Self-Esteem and Procrastination: Evidence from a Field Experiment

Andrea Burro*

This version: September 13, 2026

Click here for the most recent version

Abstract

This study presents causal evidence on the effect of self-esteem on procrastination, using a field experiment with college students. Leveraging a novel high-frequency dataset from a learning app, I randomly assign students to a self-esteem booster, a goal-setting intervention, or a control group, with treatments delivered through weekly notifications. I find that students who receive the self-esteem booster procrastinate less, while the goal-setting intervention has no effect. This effect persists in the months after the intervention, supporting a habit-formation mechanism. The results differ by gender: the self-esteem booster reduces procrastination among male students but reduces study effort among female students. Finally, students in the self-esteem booster group also report higher self-esteem and lower procrastination in surveys.

*Department of Economics, University of St. Gallen. E-mail: andrea.burro@unisg.ch.
I am grateful to Beatrix Eugster for extensive discussions and her mentorship. I thank Palaash Bhargava, Luca Braghieri, Giovanni Burro, Bruno Caprettini, Michela Carlana, Matteo Fossi, Vincenzo Galasso, Sam Kang, John List, Paolo Piacquadio, Julia Seither, Sofia Shchukina, Sofía Sierra Vásquez, Giuseppe Sorrenti, and Jeremia Stalder; and seminar participants at the University of St. Gallen, European Economic Review Summer School, Chicago School in Experimental Economics, Advances with Field Experiment Conference, North-American Economic Science Association Conference, University of Chicago Booth Behavioral Economics Lab Meeting, University of Chicago Experimental Meeting, SKILS seminar, University of Chicago Behavioral Insights and Parenting Lab, LEAP lunch seminar, Milan PhD Economics Workshop, Swiss Society of Economics and Statistics Congress, European Economic Association Congress, European Association for Labour Economists Conference, European Economic Science Association Meeting, Swiss Education Network Workshop for helpful comments and suggestions. The study was funded under the "Green Box programme", which is financed by the University of St. Gallen. The experiment was pre-registered on the AEA RCT Registry on September 19, 2024 (RCT ID: AEARCTR-0014369) and approved by the Ethics Committee of the University of St. Gallen on May 31, 2024 (reference number: HSG-EC-202240508).

1 Introduction

In everyday life, people procrastinate on many tasks: writing a dissertation, going to the gym, paying taxes. Several studies have investigated how people *"put off an impending task to a later time even when such practice results in counterproductive and needless delay"* (Schraw, Wadkins, and Olafson 2007).

Universities are an ideal setting to study procrastination. Approximately 46% of college students identify as serious procrastinators (Solomon and Rothblum 1984), with 80% to 95% having experienced procrastination at some point in their academic lives and 50% regularly procrastinating on academic tasks (Steel 2007). Procrastinators are also more likely to have mental health issues and worse academic performance, although evidence mostly comes from correlational studies (De Paola and Scoppa 2015; Xu 2021; Arnold 2023; De Paola, Gioia, and Scoppa 2023). Understanding how to reduce procrastination therefore has direct implications for productivity and human capital accumulation.

Procrastination has traditionally been studied as a problem of self-control, in which people procrastinate because they undervalue the potential future benefits of completing an unpleasant task compared to the immediate costs of doing it. This has been modeled in the form of quasi-hyperbolic discounting in β-δ models (e.g., Laibson 1997; O'Donoghue and Rabin 1999; Hyndman and Bisin 2022) or as a dual-self problem (Fudenberg and Levine 2006). These models suggest that commitment devices, such as binding deadlines and goal setting, should mitigate procrastination. However, empirical evidence on these interventions is mixed, suggesting that commitment devices alone may be insufficient to overcome self-control problems (Burger, Charness, and Lynham 2011; Bisin and Hyndman 2020; John 2020).

Research in psychology has shown a link between self-esteem and procrastination. In particular, procrastination can be used as a self-handicapping strategy (Ferrari and Tice 2000). When individuals face tasks whose outcomes reflect on their ability, delaying effort can be used to protect self-image: failure following low effort is attributed to a lack of effort rather than a lack of ability. This mechanism predicts that interventions that boost self-esteem should reduce self-handicapping procrastination.

In this paper, I test these predictions by estimating the causal effect of self-esteem on procrastination through a field experiment with college students. I leverage a novel dataset from a learning app to measure procrastination and combine it with university administrative data and survey data. Students in the experiment are randomized into three groups: the *Booster* group, the *Goals* group, and the *Control* group. Students in the *Booster* group receive weekly notifications containing positive messages about their study time. Students in the *Goals* group receive notifications about setting goals for their study time. Finally, students in the *Control* group do not receive additional notifications.

I find that boosting self-esteem increases effort in the app. The *Booster* treatment increases the number of correct answers, total answers, and distinct questions answered. On the other hand, the *Goals* treatment has no significant effect on usage. Looking at outcomes one semester after the intervention, students in the *Booster* group continue to exhibit higher effort, suggesting a habit-formation mechanism. Heterogeneity analysis reveals that the effect of the self-esteem booster is driven by male students, with female students reacting to the treatment by reducing their study effort. These results are in line with previous interventions in education (Smithers 2015; Clark et al. 2020; Delavande et al. 2023; Campos-Mercade, Thiemann, and Wengström 2025).

The main contribution of this paper is that, to my knowledge, it is the first study to estimate the causal effect of self-esteem on procrastination. Previous literature relied on correlational measures (e.g., Lay 1986; Tice and Baumeister 1997; Schuenemann et al. 2022), which may yield biased estimates due to reverse causality if higher procrastination reduces self-esteem. I address this by boosting self-esteem in a controlled experiment. This design relates to interventions by Hall, Zhao, and Shafir (2014), who find positive effects of self-esteem on cognitive performance, and Bursztyn et al. (2018), where increasing self-esteem reduces the demand for status goods.

I also contribute to the literature on self-control problems, which are prevalent across different domains. DellaVigna and Malmendier (2006) examine gym members; Cadena et al. (2011) and Kaur, Kremer, and Mullainathan (2015) study workers; Thaler and Benartzi (2004) and Ashraf, Karlan, and Yin (2006) analyze the phenomenon in the context of savings, and Shah and Mukherjee (2026) in that of retirement planning. I extend this work by showing that boosting self-esteem can serve as an effective substitute for commitment devices[1].

1. Clark et al. (2020) and Van Lent and Souverijn (2020) find positive effects of goal-setting interventions on academic performance. Similarly, Ariely and Wertenbroch (2002) observe improved performance among MBA students with external deadlines. However, the results do not replicate in a recent lab experiment (Hyndman and Bisin 2026), and non-binding target dates appear to work better than binding ones at nudging early completion (Dirkmaat et al. 2023). Burger, Charness, and Lynham (2011) and Campos-Mercade, Thiemann, and Wengström (2025) and Campos-Mercade and Wengström (2025) find negative effects of using commitment devices on performance. More generally, setting goals and providing feedback have different effects that depend on their design and on individuals' beliefs (Aulagnon et al. 2025; Strulov-Shlain and Wellsjo 2025; Abel et al. 2026; Brade, Himmler, and Jäckle 2026).

The granularity of the data allows me to estimate the evolution of the treatment effect over time. This is particularly relevant for procrastination, given its dynamic nature. In a lab experiment, recording behaviors over long periods captures these dynamics (e.g., Augenblick and Rabin 2019; Fedyk 2024; Cerrone et al. 2026), but requires substantial resources and increases experimental fatigue. Using field data, Burger, Charness, and Lynham (2011) track study patterns over a semester but require monitored library logins, which could introduce demand effects. By contrast, participants in my setting likely perceive the notifications as a standard app update, helping to mitigate such effects. In addition, I observe repeated measures of study behaviors, while other field experiments were conducted in settings in which individuals can only procrastinate once (e.g., submitting a term paper in Ariely and Wertenbroch 2002 or completing an experimental task in Bisin and Hyndman 2020).

Finally, I examine both the short-run effects during the intervention and the long-run persistence one semester later, after all students had switched back to the standard version of the app. While previous studies rarely estimate long-term effects, my setting allows me to shed light on the mechanisms of habit formation, similar to the approach of Giuntella, Saccardo, and Sadoff (2026) to study the effect of sleep on academic performance.

I provide a comprehensive analysis of student productivity along two dimensions. First, focusing only on traditional outcomes (e.g., grades or dropouts) would provide partial evidence, given the role that effort plays in academic performance. In this sense, my study is similar to Cotton et al. (2026), Gneezy et al. (2019), and Ersoy (2023), which focus on students' effort rather than only on performance. Second, previous studies have mainly looked at beliefs. Training in emotion-focused strategies shows effectiveness in reducing perceived procrastination in a randomized experiment (Eckert et al. 2016). However, these results are based on validated scales to assess procrastination, which may not accurately reflect actual behavior (Piedmont et al. 2000). In this study, I am able to overcome previous data limitations and identify effects on both perceived and actual procrastination by having access to data on study effort in a high-stakes setting.

The results of this study have practical implications for educators. They suggest that interventions providing periodic positive feedback or boosting self-esteem can effectively reduce procrastination and drive academic productivity. Importantly, these interventions do not require expensive ed-tech infrastructure; they can be implemented through standard teacher-student interactions. As such, they offer a feasible and scalable tool to foster human capital accumulation across different educational settings and grade levels (List 2022).

The remainder of the paper is structured as follows. Section 2 outlines the theoretical framework. Section 3 describes the setting. Section 4 presents the experimental design. Section 5 describes the data. The main results are presented in Section 6, while validity checks and mechanisms are discussed in Section 7. Finally, Section 8 concludes.

2 Theoretical Framework

In this section, I introduce a theoretical framework to motivate the experimental design and illustrate the mechanism linking self-esteem to procrastination. Specifically, I bridge the gap between the traditional β-δ models of time-inconsistent preferences (e.g., Laibson 1997; O'Donoghue and Rabin 1999, 2001, 2008) and recent findings on "ego utility" and fragile self-esteem (e.g., Bénabou and Tirole 2002; Kőszegi 2006; Kopylov 2012; Kőszegi, Loewenstein, and Murooka 2022). For a review of the present-bias literature, see O'Donoghue and Rabin (2015).

2.1 Traditional β-δ model

Consider a student facing an intertemporal choice regarding study effort. Assume the student must choose between studying for seven hours on March 1 (Option A) or eight hours on March 14 (Option B). In a standard model of time-inconsistent preferences, the student prefers the "efficient" outcome (Option A) when asked on February 1; however, once March 1 arrives, these preferences reverse, leading him to delay the task. This phenomenon, known as present bias, implies that individuals give stronger relative weight to the present moment relative to any future moment. Formally, this is modeled using (β, δ) preferences. The utility of the student at time t, evaluating a stream of payoffs $(u_t, u_{t+1}, \ldots, u_T)$, is given by:

$$U_t = u_t + \beta \sum_{\tau=1}^{T-t} \delta^\tau u_{t+\tau}$$

where $0 < \beta < 1$ captures the present bias (short-run impatience) and $0 < \delta \leq 1$ represents the standard exponential discount factor (long-run patience). For the sake of simplicity, assume $\delta = 1$, so that the choice is only driven by present bias rather than by long-run discounting. The parameter β generates the preference reversal: from the perspective of February, both March 1 and March 14 are in the future and are discounted equally by β, so the student chooses the lower cost (7 hours). However, when March 1 arrives, the cost of studying is immediate (u_t) and is not discounted by β, while the future cost on March 14 is discounted by β.

Standard theory distinguishes between naïfs (who are unaware of their future self-control problems) and sophisticates (who are fully aware). A sophisticated student, anticipating his future reversal, would ideally use a commitment device to force himself to study on March 1.

2.2 A β-δ model with self-esteem dependence

Consider now a variation of the decision problem. Before the student decides whether to study on March 1, he receives a phone notification that he failed a previous, unrelated exam. Under the standard β-δ model, preferences are exogenous and stable; this "sunk-cost" news regarding a past event should not alter his discount factor or his optimal decision regarding the next exam.

However, psychological research suggests that decision-making is state-dependent, specifically regarding mood and self-esteem (Kőszegi, Loewenstein, and Murooka 2022). Moreover, Kőszegi (2006) argues that individuals derive utility not only from consumption or leisure but also from their ego utility: their beliefs about their own ability.

This introduces the mechanism of self-handicapping. Assume the student believes the outcome of an exam depends on both ability (a) and effort (e). A failure after high effort can be interpreted as a signal of low ability ($\downarrow a$), which is psychologically painful. Conversely, a failure after procrastination (low effort) creates a "noisy" signal; the failure can be attributed to lack of effort rather than lack of ability, therefore protecting the student's self-image.

To capture this psychological mechanism in reduced form, I let the present-bias parameter of the traditional β-δ model depend on the agent's current state of self-esteem, denoted by s. The utility function is rewritten as:

$$U_t(s) = u_t + \beta(s) \sum_{\tau=1}^{T-t} \delta^{\tau} u_{t+\tau}$$

Here, $\beta(s)$ represents the "self-esteem dependent" present bias[2]. I assume that $\beta'(s) > 0$.

High Self-Esteem: As s increases, $\beta(s) \to 1$. The individual feels secure in their ability, has less need for ego-protection (self-handicapping), and exhibits behavior closer to standard exponential discounting.

Low Self-Esteem: As s decreases, $\beta(s)$ falls, representing in reduced form the pull toward delay created by the immediate value of "protecting the ego" (making the signal of ability noisier). This increase in impatience can therefore lead to procrastination.

To capture the "fragility" in self-esteem described by Kőszegi, Loewenstein, and Murooka (2022), I assume s shifts between two states: a confident state $\bar{s}$ and a fragile state $\underline{s}$. If an individual receives a negative shock (e.g., the failed exam notification), s shifts from $\bar{s}$ to $\underline{s}$. Consequently, $\beta(\bar{s})$ drops to $\beta(\underline{s})$. The cost of immediate effort now outweighs the discounted future benefits not just because of laziness, but because the immediate "ego cost" of potentially proving oneself a low-ability type is too high. Therefore, contrary to the standard model, an exogenous shock to self-esteem exacerbates present bias, making self-handicapping

2. Evidence that β is not a fixed trait is also provided by Augenblick, Niederle, and Sprenger (2015), who documented that the same individuals exhibit present bias in the domain of real effort, but not in the domain of monetary payoffs.

procrastination the optimal strategy for the agent to preserve their remaining self-esteem.

The model therefore yields a key testable prediction: boosting self-esteem should shift some agents from $\beta(\underline{s})$ to $\beta(\bar{s})$, reducing their propensity to procrastinate. My experiment is designed to test this hypothesis.

3 Background

3.1 Academic Setting

The university. The University of St. Gallen is a public Swiss institution offering bachelor's, master's, and doctoral programs in business, law, economics, and computer science. It enrolls roughly 10,000 students and is recognized as the leading business school in Switzerland and among the top 10 in Europe, according to the Financial Times (FT European Business School Rankings 2025). The university is located in the city of St. Gallen, capital of the canton (state) of St. Gallen in eastern Switzerland; the canton has a population of approximately 540,000.

The assessment year. Admission to the bachelor's program requires completion of the first year, known as the *assessment year*, which is offered in two tracks based on the language of instruction: German or English. All students take courses in business administration, economics, law, and mathematics. Each subject consists of a Fall semester course (September–December) and a Spring semester course (February–May), and each semester course includes a two-week midterm break.

Grades range from 1 to 6, with 6 as the highest and 4 as the passing grade. Grades are weighted by course credits to calculate total weighted credit points. Students failing an exam receive negative credit points proportional to the shortfall from the passing grade. To pass the assessment year, students must accumulate at least 240 weighted credit points and have no more than 12 negative points. Students may repeat the year once if they fail, but credits cannot be carried over; failing a second time prevents them from enrolling in the same program anywhere in the country. Therefore, this high-stakes framework makes the assessment year a particularly suitable context for investigating study effort.

3.2 Data Partner

Brian AG is a Swiss company that provides educational services to universities and schools through its learning platform, Brian, where professors can upload study materials (e.g., flashcards, questions) that students access via smartphones and tablets.

The app uses gamification to increase student engagement with the topics (see Figure A.1 for an example of the interface). In particular, the interface of the app resembles a journey, with different milestones representing each topic. To complete a milestone, students need to answer 20 questions. Students are allowed to answer as many questions as they want, and they can also answer the same question multiple times. In the app, students see real-time feedback on whether the answer is correct or not, and they also see their progress in the app (i.e., the share of answered questions), the total number of correct answers, and the ranking for the course, computed based on the points. Students register with their university email and select a nickname for their profile; therefore, a student knows her position in the ranking but only knows the position of the peers whose nicknames she knows.

Using Brian is recommended by lecturers but is voluntary and does not influence course grades. Introduced in 2020 in the Business Administration A course, the app remains most extensively used in this course, and students report it as helpful in course evaluations. In this study, I analyze study behavior in Business Administration A and Economics A during the Fall semester, and in Business Administration B during the Spring semester.

4 Experimental Design

4.1 Treatments

Students were randomly assigned to three conditions: *Booster*, *Goals*, and *Control*. I used a simple randomization method to assign all eligible students to the three treatment arms (see Section 5 for a description of the experimental sample).

Booster. Students in the self-esteem booster group received the standard version of the app until the start of the experiment (the last Sunday of the term break). After that, they received a positive notification on their app usage every Sunday at 1 p.m. for 11 weeks (up to the last Sunday before the Business Administration A exam). In particular, they received the following message *"Great job! You studied X minutes last week for Business Administration. Keep up the amazing work!"*, as shown in Figure A.2. If students never opened the app during the week, they did not receive the notification.

Goals. Students in this treatment arm received the standard version of the app until the midterm break. After that, every Sunday at 1 p.m. for 11 weeks, they received a notification prompting them to set a goal for their app use the following week. In particular, they received the following message *"Take control of your week. How many minutes do you plan to study on Brian next week in Business Administration?"*. If they clicked on the notification, an input box would open in the app, allowing them to set their goal for the week. Figure A.3 displays the notification and the input box.

Control. Students in the control group received the standard version of the app, without additional notifications. This group provides a baseline to identify the causal effect of the treatments.

4.2 Timeline

Figure A.5 presents the experimental timeline. The experiment took place during the 2024/2025 academic year, covering both the Fall and Spring semesters. The Fall semester began in mid-September and continued until Christmas, with a two-week break at the end of October. The corresponding exam period started in mid-January and ended in mid-February. The Spring semester ran from mid-February to mid-May, with a two-week break at the end of March. The Spring exam period began in mid-June and concluded in mid-July.

I complemented the analysis with three survey waves. The baseline survey was conducted from the end of September until mid-October, the midline survey from mid-December until mid-January, and the endline survey from the end of April until the end of May.

The intervention started on the last Sunday of the term break in the Fall semester and lasted for 11 weeks, until the last week before the Business Administration A exam. During the Spring semester, no intervention was conducted; therefore, all students returned to the standard version of the app. Short-run effects are defined relative to the Fall semester, immediately following the intervention, while long-run effects are defined relative to the Spring semester.

5 Data

Brian. I used app usage data, including the number of questions answered, the number of correct answers, the time spent on each answer, the ranking, and the time and day when students answered each question. Outcome variables pool app activity across all the courses using the app: Business Administration A and Economics A in the Fall semester, and Business Administration B in the Spring semester.

Surveys. Participation in the surveys was voluntary. Students could complete the surveys by scanning a QR code shown in class during lecture time. The survey was advertised by faculty members, and the link to the survey was also available on the university platform where students access study materials. Reminders were sent via email only for the endline survey. To incentivize participation, in each wave, 3 students who completed the survey were randomly selected to win 100 CHF (around 120 USD) each.

I collected the following information: demographics, parental education, overconfidence, and expected grade in Business Administration. Moreover, I used validated scales to measure self-esteem (Rosenberg's Self-Esteem Scale, Rosenberg 1965), procrastination (a short version of the General Procrastination Scale, Sirois, Yang, and van Eerde 2019), and time management (a short version of the Time Management Behaviors Scale, Peeters and Rutte 2005). Each scale produces an aggregate score by summing item responses, after reverse-coding negatively worded items, with higher values indicating higher levels of the underlying trait. The General Procrastination Scale includes items such as "In preparing for some deadlines, I often waste time by doing other things" or "I often find myself performing tasks that I had intended to do days before". Rosenberg's Self-Esteem Scale includes items like "I feel that I have a number of good qualities" or "On the whole, I am satisfied with myself". Finally, the Time Management Behaviors Scale contains items like "I feel in control of my time" or "I finish top-priority tasks before going on to less important ones". For the text of the surveys, see Section A.3.

University. Data on place of residence, gender, and language of study were obtained from the university's administrative records.

5.1 Sample Description

All students in the 2024/2025 assessment year (N=1,825) were eligible[3]. Exclusions were applied to students who had not downloaded the app at the time of randomization or who had notifications disabled. The randomization sample consists of 1,646 students, representing 90% of the eligible population.

Table A.1 reports summary statistics for the experimental sample and the rest of the students enrolled in the assessment year. On average, students in the experiment are more likely to

3. I was not able to link one student with administrative data from the university; therefore, the overall sample I used in the analysis consists of 1,824 students.

be enrolled in the German track (63% versus 44%), but other demographic characteristics such as gender and place of residence are balanced. As expected, students in the experiment have a higher baseline usage of the app. This reflects the inclusion criteria, as only students who had downloaded the app and enabled notifications were included in the experiment. Indeed, only 2 out of 178 students in the non-experimental sample had used the app before the randomization.

Summary statistics for the experimental sample are reported in Table 1. On average, a large share of the students (83%) had a Swiss address as their residence when they applied to university, and 63% are enrolled in the German track. Finally, only 38% of the sample consists of female students. Regarding the baseline survey, 70% of the students have a parent with at least a bachelor's degree, and the average expected grade in Business Administration A is 4.9 (in Switzerland 6 is the highest grade and 4 is the passing grade).

There is no evidence of selection into treatments, as reported in Table 2. For each baseline characteristic, I regress the variable on the treatment dummies and report the p-value of the F-test for joint significance. Both administrative covariates (gender, place of residence, and language of study) and survey measures (self-esteem, procrastination, time management, expected grade in Business Administration A, and parental education) are statistically balanced across groups. Because the baseline survey was conducted prior to randomization and students were randomly assigned to the treatment arms, the groups are comparable at baseline.

Students could disable the notifications or uninstall the app. A small share of students did so during the experiment, as displayed in Figure A.4. In particular, in the last week of the experiment, only 6% in the *Booster* group and around 1% in the *Goals* group had disabled notifications or uninstalled the app. This figure reflects opt-out only; Section 7 separately discusses week-by-week non-receipt of the *Booster* message driven by zero app use (Figure A.7).

The survey take-up rate is relatively low: 440 students participated in the three waves, representing approximately 26% of the experimental sample. Across the three waves, there is no evidence that the composition of respondents differs by treatment arm, demographics, or baseline app usage, as displayed in Table A.2.

The only significant difference across waves is a higher proportion of German-track students in the baseline sample (72%) compared to the midline (57%) and endline (58%) samples.

Table 1: Summary statistics for experimental sample

Variable	Count	Mean	SD	Min	Max
Female	1646	0.38	0.48	0.00	1.00
Swiss	1646	0.83	0.38	0.00	1.00
German Track	1646	0.63	0.48	0.00	1.00
Brian User	1646	1.00	0.00	1.00	1.00
Total Answers	1646	812.07	965.56	0.00	7203.00
Total Correct Answers	1646	580.87	735.58	0.00	5469.00
Total Distinct Answers	1646	430.88	443.08	0.00	3845.00
Procrastination	274	26.78	6.16	13.00	45.00
Self-Esteem	274	37.68	6.75	18.00	50.00
Time Management	272	34.14	6.20	15.00	49.00
Expected Grade	271	4.90	0.47	3.50	6.00
Parental education: higher than bachelor	270	0.70	0.46	0.00	1.00

This table reports summary statistics for baseline variables for individuals in the experiment. The variables refer to the period before the randomization. Procrastination, self-esteem, time management, the expected grade in Business Administration A, and parental education refer to the baseline survey. Students that failed the attention check in the baseline survey and that did not answer the questions on procrastination and self-esteem were excluded from the sample.

Table 2: Balance Table: students in the experimental sample by treatment arm

Variable	Booster	Goals	Control	P value (F test)
Female	0.38	0.37	0.38	0.968
Swiss	0.82	0.82	0.84	0.500
German Track	0.63	0.65	0.61	0.310
Total Answers	821.45	776.75	840.96	0.525
Total Correct Answers	591.34	560.71	591.99	0.718
Total Distinct Answers	438.03	416.90	438.71	0.645
Procrastination	26.45	26.79	27.16	0.734
Self-Esteem	37.30	38.38	37.43	0.512
Time Management	34.72	34.48	33.18	0.194
Expected Grade	4.92	4.85	4.92	0.483
Parental education: higher than bachelor	0.68	0.68	0.72	0.809
N (experimental sample)	548	574	524	
N (baseline survey)	100	84	90	

This table reports means for baseline variables for individuals in the experiment by treatment arm. The variables refer to the period before the randomization. Procrastination, self-esteem, time management, the expected grade in Business Administration A, and parental education refer to the baseline survey.

6 Results

6.1 Effect of self-esteem booster on study effort

To assess the effect of the treatment arms, I estimate the following regression:

$$Y_{it} = \alpha + \beta_1 \mathbf{1}\{\text{Booster}_i\} + \beta_2 \mathbf{1}\{\text{Goals}_i\} + \gamma \mathbf{X}_i + \omega_t + \varepsilon_{it},$$

where Y_{it} denotes the outcome of student i in week t. Outcomes include (i) the number of correct answers, (ii) the total number of answers, and (iii) the number of distinct questions answered in the app. $\mathbf{X}_i$ is a vector of student-level controls measured at baseline, including gender, country of residence at enrollment, language of study, and baseline usage of the app. Week fixed effects (ω_t) are included in the model, and the standard errors are clustered at the student level.

The coefficients β_1 and β_2 capture the intention-to-treat (ITT) effects of assignment to the *Booster* and *Goals* interventions relative to the *Control* group. Overall, there is no evidence that *Goals* increases effort in the app, as displayed in Table 3. On the other hand, students assigned to the *Booster* group give more correct answers, more total answers, and more distinct answers in the app. The estimated effects across these outcomes are consistently around 0.04 standard deviations.

Table 3: Effects of Treatment on Weekly Outcomes

	Correct Answers	Total Answers	Distinct Answers
Goals	0.009	0.013	0.002
	(0.016)	(0.016)	(0.017)
Booster	0.042**	0.040***	0.049***
	(0.017)	(0.016)	(0.018)
Observations	70778	70778	70778
Week FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Clustered SEs	Yes (Student)	Yes (Student)	Yes (Student)

* p <0.1, ** p <0.05, *** p <0.01
Each column reports OLS estimates with week fixed effects and standard errors clustered at the student level. All regressions control for student gender, nationality, language of study, and baseline usage of the app. Outcome variables are standardized.

Shifting levels of effort might have no effect on procrastination if the effect is driven by students studying more on one specific day. Therefore, I run a separate OLS regression per week to estimate the dynamics of the ITT[4]. Figure 1 shows the evolution of the effect of the *Booster* on the number of answers across the weeks when the intervention took place. Despite negative or null effects during the beginning of the intervention (when students were exposed to a low number of overall notifications compared to subsequent weeks) and during the Christmas holidays, the coefficients are positive and statistically significant for almost every week, suggesting that students in the *Booster* group were procrastinating less. In particular, students in this group give on average 50 more answers per week in the last two weeks of the intervention, when the exposure to the treatment was the highest.

When decomposing this effect between the short run, when students are exposed to the treatment, and the long run, when all students switch back to the *Control* group, the overall effect is driven by the long run. Comparing the estimates in Figure 2, there is evidence of a

4. Given that students could disable the notifications, the estimated effects should be interpreted as ITT effects.

habit-formation mechanism, with students in the *Booster* group increasing their effort in the app when exposed to the intervention (although not statistically significant) and continuing to use the app more in the Spring semester.

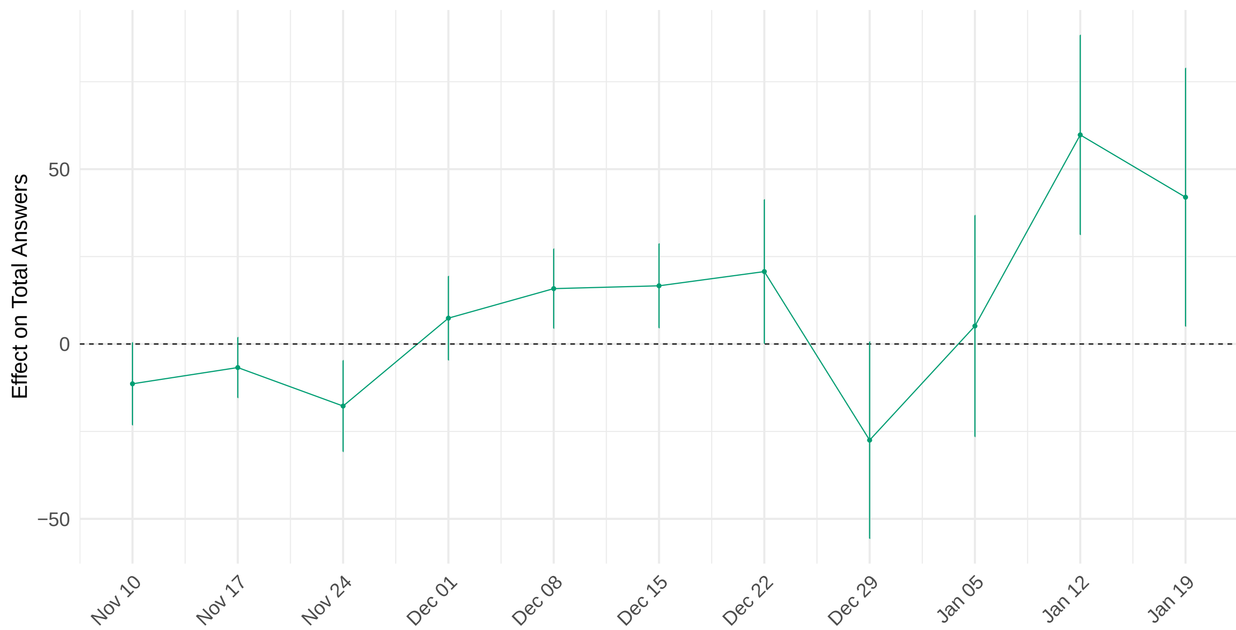

Figure 1: Effect of the Booster treatment on the number of Total Answers, estimated separately for each week of the experiment. The estimates were derived using OLS regressions with robust standard errors. All regressions include controls for gender, language of study, and baseline usage of the app.

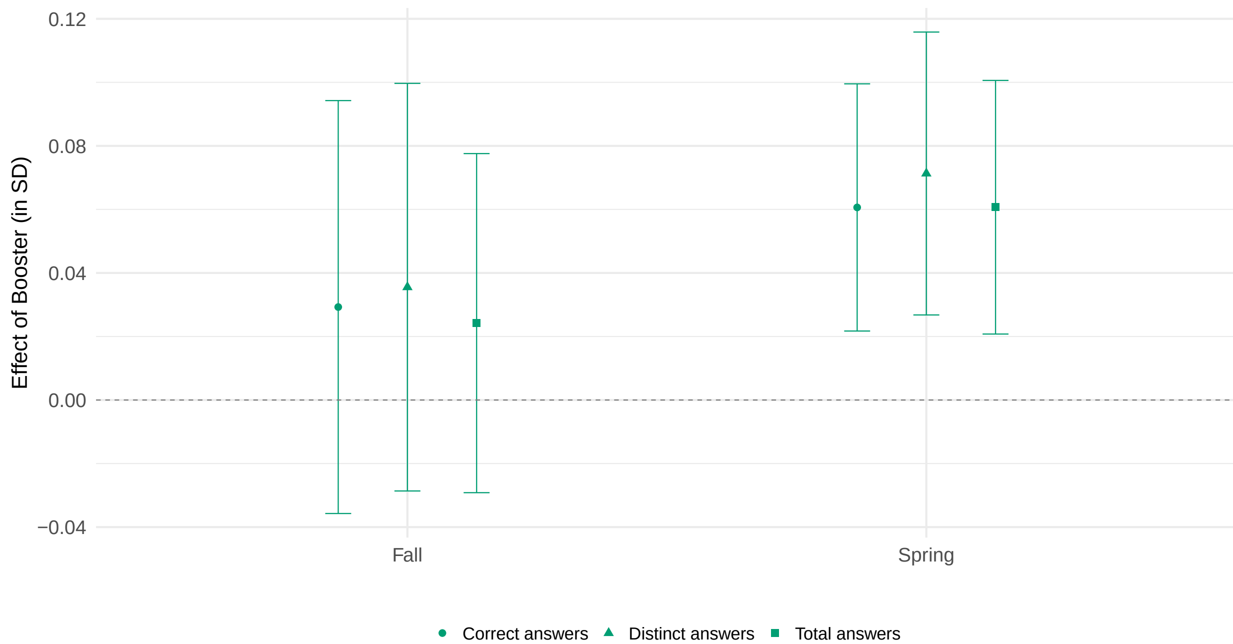

Figure 2: Effect of Booster on number of total answers, number of correct answers, and number of distinct answers. Each coefficient is the OLS estimate with week fixed effects and standard errors clustered at the student level. All regressions control for student gender, nationality, language of study, and baseline usage of the app. Outcome variables are standardized. The effects are estimated separately by semester: in the Fall semester the intervention was implemented, and in the Spring semester all students switched back to the control group.

6.2 Effect heterogeneity

To explore gender differences, I conducted a preregistered heterogeneity analysis by estimating interactions between treatment indicators and a female dummy:

$$Y_{it} = \alpha + \beta_1 \mathbf{1}\{\text{Booster}_i\} + \beta_2 \mathbf{1}\{\text{Goals}_i\} + \phi \mathbf{1}\{\text{Female}_i\} + \delta_1 \Big[\mathbf{1}\{\text{Booster}_i\} \times \mathbf{1}\{\text{Female}_i\}\Big]$$
$$+ \delta_2 \Big[\mathbf{1}\{\text{Goals}_i\} \times \mathbf{1}\{\text{Female}_i\}\Big] + \gamma \mathbf{X}_i + \omega_t + \varepsilon_{it},$$

where Y_{it} denotes the outcome of student i in week t. $\mathbf{X}_i$ is a vector of student-level controls measured at baseline, including country of residence at enrollment and language of study. Week fixed effects (ω_t) are included in the model, and the standard errors are clustered at the student level.

Results in Table 4 highlight substantial heterogeneity. Overall, the effect is driven by male students, who react to the self-esteem booster by increasing their effort in the app, while the intervention has negative effects on female students. The *Goals* treatment has no effect on male students, while it has negative effects on female students. These results are in line with previous interventions in education that find differential effects by gender (Smithers 2015; Clark et al. 2020; Delavande et al. 2023; Campos-Mercade, Thiemann, and Wengström 2025).

On the other hand, there is limited evidence of different effects based on whether the student was registered at a Swiss address at enrollment, as displayed in Table A.3. The interaction with Booster is marginally significant for correct answers ($p < 0.1$), but not for total or distinct answers, and none of the Goals interactions are significant. This heterogeneity analysis was preregistered, based on the institutional setting, where students coming from a Swiss high school are automatically admitted to the university, while other students need to pass an entry test. Therefore, students coming from a non-Swiss high school might be more motivated or have higher ability. In this analysis, I used place of residence as a proxy for having a Swiss high-school diploma.

Table 4: Effect heterogeneity by gender of treatment on weekly outcomes

	Correct Answers	Total Answers	Distinct Answers
Goals	0.027	0.027	0.005
	(0.027)	(0.026)	(0.029)
Booster	0.076***	0.066***	0.076**
	(0.027)	(0.024)	(0.030)
Goals × Female	−0.077*	−0.075*	−0.049
	(0.041)	(0.039)	(0.046)
Booster × Female	−0.095**	−0.082**	−0.076
	(0.042)	(0.039)	(0.047)
Observations	70778	70778	70778
Week FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Clustered SEs	Yes (Student)	Yes (Student)	Yes (Student)

* p <0.1, ** p <0.05, *** p <0.01
Each column reports OLS estimates with week fixed effects and standard errors clustered at the student level. All regressions control for nationality and language of study. Outcome variables are standardized.

7 Validity Checks and Mechanisms

7.1 Effects on beliefs

The self-esteem booster is motivated by the idea that students with low self-esteem may use procrastination as a self-protective mechanism. By increasing self-esteem, the intervention is expected to reduce procrastination. To verify whether the observed effects operate through this mechanism, I examine self-reported measures of self-esteem and procrastination.

I first conduct a manipulation check by comparing self-esteem scores across the *Booster*, *Goals*, and *Control* groups. As shown in Figure 3, self-esteem was balanced across groups at baseline. Following the intervention, students in the *Booster* group reported significantly higher self-esteem in the midline survey (approximately an 8% increase in the score), but the effect was not statistically significant in the endline survey. This pattern is consistent with the intervention being delivered over a single semester and the endline survey capturing a semester where all students had experienced the same conditions. By comparing the control averages across survey waves, it seems that the effect of the self-esteem booster is to counteract the strong reduction in self-esteem between the start and the end of the Fall semester (from 37.43 to 34.91).

I then examine self-reported procrastination in the surveys. Figure 4 illustrates the evolution of reported procrastination across the three survey waves, with students in the *Booster* group reporting lower levels of procrastination in the endline survey. In particular, the effect corresponds to around an 11% reduction relative to the control mean.

Finally, students in the *Goals* group follow similar trends compared to those in the *Control* group, both in terms of reported self-esteem and procrastination. Therefore, only the self-esteem booster has an effect on both behaviors in the app and beliefs in the surveys.

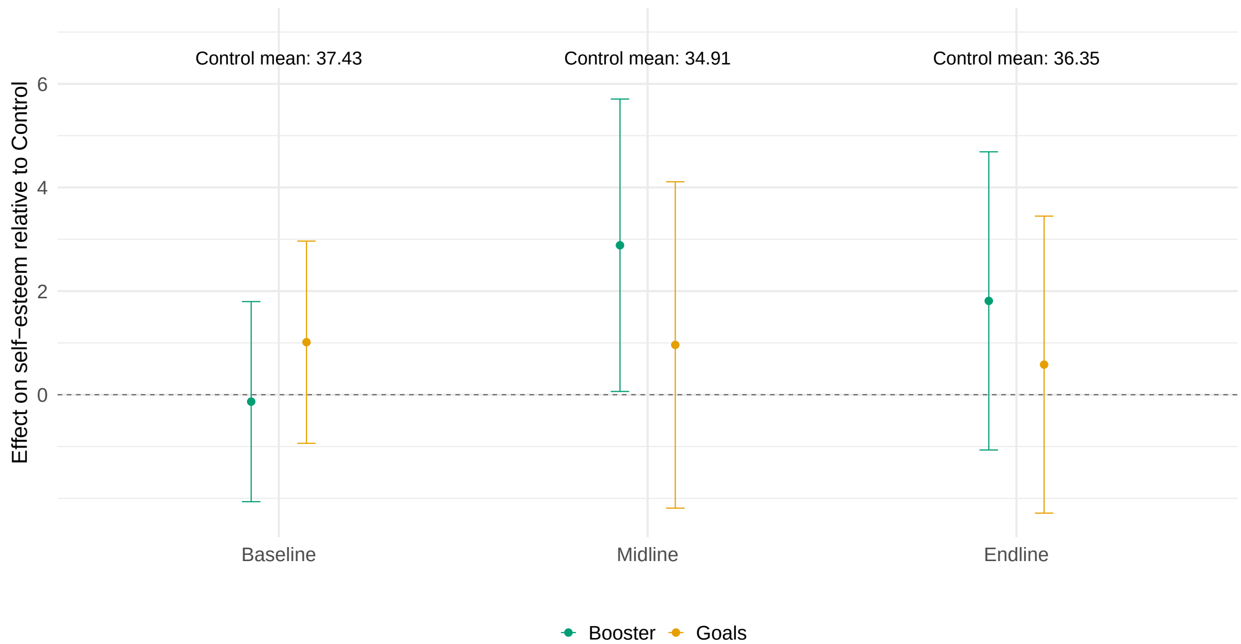

Figure 3: Effect of Booster and Goals on reported self-esteem. Each coefficient is the OLS estimate with robust standard errors. Reported self-esteem is measured using the Rosenberg Self-Esteem Scale (Rosenberg 1965). For the full text of the survey, refer to Section A.3.

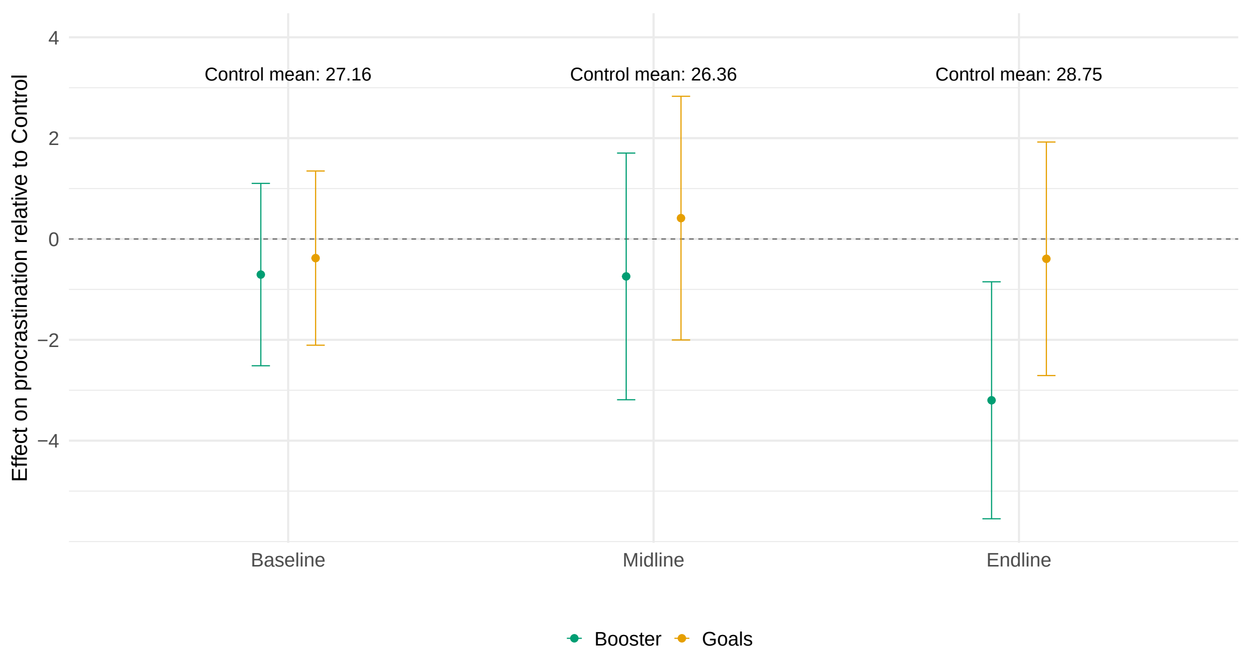

Figure 4: Effect of Booster and Goals on reported procrastination. Each coefficient is the OLS estimate with robust standard errors. Reported procrastination is measured using the Short General Procrastination Scale (Sirois, Yang, and van Eerde 2019). For the full text of the survey, refer to Section A.3.

7.2 Spillovers

The estimation of the ITT could lead to biased results in the presence of spillovers. This could happen if students study in groups and the treatment assignment of one student affects the effort of another student in the same group. For example, a student in the *Control* group may become discouraged if he sees a peer in the *Booster* group receiving a notification praising his effort, leading the *Control* student to question his own performance and potentially decrease his study effort.

To test for spillovers, I leverage randomly formed groups to which students were assigned during a mandatory orientation week. These groups have been shown to be relevant for friendship networks, major choices, GPA, and dropout rates (Thiemann 2022). Students were assigned to these groups of average size 24 based on a stratified randomization by gender and whether or not they completed high school in Switzerland. During the mandatory orientation week, they spent 60 hours together and had to prepare a presentation on a case study that took place on the last day of the week.

Table A.4 presents the results of a regression of study effort on the treatment indicators, controlling for the share of treated peers within the orientation week group. Given that these coefficients are not statistically significant, there is no evidence supporting the presence of spillovers.

7.3 Salience of the app and engagement

Another potential concern is that the observed effects are not driven by increased self-esteem, but rather by greater salience of the app. Students in the *Booster* group receive weekly notifications, which could increase exposure to the app relative to the *Control* group and mechanically boost engagement. However, students in the *Goals* group also receive weekly notifications, so any salience effect should similarly affect this group.

The different effects of the *Goals* treatment compared to *Booster* suggest that the positive effect of the self-esteem booster is unlikely to be due only to increased app salience. Moreover, students in the *Booster* group received on average fewer notifications than those in *Goals*, given the design of the intervention. In particular, students in *Booster* only received the notification if they studied for at least one minute in the app, while those in the *Goals* group received it independently of their effort. However, results are different both in terms of observed behaviors (in Table 3) and beliefs (in Figure 3 and Figure 4).

To further explore this mechanism, Figure A.6 plots the average number of answers per hour on Sundays, when notifications are sent, in the weeks when the intervention took place. The post-notification usage trajectories show opposing, non-statistically significant trends: after receiving the notification, usage increases for students in the *Booster* group but decreases for those in the *Goals* group. These patterns indicate that the effect of the self-esteem booster is not simply a mechanical consequence of app reminders.

7.4 Self-esteem booster and treatment intensity

Being assigned to the *Booster* treatment does not automatically translate into receiving 11 self-esteem boosters during the experiment. In fact, if a student never opened the app for one week, no notification would be sent. Moreover, the perceived impact of the message may vary with weekly effort. For example, receiving a message praising a student for studying only four minutes may not effectively boost self-esteem.

Figure A.7 displays, for each week, the share of students in the *Booster* group with zero app use, who therefore did not receive that week's notification, and the share with low, but positive, app use (under five minutes). The share with positive low effort is small and stable over time, whereas the share with zero effort is substantially higher and decreases over the semester. If these subgroups attenuate the treatment effect, the estimated ITT should be interpreted as a lower bound for both app-based study effort and self-reported survey outcomes.

8 Conclusion

This study investigates the relationship between self-esteem and procrastination by testing whether an intervention that increases self-esteem can reduce procrastination. I conducted a randomized controlled trial with first-year students at the University of St. Gallen, in collaboration with Brian AG, a learning platform.

The results indicate that the self-esteem booster increases study effort in the app, with the effect mainly driven by male students. Students in the *Booster* group also exhibit more effort in the six months after the intervention. However, students in the *Goals* group do not react to the treatment by increasing their effort.

These findings have important implications for the design of educational interventions. Interventions intended to increase academic productivity need to account for gender-specific responses, as effects are heterogeneous across students. In addition, the results suggest that low-cost interventions, such as periodic positive feedback, can have positive long-term effects on student effort.

One limitation of this study is that the analysis focuses on online learning. Future studies should also analyze offline learning to provide a more comprehensive framework. However, given the proliferation of learning platforms, I expect future generations to spend an increasing amount of time using new learning tools, making this context highly relevant. In an ideal setting, a researcher would have access to both online and offline learning data. In my setting, this could have been possible if library access data were available, assuming students primarily study there. However, since students do not scan their student cards to access the library, it is not possible to objectively quantify offline effort in this study.

Regarding external validity, future studies should replicate these findings. While the theoretical framework suggests that these results should generalize, individuals in other settings might respond differently for two reasons. First, first-year college students have recently transitioned to a new learning environment; they are still adapting their learning strategies and may be more willing to change them (Dai, Milkman, and Riis 2014). Second, compared to other populations, college students may attach greater weight to academic performance in their self-image, given the strong implications of a degree for future earnings and social status.

As already discussed in Section 7, not all students assigned to the *Booster* group received full exposure to the treatment. Future interventions incorporating teacher-delivered positive messages could overcome this problem, as face-to-face interactions might have higher treatment compliance compared to in-app notifications.

Finally, this study contributes to the economic literature by demonstrating a causal effect of self-esteem on effort. A natural extension would be to investigate the effect of boosting self-esteem in other settings where individuals might engage in self-handicapping strategies to protect their self-image, such as newly hired workers or freelancers working on a first project with a new client. For instance, a freelancer faces a trade-off: high effort increases the chance of future work, but low effort provides a self-protective excuse (blaming effort rather than ability) if the contract is not renewed.

References

Abel, Martin, Tomoko Harigaya, Michael Kremer, and Jessica Zhu. 2026. "The Effect of Reminders for Self-Set Goals on Productivity." *Journal of Development Economics* 182:103822.

Ariely, Dan, and Klaus Wertenbroch. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment." *Psychological Science* 13 (3): 219–224.

Arnold, Ivo J. M. 2023. "The Link between Procrastination and Graduation Rates: Evidence from the ALEKS Learning Platform." *Education Economics* 31 (3): 275–287.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines." *The Quarterly Journal of Economics* 121 (2): 635–672.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger. 2015. "Working over Time: Dynamic Inconsistency in Real Effort Tasks." *The Quarterly Journal of Economics* 130 (3): 1067–1115.

Augenblick, Ned, and Matthew Rabin. 2019. "An Experiment on Time Preference and Misprediction in Unpleasant Tasks." *The Review of Economic Studies* 86 (3): 941–975.

Aulagnon, Raphaëlle, Julian Cristia, Santiago Cueto, and Ofer Malamud. 2025. "Streaks to Success: The Effects of Highlighting Streaks on Student Effort and Learning." *Economics of Education Review* 109:102721.

Bénabou, Roland, and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *The Quarterly Journal of Economics* 117 (3): 871–915.

Bisin, Alberto, and Kyle Hyndman. 2020. "Present-Bias, Procrastination and Deadlines in a Field Experiment." *Games and Economic Behavior* 119:339–357.

Brade, Raphael, Oliver Himmler, and Robert Jäckle. 2026. "No Student Left behind? Relative Feedback and University Completion." *Journal of Economic Behavior & Organization* 241:107383.

Burger, Nicholas, Gary Charness, and John Lynham. 2011. "Field and Online Experiments on Self-Control." *Journal of Economic Behavior & Organization* 77 (3): 393–404.

Bursztyn, Leonardo, Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao. 2018. "Status Goods: Experimental Evidence from Platinum Credit Cards." *The Quarterly Journal of Economics* 133 (3): 1561–1595.

Cadena, Ximena, Antoinette Schoar, Alexandra Cristea, and Héber M. Delgado-Medrano. 2011. *Fighting Procrastination in the Workplace: An Experiment.* Technical report 16944. National Bureau of Economic Research.

Campos-Mercade, Pol, Petra Thiemann, and Erik Wengström. 2025. *Performance Incentives in Education: The Role of Goal Mismatch.* Technical report 2025:5.

Campos-Mercade, Pol, and Erik Wengström. 2025. *Threshold Incentives and Academic Performance.*

Cerrone, Claudia, Anujit Chakraborty, Hyok Jung Kim, and Leonhard K. Lades. 2026. "Estimating Present Bias and Sophistication over Effort and Money." *Experimental Economics,* 1–23.

Clark, Damon, David Gill, Victoria Prowse, and Mark Rush. 2020. "Using Goals to Motivate College Students: Theory and Evidence From Field Experiments." *The Review of Economics and Statistics* 102 (4): 648–663.

Cotton, Christopher S., Brent R. Hickman, John A. List, Joseph Price, and Sutanuka Roy. 2026. "Why Don't Struggling Students Do Their Homework? Disentangling Motivation and Study Productivity as Drivers of Human Capital Formation." *Journal of Political Economy* 134 (1): 86–149.

Dai, Hengchen, Katherine L. Milkman, and Jason Riis. 2014. "The Fresh Start Effect: Temporal Landmarks Motivate Aspirational Behavior." *Management Science* 60 (10): 2563–2582.

De Paola, Maria, Francesca Gioia, and Vincenzo Scoppa. 2023. "Online Teaching, Procrastination and Student Achievement." *Economics of Education Review* 94:102378.

De Paola, Maria, and Vincenzo Scoppa. 2015. "Procrastination, Academic Success and the Effectiveness of a Remedial Program." *Journal of Economic Behavior & Organization,* Behavioral Economics of Education, 115:217–236.

Delavande, Adeline, Emilia Del Bono, Angus Holford, Sonkurt Sen, and Vedran Lesic. 2023. *Expectations about the Productivity of Effort and Academic Outcomes: Evidence from a Randomized Information Intervention.* SSRN Scholarly Paper, 4483970. Social Science Research Network: 4483970.

DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review* 96 (3): 694–719.

Dirkmaat, Thomas, Kirsten I. M. Rohde, Evelien van de Veer, Bram van Dijk, and Xiao Yu. 2023. "Managing "Last Moment Behavior": Non-binding Target Dates to Reduce the Spikes in Task Completion at Deadlines." *Journal of Public Economics* 219:104813.

Eckert, Marcus, David D Ebert, Dirk Lehr, Bernhard Sieland, and Matthias Berking. 2016. "Overcome Procrastination: Enhancing Emotion Regulation Skills Reduce Procrastination." *Learning and Individual Differences* 52:10–18.

Ersoy, Fulya. 2023. "Effects of Perceived Productivity on Study Effort: Evidence from a Field Experiment." *Journal of Economic Behavior & Organization* 207:376–391.

Fedyk, Anastassia. 2024. "Asymmetric Naïveté: Beliefs About Self-Control." *Management Science.*

Ferrari, Joseph R., and Dianne M. Tice. 2000. "Procrastination as a Self-Handicap for Men and Women: A Task-Avoidance Strategy in a Laboratory Setting." *Journal of Research in Personality* 34 (1): 73–83.

Fudenberg, Drew, and David K. Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review* 96 (5): 1449–1476.

Giuntella, Osea, Silvia Saccardo, and Sally Sadoff. 2026. "Sleep: Educational Impact and Habit Formation." *Journal of Political Economy* 0 (0): 000–000.

Gneezy, Uri, John A. List, Jeffrey A. Livingston, Xiangdong Qin, Sally Sadoff, and Yang Xu. 2019. "Measuring Success in Education: The Role of Effort on the Test Itself." *American Economic Review: Insights* 1 (3): 291–308.

Hall, Crystal C., Jiaying Zhao, and Eldar Shafir. 2014. "Self-Affirmation Among the Poor: Cognitive and Behavioral Implications." *Psychological Science* 25 (2): 619–625.

Hyndman, Kyle, and Alberto Bisin. 2022. "Procrastination, Self-Imposed Deadlines and Other Commitment Devices." *Economic Theory* 74 (3): 871–897.

———. 2026. "Replication of "Procrastination, Deadlines, and Performance: Self-Control by Precommitment"." *Psychological Science,* 09567976261460772.

John, Anett. 2020. “When Commitment Fails: Evidence from a Field Experiment.” *Management Science* 66 (2): 503–529.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2015. “Self-Control at Work.” *Journal of Political Economy* 123 (6): 1227–1277.

Kopylov, Igor. 2012. “Perfectionism and Choice.” *Econometrica* 80 (5): 1819–1843.

Kőszegi, Botond. 2006. “Ego Utility, Overconfidence, and Task Choice.” *Journal of the European Economic Association* 4 (4): 673–707.

Kőszegi, Botond, George Loewenstein, and Takeshi Murooka. 2022. “Fragile Self-Esteem.” *The Review of Economic Studies* 89 (4): 2026–2060.

Laibson, David. 1997. “Golden Eggs and Hyperbolic Discounting.” *The Quarterly Journal of Economics* 112 (2): 443–477. JSTOR: 2951242.

Lay, Clarry H. 1986. “At Last, My Research Article on Procrastination.” *Journal of Research in Personality* 20 (4): 474–495.

List, John A. 2022. *The Voltage Effect: How to Make Good Ideas Great and Great Ideas Scale.* Penguin UK.

O’Donoghue, Ted, and Matthew Rabin. 1999. “Doing It Now or Later.” *American Economic Review* 89 (1): 103–124.

———. 2001. “Choice and Procrastination.” *The Quarterly Journal of Economics* 116 (1): 121–160.

———. 2008. “Procrastination on Long-Term Projects.” *Journal of Economic Behavior & Organization* 66 (2): 161–175.

———. 2015. “Present Bias: Lessons Learned and to Be Learned.” *American Economic Review* 105 (5): 273–279.

Peeters, Miranda A. G., and Christel G. Rutte. 2005. “Time Management Behavior as a Moderator for the Job Demand-Control Interaction.” *Journal of Occupational Health Psychology* 10 (1): 64–75.

Piedmont, Ralph L., Robert R. McCrae, Rainer Riemann, and Alois Angleitner. 2000. “On the Invalidity of Validity Scales: Evidence from Self-Reports and Observer Ratings in Volunteer Samples.” *Journal of Personality and Social Psychology* 78 (3): 582–593.

Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image.* Princeton University Press. JSTOR: j.ctt183pjjh.

Schraw, Gregory, Theresa Wadkins, and Lori Olafson. 2007. “Doing the Things We Do: A Grounded Theory of Academic Procrastination.” *Journal of Educational Psychology* 99 (1): 12.

Schuenemann, Laura, Viviane Scherenberg, Maria von Salisch, and Marcus Eckert. 2022. ““I’ll Worry About It Tomorrow” – Fostering Emotion Regulation Skills to Overcome Procrastination.” *Frontiers in Psychology* 13.

Shah, Rohan, and Anita Mukherjee. 2026. “Procrastination in Personal Finance: Implications for Estate Planning and Retirement Satisfaction.” *Journal of Behavioral and Experimental Finance* 49:101137.

Sirois, Fuschia M., Sisi Yang, and Wendelien van Eerde. 2019. “Development and Validation of the General Procrastination Scale (GPS-9): A Short and Reliable Measure of Trait Procrastination.” *Personality and Individual Differences* 146:26–33.

Smithers, Samuel. 2015. “Goals, Motivation and Gender.” *Economics Letters* 131:75–77.

Solomon, Laura J., and Esther D. Rothblum. 1984. "Academic Procrastination: Frequency and Cognitive-Behavioral Correlates." *Journal of Counseling Psychology* 31:503–509.

Steel, Piers. 2007. "The Nature of Procrastination: A Meta-Analytic and Theoretical Review of Quintessential Self-Regulatory Failure." *Psychological Bulletin* 133 (1): 65–94.

Strulov-Shlain, Avner, and Alexandra Steiny Wellsjo. 2025. *Goals, Expectations, and Performance.* SSRN Scholarly Paper, 5126021. Social Science Research Network: 5126021.

Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112 (S1): S164–S187.

Thiemann, Petra. 2022. "The Persistent Effects of Short-Term Peer Groups on Performance: Evidence from a Natural Experiment in Higher Education." *Management Science* 68 (2): 1131–1148.

Tice, Dianne M, and Roy F. Baumeister. 1997. "Longitudinal Study of Procrastination, Performance, Stress, and Health: The Costs and Benefits of Dawdling." *Psychological Science* 8 (6): 454–458.

Van Lent, Max, and Michiel Souverijn. 2020. "Goal Setting and Raising the Bar: A Field Experiment." *Journal of Behavioral and Experimental Economics* 87:101570.

Xu, Shuai. 2021. "Academic Procrastination of Adolescents: A Brief Review of the Literature." *Psychology and Behavioral Sciences* 10 (6): 198.

A.1 Tables

Table A.1: Balance Table: students in the experimental sample vs. other students

Variable	Experimental Sample	Other Students	Diff.
Female	0.38	0.32	0.06
	(0.01)	(0.04)	(0.04)
Swiss	0.83	0.83	-0.00
	(0.01)	(0.03)	(0.03)
German Track	0.63	0.44	0.19***
	(0.01)	(0.04)	(0.04)
Brian User	1.00	0.01	0.99***
	(0.00)	(0.01)	(0.01)
Total Answers	812.07	0.26	811.81***
	(23.80)	(0.19)	(23.80)
Total Correct Answers	580.87	0.13	580.73***
	(18.13)	(0.10)	(18.13)
Total Distinct Answers	430.88	0.25	430.63***
	(10.92)	(0.18)	(10.92)
Number of students	1646	178	

This table reports means and standard errors (in parentheses) for baseline variables for individuals in the experiment and individuals not in the experiment. The variables refer to the period before the randomization. Only students that had downloaded the app and had notifications enabled were included in the experiment.

Table A.2: Balance Table: students in the experimental sample by survey wave

Variable	Baseline	Midline	Endline	P value (F test)
Female	0.46	0.45	0.44	0.908
	(0.03)	(0.04)	(0.04)	
Swiss	0.78	0.82	0.83	0.292
	(0.03)	(0.03)	(0.03)	
German Track	0.72	0.57	0.58	0.001***
	(0.03)	(0.04)	(0.04)	
Total Answers	987.88	952.66	929.44	0.830
	(66.99)	(87.17)	(67.57)	
Total Correct Answers	711.18	666.95	660.60	0.754
	(50.37)	(65.55)	(51.71)	
Total Distinct Answers	506.55	517.53	510.09	0.976
	(30.61)	(39.96)	(30.42)	
Treatment: Booster	0.36	0.34	0.34	0.869
	(0.03)	(0.04)	(0.03)	
Treatment: Goals	0.31	0.30	0.36	0.426
	(0.03)	(0.04)	(0.04)	
Treatment: Control	0.33	0.35	0.30	0.564
	(0.03)	(0.04)	(0.03)	
N	274	128	186	

This table reports means and standard errors (in parentheses) for baseline variables for individuals in the experiment by survey wave. The variables refer to the period before the randomization.

Table A.3: Effect heterogeneity by place of residence of treatment on weekly outcomes

	Correct Answers	Total Answers	Distinct Answers
Goals	−0.009	−0.009	0.008
	(0.051)	(0.045)	(0.056)
Booster	−0.036	−0.030	−0.012
	(0.049)	(0.044)	(0.056)
Goals × Swiss	0.007	0.007	−0.026
	(0.056)	(0.050)	(0.061)
Booster × Swiss	0.092*	0.078	0.071
	(0.054)	(0.049)	(0.061)
Observations	70778	70778	70778
Week FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Clustered SEs	Yes (Student)	Yes (Student)	Yes (Student)

* p <0.1, ** p <0.05, *** p <0.01
Each column reports OLS estimates with week fixed effects and standard errors clustered at the student level. All regressions control for gender and language of study. Outcome variables are standardized. Swiss is a dummy variable indicating that the student was registered at a Swiss address at enrollment.

Table A.4: Effects of Treatment on Weekly Outcomes

	Correct Answers	Total Answers	Distinct Answers
Goals	0.010	0.014	−0.006
	(0.019)	(0.020)	(0.021)
Booster	0.045**	0.042**	0.057**
	(0.020)	(0.019)	(0.022)
Share of treated peers (Booster)	0.028	0.070	−0.045
	(0.114)	(0.121)	(0.123)
Share of treated peers (Goals)	−0.137	−0.135	−0.144
	(0.104)	(0.096)	(0.111)
Observations	51170	51170	51170
Week FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Clustered SEs	Yes (Student)	Yes (Student)	Yes (Student)

* p <0.1, ** p <0.05, *** p <0.01
Each column reports OLS estimates with week fixed effects and standard errors clustered at the student level. All regressions control for student gender, nationality, language of study, and baseline usage of the app. Outcome variables are standardized. The share of treated peers is computed by using the randomly formed groups during the mandatory orientation week. Only students that attended the orientation week in the academic year 2024/2025 were included in the analysis.

A.2 Figures

Figure A.1: Brian's interface showing the course page, an example of a question, and the home page.

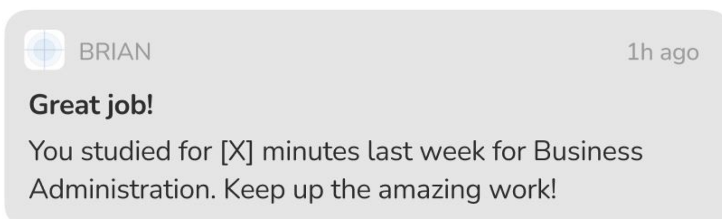

Figure A.2: Text of notification of the Booster treatment.

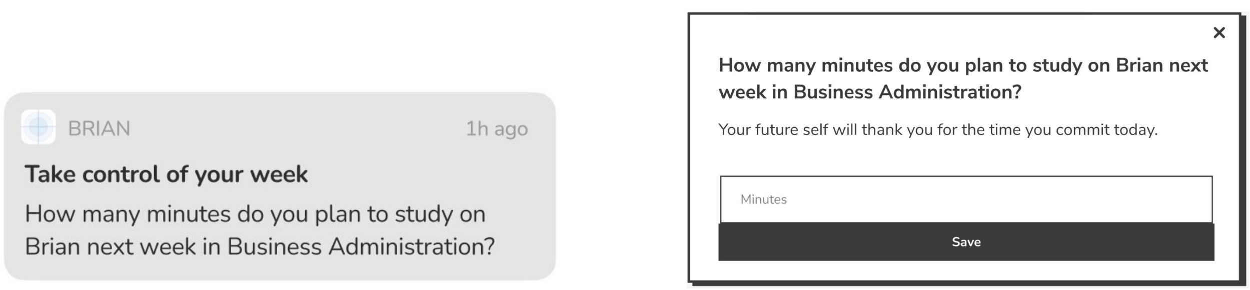

Figure A.3: Text of notification of the Goals treatment and input box.

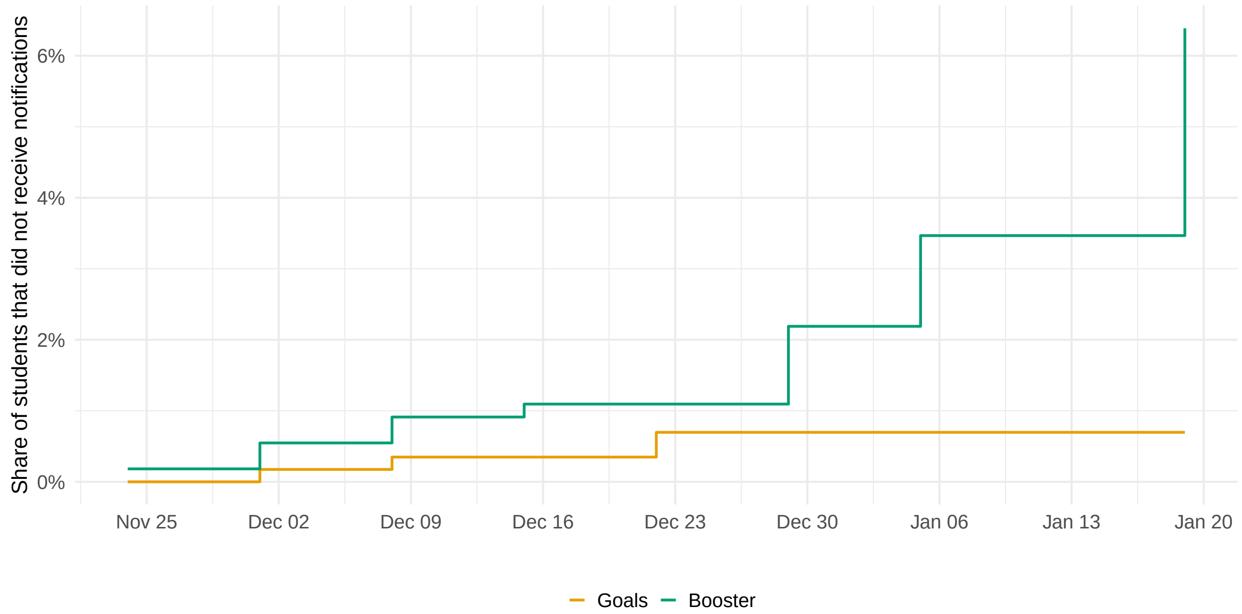

Figure A.4: Cumulative share of students in the treatment arms that did not receive notifications by week of the experiment.

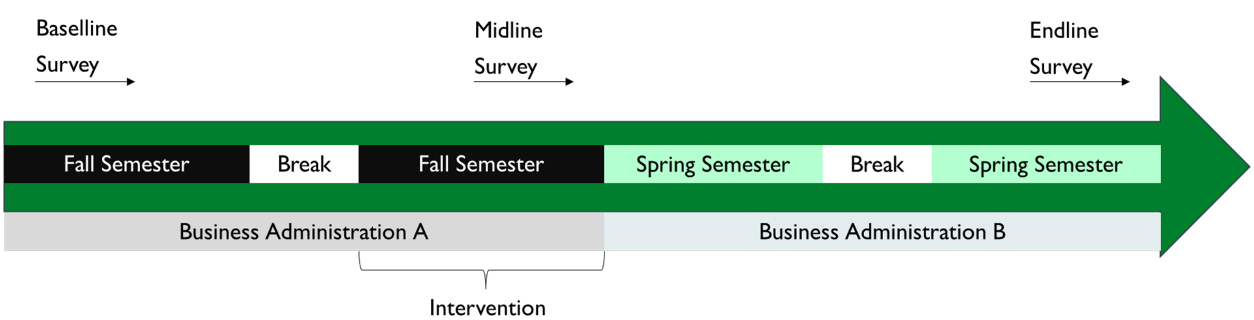

Figure A.5: Timeline of the Experiment.

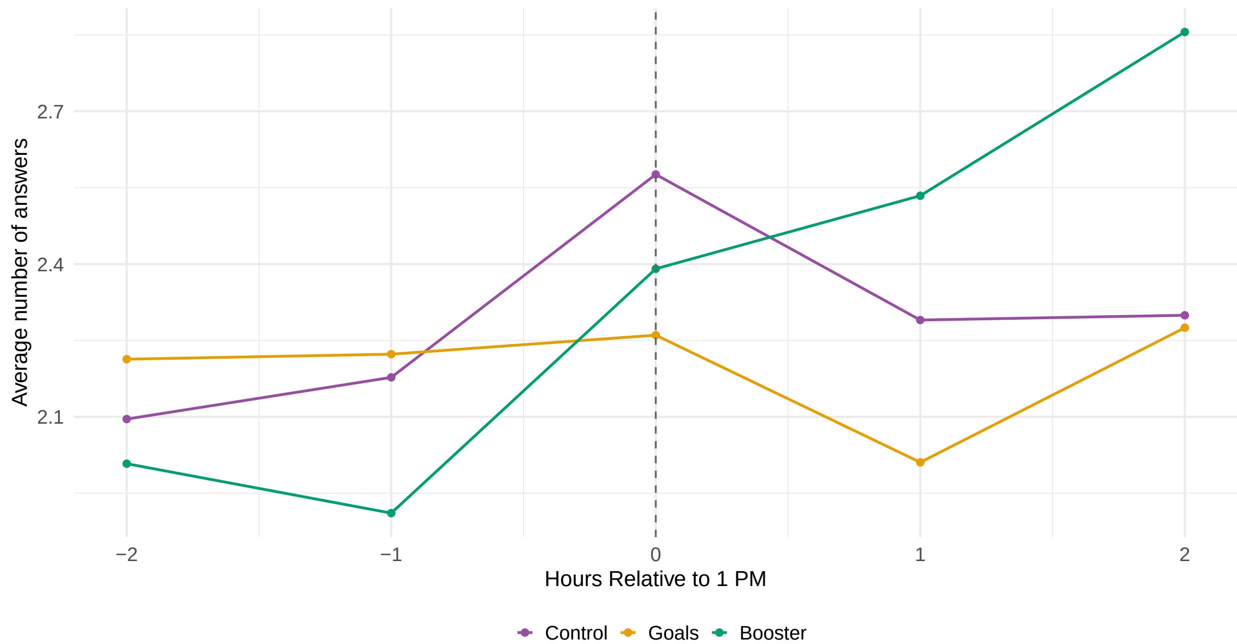

Figure A.6: Average number of answers on Sundays by hour relative to the time of the notification. Only Sundays during the experiment are included.

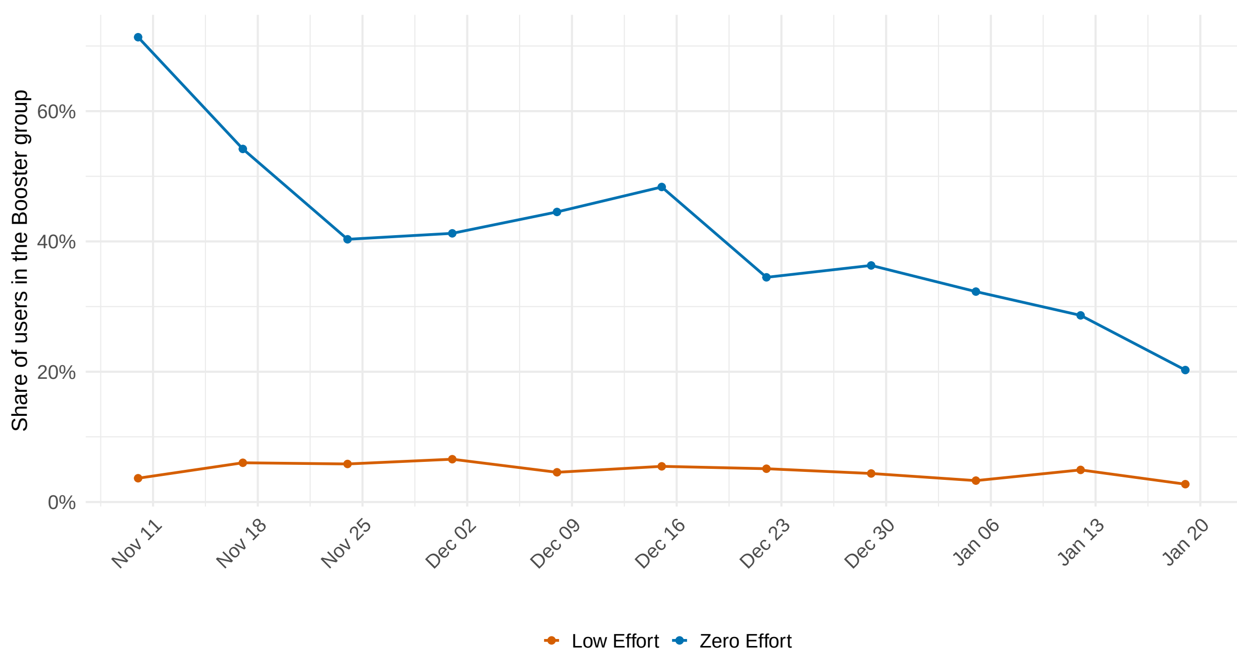

Figure A.7: Evolution of the share of users in the *Booster* treatment with zero app use (0 minutes) or low, but positive, app use (less than 5 minutes) by week of the experiment.

A.3 Surveys

Here I report the text of the survey. The order of the questions related to procrastination, self-esteem, time management, and the attention check was randomized. Self-esteem is measured using the Rosenberg Self-Esteem Scale (Rosenberg 1965); procrastination is measured using the Short General Procrastination Scale (Sirois, Yang, and van Eerde 2019); time management is measured using the Short Time Management Behaviors Scale (Peeters and Rutte 2005).

- What is your matriculation number?
- What is your date of birth?
- **Procrastination**: On a scale from 1 to 5, where 1 means "Strongly Disagree" and 5 means "Strongly Agree", please indicate your level of agreement with the following statements:
  - In preparing for some deadlines, I often waste time by doing other things
  - I am continually saying "I'll do it tomorrow"
  - I often have a task finished sooner than necessary
  - I generally delay before starting work I have to do
  - I usually accomplish all the things I plan to do in a day
  - I usually take care of all the tasks I have to do before I settle down and relax for the evening

  - Even with jobs that require little else except sitting down and doing them, I find they seldom get done for days
  - I often find myself performing tasks that I had intended to do days before
  - I usually buy even essential items at the last minute

- **Self-Esteem**: On a scale from 1 to 5, where 1 means "Strongly Disagree" and 5 means "Strongly Agree", please indicate your level of agreement with the following statements:
  - I feel that I am a person of worth, at least on an equal plane with others
  - I feel that I have a number of good qualities
  - All in all, I am inclined to feel that I am a failure
  - I am able to do things as well as most other people
  - I feel I do not have much to be proud of
  - I take a positive attitude towards myself
  - On the whole, I am satisfied with myself
  - I wish I could have more respect for myself
  - I certainly feel useless at times
  - At times I think I am no good at all

- **Time Management**: On a scale from 1 to 5, where 1 means "Seldom true" and 5 means "Very often true", please indicate the extent to which the following statements apply to you:
  - I feel in control of my time
  - I review my daily activities to see where I am wasting my time
  - I break complex, difficult projects down into smaller manageable tasks
  - I set short-term goals for what I want to accomplish in a few days or weeks
  - I set deadlines for myself when I set out to accomplish a task
  - I look for ways to increase the efficiency with which I perform my work activities
  - I set priorities to determine the order in which I will perform tasks each day
  - I review my goals to determine if they need to be adjusted
  - I finish top-priority tasks before going on to less important ones
  - I make a list of things to do each day and check off each task as it is accomplished
- This is a question to check whether you are paying attention and reading the questions carefully. Please select both "strongly disagree" and "strongly agree" to continue
- What grade do you expect to receive in Business Administration? Please report a number from 1 to 6

- On a scale from 1 to 5, where 1 means “Not at all certain” and 5 means “Extremely certain”, how certain are you about your previous answer?
- In relation to your grade in the Business Administration course, how do you think you will rank compared to others in your cohort?
- On average, how many minutes per day do you spend studying Business Administration? Please provide a specific number.
- What is the highest level of education of your father?
- What is the highest level of education of your mother?

Note that there were some changes in the text of the surveys. In particular, the baseline survey also asked whether the student was employed or not. In both the baseline and midline surveys, parental education was recorded through a single question, without distinguishing between the levels of maternal and paternal education. For a separate research project, in the midline and endline surveys, some additional questions were included. The questions asked about friends at the university, social comparison, and parental expectations. The questions about the expected grade and ranking in the course were related to Business Administration A in the baseline and midline surveys, and to Business Administration B in the endline survey.