.F., Yeager, D., 2017. Increasing perseverance in math: evidence from a field experiment in Norway. J. Econ. Behav. Organ. 146, 1–15.
Broda, M., Schneider, J.Y., Yeager, D.S., Walton, G.M., Diemer, M., 2018. Reducing inequality in academic success for incoming college students: a randomized trial of growth mindset and belonging interventions. J. Res. Educ. Eff. 11 (3), 317–338.
Canning, E.A., Harackiewicz, J.M., 2015. Teach it, don't preach it: the differential effects of directly-communicated and self-generated utility-value information. Motiv. Sci. 1 (1), 47–71.
Castleman, B., Page, L., 2015. Summer nudging: can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? J. Econ. Behav. Organ. 115, 144–160.
Castleman, B., Page, L., 2016. Freshman year financial aid nudges: an experiment to increase FAFSA renewal and college persistence. J. Hum. Resour. 51 (2), 389–415.
Castleman, B., Meyer, K., 2016. Can Text Message Nudges Improve Academic Outcomes in College? Evidence From a West Virginia Initiative. University of Virginia, Charlottesville, VA Center for Education Policy and Workforce Competitiveness Working Paper # 43.
Clark, D., Gill, D., Prowse, V., Rush, M., 2017. Using Goals to Motivate College Students: Theory and Evidence From Field Experiments. National Bureau of Economic Research, Cambridge, Mass. NBER Working Paper # 23638.
Clotfelter, C.T., Hemelt, S.W., Ladd, H.F., 2018. Multifaceted aid for low-income students and college outcomes: evidence from North Carolina. Econ. Inq. 56 (1), 278–303.
Coleman, J.S., 1961. The Adolescent Society: The Social Life of the Teenager and Its Impact on Education. The Free Press, NY.
Damgaard, M.T., Nielsen, H.S., 2018. Nudging in education. Econ. Educ. Rev. 64, 313–342.
David, E., 2019. Rising Suicide Rates at College Campuses Prompt Concerns over Mental Health Care. ABC News ABC News Network, 9 Oct. abcnews.go.com/Health/rising-suicide-rates-college-campuses-prompt-concerns-mental/story?id=66126446.
De Neve, J.E., Oswald, A.J., 2012. Estimating the influence of life satisfaction and positive affect on later income using sibling fixed effects. Proc. Natl. Acad. Sci. 109 (49), 19953–19958.
Di Tella, R., MacCulloch, R.J., Oswald, A.J., 2001. Preferences over inflation and unemployment: evidence from surveys of happiness. Am. Econ. Rev. 91 (1), 335–341.
Dobronyi, C.R., Oreopoulos, P., Petronijevic, U., 2019. Goal setting, academic reminders, and college success: a large-scale field experiment. J. Res. Educ. Eff. 12 (1), 38–66.
Dynarski, S., Libassi, C.J., Michelmore, K., Owen, S., 2018. Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students. National Bureau of Economic Research, Cambridge, MA NBER Working Paper No. 25349.
Eisenberg, D., Golberstein, E., Hunt, J.B., 2009. Mental health and academic success in college. B.E. J. Econ. Anal. Policy. 9 (1), 1–35.
Evans, W.N., Kearney, M.S., Perry, B., Sullivan, J.X., 2017. Increasing Community College Completion Rates Among Low-Income Students: Evidence From a Randomized Controlled Trial Evaluation of a Case Management Intervention. National Bureau of Economic Research, Cambridge, Mass. NBER Working Paper # 24150.
Golberstein, E., Eisenberg, D., Downs, M.F., 2016. Spillover effects in health service use: evidence from mental health care using first-year college housing assignments. Health Econ. 25, 40–55.
Gollwitzer, P.M., Schaal, B., 1998. Metacognition in action: the importance of implementation intentions. Person. Soc. Psychol. Rev. 2 (2), 124–136.
Hastings, J., Neilson, C.A., Zimmerman, S.D., 2018. The Effects of Earnings Disclosure on College Enrollment Decisions. National Bureau of Economic Research, Cambridge, Mass. NBER Working Papers # 21300.
Hefner, J., Eisenberg, D., 2009. Social support and mental health among college students. Am. J. Orthopsychiatr 79 (4), 491–499.
Hofmann, S., Muhlenweg, A., 2018. Learning intensity effects in students' mental and physical health – evidence from a large scale natural experiment in Germany. Econ. Educ. Rev. 67, 216–234.
Jacob, B., McCall, B., Stange, K., 2018. College as country club: do colleges cater to students' preferences for consumption? J. Labor Econ. 36 (2), 309–348.

Kling, J.R., Liebman, J.B., Katz, L.F., 2007. Experimental analysis of neighborhood effects. Econometrica 75 (1), 83–119.
Lavecchia, A., Liu, H., Oreopoulos, P., 2016. Behavioral economics of education: progress and possibilities. Handb. Econ. Educ. 5, 1–74.
List, J.A., Shaikh, A.M., Xu, Y., 2019. Multiple hypothesis testing in experimental economics. Exp. Econ. doi:10.1007/s10683-018-09597-5.
Luttmer, E.F.P., 2005. Neighbors as negatives: relative earnings and well-being. Q. J. Econ. 963–1002 August.
Oreopoulos, P., Ford, R., 2016. Keeping College Options Open: A Field Experiment to Help all High School Seniors Through the College Application Process. National Bureau of Economic Research, Cambridge, Mass. NBER Working Paper # 22320.
Oreopoulos, P., Patterson, R.D., Petronijevic, U., Pope, N., 2020. Low-touch attempts to improve time management among traditional and online college students. J. Hum. Resour. forthcoming.
Oreopoulos, P., Petronijevic, U., 2018. Student coaching: how far can technology go? J. Hum. Resour. 53 (2), 299–329.
Oreopoulos, P., Petronijevic, U., 2019. The remarkable unresponsiveness of college students to nudging and what we can learn from it, NBER Working Paper #26059.
Oswald, A.J., Wu, S., 2010. Objective confirmation of subjective measures of human well-being: evidence from the U.S.A. Science 327, 576–579.
Oswald, A.J., Eugenio, P., Sgroi, D., 2015. Happiness and productivity. J. Labor Econ. 33 (4), 789–822.
Pope, D.G., Pope, J.C., 2009. The impact of college sports success on the quantity and quality of student applications. South Econ. J. 75 (3), 750–780.
Reilly, K., Record numbers of college students are seeking treatment for depression and anxiety — but schools can't keep up. Time, Time, 19 Mar. 2018, time.com/5190291/anxiety-depression-college-university-students/.
Scrivener, S., Weiss, M.J., 2013. More graduates: two-year results from an evaluation of accelerated study in associate programs (ASAP) for developmental education students. Policy Brief, MDRC. https://files.eric.ed.gov/fulltext/ED546636.pdf.
Senik, C., 2004. When information dominates comparison learning from Russian subjective panel data. J. Publ. Econ. 88, 2099–2123.
Shapiro, D., Dundar, A., Huie, F., Wakhungu, P., Bhimdiwala, A., Wilson, S.E., 2019. Completing College: Eight Year Completion Outcomes for the Fall 2010 Cohort (Signature Report No. 12c). National Student Clearinghouse Research Center, Herndon, VA.
Staglin, G., 2019. Addressing Mental Health Challenges on College Campuses. Forbes, Forbes Magazine 4 Oct. www.forbes.com/sites/onemind/2019/10/04/addressing-mental-health-challenges-on-college-campuses/#36b2b511400d .
Steele, C.M., 1997. A threat in the air: how stereotypes shape intellectual identity and performance. Am. Psychol. 52 (6), 613–629.
Walton, G., 2014. The new science of wise psychological interventions. Curr. Dir. Psychol. Sci. 23 (1), 73–82.
Walton, G., Cohen, G., 2011. A brief social-belonging intervention improves academic and health outcomes of minority students. Science 331, 1447–1451.
Yeager, D., Walton, G., 2011. Social-Psychological interventions in education: they're not magic. Rev. Educ. Res. 81 (2), 267–301.
Walton, G., Logel, C., Peach, J., Spencer, S., Zanna, M., 2015. Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. J. Educ. Psychol. 107 (2), 468–485.
Yeager, D., Romero, C., Paunesku, D., Hulleman, C.S., Schneider, B., Hinojosa, C., Lee, H.Y., O'Brien, J., Flint, K., Roberts, A., Trott, J., 2016a. Using design thinking to improve psychological interventions: the case of the growth mindset during the transition to high school. J. Educ. Psychol. 108 (3), 374–391.
Yeager, D.S., Walton, G.M., Brady, S.T., Akcinar, E.N., Paunesku, D., Keane, L., Kamentz, D., Ritter, G., Duckworth, A.L., Urstein, R., Gomez, E.M., Markus, H.R., Cohen, G.L., Dweck, C.S., 2016b. Teaching a lay theory before college narrows achievement gaps at scale. Proc. Natl. Acad. Sci. doi:10.1073/pnas.1524360113.
Zivin, K, Eisenberg, D., Gollust, S.E., Golberstein, E., 2009. Persistence of mental health problems and needs in a college student population. J. Affect. Disord. 117, 180–185.