The Efficacy of Growth and Belonging Mindset Interventions Among College Students*

Abid Nahiyan Alam, Queen's University

Philip Oreopoulos, University of Toronto

Uros Petronijevic, York University

September 2025

Preliminary and Incomplete. Please do not cite or circulate.

Abstract

Using four large-scale experiments across two major Canadian universities, we experimentally evaluate the effects of growth mindset and social belonging interventions on student outcomes. In a sample of nearly 12,000 students, we find no immediate or dynamic effects on student grades and no effect on persistence through university. We further combine survey and administrative data with machine learning methods to explore treatment effect heterogeneity, finding no evidence of meaningful variation in treatment effects across student subgroups. Despite the recent promise of these light-touch interventions, our findings indicate that further research is required to identify the contexts in which their benefits generalize.

*We are grateful to...All remaining errors are our own. Contact information: Alam, Queens University, abid.alam@queensu.ca; Oreopoulos, University of Toronto, philip.oreopoulos@utoronto.ca; Petronijevic, York University, upetroni@yorku.ca.

1 Introduction

Academic outcomes such as grades, graduation rates, and human capital gains, vary significantly between students belonging to different racial, ethnic, and socioeconomic groups, even when holding constant prior academic preparation (Yeager et al., 2016). A large and growing body of social-psychology research posits that these gaps exist, in part, because students face unique challenges such as negative stereotypes about their abilities, group-based threats, or under-representation (Murphy et al., 2007; Olson and Dweck, 2008; Espenshade and Radford, 2009). These challenges may give rise to personal narratives involving deeply-rooted doubts about an individual's sense of belonging or perception of potential, resulting in reduced effort and achievement (Coleman, 1961; Akerlof and Kranton, 2002; Chrobak, 2024). Another area of social-psychology research attributes performance gaps to differences in students' beliefs about the importance of effort and struggling through challenges in the process of improving academic ability (Dweck, 2006; Delavande et al., 2023).

Over the last two decades, social-psychological interventions, such as social-belonging and growth mindset interventions, have become an increasingly common tool for improving student outcomes or reducing student achievement gaps in all levels of academia. Social-belonging interventions aim to make individuals feel that they belong by counteracting the effects of implicit signals that individuals may receive (Rattan et al., 2015). They do so by making individuals believe that the difficulties they face do not represent any specific inadequacy unique to them but rather are faced by everyone and, over time and with effort, can be overcome. These interventions are hypothesized to improve academic achievement by changing the way students view doubts about belonging from something inherently unique to them, to something which is transitory, and a direct consequence of the challenges that are common to all students. Since doubts about belonging can be described as "personal narratives" (Chrobak, 2024), they can be altered using so-called "psychologically wise" interventions, which are designed to change individuals' thoughts and feelings to allow them to succeed (Walton, 2014; Walton and Wilson, 2018).

Growth mindset interventions, on the other hand, are motivated by Mindset Theory,

which posits that individuals can approach challenges with either a growth or fixed mindset (Dweck, 2006). A growth mindset is often viewed as an important non-cognitive skill in which individuals view intelligence as malleable and effort being productive in improving one's ability to perform a given task. A fixed mindset, in contrast, leads individuals to view intelligence as largely predetermined, thereby leading to less effort or avoidance of challenges altogether (Dweck, 2007, 2009). Growth mindset interventions therefore aim to counteract the effects of a fixed mindset by teaching participants that intelligence is fluid and that, through effort exertion, individuals can improve their capabilities and get closer toward their goals (Yeager and Walton, 2011; Yeager and Dweck, 2012).

In this paper, we evaluate the generalizibility of social-belonging and growth mindset interventions in large-scale randomized experiments between 2015 and 2019 in two of the largest Universities in Ontario, Canada: the University of Toronto and York University. Prior studies estimate wide-ranging impacts of social belonging interventions on academic outcomes, with positive effects for racial or ethnic minority students (Walton and Cohen, 2007, 2011; Binning et al., 2020), first-generation students (Stephens et al., 2014), ESL students (LaCosse et al., 2020) and female students (Walton et al., 2015), but no effect in the broader population (Broda et al., 2018). Student heterogeneity also plays an important role in the literature on growth mindset interventions: some studies find positive effects for at-risk and low-SES students (Broda et al., 2018; Kim et al., 2022; Yeager et al., 2016b; Damgaard and Nielsen, 2018; Sisk et al., 2018; Blackwell et al., 2007; Paunesku et al., 2015; Bettinger et al., 2018; Yun et al., 2019; Yeager et al., 2019), while others find no effects or mixed results for the broader student population (Kim et al., 2022; Brougham and Kashubeck-West, 2017; Sisk et al., 2018; Kaijanaho and Tirronen, 2018; Warren et al., 2019; Bosch, 2021). Overall, both types of intervention have shown some efficacy among specific student subgroups but the generalizabilty of these results and the effectiveness of these interventions in the broader student population remain unsettled issues.

We contribute fresh evidence to the literature by studying the effects of a standard belonging intervention, a standard growth mindset intervention, and two context-specific,

novel extensions of each type of intervention. We conduct these studies across two major Canadian universities covering three cohorts of students in first-year economics classes. The extensions of each intervention are tailored to the characteristics of the student body: we modify the standard belonging intervention to specifically address the belonging concerns of first-year international students and we modify the growth mindset intervention with materials specifically relevant to students' field of study.[1]

We evaluate the effects of each of the four interventions on short- and long-run GPA, persistence rates through university, and five-year graduation rates in entire sample and across student subgroups. We explore treatment effect heterogeneity in two ways, first using the prior literature to identify student subgroups across which we estimate separate treatment effects. We then use causal forest methods—a hands-off, data-driven approach—to identify any potential variation in treatment effects across student types (Wager and Athey, 2018). Being data-driven, causal forest methods allow the data to potentially identify novel subgroups in which these interventions may be effective.

Taken together, we find no evidence that any of the four interventions affect student academic outcomes, both in the full sample and across many different student subgroups. That is, we estimate null effects on short- and long-term GPA, persistence through university, and the likelihood of graduating in five years. Pooling all interventions together, we rule out treatment effects larger than 4.5 percent of a standard deviation on contemporaneous grades and 4.3 percent of a standard deviation on persistence. These null effects span across all subgroups, as we find no evidence of treatment effect heterogeneity by English proficiency status, gender, first-generation status, and many other groups we are able to identify with our unique survey data. We are also unable to detect any meaningful treatment effect variation for any intervention using the causal forest approach.

Considering the extensive and expanding literature on similar social psychology inter-

[1]We developed the two modified interventions with leading researchers in the areas: the belonging intervention for international students resulted form our collaboration with Christine Logel and Greg Walton, while the field-of-study growth mindset was designed in collaboration with David Yeager. These interventions add to the growing dialogue on mindset interventions needing to be customized before implementation based on contextual factors.

ventions, our study makes three primary contributions: (i) our unique survey data allows for a comprehensive analysis of treatment effects; (ii) we explicitly evaluate versions of the standard interventions tailored to the specific institutional contexts students encounter; and (iii) taken together, our findings contribute to the broader scientific effort to assess the generalizability of these soft-touch social psychology iinterventions.

On the data front, we estimate dynamic treatment effects on longer-run outcomes, including grades throughout the first four years of university, persistence through university, and five-year graduation rates, standing in contrast to the majority of prior studies that evaluate only the short-term impacts. Our unique baseline survey data, paired with state-of-the-art machine-learning methods, also allow us to fully explore heterogeneity in treatment effects, a recurring theme in much of the prior literature. In addition to more common demographic characteristics, the causal forest approach allows for potential treatment effect variation across students who differ along work-for-pay responsibilities, planned weekly study hours, performance goals, graduate school intentions, and procrastination tendencies.

In terms of experimental design, we embrace the notion put forward in the prior literature that the challenges to academic success students face are often setting specific. Prior work, for example, has found that growth mindset interventions are more effective in institutions where the norms align with the message of the intervention and that the effectiveness of the message requires a match in cultural relevance, norms, and values between the institution and specific student subgroups (Yeager et al., 2016a; Stephens et al., 2012).[2] In addition to standard belonging and mindset interventions, we therefore design and estimate the effects of a customized version of each, focusing on a belonging intervention for international students and a growth mindset intervention for the particular field of study of most students in our sample.

Our study—which can be considered a "conceptual" replication of social belonging and

[2]Data-driven research has also identified the importance of student-level factors such as the degree of participation in intervention exercises and self-reported motivations to predict the effectiveness of such interventions (Bosch, 2021). Moreover, information interventions in the form of personalized academic feedback are effective for subgroups such as first-generation students, but are dependent on the instructional mode (Aucejo and Wong, 2025).

growth mindset interventions (Shrout and Rodgers, 2018)—supports the scientific process underlying efforts to assess the generalizability of these soft-touch social psychology interventions, the importance of which has been well documented (Ioannidis, 2012; Shrout and Rodgers, 2018). In particular, the null effects we document across four interventions, two universities, three cohorts, and many student subgroups constitute an important contribution to a literature in which publication bias is a long-standing concern (Dickersin et al., 1987; Easterbrook et al., 1991; Goldacre, 2009; Song et al., 2010; Fanelli, 2010; Campbell et al., 2020). Over time, publication bias can lead to significant findings being overrepresented relative to equally-valid null effects, results from meta-analytic studies being misleading, and the exacerbation of the replication problems (Dickersin et al., 1987; Kozlov, 2024).[3] We therefore believe that the null effects we document, given their consistency and fidelity, are an important contribution to our understanding of how these interventions generalize and replicate.

The remainder of the paper is organized as follows. In the next section, we provide a detailed description of our experimental design and interventions. We describe our data in Section 3 and outline our empirical strategy in Section 4, where we cover the machine learning procedure that we use for heterogeneity analysis. Section 5 presents the results and Section 6 offers concluding remarks.

2 The Experimental Design

In this section, we describe the experimental environment and summarize the control and treatment exercises students completed. Full documentation for the control group module, the two belonging modules, and the two growth mindset modules can be found in Appendix XX.

[3]Such problems are limited not only to the significance of results, but also to the direction and magnitude of the effect. For example, Camerer et al. (2016), while replicating 18 experimental studies published between 2011 and 2014 in the *Quarterly Journal of Economics* and *American Economic Review*, find that only 11 of the 18 papers had a significant effect in the same direction as reported the original papers, and, on average, the replicated effect size was 66% of which was originally reported.

We conducted our experiments on the three campuses of the University of Toronto (UofT) and the Keele campus of York University between 2015 and 2019 involving three different cohorts of first-year students. UofT is the largest university in Canada, and one of the largest publicly-funded universities in North America. It is consistently ranked among the top universities in the world. Across its main downtown Toronto campus, St. George (UTSG), and its two commuter campuses in the surrounding suburbs, the Scarborough (UTSC) and Mississauga campuses (UTM), UofT has over 90,000 students. Approximately 30% of the diverse student body consists of international students from roughly 178 countries or territories as of the fall 2023 - 2024 term enrollment (University of Toronto, 2025).[4] York University is the third largest-publicly funded university in Canada and is also located in Toronto. York has a diverse student body of more than 46,000 undergraduate students, approximately 27% of whom are international students. It has two campuses: Keele, its main (commuter) campus, and Glendon, a campus focused on bilingual liberal arts. Our sample consists entirely of students from York's main campus, Keele.

By collaborating with the instructors of first-year economics courses in the fall semester of each experiment year, all incoming first-year students were invited to participate in an online survey and subsequent control group or treatment group activity. We incentivized students to complete these online modules by assigning a small participation grade of 1-2% of the total course grade for completion. The survey collected baseline data prior to students being assigned to their experimental condition, and we later combined the survey data from each university's student-level administrative data.[5] After completion of the online survey, students were randomly assigned to a control or a treatment group.

Control Group: To make the duration of the control group exercise similar to that of the different treatment exercises, control group students were given two Big Five personality

[4]For a more detailed description of the University of Toronto as it relates our experimental setup, see Oreopoulos and Petronijevic (2023), who explore the effects of different—not analyzed in the current paper—low-touch, experimental interventions over the same time period.

[5]The interventions we evaluate in this study were conducted under the umbrella of the *Student Achievement Lab (SAL)*, which we operated at UofT between 2104-15 to 2019-20 at York between 2017-18 to 2019-20. Through a collobaration between our research team and first-year course instructors, SAL implemented and evaluated many light-touch, scalable interventions over that period. For a fuller description of the SAL environment and interventions not covered in this paper, see Oreopoulos and Petronijevic (2023).

tests—one based on an absolute score (Donnellan et al., 2006), which allows students to obtain a high score in all traits; and another based on a relative score (Hirsh and Peterson, 2008), which highlights the traits that dominate a student's personality profile. The students were also asked questions about their attitude towards risk (Dohmen et al., 2011), time preferences (Andersen et al., 2008) and grit (Duckworth and Quinn, 2009). The overall exercise took 45 – 60 minutes to complete, and students were emailed a short report afterwards.[6]

Social Belonging Intervention: The social-belonging intervention was implemented in 2015 in all UofT campuses. Students randomly assigned to the (social) belonging intervention read short descriptive stories from a diverse body of upper-year students. These stories were designed to convey the message that worries about belonging are common during academic transition and, with time, they should expect these feelings to dissipate as they meet others like themselves, become more familiar with their surroundings, and settle into a routine. The stories also illustrated that there are others like them who faced and overcame similar struggles. The students were then asked to write a short script on how this message resonated with their own experiences to date, and were told that these responses might be used to help future cohorts. The design of this intervention is closely based on that used in Yeager and Walton (2011), except for the addition of additional video instructions, a minimum word count for the responses, and time-on-page restrictions. This intervention was tested in fall 2015 on all campuses of the University of Toronto.

International Belonging Intervention: The international belonging intervention was implemented and tested in fall 2017, at UTSG, UTSC, and York. This intervention was designed in collaboration with Christine Logel and Greg Walton and was inspired by the fact that a large proportion of the student body consists of international students[7]. To be assigned to this intervention, a student had to be both an international and first-year student who was randomly assigned to a treatment condition.[8] International students face unique chal-

[6]More details about the control group exercises and analyses with the data gathered from them can be found in Beattie et al. (2018).

[7]Approximately 24% of undergraduate student body at UofT, and 15.6% of the undergraduate student body at York are international students (University of Toronto Office of Planning and Budget, 2024; Office of Institutional Planning and Analysis, York University, 2025).

[8]Students who were either not an international student or not a first-year student, but were selected for

lenge because they not only have to adjust to transitioning to a new academic environment, but also must navigate living in a country, away from their parents, friends, and support systems. We conducted focus groups, and following an iterative process outlined in Walton et al. (2017), the stories from the 2015 social-belonging mindset intervention were modified to incorporate distinct themes that were highlighted. These included, among others, feeling homesick, being worried about missing family and their old friends; being Worried about making meaningful friendships; being intimidated by professors; being too intimidated to ask questions in class or talk with professors; and feeling uneasiness from a lack of guidance on how to study.

Like the social belonging intervention, students viewed short stories from anonymous upper-year students describing their own experiences, and how they eventually overcame them. Some of the stories were reported as being from domestic students to demonstrate that challenges during transition to university, and concerns about belonging, are common among all students, and not something uniquely faced by international students. The participating students were shown a summary from a "current student survey" that concluded,"most students worry at first about whether they belong at university, but, after some time, they overcome these concerns and come to feel at home in their new environment". The students were then asked to reflect on these stories and describe their own experiences about what makes students unsure about belonging during academic transitions and were told that their responses might be used anonymously to help future students.

Growth Mindset Intervention: The growth mindset intervention was implemented in 2019 at York. Students who were randomly assigned first read information about the brain and how it learns new information. The content emphasized that by exerting effort through practicing and learning, the brain can be developed and abilities can be improved. In addition, the messages emphasized that working on strategies along with exerting effort can be beneficial. This was followed by scripts that conveyed personal experiences and responses from upper-year students who received the same information. After going through the articles

treatment, were assigned to the economics growth mindset treatment. See more details below.

and responses, students were asked to reflect on the information presented to them and were again informed that these responses may be used in to help future incoming students.

Economics Growth Mindset: The economics growth mindset intervention was implemented in 2017 in UTSG, UTSC, and York. To be assigned to this intervention, a student had to be either a domestic student or not in their fist year (that is, they could not be both an international and first-year student) and had to be randomly selected for a treatment condition. This intervention was designed in collaboration with David Yeager, and the intervention focused on three messages. First, while it sometimes it is difficult to see how the ideas and tools from a first-year economics class can be applied in the real world, these courses do in fact train students to think more logically about a wide variety of important real-world problems and prepare them for tackling problems with more complexity later on. Second, some problems in the course are tricky and challenging, but working through new and difficult material is how we learn best, even when we make mistakes or struggle along the way. Third, professors ask challenging questions not because they want some students to do poorly, but because they want all students to push themselves to improve. Like the other mindset interventions, the students were asked to reflect on the messaging and related stories through writing, and were told that their responses might be used to help future students.

3 Data

Our primary sample includes students enrolled in introductory economics courses in the fall term of the experiment year from all UofT campuses in 2015; UTSG, UTSC and York in 2017; and York in 2019. Table 1 the distribution of students across years, campuses, and treatment conditions. Altogether, our sample consists of 11,947 students.

We linked data gathered from the baseline survey with student-level administrative data using unique student identifiers. Across both data sources, we observe rich demographic information such as student age, gender, Canadian citizenship status, first-generation student

status, and whether English is their first language. We also observe measuresof of student expectations and time use, including grade expectations, intentions to pursue graduate studies, expected hours of work per week (for pay), and expected hours spent studying in a week. In terms of outcomes, the administrative data include students course grades for each semester they are enrolled and their eventual graduation status, allowing us to evaluate treatment effects on both short- and longer-run academic outcomes. Throughout our analysis below, we use the mean initial fall term course grade as the immediate academic outcome of interest and end-of-year mean grades in subsequent years as long-run outcomes.

We observe the students who were assigned to the social-belonging intervention from their initial enrollment until they graduate or drop out. Thus, we can study the impact of this intervention on all our outcomes. Those assigned to the economics growth mindset and international belonging interventions are observed until their potential third year in university, while those in the standard growth mindset intervention observed only in their first year.

After completion of the online survey, students were randomly assigned to either a treatment group or a control group in each experimental year. Table 2 provides descriptive statistics and balance tests for all the different treatment arms. The first column shows the mean characteristics of those in the control group from the full sample and column 2 shows the standard deviations. The average student completed their online module 9 days after it was available. Approximately half of students are male and the average student is 19 years old with an average high school admissions grade of 85 percent. Thirty-six percent of students are first generation and half are Canadian citizens. In terms of expectations and time use, two-thirds expect to earn an A average in university, 63 percent intend to earn more than a BA, and the average student expects to work 8.5 hours of week for pay and study approximately 16 hours per week for school.

Columns 3 through 11 show the differences in mean characteristics of those in different treatment groups compared to the control group along with standard errors. We estimate these differences by regressing each descriptive variable on treatment group indicators and

fixed effects for each of the campus-year-student-type categories indicated in columns 3 to 10 of Table 1. The fixed effects are necessary to account of differences in intervention design, proportions of students assigned to treatment, and characteristics of the student body across the groups. Overall, there is little evidence of differences between the treatment and control groups: only three tests result in differences significant at the 5 percent level, which is what we would expect by chance, and the magnitudes of any statistically significant differences are quite small.

4 Empirical Methodology

In this section, we describe how we estimate treatment effects, both in the full sample and across student subgroups.

4.1 Average Treatment Effects

We estimate average treatment effects for the different interventions by regressing outcome variables on treatment group indicators and fixed effects for the campus-year-student-type cluster groups indicated by the columns of Table 1. Treatment effect estimates are therefore interpreted as the average differences between treated and control students within a given cluster group. Our main specifications do not condition on other background variables for ease of interpretation, but we show in robustness checks in Appendix XX that our results are largely unchanged when we include additional control variables. We report robust standard errors for all specifications.

4.2 Informed Sub-Group Analysis

We first approach our subgroup analysis by intuitively selecting subgroups informed from existing literature. Although the literature studying the impact of growth mindset and social belonging interventions is vast, particular attention has been paid to certain student subgroups. For example, in the case of the growth mindset intervention, subgroups based

on prior preparation (in terms of measures such as GPA), low versus high beliefs in one's abilities and educational pathway (academic versus vocational) (Bettinger et al., 2018), low versus high achievers (Yeager et al., 2018), race (e.g. Latinx) (Broda et al., 2018), and gender (Delavande et al., 2023), among others, have been documented. With social belonging interventions, subgroups based on students for whom English is a second language (ESL) (LaCosse et al., 2020), low socioeconomic status (SES) (Yeager et al., 2016b), race and/or ethnic minority (Murphy et al., 2020; Walton and Cohen, 2011), first-generation students (Murphy et al., 2020; Harackiewicz et al., 2014), and students in STEM fields (Weaver et al., 2021), among others, have been considered. Keeping line with literature, based on the data collected during our experiments, we perform our subgroup analysis based on gender, whether a student is first generation, ESL students, and prior preparation (based on high school average). While we cannot perform a subgroup analysis based on factors such as race, due to data limitations, we can explore other novel subgroups based on other student characteristics, to our knowledge, have not yet been considered, for example factors based on pre-course expectations (e.g., low versus high expected grades, low vs high expected study hours, low vs high expected hours to work for pay, etc.). Similar to the estimation of average treatment effects described in the previous section, we use standard OLS regressions.

4.3 Data Driven Sub-Group Analysis

Our objective here is to estimate the effect of treatment on the outcomes of student i, conditional on observed characteristics. To estimate these conditional average treatment effects (CATE) for all the students in our sample, we use the causal forest procedure introduced in Wager and Athey (2018). To keep the discussion accessible, we provide a brief description of the method without the technical details.[9]

Causal forests are ensemble methods that require growing a large number of causal trees (Wager and Athey, 2018). Unlike standard decision trees, which split the data by grouping

[9]For details, we direct the reader to the theory in (Athey and Imbens, 2016; Wager and Athey, 2018; Athey et al., 2019) and two articles that apply the method (Davis and Heller, 2017; Hitsch et al., 2024).

observations based on the similarity of outcomes, causal trees split the data and group observations based on the similarity of treatment effects (Athey and Imbens, 2016). That is, causal trees recursively partition the covariate space using a modified splitting rule: instead of minimizing prediction error, each split is chosen to maximize the difference in estimated treatment effects between the groups. Each individual causal tree is "honest," as the sample splitting procedure ensures that the same data are not used to build the tree (partition) and predict treatment effects.[10] One can therefore think of the trees (partitions) as exogenously given, as the tree used to estimate observation $i's$ treatment effect never "sees" observation $i's$ data in the training process.

For a target $X = x_i$ (an observation with particular covariate values), the causal tree estimate of the treatment effect is the difference in the average outcomes between the treated and control observations, which are "close" to or in the same leaf as X.[11] In the causal forest procedure, to estimate the Conditional Average Treatment Effect (CATE) for an observation k with covariates X_k, X_k is passed down each causal tree in the forest. [12] Each tree assigns X_k to a specific leaf and using the estimation sample (separate from the sample used to build the tree structure, due to honesty), the tree computes the difference in average outcomes between treated and control observations within that leaf. This process is repeated across all the trees in the forest. Since each tree has different splits (each tree is grown using a random subsample of the data), each tree produces a slightly different estimate for the treatment effect.[13] The final causal forest prediction for the treatment effect for k is simply the average of the predictions across all the trees in the forest.[14]

[10]The data is split into 3 disjoint sets; a splitting set which is used to choose the splits, an estimation set which is used to estimate the treatment effects and a test set which is used to validate the estimates.

[11]A leaf in the context of causal trees is the final group of observations from the data that results from the tree making its series of splits. The observations in the same leaf have similar treatment effects.

[12]In Athey et al. (2019), causal forests are framed as a type of adaptive-kernel or locally weighted estimator. When estimating the treatment effect at a specific point $X = x_i$, more weight is given to observations that are "close" to X. This makes causal forests similar to methods like k-nearest neighbors, with the difference that instead of using distance-based kernels, causal forests define "closeness" based on how frequently two observations fall into the same leaf across the trees in the forest.

[13]Causal trees are weak learners i.e., they produce noisy estimates. By averaging the estimates from a large number of trees, causal forests provide stable estimates of treatment effects

[14]For estimating CATE using this procedure, two assumptions need to be satisfied. The first is "unconfoundedness", which stipulates that, depending on observable characteristics, the assignment of treatment is

After obtaining the CATE estimates, we create subgroups based on the strength (magnitude) of the predicted CATE, an idea similar to those found in Chernozhukov et al. (2017). We rank the observations and divide the data into four quartiles. It is important to note that these comparisons of predicted treatment effects are between different student observations, for example between student observation i and observation j. To ensure that our ranking is valid, we ensure that the model used for ranking does not see observation i or j in the training process. We then estimate the average effect of the treatment in each of the four quartiles (this is referred to as group average treatment effects or GATEs).

We use two different approaches to estimate the average treatment effect within each of the quartiles. The first approach uses a simple difference in means estimator (call this sample ATE) which is the average difference in outcomes for treated and control observations within each quartile. This is only valid in strictly randomized settings. The second is augmented inverse propensity score weighted average treatment effect (AIPW ATE). This method is reliable in both randomized and observational settings. This is because it uses observed data to predict outcomes, and then uses the probability of being assigned to treatment to re-weight the data to correct for confounding (or any issues with randomization). After estimating the group specific treatment effects, we test whether they are statistically different, while accounting for multiple hypothesis testing using the Romano-Wolf correction (Romano and Wolf, 2005a,b). Following this, We calculate the average covariate values between the quartiles to understand the difference in characteristics between those with the lowest predicted treatment effects compared to those with higher predicted treatment effects. Note that this procedure does not require us to pre-specify any subgroups, thus, our process of identifying students with different treatment effects, and assessing the heterogeneity across covariates between these individuals is data driven, and for that matter, "hands off". For robustness, we evaluate the quality of the estimates from our causal forest using the "test.calibration" function from the grf package, which was motivated by Chernozhukov et al. (2018).

random. The second is"overlap" which is the requirement that we observe the outcomes of enough treatment and control observations. Both assumptions are satisfied by virtue of our experimental design.

5 Results

In this section we present estimates of the treatment effects of the different mindset interventions on initial fall term grades, end of year 1 grades, subsequent year end grades, academic persistence, and timely graduation. Where the outcome is continuous (grades) we provide coefficient estimates from OLS regressions, and where the outcome is binary, we report coefficient estimates from linear probability models. All treatment effects are reported in the tables that follow. Alongside providing the average treatment effects for each of the different interventions (samples), we also provide pooled average treatment effect estimates for the full sample of students (i.e., the effect of whether a student was treated in any way). Lastly, we provide conditional average treatment effect estimates for all interventions from our treatment effect heterogeneity analysis.

5.1 Average Treatment Effects - Initial Fall Term Grades

Table 6 shows the average treatment effects of the different mindset interventions on initial fall term grades. The results indicate that there is no evidence of treatment effects from any of the different mindset interventions. Additionally we find no evidence of treatment effects when pooling the treatments, and analyzing the full sample of students.

5.2 Average Treatment Effects

Table 7 shows the average treatment effect estimates for all samples, and all interventions on year-end grades, and on academic persistence. The outcome variable "Persisted Year n" is simply a binary variable which takes a value of 1 if the student was treated in year 1, and was still observed in year n. "Graduated in 5 Years or Less" is a binary variable that takes a value of 1 if the student graduates in 5 years or less, and 0 otherwise. The results indicate that there were no significant treatment effects on year end grades, academic persistence, and on whether a student graduates timely, from any of the different mindset interventions. Additionally, we find no evidence of treatment effects when pooling all treatments. Overall,

we find no evidence of, immediate, mid-term or long-term treatment effects of the (social) belonging mindset interventions, no immediate or mid-term treatment effects of the economics growth, international belonging, and original belonging mindset interventions, on any of our academic outcomes.

5.3 Conditional Average Treatment Effects

Table 5 below shows the average covariate values between the four quartiles) for the full sample. This shows us the mean characteristics of individuals in the high and low treatment effect quartiles. A version of this table has been provided for each of the samples in the appendix. The variability of the covariates from this table is used to select the subgroups for conducting the subgroup analysis. Planned study hours, expected work hours (for pay), and the number of days a student took to start the initial online exercise were selected. These variables were split below and above median to create "low" vs. "high" groups . Additionally we create an indicator for first generation student (neither mother or father are university educated) and include that in the analysis.

5.4 Summary of Subgroup Results

6 Conclusion

In this paper, we present the results of a large-scale controlled experiment, conducted between 2015 and 2019 at two of Canada's largest publicly funded universities, designed to test the generalizability of two sets of soft-touch social psychology interventions, growth mindset and social belonging. Motivated by the promise that such inexpensive and scalable interventions have shown in similar contexts, in this experiment, incoming students enrolled in first-year economics courses participated in a pre-intervention survey, before being randomized into either treatment or control groups. Treatment groups across various campuses and years were subjected to a standard growth mindset intervention, a standard social belonging

intervention, and two context-specific interventions - the international belonging, and the economic growth mindset. The participating students in each group then completed an online exercise that was approximately an hour long. While the students in the control group received a variety of personality tests, those in each of the treatment groups participated in different types of exercises and received different pieces of information. Every effort was made to ensure that the student experience was similar between the treatment and control groups.

By design, the first-year economics courses at both U of T and York are very large courses for students not only economics but also from a diverse set of majors, and thus the entire sample of students who participated in the experiment are highly likely to be representative of the entire student body. Furthermore, the fact that students were incentivized using a grade requirement ensured that [XXX] of the students participated in both the survey and the subsequent online exercise. With UTSG being more representative of a typical highly ranked 4-year college from the USA, and all the other campuses being commuter campuses, we are confident that our sample of students is representative of the university student population in Canada. The response rate to the surveys, the time spent completing the online exercises, and the responses to open-ended questions indicate a strong student engagement in the study.

In line with many previous findings, we did not find evidence that any of the four interventions had an impact on short- or long-term student outcomes for the full sample of students. However, contrary to what some studies have previously found with regard to the effectiveness of these interventions for certain subgroups of students (e.g., females, racial minorities, low SES, first-generation students), we do not find evidence that the interventions were effective for any particular subgroups of students. In both our informed subgroup analysis (where we choose our subgroups based on what is conventionally done in the literature) and our casual forest-based treatment effect heterogeneity analysis (data-driven subgroup analysis), we found no evidence that any of the interventions had an impact on immediate or persistent student outcomes.

References

Akerlof, G. A. and Kranton, R. E. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of economic literature*, 40(4):1167–1201.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3):583–618.

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Aucejo, E. M. and Wong, K. (2025). The effect of feedback on student performance. *Journal of Public Economics*, 241:105274.

Beattie, G., Laliberté, J.-W. P., and Oreopoulos, P. (2018). Thrivers and divers: Using non-academic measures to predict college success and failure. *Economics of Education Review*, 62:170–182.

Bettinger, E., Ludvigsen, S., Rege, M., Solli, I. F., and Yeager, D. (2018). Increasing perseverance in math: Evidence from a field experiment in norway. *Journal of Economic Behavior & Organization*, 146:1–15.

Binning, K. R., Kaufmann, N., McGreevy, E. M., Fotuhi, O., Chen, S., Marshman, E., Kalender, Z. Y., Limeri, L., Betancur, L., and Singh, C. (2020). Changing social contexts to foster equity in college science courses: An ecological-belonging intervention. *Psychological Science*, 31(9):1059–1070.

Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, 78(1):246–263.

Bosch, N. (2021). Identifying supportive student factors for mindset interventions: A two-model machine learning approach. *Computers & education*, 167:104190.

Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., and Diemer, M. (2018). Reducing inequality in academic success for incoming college students: A randomized trial of growth mindset and belonging interventions. *Journal of Research on Educational Effectiveness*, 11(3):317–338.

Brougham, L. and Kashubeck-West, S. (2017). Impact of a growth mindset intervention on academic performance of students at two urban high schools. *Professional School Counseling*, 21(1):2156759X18764934.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436.

Campbell, K. L., Moore, J. B., and Bartholomew, J. B. (2020). The importance of publishing null results: Editorial guidelines to contribute to the reduction of publication bias in translational exercise research.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernández-Val, I. (2017). Fisher-schultz lecture: Generic machine learning inference on heterogenous treatment effects in randomized experiments, with an application to immunization in india. *arXiv preprint arXiv:1712.04802*.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I. (2018). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. Technical report, National Bureau of Economic Research.

Chrobak, B. (2024). Narrowing (achievement) gaps in higher education with a social-belonging intervention: A systematic review. *Educational Psychology Review*, 36(1):33.

Coleman, J. S. (1961). *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education*. Free Press, New York.

Damgaard, M. T. and Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, 64:313–342.

Davis, J. M. and Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: An application to summer jobs. *American Economic Review*, 107(5):546–550.

Delavande, A., Bono, E. D., Holford, A., Sen, S., and Lesic, V. (2023). Expectations about the productivity of effort and academic outcomes: Evidence from a randomized information intervention. *Available at SSRN 4483970*.

Dickersin, K., Chan, S., Chalmersx, T., Sacks, H., and Smith Jr, H. (1987). Publication bias and clinical trials. *Controlled clinical trials*, 8(4):343–353.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association*, 9(3):522–550.

Donnellan, M. B., Oswald, F. L., Baird, B. M., and Lucas, R. E. (2006). The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. *Psychological assessment*, 18(2):192.

Duckworth, A. L. and Quinn, P. D. (2009). Development and validation of the short grit scale (grit–s). *Journal of personality assessment*, 91(2):166–174.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Dweck, C. S. (2007). Boosting achievement with messages that motivate. *Education Canada*, 47(2):6–10.

Dweck, C. S. (2009). Mindsets: Developing talent through a growth mindset. *Olympic coach*, 21(1):4–7.

Easterbrook, P. J., Gopalan, R., Berlin, J., and Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, 337(8746):867–872.

Espenshade, T. J. and Radford, A. W. (2009). No longer separate, not yet equal. In *No Longer Separate, Not Yet Equal*. Princeton University Press.

Fanelli, D. (2010). "positive" results increase down the hierarchy of the sciences. *PloS one*, 5(4):e10068.

Goldacre, B. (2009). *Bad science*. Harper Perennial.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., and Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of educational psychology*, 106(2):375.

Hirsh, J. B. and Peterson, J. B. (2008). Predicting creativity and academic success with a "fake-proof" measure of the big five. *Journal of Research in Personality*, 42(5):1323–1333.

Hitsch, G. J., Misra, S., and Zhang, W. W. (2024). Heterogeneous treatment effects and optimal targeting policy evaluation. *Quantitative Marketing and Economics*, 22(2):115–168.

Ioannidis, J. P. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, 7(6):645–654.

Kaijanaho, A.-J. and Tirronen, V. (2018). Fixed versus growth mindset does not seem to matter much: A prospective observational study in two late bachelor level computer science courses. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, pages 11–20.

Kim, S., Yun, J., Schneider, B., Broda, M., Klager, C., and Chen, I.-C. (2022). The effects of growth mindset on college persistence and completion. *Journal of Economic Behavior & Organization*, 195:219–235.

Kozlov, M. (2024). So you got a null result. will anyone publish it? *Nature*, 631(8022):728–730.

LaCosse, J., Canning, E. A., Bowman, N. A., Murphy, M. C., and Logel, C. (2020). A social-belonging intervention improves stem outcomes for students who speak english as a second language. *Science advances*, 6(40):eabb6543.

Murphy, M. C., Gopalan, M., Carter, E. R., Emerson, K. T., Bottoms, B. L., and Walton, G. M. (2020). A customized belonging intervention improves retention of socially disadvantaged students at a broad-access university. *Science advances*, 6(29):eaba4677.

Murphy, M. C., Steele, C. M., and Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological science*, 18(10):879–885.

Office of Institutional Planning and Analysis, York University (2025). Undergraduate student headcount. Data Dashboard, Quick-Facts. Accessed 19 Jun 2025 via York University OIPA Quick-Facts.

Olson, K. R. and Dweck, C. S. (2008). A blueprint for social cognitive development. *Perspectives on Psychological Science*, 3(3):193–202.

Oreopoulos, P. and Petronijevic, U. (2023). The promises and pitfalls of using (mostly) low-touch coaching interventions to improve college student outcomes. *The Economic Journal*, 133(656):3034–3070.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., and Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological science*, 26(6):784–793.

Rattan, A., Savani, K., Chugh, D., and Dweck, C. S. (2015). Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives on Psychological Science*, 10(6):721–726.

Romano, J. P. and Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469):94–108.

Romano, J. P. and Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282.

Shrout, P. E. and Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual review of psychology*, 69(1):487–510.

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., and Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? two meta-analyses. *Psychological science*, 29(4):549–571.

Song, F., Parekh, S., Hooper, L., Loke, Y. K., Ryder, J., Sutton, A. J., Hing, C., Kwok, C. S., Pang, C., and Harvey, I. (2010). Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess*, 14(8):1–193.

Stephens, N. M., Fryberg, S. A., Markus, H. R., Johnson, C. S., and Covarrubias, R. (2012). Unseen disadvantage: how american universities' focus on independence undermines the academic performance of first-generation college students. *Journal of personality and social psychology*, 102(6):1178.

Stephens, N. M., Hamedani, M. G., and Destin, M. (2014). Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological science*, 25(4):943–953.

University of Toronto (2025). Quick facts. Accessed: 2025-06-06.

University of Toronto Office of Planning and Budget (2024). 2023–2024 enrolment report. Technical report, University of Toronto. Accessed via University of Toronto Planning & Budget website.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Walton, G., Murphy, M., Logel, C., Yeager, D., and Collaborative, T. (2017). The social-belonging intervention: A guide for use and customization. *Handbook of Wise Interventions. Guilford Press: New York*.

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, 23(1):73–82.

Walton, G. M. and Cohen, G. L. (2007). A question of belonging: race, social fit, and achievement. *Journal of personality and social psychology*, 92(1):82.

Walton, G. M. and Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, 331(6023):1447–1451.

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., and Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of educational Psychology*, 107(2):468.

Walton, G. M. and Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological review*, 125(5):617.

Warren, F., Mason-Apps, E., Hoskins, S., Devonshire, V., and Chanvin, M. (2019). The relationship between implicit theories of intelligence, attainment and socio-demographic factors in a uk sample of primary school children. *British Educational Research Journal*, 45(4):736–754.

Weaver, J. P., DeCaro, M. S., and Ralston, P. A. (2021). Limited support for use of a social-belonging intervention with first-year engineering students. *Journal for STEM Education Research*, 4:47–72.

Yeager, D. S. and Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational psychologist*, 47(4):302–314.

Yeager, D. S., Hanselman, P., Paunesku, D., Hulleman, C., Dweck, C., Muller, C., and Duckworth, A. (2018). Where and for whom can a brief, scalable mindset intervention improve adolescents' educational trajectories. *PsyArXiv Preprints*.

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., et al. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774):364–369.

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A., et al. (2016a). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of educational psychology*, 108(3):374.

Yeager, D. S. and Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of educational Research*, 81(2):267–301.

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., et al. (2016b). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences*, 113(24):E3341–E3348.

Yun, J., Kim, S., Schneider, B., and Lee, H. (2019). Challenges to causal attribution and estimation in randomized light-touch interventions: Learning from the spartan persistence project.

A Main Tables

Table 1: Random Assignment to Treatment and Control Group Across Year and Campus

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Year		2015	2017	2017	2017	2017	2017	2017	2019
Campus		All UofT	UTSG	UTSG	UTSC	UTSC	York	York	York
Sample		First Year	First Year & International	Not (First Year & International)	First Year & International	Not (First Year & International)	First Year & International	Not (First Year & International)	All
Original Belonging Mindset		972							
Economics Growth Mindset				628		398		898	
International Belonging Mindset			385		257		284		
Original Growth Mindset									1809
Control		1455	421	661	248	388	289	961	1893
Total	11947	2427	806	1289	505	786	573	1859	3702

Notes : This table displays the number of students enrolled in a first-year economics course assigned to each treatment by year and campus. UofT, University of Toronto (all campuses), UTSG, University of Toronto at St. George (downtown) campus; UTSC, University of Toronto at Scarborough campus; York, York University at Keele St. Campus.

Table 2: Descriptive Statistics and Balance Tests

	(1) Control Mean [Std. Dev]		(3) Original Belonging Mindset		(5) Economic Growth Mindset		(7) International Belonging Mindset		(9) Original Growth Mindset		(11) Obs.
Num. Days Since Sept 1. Started	9.22	[5.96]	0.006	[0.238]	0.041	[0.169]	-0.266	[0.234]	0.001	[0.188]	11,879
Male	0.53	[.50]	0.021	[0.021]	0.015	[0.016]	-0.006	[0.023]	-0.031	[0.016]*	11,703
English Mother Tongue	0.466	[.50]	0.017	[0.021]	0.031	[0.016]**	0.005	[0.016]	0.001	[0.017]	11,588
Age on Sept 1. of Experiment Year	19.25	[3.05]	0.024	[0.056]	-0.162	[0.107]	-0.102	[0.059]*	-0.102	[0.122]	11,588
HS Admission Grade Average	85.26	[7.73]	0.125	[0.304]	0.035	[0.253]	-0.435	[0.357]	0.143	[0.392]	7,595
First Generation Student	0.36	[.48]	0.02	[0.019]	-0.016	[0.015]	-0.003	[0.021]	0.019	[0.016]	11,872
Expects an Overall Grade of A	0.66	[.47]	0.019	[0.019]	-0.031	[0.016]**	-0.028	[0.019]	-0.008	[0.016]	11,860
Intends to Earn more than BA	0.63	[.48]	0.013	[0.019]	0.007	[0.015]	-0.017	[0.021]	-0.014	[0.016]	11,871
Hours/Week Expecting to Work	8.56	[10.08]	0.066	[0.402]	-0.052	[0.325]	-0.47	[0.430]	0.636	[0.350]*	11,814
Hours/Week Expecting to Study	15.82	[11.65]	0.237	[0.447]	-0.344	[0.387]	0.156	[0.544]	0.292	[0.371]	11,861
Think About Future Goals (1-7)	5.67	[1.23]	-0.001	[0.050]	0.037	[0.040]	-0.086	[0.052]*	0.023	[0.058]	10,030
Tend to Study at Last Minute (1-7)	4.07	[1.54]	-0.067	[0.064]	0.037	[0.050]	0.102	[0.070]	0.026	[0.051]	11,814
Difficult Transition to University (1-7)	4.58	[1.62]	-0.113	[0.069]	0.124	[0.052]**	0.05	[0.069]	0.008	[0.076]	10,030
Canadian	0.49	[.50]	0.009	[0.021]	0	[0.014]	0.013	[0.011]	0.021	[0.017]	11,806

Notes : This table presents means and standard deviations for key variables across control and treatment groups. Values in the square brackets show standard errors. Differences between each treatment category and the control group are reported with one and two asterisks indicating statistical significance at the 10% and 5%, respectively.

Table 3: Estimated Treatment Effects on Initial Fall Term Grades [0-100]

	Outcome Variable						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Missing Fall Grade	Fall Term Grade	Grade >50	Grade >60	Grade >70	Grade >80	Grade >90
Original Belonging Mindset	0.018	0.932	0.012	0.045	0.037	0.022	0.011
	[0.015]	[0.608]	[0.011]	[0.017]***	[0.022]*	[0.020]	[0.009]
Economics Growth Mindset	-0.005	-0.651	-0.006	-0.015	-0.009	0.009	0.009
	[0.009]	[0.593]	[0.011]	[0.014]	[0.017]	[0.013]	[0.006]
International Belonging Mindset	0.016	0.498	0.015	0.008	-0.022	0.006	0.008
	[0.011]	[0.833]	[0.015]	[0.020]	[0.023]	[0.020]	[0.010]
Original Growth Mindset	-0.001	0.481	0.009	0.001	0	0.005	0.015
	[0.008]	[0.603]	[0.012]	[0.015]	[0.017]	[0.014]	[0.007]**
Pooled Treatment Effect	0.004	0.186	0.005	0.005	0	0.009	0.011
	[0.005]	[0.328]	[0.006]	[0.008]	[0.010]	[0.008]	[0.004]***
Control Mean [Std. Dev]	0.96 [.294]	67.22 [17.02]	0.88 [.33]	0.76 [.43]	.51 [.50]	.22 [.41]	.03 [.18]
Sample Size	11,947	10,807	10,807	10,807	10,807	10,807	10,807

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1. Control means and standard deviations are also shown at the bottom. The values within the square brackets are standard errors. One, two and three asterisks indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 4: Estimated Treatment Effects on Persistent Academic Outcomes

	Outcome Variable							
	(1)	(2)	(3)	(4)	(5)	(6)	(6)	(8)
	Final Grade Year 1	Persisted Year 2	Final Grade Year 2	Persisted Year 3	Final Grade Year 3	Persisted Year 4	Final Grade Year 4	Graduated in 5 Years or Less
Original Belonging Mindset	0.431	-0.02	0.648	0.002	0.73	0.001	-0.012	0.012
	[0.552]	[0.013]	[0.563]	[0.016]	[0.549]	[0.017]	[0.567]	[0.021]
Economic Growth Mindset	-0.678	0.013	-0.684	-0.002	-0.149			
	[0.527]	[0.013]	[0.557]	[0.015]	[0.515]			
International Belonging Mindset	0.598	-0.007	-1.15	0.009	-0.104			
	[0.762]	[0.014]	[0.706]	[0.018]	[0.586]			
Original Growth Mindset	0.508							
	[0.518]							
Pooled Treatment Effect	0.113	-0.001	-0.412	0.002	0.119	0.001	-0.012	0.012
	[0.288]	[0.008]	[0.351]	[0.009]	[0.319]	[0.017]	[0.567]	[0.021]
Control Mean [Std. Dev.]	67.67 [15.24]	.85 [.35]	68.24 [14.48]	.76 [.43]	71.85 [12.95]	.78 [.41]	73.09 [11.18]	.65 [.48]
Sample Size	11,402	8,245	7,011	8,245	6,243	2,427	1,903	2,147

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1. Control means and standard deviations are also shown at the bottom. The values within the square brackets are standard errors.

Table 5: Estimated Treatment Effects on Final Grade Year 1 [0-100] in Subgroups

	Subgroups							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	High Panned Study Hours	Low Planned Study Hours	High Expected Work Hours	Low Expected Work Hours	High Days to Start Exercise	Low Days to Start Exercise	First Generation Student	Not First Generation Student
Original Belonging Mindset	-0.067	1.229	0.621	0.360	0.473	0.283	-0.067	0.704
	[0.664]	[0.974]	[0.805]	[0.754]	[0.651]	[1.014]	[1.021]	[0.652]
Economics Growth Mindset	0.075	-1.282	-0.872	-0.449	0.318	-1.400	-0.435	-0.903
	[0.732]	[0.749]*	[0.751]	[0.728]	[0.800]	[0.683]**	[0.899]	[0.648]
Original Growth Mindset	0.489	0.489	0.190	1.029	1.537	-0.453	0.187	0.829
	[0.785]	[0.684]	[0.663]	[0.834]	[0.783]**	[0.667]	[0.809]	[0.676]
International Belonging Mindset	1.490	-0.593	-0.401	1.430	0.971	0.207	2.677	-0.319
	[0.955]	[1.211]	[1.049]	[1.094]	[1.113]	[1.017]	[1.558]*	[0.855]
Pooled Treatment Effect	0.384	-0.164	-0.164	0.426	0.796	-0.624	0.250	0.031
	[0.386]	[0.425]	[0.397]	[0.419]	[0.401]**	[0.405]	[0.499]	[0.352]
Sample Size	5,664	5,630	6,018	5,276	5,874	5,420	4,056	7,238

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 across subgroups identified using differences in baseline characteristics based on conditional average effect estimates from the causal forest model. The values within the square brackets are standard errors. The control mean and standard deviation are also listed below. One, two and three asterisks indicate statistical significance at the 10%, 5% levels, and 1% levels respectively.

Table 6: Mean Characteristics of Low versus High Predicted Treatment Effect Students

	Quartiles of Predicted CATE Distribution			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Hours/Week Expecting to Study	12.114	14.505	17.620	19.013
Tends to Study at Last Minute (1-7)	4.569	4.201	3.819	3.753
Hours/Week Expecting to Work for Pay	8.977	9.057	7.711	8.399
Intends to Earn more than BA	0.598	0.644	0.633	0.648
Expects an Overall Grade Average of A	0.642	0.655	0.657	0.671
Male	0.543	0.537	0.516	0.517
English Mother Tongue	0.466	0.479	0.484	0.455
Canadian Citizen	0.513	0.518	0.514	0.495
Father University Educated	0.521	0.566	0.580	0.583
Mother University Educated	0.458	0.479	0.488	0.511
Age on Sept 1. of Experiment Year	19.136	19.345	19.285	19.087
Num. Days Since Sept 1. Started Exercise	6.823	8.049	9.482	11.841

Notes : This table shows the mean baseline characteristics of students in the different quartiles of the predicted (pooled) treatment effect distribution. The predicted treatment effects were obtained from a causal forest model trained on baseline covariates and final grade at the end of the first year as the outcome.

B Intuitive Subgroup Tables

Table 7: Pooled Social Belonging Interventions

	Male	Canadian	ESL	1st Gen
Treated	0.101	1.321*	0.103	1.123
	[0.753]	[0.720]	0.744	[0.891]
Num.Obs.	1876	1189	1355	1276

Notes : This table reports estimated pooled treatment effects for the two belonging interventions separately for four subgroups selected intuitively informed by existing literature. Each column presents results from a regression of the average final grade at the end of the first year on the treatment indicator plus campus-year fixed effects for each subgroup. Values within the square brackets are standard errors. One asterisk indicates statistical significance at the 10% level.

Table 8: Pooled Growth Mindset Interventions

	Male	Canadian	ESL	1st Gen	HS AG
Treated	-0.002	0.135	-0.038	0.721	1.071*
	(0.612)	(0.499)	(0.565)	(0.720)	(0.529)
Num.Obs.	2785	3697	3054	1895	2418

Notes : This table reports estimated pooled treatment effects for the two growth mindset interventions separately for four subgroups selected intuitively informed by existing literature. Each column presents results from a regression of the average final grade at the end of the first year on the treatment indicator plus campus-year fixed effects for each subgroup. Values within the square brackets are standard errors. Two asterisks indicates statistical significance at the 5% level.

Table 9: All Interventions Pooled - No Internationals

	Male	Canadian	ESL	1st Gen	HS AG
Treated	-0.137	0.319	0.021	0.505	0.941*
	(0.533)	(0.431)	(0.484)	(0.609)	(0.432)
Num.Obs.	3623	4644	3956	2526	3498

Notes : This table reports estimated pooled treatment effects for all interventions separately for four subgroups selected intuitively informed by existing literature. Each column presents results from a regression of the average final grade at the end of the first year on the treatment indicator plus campus-year fixed effects for each subgroup. Values within the square brackets are standard errors. One asterisk indicates statistical significance at the 5% level. The sample for these results exclude all international students.

Table 10: All Interventions Pooled - Full Sample

	Male	Canadian	ESL	1st Gen
Treated	-0.356	0.201	-0.076	0.026
	(0.424)	(0.395)	(0.422)	(0.616)
Num.Obs.	5967	5761	5321	2695

Notes : This table reports estimated pooled treatment effects for all interventions separately for four subgroups selected intuitively informed by existing literature. Each column presents results from a regression of the average final grade at the end of the first year on the treatment indicator plus campus-year fixed effects for each subgroup. Values within the square brackets are standard errors.The results are for the full sample of students.

C Descriptive Statistics and Balance Tests for Other Samples

Table 11: Descriptive Statistics and Balance Test (Full Sample - No Internationals)

	(1)	(2)	(3)	(4)	(5)
	Control Mean		**Pooled Treatment**		**Sample Size**
	[Std. Dev]		**Difference**		
Num. Days Since Sept 1. Started Exercise	10.924	[.140]	0.011	[0.134]	6,593
Male	0.524	[0.017]	-0.002	[0.012]	6,593
English Mother Tongue	0.559	[0.017]	0.022	[0.012]*	6,593
Age on Sept 1 of Experiment Year	18.581	[0.058]	-0.026	[0.045]	6,593
High School Admission Grade Average	89.613	[0.210]	0.146	[0.183]	6,593
First Generation Student	0.243	[0.015]	0.001	[0.012]	6,593
Expects an Overall Grade Average of A	0.640	[0.017]	-0.005	[0.012]	6,593
Intends to Earn More Than BA	0.707	[0.016]	0.008	[0.012]	6,593
Hours/Week Expecting to Work for Pay	5.618	[0.290]	0.019	[0.232]	6,593
Hours/Week Expecting to Study	18.095	[0.425]	0.309	[0.282]	6,593
Tend to Study at Last Minute (1-7)	4.093	[0.053]	-0.031	[0.038]	6,593
Canadian Student	0.692	[0.016]	0.026	[0.011]**	6,593

Notes : This table presents means and standard deviations for key variables across control and treatment groups excluding international students. Values in the square brackets show standard errors. Difference between the treatment and control group are reported with asterisks. One and two asterisks indicate statistical significance at the 10% and 5% levels, respectively.

Table 12: Descriptive Statistics and Balance Tests (Growth Sample)

	(1)	(2)	(3)	(4)	(5)
	Control Mean		Pooled Treatment		Sample
	[Std. Dev]		Difference		Size
Num. Days Since Sept 1. Started Exercise	8.21	[5.75]	-0.089	[0.155]	4,799
Male	0.584	[0.493]	-0.007	[0.014]	4,799
English Mother Tongue	0.63	[0.483]	0.014	[0.014]	4,799
Age on Setp. 1 of Experiment Year	18.87	[2.07]	-0.046	[0.059]	4,799
High School Admission Grade Average	84.14	[8.30]	0.136	[0.224]	4,799
First-Generation Student	0.399	[0.488]	-0.008	[0.014]	4,799
Expects an Overall Grade Average of A	0.602	[.490]	-0.009	[0.014]	4,799
Intends to Earn More than BA	0.607	[0.488]	0.012	[0.014]	4,799
Hours/Week Expecting to Work for Pay	8.452	[9.486]	0.222	[0.274]	4,799
Hours/Week Expecting to Study	14.865	[11.654]	0.482	[0.336]	4,799
Tend to Study at Last Minute (1-7)	4.138	[1.516]	-0.011	[0.045]	4,799
Canadian Student	0.761	[.426]	0.020*	[0.012]	4,799

Notes : This table presents means and standard deviations for key variables across control and treatment groups for the students in either of the two growth mindset interventions. Values in the square brackets show standard errors. Difference between the treatment and control group are reported with asterisks. One asterisk indicates statistical significance at the 10% level.

Table 13: Descriptive Statistics and Balance Tests (Belonging Sample)

	(1)	(2)	(3)	(4)	(5)
	Control Mean		Pooled Treatment		Sample
	[Std. Dev]		Difference		Size
Num. Days Since Sept 1. Started Exercise	10.404	[5.487]	-0.065	[0.164]	4,154
Male	0.448	[0.497]	0.011	[0.016]	4,154
English Mother Tongue	0.335	[0.472]	0.011	[0.014]	4,154
Age on Sept 1. of Experiment Year	18.345	[1.277]	-0.031	[0.039]	4,154
First-Generation Student	0.304	[.460]	0.004	[0.014]	4,154
Difficulty in Transitioning to University (1-7)	4.71	[1.610]	-0.028	[0.050]	4,154
Future Goals (1-7)	5.646	[1.160]	-0.045	[0.037]	4,154
Expects an Overall Grade Average of A	0.734	[0.442]	0.004	[0.014]	4,154
Intends to Earn More Than BA	0.686	[4.643]	-0.004	[0.014]	4,154
Hours/Week Expecting to Work for Pay	7.333	[9.660]	-0.290	[0.296]	4,154
Hours/Week Expecting to Study	17.601	[11.281]	0.275	[0.354]	4,154
Tend to Study at Last Minute (1-7)	3.997	[1.531]	-0.014	[0.048]	4,154
Canadian Student	0.297	[0.457]	0.012	[0.013]	4,154

Notes : This table presents means and standard deviations for key variables across control and treatment groups for the students in either of the two belonging interventions. Values in the square brackets show standard errors. Difference between the treatment and control group are reported with asterisks.

D Average Treatment Effects for Other Samples

Table 14: Treatment Effects on Overall First-Year Grades (Full Sample - No Internationals)

	Outcome Variable	
	(1)	**(2)**
	Final Grade Year 1	**Final Grade Year 1**
Original Belonging Mindset	0.503	0.205
	[0.653]	[0.570]
Economics Growth Mindset	-0.192	-0.391
	[0.588]	[0.523]
Original Growth Mindset	1.030	0.789
	[0.703]	[0.621]
Pooled Treatment Effect	0.394	0.123
	[0.372]	[0.331]
Controls?	**N**	**Y**
Control Mean [Std. Dev]	67.763 [14.78]	67.763 [14.78]
Sample Size	6,593	6,593

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 for the full sample of students excluding international students. The control mean and standard deviation are also listed below. The values within the square brackets are standard errors.

Table 15: Treatment Effects on Overall First-Year Grades (Growth Sample)

	Outcome Variable	
	(1)	(2)
Original Belonging Mindset		
Economics Growth Mindset	-0.192	-0.411
	[0.588]	[0.523]
International Belonging Mindset		
Original Growth Mindset	1.030	0.836
	[0.703]	[0.620]
Pooled Treatment Effect	0.31	0.101
	[0.451]	[0.400]
Controls?	Y	N
Sample Size	4,799	4,799

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 for the the students in either of the two growth interventions. The control mean and standard deviation are also listed below. The values within the square brackets are standard errors.

Table 16: Treatment Effects on Overall First Year Grades (Belonging Sample)

	Outcome Variable	
	(1)	(2)
	Final Grade Year 1	Final Grade Year 1
Original Belonging Mindset	0.431	0.281
	[0.552]	[0.524]
Economics Growth Mindset	-	-
Original Growth Mindset	-	-
International Belonging Mindset	0.598	0.518
	[0.762]	[0.669]
Pooled Treatment Effect	0.506	0.386
	[0.457]	[0.437]
Controls?	N	Y
Control Mean [Std. Dev]	67.623 [14.812]	67.623 [14.812]
Sample Size	4,154	4,154

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 for the the students in either of the two belonging interventions. The control mean and standard deviation are also listed below. The values within the square brackets are standard errors.

E CATE Density Plots for Other Samples

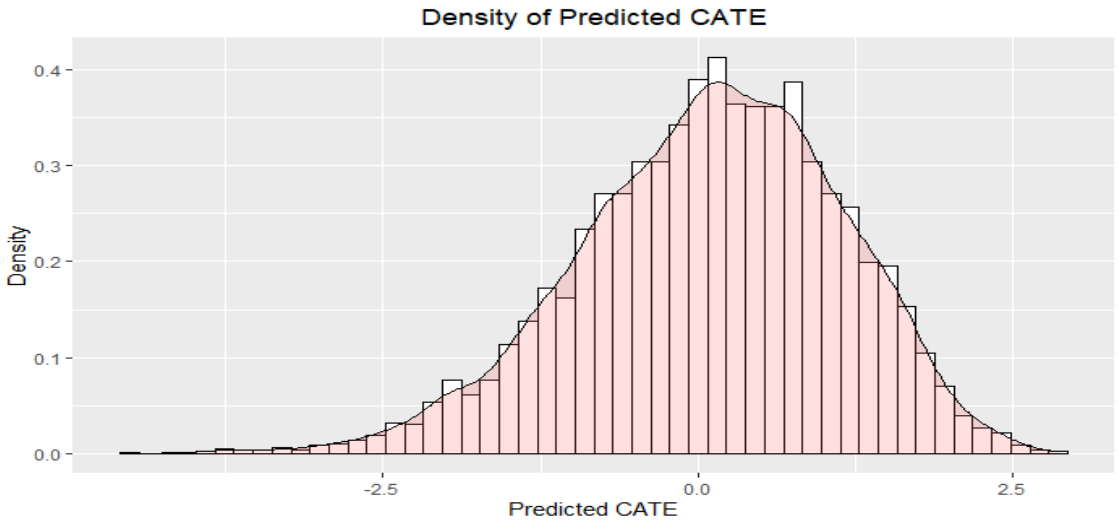


Figure 1: The Density of Predicted CATE for the Full Sample

Notes: This figure displays the distribution of predicted conditional average treatment effects across the full sample.Values are binned for clarity.

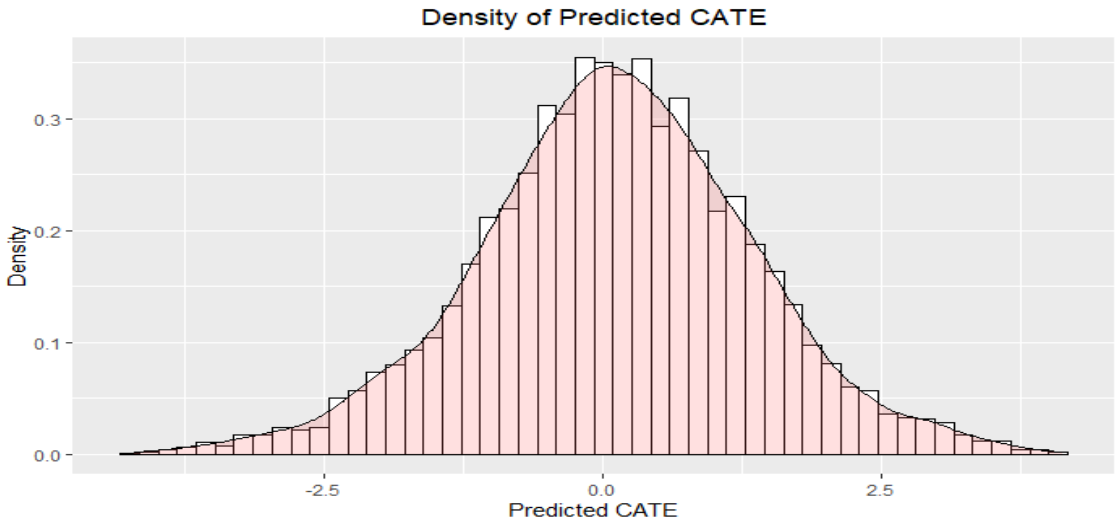


Figure 2: The Density of Predicted CATE for the Growth Sample

Notes: This figure displays the distribution of predicted conditional average treatment effects for the sample of students assigned to either of the two growth mindset interventions.Values are binned for clarity.

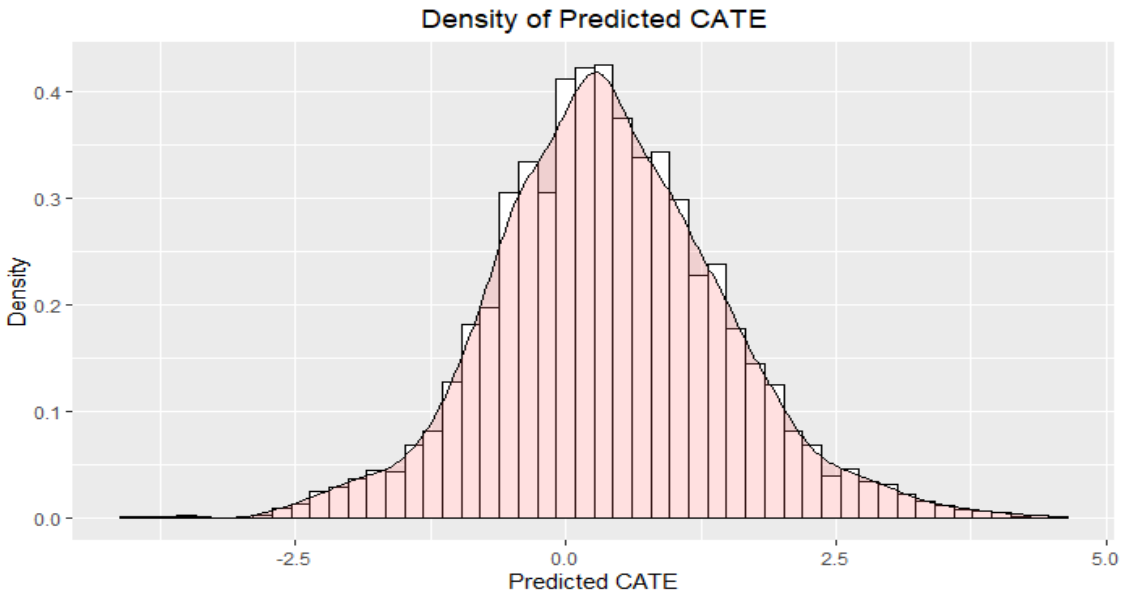


Figure 3: The Density of Predicted CATE for the Belonging Sample

Notes: This figure displays the distribution of predicted conditional average treatment effects for the sample of students assigned to either of the two belonging interventions.Values are binned for clarity.

F Sample ATE and AIPW ATE Quartile Plots for Other Samples

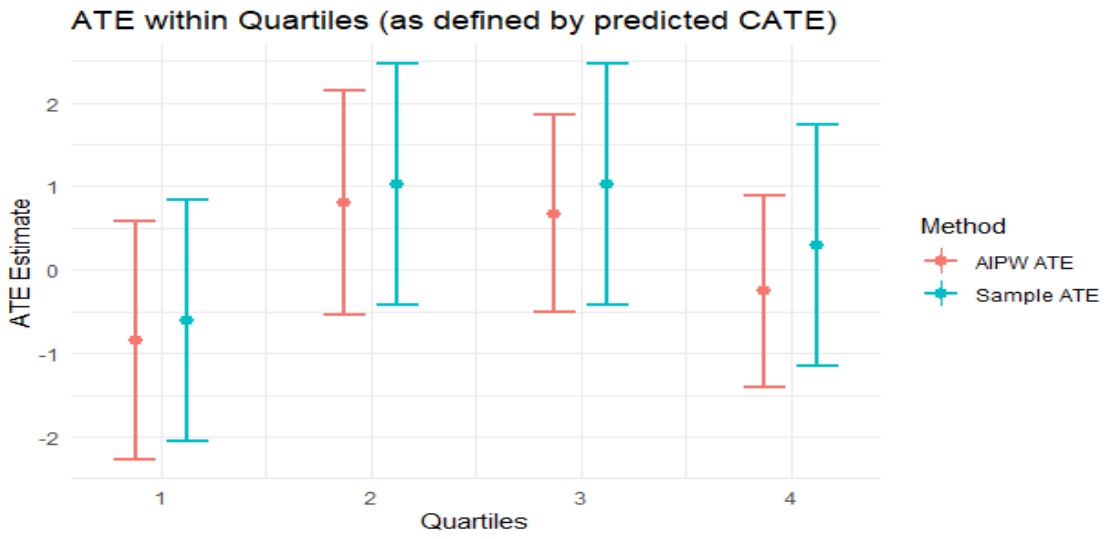


Figure 4: Sample ATE vs AIPW ATE Quartile Plot for the Full Sample - No Internationals

Notes : This figure shows the average pooled treatment effect estimates within each quartile for the full sample excluding international students estimated using two methods. AIPW, augmented inverse propensity weighting; Sample ATE, regular average treatment effect. Quartile 4 represents those with the highest predicted treatment effects from a causal forest model. Monotonicity in this figure provides evidence for the existence of treatment effect heterogeneity.

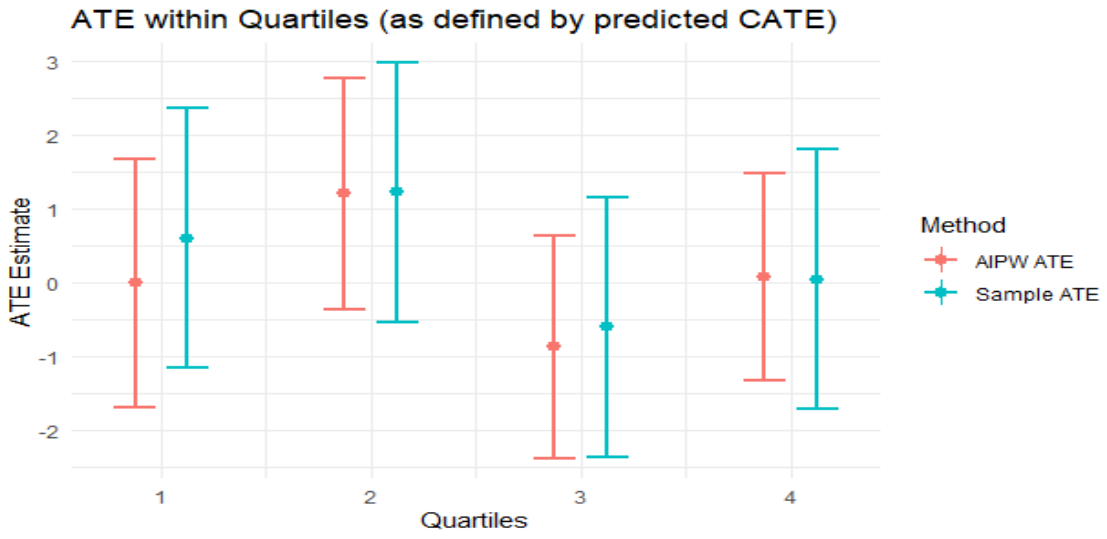


Figure 5: Sample ATE vs AIPW ATE Quartile Plot for the Growth Sample

Notes : This figure shows the average pooled treatment effect estimates within each quartile for the sample of students assigned to either of the two growth mindset interventions estimated using two methods. AIPW, augmented inverse propensity weighting; Sample ATE, regular average treatment effect. Quartile 4 represents those with the highest predicted treatment effects from a causal forest model. Monotonicity in this figure provides evidence for the existence of treatment effect heterogeneity.

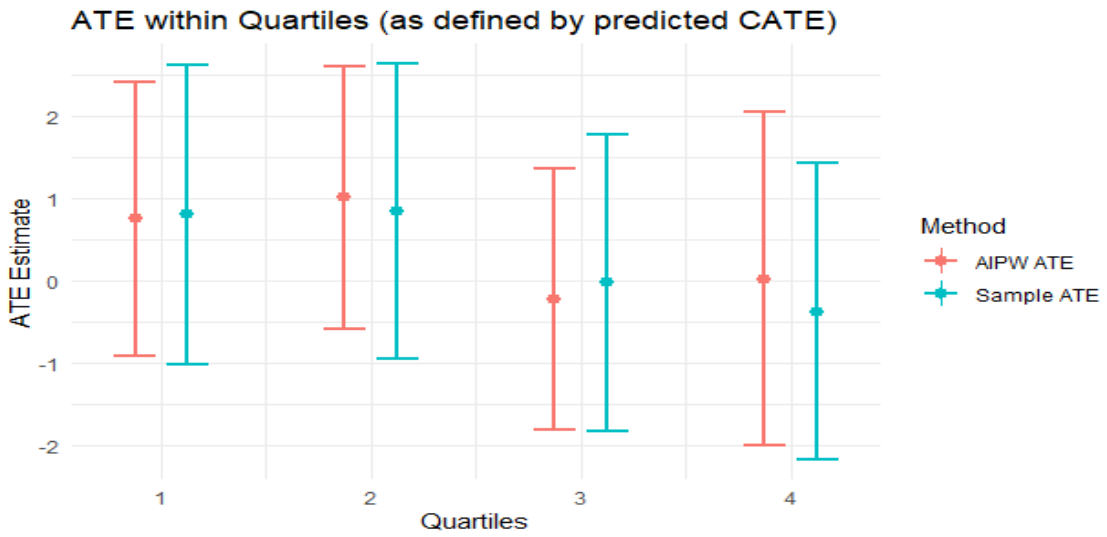


Figure 6: Sample ATE vs AIPW ATE Quartile Plot for the Belonging

Notes : This figure shows the average pooled treatment effect estimates within each quartile for the sample of students assigned to either of the two belonging interventions estimated using two methods. AIPW, augmented inverse propensity weighting; Sample ATE, regular average treatment effect. Quartile 4 represents those with the highest predicted treatment effects from a causal forest model. Monotonicity in this figure provides evidence for the existence of treatment effect heterogeneity.

G Mean Characteristics Between Quartiles and Subgroup Treatment Effects for Other Samples

Table 17: Mean Characteristics between Quartiles (Full Sample - No Internationals)

	Quartile of Estimated CATE Distribution				
Variables	**Quartile 1**	**Quartile 2**	**Quartile 3**	**Quartile 4**	**Observations**
Hours/Week Expecting to Study	12.816	15.110	16.758	19.302	6,593
Tend to Study at Last Minute (1 - 7)	4.071	4.063	4.104	4.110	6,593
Hours/Week Expecting to Work for Pay	8.587	8.904	8.192	6.959	6,593
Intends to Earn more than BA	0.591	0.611	0.632	0.656	6,593
Expects an Overall Grade Average of A	0.585	0.625	0.621	0.648	6,593
Male	0.618	0.538	0.533	0.510	6,593
English Mother Tongue	0.654	0.620	0.568	0.558	6,593
Canadian Citizen	0.767	0.711	0.687	0.652	6,593
Mother University Educated	0.393	0.461	0.478	0.516	6,593
Father University Educated	0.466	0.512	0.552	0.607	6,593
Age on Sept 1 of Experiment Year	18.878	18.740	18.520	18.423	6,593
Num. Days Since Sept 1. Started Exercise	7.316	7.888	9.612	11.652	6,593
High School Admission Grade Average	80.781	84.346	86.322	88.375	6,593

Notes : This table shows the mean baseline characteristics of students in the different quartiles of the predicted treatment effect distribution for the sample of students excluding international students. The predicted treatment effects were obtained from a causal forest model trained on baseline covariates and final grade at the end of the first year as the outcome.

Table 18: Treatment Effects on Year 1 Grade by Subgroups (Full Sample - No Internationals)

	Subgroups							
	(1) High HS Gradw	(2) Low HS Grade	(3) High Planned Study Hours	(4) Low Planned Study Hours	(5) Male	(6) Female	(7) First Generation Student	(8) Not First Generation Student
Original Belonging Mindset	0.947	-0.601	0.242	0.857	-0.665	1.628	-0.131	0.897
	[0.721]	[1.131]	[0.783]	[1.147]	[1.068]	[0.764]**	[1.114]	[0.802]
Economics Growth Mindset	0.644	-1.631	0.323	-0.604	-0.656	0.486	0.521	-0.728
	[0.653]	[0.929]*	[0.802]	[0.847]	[0.814]	[0.836]	[0.981]	[0.729]
Original Growth Mindset	1.502	0.887	0.607	0.907	0.856	1.013	0.964	1.163
	[0.951]	[0.850]	[1.048]	[0.926]	[0.929]	[1.054]	[1.061]	[0.933]
Pooled Treatment Effect	0.920	-0.416	0.366	0.223	-0.151	0.979	0.513	0.513
	[0.432]**	[0.553]	[0.499]	[0.551]	[0.532]	[0.510]*	[0.609]	[0.609]
Sample Size	3,338	3,255	3,323	3,270	3,623	2,970	2,526	4,067

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 across subgroups identified using differences in baseline characteristics based on conditional average effect estimates from the causal forest model for the sample excluding international students. The values within the square brackets are standard errors. One and two asterisks indicate statistical significance at the 10% and 5% levels respectively.

Table 19: Mean Characteristics between Quartiles (Growth Sample)

	Quartile of Predicted CATE Distribution			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Hours/Week Expecting to Study	12.814	14.702	15.335	17.563
Tend to Study at Last Minute (1-7)	4.181	4.109	4.106	4.138
Hours/Week Expecting to Work for Pay	7.640	8.769	9.478	8.349
Intends to Earn More than BA	0.586	0.622	0.619	0.625
Expects an Overall Grade Average of A	0.569	0.613	0.616	0.590
Male	0.631	0.559	0.555	0.576
English Mother Tongue	0.672	0.649	0.630	0.595
Canadian Citizen	0.813	0.766	0.778	0.725
Mother University Educated	0.382	0.451	0.481	0.489
Father University Educated	0.459	0.497	0.539	0.568
Age on Sept. 1 Started Exercise	18.839	18.833	18.844	18.875
Num. Days Since Sept. 1 Started Exercise	6.667	7.028	7.626	11.411
High School Admission Grade Average	81.042	84.050	85.270	86.493

Notes : This table shows the mean baseline characteristics of students in the different quartiles of the predicted treatment effect distribution for the sample of students in either of the two growth mindset interventions. The predicted treatment effects were obtained from a causal forest model trained on baseline covariates and final grade at the end of the first year as the outcome.

Table 20: Treatment Effects on Final Grade Year 1 by Subgroup (Growth Sample)

	Subgroups							
	(1)	(2)	(3)	(4)	(5)	(6)	(6)	(8)
	High HS Grade	Low HS Grade	High Planned Study Hours	Low Planned Study Hours	First Generation Student	Not First Generation Student	High Expected Work Hours	Low Expected Work Hours
Original Belonging Mindset	-	-	-	-	-	-	-	-
Economics Growth Mindset	0.524	-1.731	0.313	-0.677	0.521	-0.728	-0.844	0.365
	[0.627]	[1.001]*	[0.738]	[0.927]	[0.981]	[0.729]	[0.881]	[0.773]
Original Growth Mindset	2.136	0.479	0.896	0.777	0.964	1.163	0.013	2.441
	[0.976]**	[0.866]	[1.001]	[0.964]	[1.061]	[0.933]	[0.941]	[1.044]***
Pooled Treatment Effect	1.027	-0.601	0.527	-0.013	0.721	-0.002	-0.457	1.127
	[0.528]*	[0.660]	[0.594]	[0.669]	[0.720]	[0.575]	[0.643]	[0.622]*
Sample Size	2,415	2,384	2,480	2,319	1,895	2,904	2,480	2,319

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 across subgroups identified using differences in baseline characteristics based on conditional average effect estimates from the causal forest model for the sample of students in either of two growth mindset interventions. The values within the square brackets are standard errors. One, two and three, asterisks indicate statistical significance at the 10%, 5% and 1% levels respectively.

Table 21: (Mean Characteristics between Quartiles (Belonging Sample)

	Quartile of Predicted CATE Distribution			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Hours/Week Expecting to Study	12.939	16.620	19.226	21.725
Tend to Study at Last Minute (1-7)	4.470	3.893	3.892	3.725
Future Goals (1-7)	5.462	5.535	5.639	5.869
Difficulty Transitioning to University (1-7)	4.775	4.711	4.673	4.654
Hours/Week Expecting to Work for Pay	7.572	7.487	6.517	7.355
Intends to Earn more than BA	0.660	0.683	0.687	0.703
Expects an Overall Grade Average of A	0.707	0.767	0.731	0.750
Male	0.478	0.428	0.430	0.470
English	0.337	0.316	0.363	0.289
Canadian Citizen	0.288	0.301	0.328	0.228
Mother University Educated	0.502	0.547	0.554	0.520
Father University Educated	0.604	0.634	0.649	0.618
Age on Sept 1. of Experiment Year	18.437	18.209	18.204	18.545
Num. Days Since Sept 1. Started Exercise	10.474	9.976	10.837	9.595

Notes : This table shows the mean baseline characteristics of students in the different quartiles of the predicted treatment effect distribution for the sample of students in either of the two belonging interventions. The predicted treatment effects were obtained from a causal forest model trained on baseline covariates and final grade at the end of the first year as the outcome.

Table 22: Treatment Effects on Final Grade Year 1 by Subgroup (Belonging Sample)

	Subgroups							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	High Planned Study Hours	Low Planned Study Hours	High Planned Work Hours	Low Planned Work Hours	High Days to Start Ex.	Low Days to Start Ex.	First Generation Student	Not First Generation Student
Original Belonging Mindset	0.175	0.688	0.585	0.430	0.508	0.292	-0.067	0.704
	[0.682]	[0.916]	[0.774]	[0.785]	[0.683]	[0.933]	[1.022]	[0.653]
Economics Growth Mindset								
Original Growth Mindset								
International Belonging Mindset	1.500	-0.390	0.253	0.927	1.081	0.307	2.677	-0.319
	[0.993]	[1.147]	[1.007]	[1.157]	[1.555]	[0.848]	[1.559]*	[0.855]
Pooled Treatment Effect	0.711	0.153	0.430	0.642	0.650	0.301	1.123	0.243
	[0.571]	[0.733]	[0.625]	[0.668]	[0.642]	[0.640]	[0.891]	[0.526]
Sample size	2,273	1,881	2,168	1,986	2,109	2,045	1,276	2,878

Notes : The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus campus-year fixed effects as described in table 1 across subgroups identified using differences in baseline characteristics based on conditional average effect estimates from the causal forest model for the sample of students in either of two belonging interventions. The values within the square brackets are standard errors. One asterisk indicates statistical significance at the 10% level.