Journal of Economic Behavior and Organization 153 (2018) 293–321

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

Can behavioral tools improve online student outcomes? Experimental evidence from a massive open online course☆

Richard W. Patterson

Department of Social Sciences, United States Military Academy, B117 Lincoln Hall West Point, NY 10996, USA

ARTICLE INFO

Article history:
Received 27 September 2016
Revised 24 May 2018
Accepted 29 June 2018
Available online 8 August 2018

Keywords:
Education
Self control
Commitment devices
Reminders

ABSTRACT

In order to address poor outcomes for online students, I leverage insights from behavioral economics to design three software tools including (1) a commitment device, (2) an alert tool, and (3) a distraction blocking tool. I test the impact of these tools in a massive open online course (MOOC). Relative to students in the control group, students in the commitment device treatment spend 24% more time working on the course, receive course grades that are 0.29 standard deviations higher, and are 40% more likely to complete the course. In contrast, outcomes for students in the alert and distraction blocking treatments are statistically indistinguishable from the control.

Published by Elsevier B.V.

People frequently fail to follow through on the plans they make: they fail to meet deadlines at work, finish assignments for school, go to the gym, and deposit money in their savings accounts. In higher education, only 59% of students complete the degree programs they begin,[1] and completion rates are often much lower in online programs and courses. For example, the graduation rate at the University of Phoenix, the largest provider of online degrees in the United States, is only 19%[2] and in massive open online courses (MOOCs), which allow thousands of students to simultaneously access course material, completion rates are often less than 10% (Perna et al., 2013).

The standard neoclassical economic model assumes that people make plans that maximize their intertemporal utility and that they will only deviate from their plans when doing so improves their overall well-being. Evidence from psychology and behavioral economics, however, suggests that people may systematically deviate from their plans in ways that significantly decrease their well-being. In particular, procrastination (Laibson, 1997), forgetting (Mullainathan, 2002), and limited willpower (Baumeister et al., 1998) may lead to detrimental deviations from long-run plans. In environments such as online education, where behavioral factors are likely to keep people from following their plans, interventions such as commitment devices and reminders may significantly increase plan completion and improve well-being.[3]

☆ All opinions expressed in this manuscript are those of the author and do not represent the opinions of the United States Military Academy, Department of Defense, or the United States Army. I thank the National Academy of Education, Spencer Foundation, the Russell Sage Foundation, Stanford Media X, the Cornell Institute for the Social Sciences, and the Cornell Population Center for their financial support of this research. I thank Daniel Greene for collaborating on the experimental design and implementation at Stanford University. I also thank Daniel Benjamin, Damon Clark, Ronald Ehrenberg, Tatiana Homonoff, Jordan Matsudaira, Ted ODonoghue and Cornell seminar participants for their feedback and insights. All errors are my own.

E-mail address: richard.patterson@usma.edu

[1] Source: http://nces.ed.gov/programs/digest/d13/tables/dt13_326.10.asp, 10/29/2014.

[2] Source: http://nces.ed.gov/ipeds/datacenter/institutionprofile.aspx?unitId=aeb2adadacae, October 12, 2014.

[3] While little work has been done to investigate the impact of commitment devices and reminders in education, there is evidence of the effectiveness of commitment devices and reminders in other settings. Commitment devices have been shown to significantly improve effort at work, savings behavior, and health behaviors (Ashraf et al., 2006; Bryan et al., 2010; Kaur et al., 2011). Additionally, recent studies have found significant positive impacts of reminders on savings behavior (Karlan et al., 2010) and health outcomes (Austin et al., 1994; Calzolari and Nardotto, 2012; Krishna et al., 2009).

https://doi.org/10.1016/j.jebo.2018.06.017
0167-2681/Published by Elsevier B.V.

In this study, I design time-management software tools for online students and experimentally test the impact of these tools in a statistics MOOC hosted by Stanford University. These tools include a commitment device, which enables students to pre-commit to daily time limits on distracting Internet activities; an alert tool, which generates an on-screen reminder that is triggered by distracted web browsing; and a distraction blocking tool, which allows students to block distracting websites for up to an hour when they go to the course website. If students struggle with time-management issues, the software treatments may improve student performance and well-being.

My results indicate that the commitment device significantly improves course outcomes relative to the control, alert, and distraction blocking treatments. I find that the commitment device increases course completion by 40% (11 percentage points), improves overall course performance by 0.29 standard deviations, and increases the amount of time students spend on the course website by 24% (5.5 h) relative to the control. Estimates for the impact of the distraction blocking treatment on course outcomes are also positive but smaller in magnitude than the commitment device and are not statistically significant. The alert treatment, however, has no measurable impact on course outcomes. I also find that the differences between the commitment and control are most pronounced in the first weeks of the course and are largest among students who were predicted to do well in the course, given their observable characteristics. In all, this study suggests that time-management issues play a significant role in poor performance among online students, and that commitment devices can have a significant impact on student performance.

This study contributes to the existing literature in several ways. First, this is one of the first studies to test whether tools from behavioral economics can improve completion rates in online education. Second, this study adds insight into the mechanisms driving poor outcomes for online students. Third, by simultaneously testing multiple behavioral tools, this study informs the relative efficacy of interventions intended to address different sources of time-management issues.

1. Background and motivation

1.1. Online education

Online courses are quickly becoming a mainstay in higher education. Between 2002 and 2012, the percentage of universities offering online courses grew from 72% to 87%, the percentage of students taking online courses grew from 9% to 34%, (Allen and Seaman, 2013) and the percentage of undergraduate students enrolled in distance or online-only degree programs grew from 2% to 11% (Ginder and Stearns, 2014). In addition to online courses becoming a large component of accredited degree programs, a number of selective universities, such as Harvard, Stanford, and Cornell now offer Internet-based massive open online courses (MOOCs) to a global population. MOOCs are designed to accommodate thousands of students and have the potential to dramatically broaden access to high-quality instruction. MOOCs typically have open enrollment, are free to join, and have no penalty for dropping out. To date, nearly 8 million students have enrolled in MOOCs to learn about a range of subjects; including science, business, mathematics, information technology, arts, and humanities (Waldrop, 2013).

While the potential benefits of online eduction are large, completion rates are often very low. For example, Xu and Jaggars (2011) find that observationally equivalent community college students are 10–15 percentage points less likely to complete online courses than traditional courses. At the University of Phoenix, the largest provider of online degrees in the United States, the graduation rate for full-time online students is only 19%.[4] In MOOCs, completion rates are often even lower. Perna et al. (2013) examined the completion rates for approximately 1 million students from 16 University of Pennsylvania MOOCs and found that only 6% of students completed the course in which they enrolled.[5]

Although the low completion rates in MOOCs and other online courses are striking, they do not necessarily indicate that students are behaving irrationally. With no cost of enrollment and no penalty for dropping out, many students may be enrolling in courses they do not intend to finish. However, there is evidence that suggests that many students drop out of courses they would have liked to finish and that behavioral factors may contribute to high dropout rates. For example, Wilkowski et al. (2014) examine completion behavior in a MOOC hosted by Google, and find that less than 25% of students who report a goal of earning a certificate of completion ultimately finish the course. Additionally, a number of studies find that students report self-regulation and time-management problems as primary reasons for failure in online courses (Doherty, 2006; Winters et al., 2008). While issues of self-regulation and time-management are likely to impact all students, aspects of the online learning environment may make students particularly susceptible to issues with time-management. Specifically, characteristics of the online course environment, such as anonymity (e.g. Kast et al., 2012) and unstructured scheduling (e.g. Ariely and Wertenbroch, 2002), make students prone to behaviors that could limit their ability to achieve their course goals. Given the disparity between desired and realized outcomes for online students, identifying and addressing behavioral barriers to online academic success could provide significant benefits to students.

[4] This graduate rate accounts for all graduations within 6 years. Source: http://nces.ed.gov/ipeds/datacenter/institutionprofile.aspx?unitId=aeb2adadacae, October 12, 2014.

[5] Perna et al. (2013) define completion by scoring at least an 80% in the course. The authors also find that only 9% of students accessed the last lecture in the course in which they enrolled.

1.2. Time-inconsistent preferences and commitment devices

One reason online students may fail to achieve their long-run course goals is that they behave impatiently and procrastinate their coursework. Economic models of intertemporal choice such as present-biased preferences (Laibson, 1997; O'Donoghue and Rabin, 1999) and dual-self models of self-control (Fudenberg and Levine, 2006) predict this type of impatient, time-inconsistent behavior.[6] The influence of time-inconsistent behavior may be particularly important in education settings where the benefits of increased effort are often realized far in the future. For example, Levitt et al. (2012) find evidence of time-inconsistency among high-school students who perform significantly better on standardized tests when they are offered a financial incentive that is delivered immediately following the test, but perform no better than a control group when the financial incentive is delayed by just a month.

While impatience may lead to detrimental outcomes for online students, both theory and evidence from the field suggest that commitment devices can help people who are aware of their time-inconsistent behavior to bring their short-run behavior in line with their long-run interests.[7] Commitment devices can increase the likelihood that an individual will behave patiently by making future procrastination more difficult or costly (Bryan et al., 2010). Commitment devices have been shown to significantly increase desired long-run behaviors including effort at work (Kaur et al., 2011), savings behavior (Ashraf et al., 2006; Thaler and Benartzi, 2004), and smoking cessation (Giné et al., 2010). While there is limited evidence of the impact of formal commitment devices in education, Ariely and Wertenbroch (2002) find that students hired to proofread multiple papers over a three week span performed significantly better when given the option to set binding intermediate deadlines.[8] If present-biased preferences are a significant detriment to performance in online education, providing online students with formal commitment devices may help them achieve their course goals.

1.3. Limited memory and reminders

In addition to behaving impatiently, online students may forget about their coursework. If online students experience failures of prospective memory, or forget about their prior intentions or plans, they may not achieve their course goals.[9] Economic models of limited prospective memory and inattention (e.g. Ericson, 2014; Karlan et al., 2010; Mullainathan, 2002; Taubinsky, 2014) predict that people will forget to follow through on their plans in ways that significantly reduce their well-being.

A simple way to address limited memory is to provide individuals access to reminder technologies.[10] Reminders have been shown to increase college matriculation repayment of loans (Cadena and Schoar, 2011), savings accounts deposits (Karlan et al., 2010), medication adherence (Zogg et al., 2012), and exercise (Calzolari and Nardotto, 2012). Given that online students must have access to a computer in order to complete their work, it is likely that they already have access to computerized reminder technologies (e.g. email, calendar, reminder software) which may limit the impact of additional reminders. However, if available reminder technologies are difficult to use or if students are over-confident in their ability to remember their plans, providing reminders may be an effective way to help students achieve their course goals.[11]

1.4. Limited willpower

Another factor that may limit students' ability to complete their goals is limited willpower. Theories in economics (e.g. Fudenberg and Levine, 2012; Ozdenoren et al., 2012) and psychology (e.g. Baumeister and Vohs, 2003) model willpower as a depletable resource and suggest that resisting temptation reduces one's subsequent ability to exercise willpower. For instance, Baumeister et al. (1998) find that subjects who were required to resist the temptation to eat chocolate in the first stage of an experiment exerted significantly less effort on a puzzle task in the second stage of the study. In a laboratory experiment that has similar elements to this study (participants work on a computerized task with the temptation of internet distractions), Houser et al. (2018) find that individuals often delay the use of a commitment device or delay giving in to the temptation to browse the internet, suggesting that self control is depleted over time. If students have limited willpower, exposure to factors that tax willpower may leave students too fatigued to complete the course tasks they start. In the context of this study, providing a mechanism to eliminate the temptation of entertaining or distracting websites may increase the willpower students have available to devote to the course.

---

[6] Time-inconsistent preferences describe a situation where the value of trade-offs between two different moments changes over time (Laibson, 1997). Perhaps the most common form of time-inconsistent preferences is procrastination—where individuals behave more impatiently in the moment than they would have liked to from a prior perspective.

[7] If people are naïve about their time-inconsistent preferences and mistakenly believe that they will behave patiently in the future, then they are unlikely to seek out and use commitment devices.

[8] Students who were given equally spaced deadlines for each paper, however, outperformed both those given one deadline or the option to set multiple deadlines. This evidence is consistent with the students exhibiting some level of naïvete about their time-inconsistent preferences.

[9] See McDaniel and Einstein (2007) for a review of prospective memory.

[10] While most models of limited memory (e.g. (Karlan et al., 2010; Taubinsky, 2014) predict that reminders will increase plan completion, Ericson (2014) suggests that reminders may reduce plan completion among present-biased individuals under certain circumstances.

[11] There is evidence that people are overconfident in their ability to remember their plans. For example, Ericson (2011) finds that MBA students significantly overestimate their ability to remember to claim a payment in six months. Students' decisions suggest an expectation of claiming payments 76% of the time, while only 53% of students actually claim the payment.

2. Experimental design and population

2.1. Experimental context and population

Participants for this study were recruited from enrollees in a nine-week Stanford statistics massive open online course (MOOC) which was held in 2014. This completely online course was administered by Stanford University on the Stanford OpenEdX platform.[12] Although the course was administered by Stanford, course enrollment was free, open to anyone worldwide, and provided no formal credit at Stanford University. Students, however, could receive a completion certificate or certificate with distinction by scoring at least 60% or 90% in the course, respectively. Scores for the course were composed of a multiple-choice final exam (45 points or 45%), nine weekly homework assignments (45 points or 45%), and participation in 53 short quizzes (10 points or 10%).[13] To ease interpretation of course grades, I convert raw scores (out of 100) to z-score measures.[14] There was no limit on how fast students could complete coursework, but students needed to submit homework assignments by weekly deadlines in order to receive credit.[15] Students who quickly completed their coursework still needed to wait to take the final exam, which was only made available during the final week of the course. The course content was primarily delivered in approximately 60 downloadable lecture videos that typically lasted between 10 and 20 min and covered a number of topics in statistics including basic statistical measures, probability distributions, statistical inference, statistical tests, and regression analysis.[16] Stanford tracked the time students spent working on the course and these data were added to the course grade and assignment submission data to construct the academic outcomes that I analyze in this study.

Students were recruited via email and were told that the study would test whether computerized time-management tools could help students use their time more effectively and complete courses more quickly.[17] Students were also incentivized to participate with $12 in Amazon.com gift cards—$5 for completing the enrollment survey and installing time management software and $7 for using software and completing a post-study survey.[18] My primary sample consists of the 657 students who participated in the MOOC, completed a pre-study survey, and installed software prior to the first course assignment deadline (a participation rate of 18%).[19] This analysis excludes 120 students who completed the pre-study survey and installed software prior to the first assignment deadline, but never visited the course website. Assignment to treatment condition was uncorrelated with whether students ever visited the course website ($F=0.5$, $p=0.68$). Participants in this study were randomly assigned to one of four treatment groups: (1) control, (2) commitment device, (3) alert, and (4) distraction blocking.

Appendix Table A.1 reports descriptive statistics for participating students.[20] Randomization appears to successfully generate balance across treatment groups, with only 2/44 variables differing by treatment at the 5% level.[21] Panel A of Appendix Table A.1 shows that participating students were highly educated (85% of students are college graduates) and geographically dispersed, with only 28% of students taking the course from the United States. International students predominately took the class from Europe (24%), Asia (20%), and Africa (13%). Additionally, Panel B of Appendix Table A.2 reveals that students in this study had ambitious course goals, with a majority stating their goal was to finish the course on-time for a certificate of completion (67%) or to finish all coursework at their own pace (21%). The most commonly reported reasons for taking the course were general interest in the topic (94%), personal growth (93%), and relevance to work, school, or research (92%). Panel C of Appendix Table A.1 reports variables related to self-control and indicates that, on average, students wanted to decrease the time they spent on distracting websites each day by one hour. Although randomization ensures that estimates of treatment effects are internally valid, selection into study participation may influence the generalizability of the estimates.

[12] OpenEdX is an open source version of the MOOC platform developed by EdX. While the platform is open source and freely available to all, Stanford retains control of all content, data, and licensing associated with the course.

[13] Students were allowed to take quizzes as many times as they wanted but were only allowed to submit each homework assignment and final exam once. The lowest grade among the nine homework assignments was dropped. All quiz, homework, and test questions were either multiple-choice or numerical entry and were computer graded.

[14] The course grade z-scores are calculated using raw scores from all students who participate in the course, not just those of study participants. The z-score (mean=0, sd=1) measures the standard deviations away from mean performance for each student.

[15] Students received zero points for assignments they failed to submit on time.

[16] Supplemental readings and transcripts of lecture videos were also available to students.

[17] A copy of the recruitment email can be seen in Appendix Fig. A.1.

[18] All study procedures were approved by both Cornell and Stanford University Institutional Review Boards (IRBs) and all students provided informed consent in order to participate.

[19] The 18% participation rate is calculated among the 3630 students who enrolled in the course prior to the start date and visited the course at some point during the semester, and excludes individuals who never visit the course website. 240 additional students enrolled in the study after the first week. A majority of these 240 students came from 2612 students who enrolled in the course after the start date and were recruited to join the study at the beginning of the second week. I focus my analysis on the 657 students who are treated in each week during the course and for whom I am able to analyze a balanced panel of weekly data, but also provide results that include students who enroll during the second week of the course (n=897) in the appendix. The participation rate was higher than expected, but the overall study participation was much lower, based on more than 20,000 students enrolling in the course's previous iteration.

[20] Statistics reported in Appendix Table A.1 were collected in the pre-study survey.

[21] Specifically, students in the control group were more likely than those in the treatment groups to be from Africa (p=0.049) and indicate that they took the course because it was relevant to a job, school, or research (p=0.007). Neither of these patterns are indicative of the type of selection that could spuriously generate the observed treatment effects. The full list of control variables for the study are listed in Appendix Table A.1.

Appendix Table A.2 compares the age, gender, and education level of study participants to other students in the course and indicates that study participants are 11% more likely to be female and 68% more likely to hold a Ph.D. or M.D., but are otherwise similar to other students taking the course.

2.2. Research design

Participants in this study were randomly assigned to one of four treatment groups: (1) control, (2) commitment device, (3) alert, or (4) distraction blocking. Students were assigned to treatment conditions at the individual level by a random number generator embedded in the pre-survey software. To ensure that participants did not differentially select themselves into the study based on the treatment conditions, all students installed the same basic version of the software and were not informed of their software functionality until after they had successfully installed the software and completed the enrollment survey.[22] The particular functions of the treatment software were not turned on until the course started, or the day following installation if students installed the software after the course began.[23] The software was designed for all Windows, Linux, and OSX operating systems, and had limited functionality on iOS and Chrome mobile operating systems. The software also worked with all major internet browsers including internet explorer, chrome, firefox, safari, and opera. When running, this software tracked and categorized time spent in the active application or browser window.[24] Each activity was categorized into groups such as email, shopping, news, entertainment, social networking, writing, and education and activity received a productivity score of unproductive, neutral, or productive.[25] This information collected by the software was used to execute each of the treatment conditions described below. The predicted impact of each treatment is modeled in Appendix B.

2.2.1. Control

Students assigned to the control treatment installed the most basic version of the time-management software in the study. The control software, like the software in all other treatments, tracked and categorized the student's computer activity. All study participants, including those assigned to the control group, were able to view summary time-use reports of their computer use by connecting to the time-management website (for an example of the time-use report, see Appendix Fig. A.2). Students in the control group received no other software tools. Students in the control group were given access to these reports in order to justify the request to install time-tracking software, provide a comparable study experience to those in the other treatment groups, and to reduce the probability of experimenter demand effects influencing the results (e.g. Zizzo, 2010).[26] Table 1 reports student interaction with the treatment software and shows that over the course of the study, students in the control group accessed summary reports slightly more often (19 times) than those assigned to other treatments (16 times), which is significant at the 10% level. While it seems unlikely that this difference in accessing reports lead to significant differences in course outcomes, estimates of the impacts of other treatments can be considered lower bounds.[27]

2.2.2. Commitment device

In addition to having access to time-use summary reports, students assigned to the commitment device treatment were able to set a limit on distracting Internet time each day. To maximize the expected impact of the treatment, students were initially assigned a limit that corresponded to the goal stated in the pre-study survey. This approach leverages the tendency people have to stay with a default choice (Madrian and Shea, 2001, e.g) and bypasses the issue of naïve students being unwilling to initially opt into a commitment treatment. Participants in this treatment group were sent a daily email at 6:45 a.m. that informed them of their current limit and asked them whether they wished to reset their limit (see Appendix Fig. A.3 for an example of how students set their distracting limit).[28] Once students exceeded their set limit, distracting websites were blocked (blocked screen shown in Appendix Fig. A.4). After exceeding their limit, students were only able to unblock websites on a site-by-site basis and needed to indicate a reason for unblocking each site.

The commitment device has the potential to address issues of present-biased preferences by allowing students to make future distracting computer use more costly. The commitment device makes distracting computer use more costly after

[22] I worked with RescueTime, a company that makes time-tracking software, to develop the software tools used in this study. RescueTime implemented the design for each tool and provided software support throughout the study.

[23] Students who did not complete the enrollment survey or were unable to install software at the time of the survey had software functionality turned on the day after installation, but did not receive messaging explaining their treatment condition.

[24] The software was programmed to automatically run when the participant's computer was turned on. The software could not be closed from any menu option and could only be turned off by manually quitting the application from the computer's task manager/activity monitor function. Activities were tracked at the application and web domain level, and keystrokes or actions taken within an application or pages within a web domain were not recorded. If multiple applications or browser tabs were open, the activity was attributed to the application or webpage with the most recent action. When a person stopped interacting with an application or website the software stopped tracking activity even when the application or website remained open.

[25] These categorizations and productivity scores were defined by RescueTime defaults. These defaults were set by an algorithm that combined website query information with aggregated user scores.

[26] Experimenter demand effects refer to experimental subjects changing behavior in order to conform with what an experimenter's apparent expectations.

[27] Considering treatments effects as lower bounds assumes that providing students access to information about productivity does not have a negative effect on performance.

[28] Time of email was according to the timezone registered by the participant's IP address.

Table 1
Treatment summary statistics.

	Control	Commitment	Alert	Blocking	Total
Software summary					
Days logged on software	38.32	33.75	36.75	37.45	36.60
	(24.08)	(24.84)	(24.01)	(25.13)	(24.52)
Avg hours productive	1.74	1.71	1.62	1.88	1.74
	(1.49)	(1.42)	(1.34)	(1.47)	(1.43)
Avg hours unproductive	0.94	0.90	1.00	0.90	0.93
	(0.97)	(0.93)	(1.33)	(1.03)	(1.08)
Avg hours on course	0.16	0.19	0.15	0.15	0.16
	(0.24)	(0.24)	(0.20)	(0.21)	(0.22)
Times visited RescueTime	18.59	15.53	17.07	15.35	16.67
	(17.25)	(15.42)	(16.99)	(15.68)	(16.39)
Commitment Device					
Commitment emails sent	–	67.18	–	–	–
	–	(8.32)	–	–	–
Commitment (hours)	–	2.69	–	–	–
	–	(2.71)	–	–	–
Times commitment exceeded	–	4.06	–	–	–
	–	(8.25)	–	–	–
Times commitment changed	–	0.92	–	–	–
	–	(1.79)	–	–	–
Avg change (hours)	–	2.61	–	–	–
	–	(2.86)	–	–	–
Alerts					
Alerts sent	–	–	48.19	–	–
	–	–	(77.64)	–	–
Distraction Blocking					
Times prompted	–	–	–	9.22	–
	–	–	–	(10.33)	–
Times initiated	–	–	–	1.93	–
	–	–	–	(3.55)	–
Average duration	–	–	–	38.24	–
	–	–	–	(15.83)	–
Observations	170	160	166	161	657

Notes: Standard deviations in parentheses. “Days logged on software” is a count of the number of days the time-management software tracked any time use. Summaries for hours of productive, unproductive, and course time exclude days for which the software was inactive.

students exceed their limit by increasing the difficulty of accessing distracting websites. The commitment device is also likely to make distracting computer use more costly prior to the limit being exceeded by creating a trade-off between current and future distracting time. Furthermore, the commitment device may make distracting computer time more costly if students see their distracting limit as a goal and experience disutility if they exceed their limit. The commitment device may also address issues of limited memory by providing students with a daily email and issues of limited willpower by blocking distracting websites.[29]

Column 2 of Table 1 summarizes student use of the commitment software. Over the duration of the study, students in the commitment device treatment set an average limit of 2.7 h and students only exceeded this limit an average of four times during the study. Although students had the flexibility to change their distracting limit on a daily basis, students rarely did only changing their limit an average of one time during the study.

2.2.3. Alert

Students in the alert treatment triggered an on-screen reminder after each half-hour they spent on distracting websites (see Appendix Fig. A.5 for example).[30] This alert reported the amount of time the student spent on distracting websites and provided students with a link to the course website. The purpose of this design was to deliver a reminder that was salient, unlikely to disrupt productive activity, and most likely to occur when the student had time available to work on the course. By providing targeted reminders to students, the alert treatment has the potential to address issues of limited memory.[31] In addition to providing students a reminder about the course, the alert treatment also provided students feedback about the time they had spent on distracting websites that day. This feedback could either increase or decrease student productivity as students might find the feedback motivating or discouraging. Table 1 indicates that students in this treatment received alerts on a regular basis, receiving an average of 48 alerts during the course.

[29] See Appendix B for more detailed predictions about the commitment device treatment.
[30] The alert opened in a new web browser window that occupied a significant portion of the student's screen.
[31] See Appendix B for a more detailed predictions regarding the alert treatment.

2.2.4. Distraction blocking

Students assigned to the distraction blocking treatment were prompted with an option to block websites for 15, 30, or 60 min when they went to the course website (see Appendix Fig. A.6 for example). This distraction blocking prompt was delivered to students at most once per day and occurred the first time a student went to the course website each day. In contrast to the commitment device, which allows students to block distracting sites in the future, the distraction blocking tool only allows students to immediately block distracting sites. Additionally, students were required to visit the course website in order to interact with the distraction blocking tool. As a result, the distraction blocking tool may address issues of limited willpower, but is unlikely to address issues of present-bias preferences or limited memory.[32] Table 1 shows that students in the distraction blocking treatment were prompted to start a distraction-free study session 9.2 times during the course, and chose to initiate a distraction-free study session an average of 1.9 times during the course. When the students did initiate a distraction-free study session, the average duration was 38 min.

3. Results

3.1. Impact on aggregate course outcomes

In this Section 1 test whether the commitment device, alert, and distraction blocking tools impact student effort and performance. Measures of effort include number of homework assignments submitted and hours logged on the course website, while measures of student performance include standardized course score (z-score) and course completion.[33] To evaluate the impact of treatments on student outcomes, I estimate:

$$y_i = a + \sum_{j=1}^{3} \gamma_j T_{ij} + \nu \mathbf{X}_i + \epsilon_i \tag{1}$$

where y_i is a measure of effort or academic performance for individual *i; T_{ij}* is an indicator of the treatment assignment for individual i; and $\mathbf{X}_i$ is a vector of student characteristics collected in the pre-study survey including age, education, income, location, course goals and objectives, previous course experience, and reported measures of self-control.[34]

The results of the estimation of Eq. (1) are presented in Table 2.[35] Table 2 reports estimates with two sets of *p*-values: those generated with standard ordinary least squares regressions with robust standard errors and those that apply a Bonferroni adjustment to the *p*-values to account for the fact that three treatments are being simultaneously tested in each specification.[36] First I estimate the impact of treatment assignment on the amount of time students spend on the course website. This measure of effort has the advantage of incorporating all course activities, not just those that are graded. Column 1 of Table 2 shows that the commitment device increased time spent on the course website by 5.5 h, or 24%, relative to the control (significant at the 10% level). Students in the commitment device treatment also spent significantly more time on the course website than students in the alert treatment (8.8 h, significant at the 1% level) and distraction blocking treatment (4.6 h, significant at the 10%). Although imprecisely estimated, students in the alert treatment actually spent 3.3 fewer hours working on the course than those assigned control while the students in the distraction blocking treatment spent 0.8 more hours than those in the control group. Neither the alert treatment nor the distraction blocking treatment led to statistically significant changes in time spent on the course.

One potential weakness of using course time as a measure of effort is that it cannot account for any impact the treatments have on how effectively students spend their time. If the treatments lead students to spend their time more efficiently when going to the course, then the estimated of impact the treatments may be biased downward. Homework submissions provide an additional measure of course effort that is not subject to this potential bias of the course time measure. Column 2 of Table 2 presents estimates of the impact of the treatments on the number of homework assignments submitted. Consistent with the course time results, I find that the commitment device has a significant impact on homework submissions, increasing the number of homework assignments submitted by 0.91—an increase of 27% relative to the control (significant

[32] See Appendix B for a more comprehensive discussion of predictions concerning the distraction blocking study treatment.

[33] Time spent on course is calculated by Stanford from web activity logs. This calculation is likely an over estimate of actual course time as Stanford counts all time between course events that are no longer than 30 min as time spent on course. The last event in any session is counted as lasting for 500 s. Z-scores were constructed using the data from all students enrolled in the MOOC. Completion is defined by meeting the 60% score threshold for earning a certificate of completion.

[34] While successful randomization insures unbiased estimates of the treatment effects, inclusion of controls reduces the residual variation in the estimate of equation (1) and therefore increases the precision of the estimates of the treatment effects. The full vector of control variables includes *age*^{2} and all variables are listed in Appendix Table A.1.

[35] Appendix Table A.3 reports estimates that include participants who join the study after the first week. The specifications estimated in Appendix Table A.3 are not statistically significant, but are consistent with those presented in Table 2.

[36] Appendix Table A.4 includes two additional approaches to constructing *p*-values. The first constructs empirical *p*-values constructed from a randomization-based estimation procedure. This approach yields *p*-values that are similar in significance and magnitude to those generated by a simple OLS regressions with robust standard errors. The second additional approach accounts for the fact that not only am I testing three treatments simultaneously in each specification, but that I am examining the effects of the treatments on four separate outcomes. To adjust my p-values to account for all 12 treatment/outcome combinations, I construct 12-way multiple hypothesis corrected *p*-values using a methodology proposed by Romano and Wolf (2016). Employing this methodology yields *p*-values that are no longer significant at conventional levels for any treatment/outcome combination.

Table 2
Impact of treatments on course outcomes.

	Course effort (Hours)	Homework submitted	Course grade (Z-score)	Course completion
Commitment device	5.491	0.909	0.291	0.107
	(3.085)	(0.403)	(0.148)	(0.0497)
Standard *p*-values	0.076	0.024	0.050	0.031
Bonferroni-adjusted *p*-values	0.211	0.070	0.143	0.090
Alert	−3.339	0.267	0.0109	0.0108
	(2.597)	(0.415)	(0.150)	(0.0503)
Standard *p*-values	0.199	0.520	0.942	0.839
Bonferroni adjusted *p*-values	0.486	0.889	1.000	0.996
Distraction blocking	0.848	0.577	0.0966	0.0135
	(2.826)	(0.412)	(0.149)	(0.0498)
Standard *p*-values	0.764	0.161	0.518	0.787
Bonferroni adjusted *p*-values	0.987	0.409	0.889	0.990
Dep var mean	22.38	3.730	0.711	0.289
Demographics	y	y	y	y
Course variables	y	y	y	y
Self-control variables	y	y	y	y
Observations	657	657	657	657

Robust standard errors in parentheses. Standard *p*-values are constructed from t-tests of regression coefficients with robust standard errors. Bonferroni-adjusted *p*-values adjust p-values to account for the simultaneous testing of all three treatments. **Demographic variables** include gender, age, age^2, education, income, continent, and indicators for missing age and income variables. **Course variables** include course goals: *finish for certificate, finish at own pace, complete some assignments, or watch some videos*, reasons taking the course: *general interest, relevant to school/work/research, career change, fun, try online course, improve English*, type of computer: *personal laptop, personal desktop, work computer*, previous online courses started, previous online courses finished, previous statistics courses taken, interest level in course, expected course hours, and importance of finishing course. **Self-control variables** include distracting time goal, desired change in distracting time, self-reported difficulty breaking habits, distractibility, ability to resist temptation, level of self-discipline, and take actions that are regretted in long run.

at the 5% level). While the impact of the alert and distraction blocking treatments on homework submission patterns are smaller than those estimated for the commitment device and statistically indistinguishable from the control, estimated effects for both groups are positive (0.27 and 0.58 additional homework assignments, respectively) and large effects cannot be ruled out for these groups.

The impact of the treatments on student academic outcomes corresponds closely with those estimated for effort. Column 3 of Table 2 shows that the commitment device improves total course performance by 0.29 standard deviations, which is significant at the 5% level. To provide some context, this is roughly the same difference in course performance observed between students with Ph.D.s or M.D.s and students with bachelor's degrees (0.28 standard deviations, significant at the 1% level). In contrast, the alert treatment has essentially no measured influence on course performance (an increase of 0.01 standard deviations) and the estimated impact of the distraction blocking treatment is one-third the size of the commitment device (0.10 standard deviations) and statistically indistinguishable from the control.

Finally, column 4 of Table 2 indicates that the commitment device has a large impact on course completion, increasing completion rates by 40% or 11 percentage points (significant at the 5% level). The alert and distraction blocking treatments, however, have no measurable impact on completion, with point estimates that are close to zero (both associated with 1 percentage point increase in completion) and that are significantly smaller than the estimated impact of the commitment device (both significant at the 10% level).

In total, the reported results in Table 2 indicate that the commitment treatment has a significant impact on both course effort and outcomes. In contrast, neither the alert nor the distraction blocking treatment have a significant impact on either student effort or performance. Given that the alert is designed to address limited attention, the distraction blocking tool is designed to address limited willpower, and the commitment device is designed to address limited attention, limited willpower, and present-biased preferences, these results are consistent with students procrastinating or exhibiting present-biased preferences, having limited attention, or having limited willpower. Importantly, the effects of the alert and distraction blocking tools are imprecisely estimated and neither their efficacy nor the potential roles of limited memory and limited willpower can be ruled out. Furthermore, it is possible that the reminder, distraction blocking, and commitment aspects of the commitment tool interact in ways that only make the tool effective when all these aspects are combined. Nevertheless, the results do indicate that the commitment device is more effective than the other treatments in improving course outcomes.

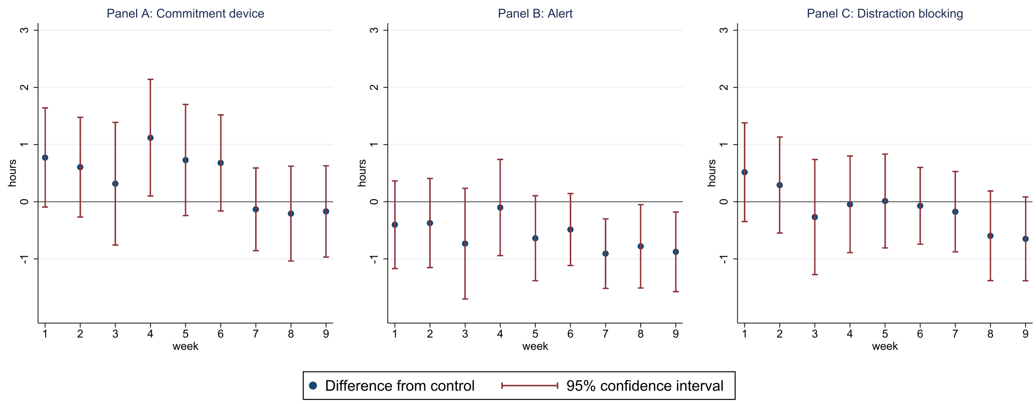

Fig. 1. Effort (in Hours), by Week.
Panels A, B, and C show estimated differences in weekly hours spent on course between treatment and control ($\gamma_j + \lambda_{jt}$ in Eq. 2) for commitment device, alert, and distraction blocking treatments, respectively. Bounds represent 95% confidence intervals. Estimates are generated from an OLS panel estimation with controls for demographic, course, and self-control variables. Standard errors are clustered at the individual level.

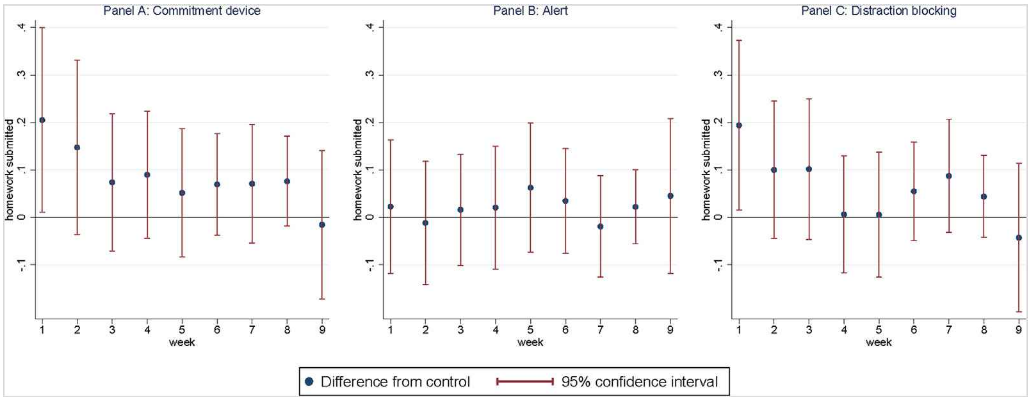

Fig. 2. Homework Assignments Submitted, by Week.
Panels A, B, and C show estimated differences in weekly homework submissions between treatment and control ($\gamma_j + \lambda_{jt}$ in Eq. 2) for commitment device, alert, and distraction blocking treatments, respectively. Bounds represent 95% confidence intervals. Estimates are generated from an OLS panel estimation with controls for demographic, course, and self-control variables. Standard errors are clustered at the individual level.

3.2. Timing of treatment effects

How the software tools impact student effort over time has important implications for how to interpret and generalize the results of this study. If the differences in student effort between treatments and control are present throughout the duration of the course, then this suggests that treatments may be effective in addressing long-run behavioral issues. However, if differences in effort between treatment and controls are only observed in the first few weeks of the course, then the impact of the software tools may not generate persistent long-run effects for students. Because student interaction with the treatment software is observed, I am able to examine how patterns in software use compare to trends in course effort. To investigate how the software treatments impact course effort over time I estimate:

$$y_{it} = a + \sum_{j=1}^{3} \gamma_j T_{ij} + \sum_{t=2}^{9} \theta_t week_{it} +, \sum_{j=1}^{3} \sum_{t=2}^{9} \lambda_{jt} T_{ij} * week_{it} + \nu \mathbf{X}_i + \epsilon_{it} \tag{2}$$

where y_{it} is a measure of effort for individual i in week t; T_{ij} is an indicator of individual treatment assignment; $week_{it}$ is an indicator for the week in which the academic outcome was observed for individual i; $T_{ij}{*}week_{it}$ is the interaction between treatment assignment and the week of the course, and other variables are as previously specified. Standard errors are clustered at the individual level. Results of this estimation for time spent on course and homework submissions are graphically presented in Figs. 1 and 2, respectively. The points in Figs. 1–2 represent estimated differences in course hours and homework submissions between treatments and control in each week ($\gamma_j + \lambda_{jt}$) with bounds indicating 95% confidence intervals. To interpret the results of these estimations it is important to note two things—first, the weekly differences between each treatment and control ($\gamma_j + \lambda_{jt}$) capture both the persistent effects of previous treatment and contemporaneous effects of the current treatment. Second, students were able to work ahead in the course and the extent to which treatments lead

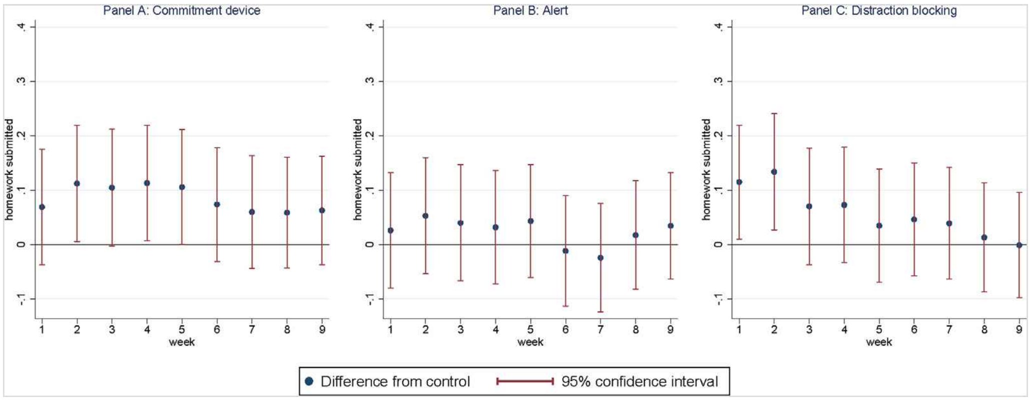

Fig. 3. On-time Homework Submissions, by Week.
Panels A, B, and C show estimated differences in whether students submitted weekly homework assignments on-time between treatment and control ($\gamma_j + \lambda_{jt}$ in Eq. 2) for commitment device, alert, and distraction blocking treatments, respectively. Bounds represent 95% confidence intervals. Estimates are generated from an OLS panel estimation with controls for demographic, course, and self-control variables. Standard errors are clustered at the individual level.

students to work ahead leads to larger differences in early weeks and smaller differences in later weeks than would have otherwise been observed. Nevertheless, these figures do provide insight into when in the course the differences in effort are observed.

While the estimates are somewhat imprecise, Figs. 1 and 2 show three interesting patterns. First, differences in effort between the commitment and control group, in terms of hours of course time and homework submissions, are largest at the beginning of the course but remain positive and significant for the majority of the course. Second, the alert treatment appears to have no positive impact on course outcomes at any point during the study. Third and finally, the differences in effort between the distraction blocking treatment and control are significant at the beginning of the course but then dissipate after the first two to three weeks.

As previously mentioned, the ability of students to work ahead in the course makes it difficult to make inference about the persistence of treatment effect. One way to identify an upper bound for the persistence of the treatment effects on homework submissions is to estimate how treatments affect on-time submission of each weekly assignment.[37] I estimate the impact of treatment on whether each weekly assignment was submitted using the same estimation strategy outlined in Eq. (2), except y_{it} is now an indicator for whether the assignment due in week t was submitted by week t. Results of this estimation are presented in Fig. 3. This plot shows similar patterns to Fig. 2. The commitment device increases the probability of each week's homework submission by approximately 10%, the alert has no significant impact on homework submissions at any time during the study, and the distraction blocking treatment significantly increases the probability that the first few weeks' homework assignments are submitted, but this difference declines over time.

3.3. Timing of software use

How students use software over time provides additional context for the treatment effects reported in Figs. 1–3. Panel A of Fig. 4 reports trends in whether software was installed in each week, and shows that students in all treatment groups, including the control, are significantly less likely to have software installed as the course progresses.[38] Given the significant differences in how frequent and intense the software experiences are by treatment, it is somewhat surprising that there are not large differences in software use by treatment (reported in Panels B, C, and D of Fig. 4). Students in the commitment device treatment are 8% less likely to be running the software relative to students in the control (significant at the 10% level), and students in the alert and distraction blocking treatments do not have statistically significant differences in software use from the control.

In addition to being less likely to have software installed over time, students in the commitment device and distraction blocking treatments who continue to use the software become less likely to utilize treatment components of their software as the course progresses. Appendix Fig. A.7, which reports patterns of student interaction with the commitment software,[39] illustrates how students make their commitments significantly less restrictive over the duration of the course. By the end of the course, students allow themselves nearly twice as much distracting time and reach their limit less than half as

[37] Because I cannot assign course time to particular assignments, I am unable to perform a similar exercise for hours spent on course.

[38] I am unable to distinguish between students who have actually uninstalled software, have turned off software, or have not used computers in a given week. However, if the student's computer does not send any time use data to the server then the student can have no interaction with the study software during the week. I therefore define software being installed as the server receiving any time-use data from the student's computer during the week.

[39] These patterns are reported for students who run the software in each week.

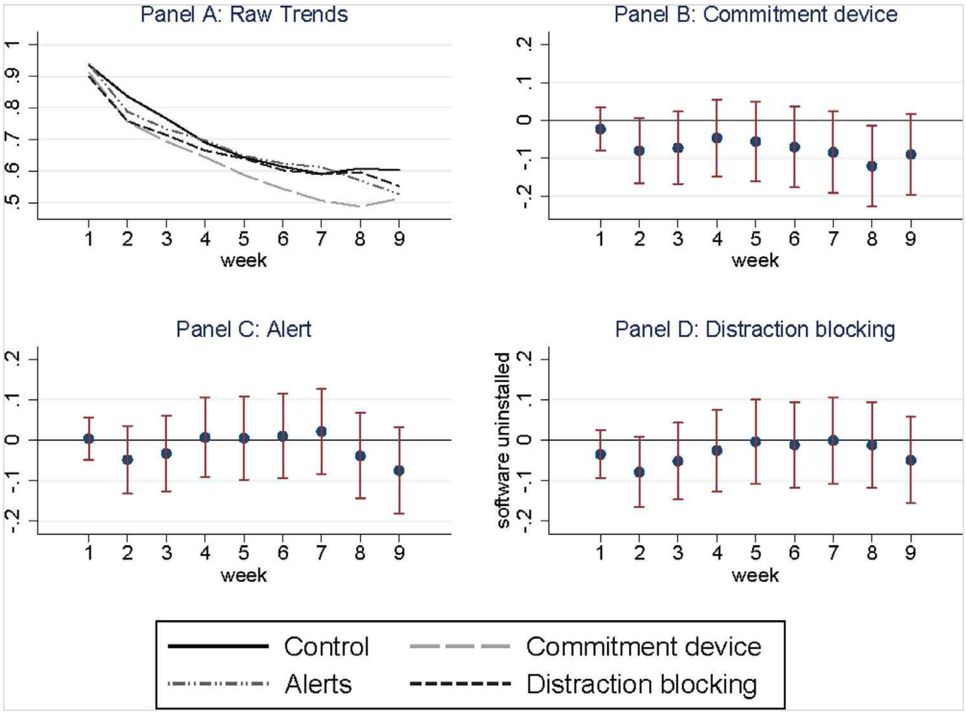

Fig. 4. Software Installed, by Week.
Panel A presents the raw trends for whether software was installed by treatment and week.
Panels B, C, and D show coefficient estimates and 95% confidence intervals for commitment device, alert, and distraction blocking treatment effects, respectively. Estimated coefficients and confidence intervals are generated from an OLS panel estimation which included *treatment*, *week*, and *treatment***week* indicators along with demographic, course, and self-control variables. Standard errors are clustered at the individual level.

often than in the first week. Appendix Fig. A.8 shows a similar drop in software engagement for students in the distraction blocking treatment. Encouragingly, the patterns in software utilization match the patterns of effort observed in Figs. 1 and 2. This consistency provides additional evidence that the differences in outcomes between treatment and control are, indeed, driven by the software treatments.

The patterns in how students engage with the study software could also help explain the greater persistence of the commitment treatment effects relative to the distraction blocking treatment effects observed in Figs. 1–3. While students in the commitment device treatment set less restrictive limits that are less likely to bind throughout the course, these students still receive an email each day asking them whether they would like to change their daily limit. Therefore, students in the commitment treatment are interacting with at least one component of the commitment tool consistently throughout the course. In contrast, students in the distraction blocking treatment only interact with the distraction blocking tool when the visit the course website. As attrition from course visits occurs throughout the semester, the frequency of interactions between those in the distraction blocking treatment and the distraction blocking software also decreases.

3.4. Heterogeneous treatment effects

To provide evidence on the behavioral mechanisms and to inform the generalizability of the results, I test whether responses to the treatments vary by student characteristics. In particular, I test whether treatment effects are larger for students with a strong desire to finish coursework on-time and whether treatment effects vary by how well students are predicted to do in the course, given characteristics that are measured prior to the course.[40]

3.4.1. Course goals

While this study focuses on the impact of the treatments on academic outcomes like assignment submission, course performance, and course completion, it is not important to all MOOC students to submit assignments or complete the course. If the treatments are most effective for students intending to finish the course, then the treatments are well-targeted and are likely to improve well-being. If, however, the treatments are most effective for students who never intended to complete

[40] I also examine whether treatment effects are larger for students who are more likely to have self-control problems given their level of agreement with statements such as: “I do things that feel good in the moment but regret later on” and “I’m good at resisting temptation.” Estimates for this analysis are imprecise and uninformative, so are not included in the main body of the paper. However, results of this analysis are reported in Appendix Table A.5.

Table 3
Heterogeneous Treatment effects by course goals.

	Course effort (Hours)	Homework submitted	Course grade (Z-score)	Course completion
Commitment*importance	9.898	1.068	0.457	0.136
	(5.813)	(0.843)	(0.306)	(0.103)
Standard p-values	0.089	0.206	0.135	0.185
Bonferroni adjusted p-values	0.423	0.749	0.581	0.707
Alert*importance	−1.659	0.804	0.311	0.115
	(5.010)	(0.847)	(0.300)	(0.0995)
Standard p-values	0.741	0.343	0.300	0.248
Bonferroni adjusted p-values	1.00	0.920	0.882	0.819
Blocking*importance	0.142	0.0808	0.194	0.0589
	(5.600)	(0.840)	(0.297)	(0.0995)
Standard p-values	0.980	0.923	0.513	0.554
Bonferroni adjusted p-values	1.00	1.00	0.987	0.992
Commitment device	−2.172	0.0793	−0.0442	0.00763
	(3.820)	(0.666)	(0.234)	(0.0788)
Standard p-values	0.570	0.905	0.851	0.923
Bonferroni adjusted p-values	0.994	1.00	1.00	1.00
Alert	−3.472	−0.263	−0.194	−0.0639
	(3.311)	(0.634)	(0.221)	(0.0703)
Standard p-values	0.295	0.679	0.382	0.364
Bonferroni adjusted p-values	0.877	0.999	0.944	0.934
Distraction blocking	0.306	0.505	−0.0369	−0.0294
	(3.721)	(0.637)	(0.220)	(0.0729)
Standard p-values	0.934	0.428	0.867	0.686
Bonferroni adjusted p-values	1.00	0.965	1.000	0.999
Dep var mean	22.38	3.730	0.289	0.711
Demographic variables	y	y	y	y
Course variables	y	y	y	y
Self-control variables	y	y	y	y
Observations	657	657	657	657

Robust standard errors in parentheses. Standard p-values are constructed from T-tests of regression coefficients with robust standard errors. The empirical p-values report the fraction simulated treatment effects that are greater than the measured treatment effects, in absolute values. Bonferroni-adjusted p-values conservatively adjust p-values to account for the simultaneous testing of all three treatments and three additional interaction effects. "Important to complete" variable is an indicator for whether students indicate that it is either very or extremely important to finish assignments and tests on-time for credit. **Demographic variables** include gender, age, age^2, education, income, continent, and indicators for missing age and income variables. **Course variables** include type of computer: *personal laptop, personal desktop, work computer*, previous online courses started, previous online courses finished, and previous statistics courses taken. **Self-control variables** include distracting time goal, desired change in distracting time, self-reported difficulty breaking habits, distractibility, ability to resist temptation, level of self-discipline, and take actions that are regretted in long run.

the course, the welfare implications are ambiguous. In the pre-study survey, students were asked how important it was to complete all the course quizzes and tests on-time. I create an indicator for students responding "very important" or "extremely important" and test whether these students are more likely to respond to the treatments.[41] To test whether response to treatment varies by student goals, I estimate:

$$y_{it} = a + \theta goal_i + \sum_{j=1}^{3} \gamma_j T_{ij} + \sum_{j=1}^{3} \lambda_j T_{ij} * goal_i + \nu \mathbf{X}_i + \epsilon_i, \tag{3}$$

where $goal_i$ is an indicator for whether student i strongly desired to finish the course, $T_{ij}*goal_i$ is the interaction between treatment assignment and desire to finish course, and all other variables are as previously specified. Table 3 reports the result of this analysis. Although the estimates are all imprecise and should be interpreted cautiously, the point estimates suggest that the response to the commitment device may driven by students for whom finishing the course material on-time is either very or extremely important. Point estimates of the interaction between the commitment treatment and goal (λ_1) are large for effort (9.9 h, significant at the 10% level), homework submissions (1.1 assignments), aggregate course performance (0.46 standard deviations), and completion (14 percentage points), while the estimates of the commitment device's impact on the students for which completion is not very important are small or even negative for effort (−2.1 h), homework (0.1 assignments), z-scores (0.04 standard deviations), and completion (1 percentage point). The estimated effects

[41] Students had 5 options to respond to this question including: (1) not important at all; (2) not very important; (3) moderately important; (4) very important; and (5) extremely important.

of the alert and distraction blocking treatments are even less precisely measured, but show similar patterns, with positive, but statistically insignificant, *alert***goal* and *blocking***goal* interaction coefficients.

3.4.2. Predicted outcomes

The above heterogeneity results describe how course objectives interact with treatments. Also of interest is how the treatments impact those expected to do better or worse in the course, given their observable pre-study characteristics. To test whether expected course performance impacted the magnitude of the treatment response, I implement a split-sample endogenous stratification estimator as is outlined by Abadie et al. (2013). This estimation strategy uses students in the control group to generate predicted outcomes for students in all treatment groups (including the control) and then estimates the treatment effects within quantiles of predicted outcomes. To overcome the bias introduced by overfitting issues that arise when a student's characteristics are used to predict their own outcomes,[42] this estimation strategy takes the following steps: (1) randomly select half the control group and use this group to estimate predicted outcomes with observable pre-study characteristics for the remainder of the students; (2) bin students into predicted outcome quantiles (excluding the students used to estimate predicted outcomes); (3) estimate treatment effects within quantile bins and store estimates; (4) iterate steps 1–3 multiple times; and (5) bootstrap standard errors.[43]

I use the above strategy to estimate the impact of the treatments on effort, homework submissions, and points scored and present the results of this estimation in Table 4. These results suggest that the impact of the commitment device has a strong positive correlation with predicted outcomes. For each outcome—course hours, homework, and grades—the estimated impact of the commitment device increases with the quintile of predicted outcome. The effects of the alert and distraction blocking treatments are also positively correlated with predicted homework and grade outcomes, but these correlations are smaller and less consistent. These results suggest that the commitment device, and, to a lesser extent, the alert and distraction blocking treatments were most helpful to students who were likely to succeed in the MOOC in the first place.

3.5. Robustness checks

While the primary estimates of the effects of the commitment device on course outcomes are large and statistically significant, they are imprecisely estimated. To test the robustness of these results, I take two approaches. First, as a randomization-based approach, I construct two-sided empirical p-values for my primary results from 10,000 simulated treatment assignments.[44] The results of this approach are reported for the primary treatment effects on course outcomes and for heterogeneity effects by course goals in the "Empirical p-value" rows of Tables 2 and 3, respectively. The empirical p-values in Table 2 show that estimates of the impacts of the commitment device on course effort (empirical p=0.046), homework submitted (empirical p=0.024), course grade (empirical p=0.046), and course completion (empirical p=0.031) are all robust to this randomization-based approach. The Empirical p-values in Table 3 also indicate that the empirical p-value estimates of the interaction between the commitment device and course goals are robust to this randomization-based estimation.

Second, I apply Bonferroni corrections to adjust for multiple hypothesis testing. In Table 2 the Bonferroni correction adjusts p-values to account for three treatments being tested simultaneously in each column and in Table 3 the Bonferroni correction adjusts p-values to account for three treatments and three interaction effects being tested in each column. The results of these adjustments reported in Table 2 indicate that estimated effects commitment device on course effort (p=0.211), homework submitted (p=0.070), course grade (p=0.143), and course completion (p=0.090) are either marginally significant or insignificant at the 10% level. Furthermore, applying the Bonferroni corrections in as reported in Table 3 makes all estimated treatment and interaction effects statistically insignificant. Although only two out of the four of the estimated impacts of the commitment device on remain statistically significant after the Bonferroni corrections, together the empirical p-value and Bonferroni correction approaches generate evidence that is consistent with the commitment device positively impacting course outcomes. However both the Bonferroni and empirical p-value approaches suggest that the heterogeneity the results should be interpreted with caution.

While both the empirical p-value and Bonferroni correction approaches suggest the commitment device is effective in improving student outcomes, neither approach addresses how likely the patterns found among students assigned to the commitment device are to be observed across outcomes. To examine the likelihood of observing the patterns found among

[42] In finite samples, predicted values for observations with large positive or negative error terms tend to be overfitted. Because of overfitting, students in the control group who have poor outcomes driven by unobservable characteristics are more likely to have poor predicted outcomes than students in the treatment group who also have poor outcomes due to unobservable characteristics. Symmetrically, students in the control group with strong outcomes due to unobservable characteristics are more likely than similar students in the treatment group to have strong predicted outcomes. As a result, estimates that include control students' own characteristics are biased towards finding positive treatment effects for weak students and negative treatment effects for strong students.

[43] Estimates reported in this paper are generated by 200 sample splits and 500 bootstrap repetitions. See Abadie et al. (2013) for more details.

[44] This approach is similar to that of Chetty et al. (2009). Specifically, I randomly re-assign study participants to counterfactual treatments and use my primary specification to estimate treatment effects on the counterfactual data. I repeat this simulation 10,000 times. I then report the fraction of simulated estimates that have more extreme values than my estimates. It is common for authors to provide one sided p-values in such an exercise, or the fraction of simulated estimates greater than the actual estimate for positive estimates and the fraction of simulated estimates less than the actual estimate for negative values. I opt for the more conservative approach of reporting the fraction of simulated estimates that have an absolute value greater than the absolute value of my estimated treatment effects.

Table 4
Heterogeneous treatment effects, predicted outcomes.

	Quintile of predicted outcome				
	(1)	(2)	(3)	(4)	(5)
Course effort (Hours)					
Commitment device	−2.031	0.272	4.958	6.512	9.362
	(3.835)	(3.159)	(3.627)	(4.633)	(6.889)
Alert	−4.223	−5.178	−6.265	−8.054	−2.609
	(3.625)	(2.482)	(2.598)	(2.994)	(3.797)
Distraction blocking	−0.618	−1.701	−1.736	−2.131	−0.003
	(3.809)	(2.956)	(3.464)	(4.262)	(6.727)
Homework Submitted					
Commitment device	−0.161	0.034	0.723	1.126	1.841
	(0.643)	(0.476)	(0.571)	(0.630)	(0.640)
Alert	−0.005	−0.037	0.060	0.308	0.784
	(0.625)	(0.452)	(0.559)	(0.600)	(0.649)
Distraction blocking	0.089	−0.075	0.322	0.848	1.345
	(0.590)	(0.457)	(0.576)	(0.621)	(0.678)
Course Grade (Z-score)					
Commitment device	−0.140	0.081	0.259	0.354	0.570
	(0.220)	(0.163)	(0.194)	(0.254)	(0.269)
Alert	−0.068	−0.132	−0.021	−0.031	0.102
	(0.215)	(0.153)	(0.185)	(0.233)	(0.256)
Distraction blocking	−0.135	−0.104	0.014	0.178	0.303
	(0.208)	(0.154)	(0.192)	(0.249)	(0.272)

Bootstrapped standard errors in parentheses. Split-sample endogenous stratification estimates reported (Abadie et al., 2013). Estimates are generated with 200 sample splits and 500 Bootstrapped repetitions. Variables used to construct predicted values include demographic, course, and self-control variables. **Demographic variables** include gender, age, age^2, education, income, continent, and indicators for missing age and income variables. **Course variables** include course goals: *finish for certificate, finish at own pace, complete some assignments, or watch some videos*, reasons taking the course: *general interest, relevant to school/work/research, career change, fun, try online course, improve English*, type of computer: *personal laptop, personal desktop, work computer*, previous online courses started, previous online courses finished, previous statistics courses taken, interest level in course, expected course hours, and importance of finishing course. **Self-control variables** include distracting time goal, desired change in distracting time, self-reported difficulty breaking habits, distractibility, ability to resist temptation, level of self-discipline, and take actions that are regretted in long run.

students assigned to the commitment device across outcomes, I employ a randomization-based method that is similar to the empirical *p*-value approach. Specifically, I randomly re-assign study participants to counterfactual treatments, use my primary specification to estimate treatment effects on the counterfactual data, and then repeat this simulation 10,000 times. I then examine how frequently each of the four coefficients from any of the three treatments have two-sided empirical *p*-values equal to or less than 0.076, which is the largest *p*-value from any of the commitment device coefficients in Table 2. With this approach, I find that only 4.6% of the simulations generate at least one treatment with coefficients with empirical *p*-values equal to or less than 0.076 for all outcomes.[45] Given the improbability of randomly observing the patterns found in the commitment device coefficients across outcomes, this randomization-based approach provides additional evidence that the commitment device improves student outcomes.

3.6. Post-study survey

Following the course, students were incentivized with a $7 Amazon.com gift card to complete a post-study survey. This survey asked students a number of questions about how the software impacted their computer use and experience in the course. The results of these survey questions are reported in Table 5. Only 52% of study participants completed the post-study survey. Also, the first row in Table 5 indicates that survey response was not constant across treatment groups—a greater portion of students in the commitment device treatment responded to the survey than those in other treatments.

[45] I also employ a less conservative approach, where I identify the how often the data generated any treatments that had coefficients with a set of empirical *p*-values that are as extreme as the *p*-values I observe in my primary results. Specifically, I identify the number of occurrences where at least one treatment coefficient has an empirical *p*-value of 0.024 or less, two coefficients have an empirical *p*-value of 0.031 or less, three coefficients have an empirical *p*-value of 0.050 or less, and all four coefficients have an empirical *p*-value of 0.076 or less. These are the p-values from the commitment device coefficients on homework, course completion, course performance and course effort, respectively. I find that 3.7% of simulations have at least one treatment that has coefficients with empirical *p*-values that are at least as extreme as the *p*-values I observe in my data.

Table 5
Post study survey responses.

	Control	Commitment	Alert	Blocking
Completed post survey	0.465	0.650	0.482	0.503
Software increased thinking about course	0.413	0.505	0.487	0.438
Software reduced distracting time	0.413	0.485	0.526	0.354
Software increased time spent on course	0.320	0.402	0.308	0.237
Software made unproductive time less enjoyable	0.280	0.510	0.449	0.291
Software improved understanding of time use	0.627	0.656	0.756	0.544
Used software to set course goals	0.189	0.271	0.205	0.175
Used software to set general goals	0.347	0.375	0.397	0.215
Software was useful	0.467	0.542	0.564	0.423
Observations	170	160	166	161

Each question was originally asked on a scale-1-Strongly Disagree, 7-Strongly Agree. I collapse question to an indicator variable for whether they agree (somewhat agree, agree, strongly agree) for interpretability.

Therefore, the results should be interpreted cautiously. Nevertheless, student responses do shed additional light on the potential mechanisms driving response to treatments. The most significant difference observed between treatment and control is that students in the commitment and alert treatments are much more likely than those in the control group to report that the treatment made unproductive time less enjoyable. Students in the commitment and alert treatments were 81% (23 percentage points, significant at the 1% level) and 61% (17 percentage points, significant at the 5% level) more likely than students in the control treatment to state that the software made unproductive time less enjoyable, respectively. That students in the commitment device treatment found unproductive time less enjoyable suggests that the commitment device worked in the way it was intended—making spending time on unproductive websites more difficult or costly. Other differences were not statistically significant, but students in the commitment device treatment were most likely to report that the software increased the time they spent on the course and to report using the software to set course goals. The results of the post-study survey are consistent with students using the commitment device to address present-biased preferences—making distracting time more costly in order to increase the amount of time spent on coursework.[46]

3.7. Generalizability of results

Although the estimated impacts of the commitment device on course outcomes are large and significant and the effects of the alert and distraction blocking tools are both insignificant, the applicability of these findings to other MOOCs and the broader online education setting depends on how the study sample and course environment relate to other populations and contexts. Regarding the applicability to other MOOCs, the highly educated and internationally diverse sample of students look quite similar to other student populations observed in other MOOC settings.[47] However, within this study it is possible that the types of students who selected into the study are differently affected by the commitment device than those who did not. Among students enrolled in the course, Appendix Table A.2 shows that students who selected to participate in the study were more likely to be female and hold a Ph.D. than those who did not select into the treatment, but were otherwise similar in observable characteristics. More striking is the difference in course outcomes between study participants and non-participants, with study participants in the control and treatment groups spending more time on coursework, submitting more assignments, earning better course scores, and completing the course at higher rates than non-participants. While these differences in course outcomes may be due a behavioral response to being monitored in the study or having the treatment software installed, it is also possible that the students who were most likely to succeed were the ones that selected into the study. Table 4 indicates that the commitment treatment is most effective for students with better predicted outcomes. If selection into the study is positively correlated with expected success, then the results from Table 4 suggest that the efficacy of the commitment device would likely be dampened if provided to all students enrolled in MOOCs.

In addition to applying the results of this study to other MOOCs, generalizing to online education settings more broadly should also be done carefully. One reason to be cautious in generalizing the results is that the highly educated study sample is not representative of online degree-seeking college students in the United States. Additionally, without tuition, formal grades, or the potential credit for a formal degree, students enrolled in this MOOC face much lower stakes than students enrolled in online degree programs. Finally, the schedules of students enrolled in this MOOC are likely to be different than many online students. Whereas this MOOC is a stand-alone course that combines synchronous and asynchronous elements,

[46] One puzzle is that students in the alert treatment were also more likely than students in the control group to report that the software made distracting sites less enjoyable, yet experienced no measurable improvements in course outcomes. One possibility is that the choice is an important aspect of the software's impact. Students in the commitment were able to choose restrictiveness of their commitments whereas students in the alert treatment had no choice for how the software would impact them. The aspect of choice in the commitment may have made it possible for students to better calibrate their interaction with the software, or lead students to have a more positive response to software disruptions.

[47] Banerjee and Duflo (2014); Breslow et al. (2013); Waldrop (2013) show that approximately 75% of MOOC students have bachelor's degrees or higher, 70% of students are from outside the United States, and students have an average age close to 30.

part-time and full-time online degree seeking students are likely to be enrolled in multiple courses at a time over the course of multiple semesters and take courses that are either more synchronous or more asynchronous than this MOOC.

Given the significant differences in between MOOCs and online degree programs, it is possible that the commitment, alert, and distraction blocking treatments have different impacts in online degree programs than in this study. For example, one reason that the alert treatment may have been ineffective in this study is that the MOOC students may have rarely planned on doing coursework when they received the alerts. Online degree seeking students who are taking multiple classes or even full-time course loads may be more likely to benefit from the alert treatment, as they are more likely to be planning on doing coursework at the times when they receive alerts. Additionally, it is possible that students in this study did not significantly benefit from the distraction blocking tool because the distraction blocking tool only operates when students open the course website and most MOOC students stop visiting the course website early in the term. Students who have paid tuition as part of an online degree program may be significantly less likely than MOOC students to stop visiting their course websites and therefore more likely benefit from the distraction blocking tool. However, it is also possible that results from this study apply to other online contexts and that the alert and distraction blocking tools have no significant impact on students in online degree programs either.

While it is difficult to predict how students in online degree programs are likely to respond to the alert and distraction blocking tools, the heterogeneity and timing analyses do provide some insight into how the impacts of the commitment tool are likely to apply in different contexts. In my heterogeneity analysis reported in Table 4, I find that when demographic, course-related, and self-control variables are considered, students with the strongest predicted outcomes are most responsive to the commitment treatment. This result generates ambiguous predictions for how students in degree seeking programs are likely to be affected by the commitment device. The commitment device may be less effective in online degree programs if the high education levels of MOOC students is an indication that MOOC students are stronger students than those in degree-seeking programs. However, the commitment device might be more effective in online degree programs if the higher stakes of degree programs act to exclude students with low probabilities of success. Related to my findings on predicted success reported in Table 3, I find that point estimates of the commitment device's impact are largest for students who indicate that finishing each assignment and test on-time is either very or extremely important. This result suggests that the treatments may have the most impact in settings where online students have a strong desire to finish their coursework.

In my timing analysis reported in Figs. 1, 2, and 4, I find that the use of study software and the positive treatment effects of the commitment device dissipate over time. One reason that software use and treatment effects may dissipate over time is related to the post-survey finding that students in the commitment device treatment found that the software made unproductive time significantly less enjoyable. It is possible that students stopped using and benefiting from the commitment tools over time because of the unpleasantness of the software experience. If the software's unpleasantness drove the dissipation in commitment software use and efficacy over a single nine-week course, it is likely that the effects of the commitment device would dissipate even further for students who take multiple courses at a time, have semesters that last longer than 9 weeks, and require multiple semesters to complete their programs.

Another explanation for the decline in commitment software use and efficacy is related to the contrasting patterns in Figs. 2 and 3. While Fig. 2 suggests that the commitment had diminishing effects on homework submissions over the course of the semester, Fig. 3 suggests that the effects of the commitment treatment on turning in Week 1 and Week 9 assignments are roughly the same. These patterns suggest that students in the commitment treatment were influenced to not only complete homeworks in early weeks, but to work ahead on future assignments. If the dissipation in software use and efficacy was due to the fact that the software led the students to complete work ahead of time (and thus they no longer needed the software in later weeks), then the treatment effects may be more persistent in online settings where students cannot work ahead or may be even greater in fully asynchronous programs where students can move onto the next course as soon as their current coursework is completed.

4. Conclusion

Low completion rates and poor student performance are among the most serious problems facing online education. This study tests whether computerized tools intended to address behavioral issues of present-biased preferences, limited memory, and limited willpower increased course completion and improved student performance in a massive open online course. The primary finding in this paper is that the commitment device, which allows students to pre-commit to the amount of distracting time they spent each day, significantly improves course outcomes, including time spent on coursework, homework submissions, overall scores, and completion rates. The most striking of these results is that the commitment device increases course completion by 40% (11 percentage points). In contrast, I find that the alert treatment, which provides students with a reminder after each half hour spent on distracting sites, has no impact on course outcomes. I also find that the distraction blocking treatment, which allows students to block distracting websites when they go to work on the course website, has generally positive estimated impacts on course performance, but these estimates are much smaller than those found for the commitment device, imprecisely estimated, and cannot be statistically distinguished from zero.

An important caveat to my findings is that the study context differs in important ways from other online education environments. Compared to online degree programs MOOCs have a much more educated student population and have much lower stakes for completion. Furthermore, even within the MOOC environment the students who are willing to participate in a time-management software study are likely to differ in important ways from those who are not. Although it is difficult to predict how the results of this study will generalize to all online students in the United States, online students in all types of programs are likely to be subject to the same types of computer distractions as the students in this study. If online students generally struggle with issues of self-control related to distracting websites, there is potential for software tools like the commitment device to have a significant positive impact on academic outcomes.

Appendix A

Table A.1
Summary statistics.

	Control	Commitment	Alert	Distraction blocking	F-stat *P*-value
Panel A- Demographic characteristics					
Age	32.094	30.169	30.367	30.161	0.295
Female	0.465	0.438	0.373	0.416	0.381
High school	0.065	0.094	0.078	0.075	0.811
Bachelors degree	0.182	0.219	0.229	0.205	0.736
Masters degree	0.394	0.338	0.271	0.317	0.115
PhD/MD	0.318	0.287	0.337	0.317	0.810
United States	0.282	0.275	0.235	0.311	0.483
Africa	0.165	0.131	0.139	0.075	0.049
Asia	0.176	0.188	0.217	0.205	0.797
Australia	0.035	0.013	0.036	0.037	0.285
Europe	0.235	0.244	0.247	0.230	0.983
North America	0.047	0.075	0.066	0.050	0.677
South America	0.059	0.075	0.060	0.093	0.631
Income: \$ 0-\$ 19,999	0.259	0.294	0.277	0.236	0.344
Income: \$ 20,000-\$ 59,999	0.253	0.225	0.229	0.335	0.145
Income: \$ 60,000-\$ 99,999	0.182	0.100	0.108	0.087	0.080
Income: \$ 100,000+	0.112	0.131	0.114	0.130	0.870
Panel B- Course experience characteristics					
Goal: complete course ontime for certificate	0.659	0.637	0.669	0.714	0.502
Goal: complete course at own pace	0.241	0.225	0.235	0.143	0.060
Goal: other	0.094	0.131	0.084	0.130	0.390
Importance of finishing material†	3.724	3.694	3.578	3.696	0.433
Expected hours on course	54.953	52.875	49.663	55.761	0.404
Reason: relevant to job, school, or research	0.959	0.919	0.904	0.913	0.007
Reason: general interest	0.941	0.938	0.952	0.925	0.556
Reason: personal growth	0.906	0.938	0.940	0.925	0.994
Reason: career change	0.288	0.306	0.349	0.335	0.578
Reason: for fun	0.612	0.613	0.608	0.615	0.968
Reason: try online course	0.447	0.450	0.530	0.460	0.235
Previous statistics courses taken	1.353	1.387	1.157	1.304	0.468
Previous online courses started	4.300	6.606	4.886	4.981	0.443
Previous online courses finished	1.859	2.731	1.867	2.093	0.649
Software installed: personal laptop	0.712	0.756	0.663	0.708	0.320
Software installed: personal desktop	0.165	0.113	0.187	0.118	0.164
Software installed: work computer	0.112	0.113	0.133	0.168	0.446
Panel C- Self-control characteristics					
Goal: distracting time	1.132	1.269	1.173	1.273	0.561
Goal: change in distracting time	−1.129	−1.015	−1.208	−1.065	0.646
Hard to break habits‡	2.871	2.950	3.096	3.087	0.070
Easily distracted‡	3.029	3.006	3.187	3.211	0.216
Able to resist tempation‡	3.147	3.081	2.976	2.957	0.332
Strong self-discipline‡	3.347	3.325	3.247	3.422	0.706
Pleasure/fun gets in way of productivity‡	2.753	2.825	2.940	2.975	0.087
Do things that regret later‡	2.547	2.556	2.699	2.640	0.397
Panel D- Outcomes					
Course effort (hours)	23.05	26.76	17.75	22.06	0.008
Homework submitted	3.433	4.194	3.642	3.957	0.268
Points (out of 100)	29.69	35.66	29.13	31.29	0.3894
Completed Course	0.269	0.356	0.267	0.267	0.234
Observations	170	160	166	161	Total=657

Notes: † 1-Not at all important, 5-Extremely important. ‡ 1-Not like me at all, 5-Very much like me.

Table A.2
Summary statistics: sample comparison.

	In sample	Out of sample	F-Stat *P*-value
Age	30.782	30.681	0.857
	(12.070)	(14.490)	
Female	0.422	0.378	0.035
	(0.494)	(0.485)	
High school	0.077	0.071	0.625
	(0.267)	(0.257)	
Bachelor's degree	0.208	0.240	0.071
	(0.406)	(0.427)	
Master's degree	0.332	0.365	0.110
	(0.471)	(0.482)	
Ph.D./M.D.	0.317	0.186	0.000
	(0.466)	(0.389)	
Effort (hours)	22.378	15.755	0.000
	(25.241)	(23.326)	
Homework assignments submitted	3.799	2.674	0.000
	(3.730)	(3.528)	
Aggregate course score (out of 100)	31.394	22.694	0.000
	(36.693)	(34.431)	
Course completion	0.289	0.213	0.000
	(0.454)	(0.409)	
Observations	657	2903	

Notes: Standard deviations (SD) in parentheses. Out of sample students include all students pre-enrolling in MOOC. Statistics come from pre-course survey administered by Stanford.

Table A.3
Impact of treatments on course outcomes.

	Primary sample				Primary Sample + Late enrollees			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Course Effort (Hours)								
Commitment device	3.712	4.619	5.486	5.491	3.282	3.560	3.721	3.940
	(3.137)	(3.159)	(3.067)	(3.085)	(2.696)	(2.720)	(2.600)	(2.618)
Alert	−5.298	−3.987	−2.745	−3.339	−5.596	−4.653	−3.710	−4.207
	(2.614)	(2.605)	(2.598)	(2.597)	(2.232)	(2.251)	(2.207)	(2.202)
Distraction blocking	−0.986	0.161	0.344	0.848	−2.862	−1.781	−1.898	−1.348
	(2.778)	(2.760)	(2.723)	(2.826)	(2.375)	(2.311)	(2.261)	(2.303)
Adjusted r-squared	0.012	0.054	0.096	0.111	0.014	0.044	0.102	0.117
Homework Submitted								
Commitment device	0.761	0.865	0.928	0.909	0.479	0.507	0.517	0.529
	(0.416)	(0.415)	(0.401)	(0.403)	(0.361)	(0.361)	(0.347)	(0.345)
Alert	0.210	0.224	0.321	0.267	−0.273	−0.279	−0.155	−0.204
	(0.408)	(0.414)	(0.414)	(0.415)	(0.350)	(0.357)	(0.345)	(0.345)
Distraction blocking	0.524	0.583	0.554	0.577	0.0433	0.130	0.0850	0.128
	(0.405)	(0.412)	(0.406)	(0.412)	(0.347)	(0.352)	(0.342)	(0.343)
Adjusted r-squared	0.002	0.031	0.083	0.090	0.002	0.027	0.099	0.113
Course Grade (Z-score)								
Commitment device	0.219	0.273	0.297	0.291	0.112	0.128	0.129	0.132
	(0.152)	(0.151)	(0.147)	(0.148)	(0.133)	(0.133)	(0.128)	(0.128)
Alert	−0.0206	−0.00818	0.0254	0.0109	−0.171	−0.176	−0.131	−0.143
	(0.146)	(0.147)	(0.148)	(0.150)	(0.125)	(0.127)	(0.124)	(0.125)
Distraction blocking	0.0589	0.0847	0.0905	0.0966	−0.0826	−0.0551	−0.0592	−0.0500
	(0.147)	(0.149)	(0.147)	(0.149)	(0.127)	(0.128)	(0.126)	(0.127)
Adjusted r-squared	0.000	0.037	0.080	0.084	0.002	0.034	0.088	0.095

(*continued on next page*)

Table A.3 (*continued*)

	Primary sample				Primary Sample + Late enrollees			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Course Completion								
Commitment device	0.0872	0.104	0.111	0.107	0.0347	0.0430	0.0457	0.0489
	(0.0510)	(0.0509)	(0.0496)	(0.0497)	(0.0443)	(0.0448)	(0.0438)	(0.0445)
Alert	−0.00234	0.00423	0.0150	0.0108	−0.0506	−0.0502	−0.0390	−0.0332
	(0.0485)	(0.0491)	(0.0498)	(0.0503)	(0.0416)	(0.0423)	(0.0428)	(0.0440)
Distraction blocking	−0.00193	0.00793	0.0131	0.0135	−0.0461	−0.0344	−0.0262	−0.0197
	(0.0488)	(0.0494)	(0.0492)	(0.0498)	(0.0427)	(0.0416)	(0.0429)	(0.0426)
Adjusted r-squared	0.002	0.028	0.063	0.066	0.003	0.024	0.062	0.066
Demographic variables	n	y	y	y	n	y	y	y
Course variables	n	n	y	y	n	n	y	y
Self-control variables	n	n	n	y	n	n	n	y
Observations	657	657	657	657	897	897	897	897

Robust standard errors in parentheses. Primary comprises all participants who enroll in course and install software in first week prior to the first homework deadline. Baseline (control) mean homework assignments submitted is 3.8 for the primary sample and 3.7 for all participants. Baseline average hours of effort is 22.4 for the primary sample and and 21.4 for all participants. **Demographic variables** include gender, age, age^2, education, income, continent, and indicators for missing age and income variables. **Course variables** include course goals: *finish for certificate, finish at own pace, complete some assignments, or watch some videos*, reasons taking the course: *general interest, relevant to school/work/research, carrier change, fun, try online course, improve English*, type of computer: *personal laptop, personal desktop, work computer*, previous online courses started, previous online courses finished, previous statistics courses taken, interest level in course, expected course hours, and importance of finishing course. **Self-control variables** include distracting time goal, desired change in distracting time, self-reported difficulty breaking habits, distractibility, ability to resist temptation, level of self-discipline, and take actions that are regretted in long run.

Table A.4
Impact of treatments on course outcomes, robustness exercise.

	Course effort (Hours)	Homework submitted	Course grade (Z-score)	Course completion
Commitment device	5.491	0.909	0.291	0.107
Standard *p*-values	0.076	0.024	0.050	0.031
Empirical *p*-values	0.046	0.026	0.046	0.031
Romano-Wolf 12-way *p*-values	0.319	0.152	0.249	0.181
Alert	−3.339	0.267	0.0109	0.0108
Standard *p*-values	0.199	0.520	0.942	0.839
Empirical *p*-values	0.221	0.500	0.939	0.826
Romano-Wolf 12-way *p*-values	0.559	0.925	0.988	0.998
Distraction blocking	0.848	0.577	0.0966	0.0135
Standard *p*-values	0.764	0.161	0.518	0.787
Empirical *p*-values	0.760	0.156	0.514	0.783
Romano-Wolf 12-way *p*-values	0.988	0.501	0.925	0.988
Dep var mean	22.38	3.730	0.711	0.289
Demographics	y	y	y	y
Course variables	y	y	y	y
Self-control variables	y	y	y	y
Observations	657	657	657	657

Standard *p*-values are constructed from *t*-tests of OLS regression coefficients with robust standard errors. Empirical *p*-values are constructed from counterfactual estimates of treatment effects from 10,000 simulations. The empirical *p*-values report the fraction simulated treatment effects that are greater than the measured treatment effects, in absolute values. **Demographic variables** include gender, age, age^2, education, income, continent, and indicators for missing age and income variables. **Course variables** include course goals: *finish for certificate, finish at own pace, complete some assignments, or watch some videos*, reasons taking the course: *general interest, relevant to school/work/research, career change, fun, try online course, improve English*, type of computer: *personal laptop, personal desktop, work computer*, previous online courses started, previous online courses finished, previous statistics courses taken, interest level in course, expected course hours, and importance of finishing course. **Self-control variables** include distracting time goal, desired change in distracting time, self-reported difficulty breaking habits, distractibility, ability to resist temptation, level of self-discipline, and take actions that are regretted in long run.

Table A.5
Heterogeneous treatment effects, present bias.

	Course effort (Hours)	Homework submitted	Course grade (Z-score)	Course completion
Commitment*present-biased	−2.249	−0.356	−0.0673	−0.220
	(6.318)	(0.833)	(0.103)	(0.303)
Standard p-values	0.722	0.669	0.514	0.468
Bonferroni adjusted p-values	0.978	0.964	0.885	0.849
Alert*present-bias	2.994	0.869	0.0754	0.189
	(5.253)	(0.833)	(0.0998)	(0.298)
Standard p-values	0.571	0.318	0.450	0.526
Bonferroni adjusted p-values	0.964	0.683	0.834	0.893
Blocking*present-bias	−0.193	−0.290	−0.0203	−0.130
	(5.612)	(0.824)	(0.0990)	(0.296)
Standard p-values	0.973	0.725	0.838	0.661
Bonferroni adjusted p-values	1.000	0.979	0.996	0.961
Commitment device	6.642	1.106	0.144	0.401
	(4.222)	(0.563)	(0.0700)	(0.206)
Standard p-values	0.116	0.050	0.040	0.052
Bonferroni adjusted p-values	0.309	0.143	0.115	0.148
Alert	−4.116	−0.0890	−0.0211	−0.0658
	(3.605)	(0.546)	(0.0662)	(0.195)
Standard p-values	0.254	0.870	0.750	0.736
Bonferroni adjusted p-values	0.585	0.998	0.984	0.982
Distraction Blocking	0.674	0.742	0.0251	0.163
	(3.842)	(0.592)	(0.0713)	(0.212)
Standard p-values	0.861	0.211	0.725	0.442
Bonferroni adjusted p-values	0.997	0.509	0.979	0.826
Present-bias	−1.701	−0.195	−0.00345	0.0111
	(4.323)	(0.599)	(0.0709)	(0.216)
Standard p-values	0.694	0.745	0.961	0.959
Bonferroni adjusted p-values	0.971	0.983	1.000	1.000
Dep var mean	22.38	3.730	0.711	0.289
Demographic variables	y	y	y	y
Course variables	y	y	y	y
Observations	657	657	657	657
R-squared	0.177	0.158	0.136	0.152

Robust standard errors in parentheses. In the pre-study survey students were asked: whether students do things in the moment that they regret later on; whether they are unable to stop themselves from doing something when they know it is wrong; whether they are good at resisting temptation; and whether they refuse things that are bad for them, even when they are fun. I combine student answers to these questions into a single index and then split the sample equally to create the *Present − bias* variable, which is an indicator for students who are most likely to exhibit present-biased preferences. **Demographic variables** include gender, age, age^2, education, income, continent, and indicators for missing age and income variables. **Course variables** include course goals: *finish for certificate, finish at own pace, complete some assignments, or watch some videos*, reasons taking the course: *general interest, relevant to school/work/research, career change, fun, try online course, improve English*, type of computer: *personal laptop, personal desktop, work computer*, previous online courses started, previous online courses finished, previous statistics courses taken, interest level in course, expected course hours, and importance of finishing course.

Invitation to participate in a Time-Management Study:

Research suggests that time management is a significant factor in how well students perform in online courses*. We invite you to participate in a study that will test whether computerized time-management tools can help you use your time more effectively and complete courses more quickly. By completing 2 10-minute surveys (One before the course begins and one after the course ends) and installing time-management software, you will receive **$12 in Amazon Gift Cards** and **FREE ACCESS** to time-management tools that have been designed to improve computer productivity.

To get more information and to sign up, please CLICK HERE.

*See Holder, B. (2007). An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs. *The Internet and higher education*, *10*(4), 245–260.

Fig. A.1. Recruitment Email.

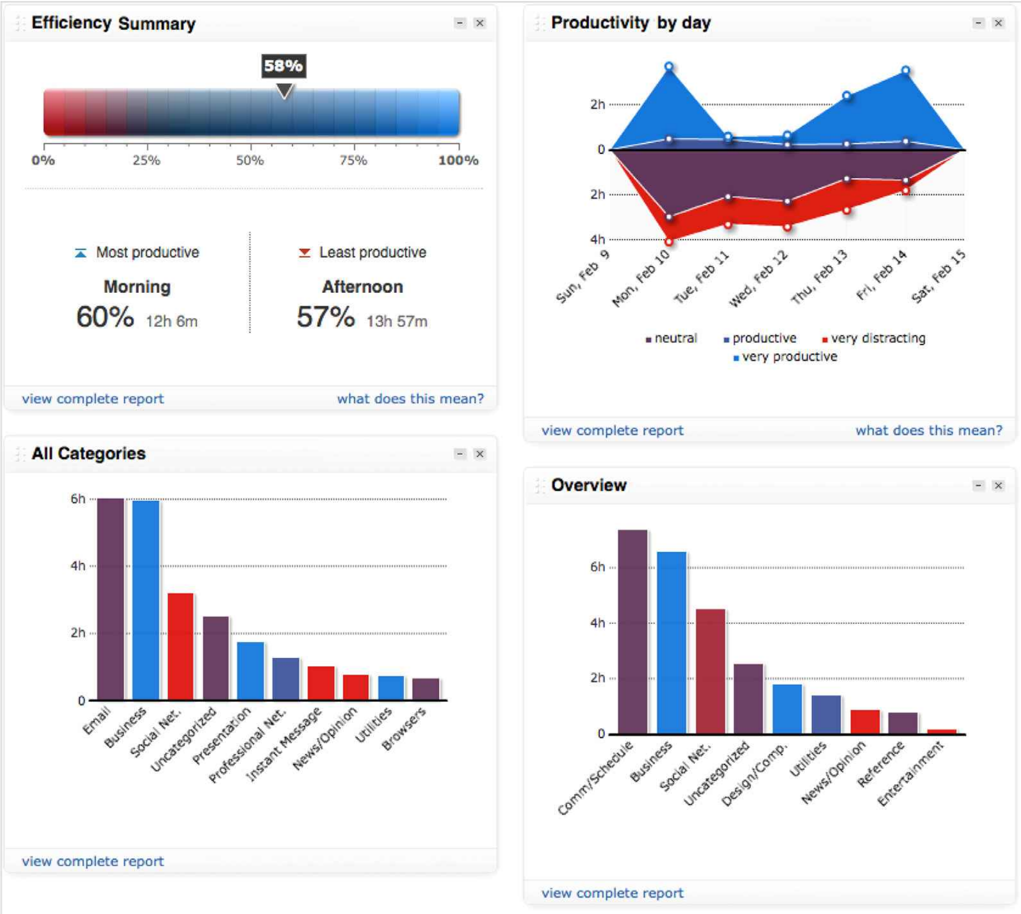

Fig. A.2. Time-Use Summary Report.

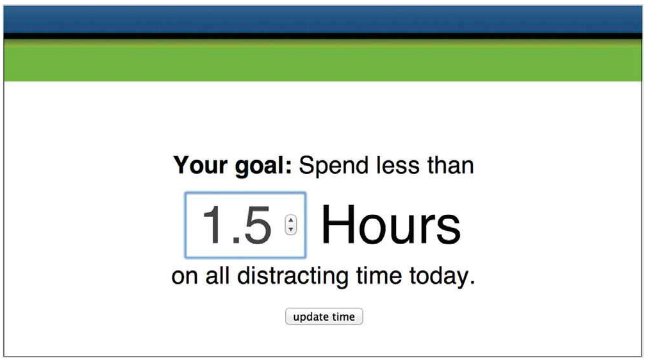

Fig. A.3. Commitment Screen.

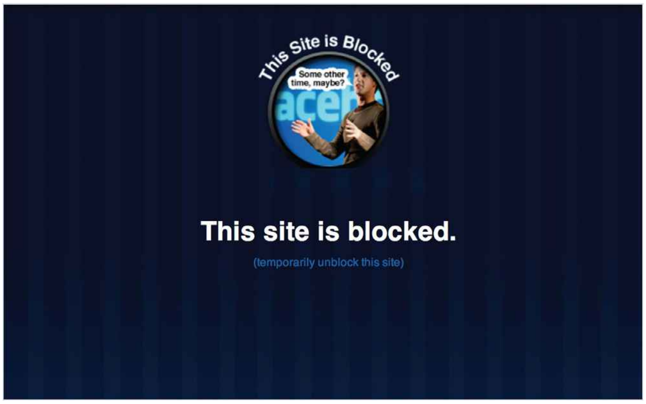

Fig. A.4. Blocked Site.

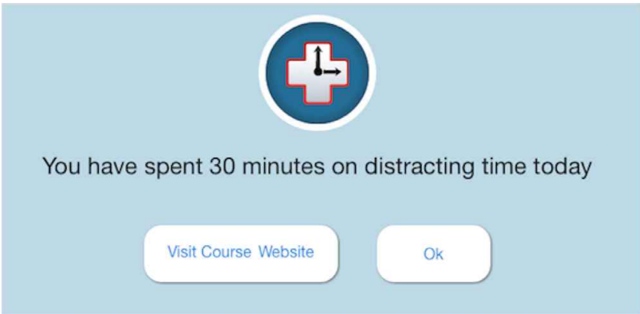

Fig. A.5. Alert.

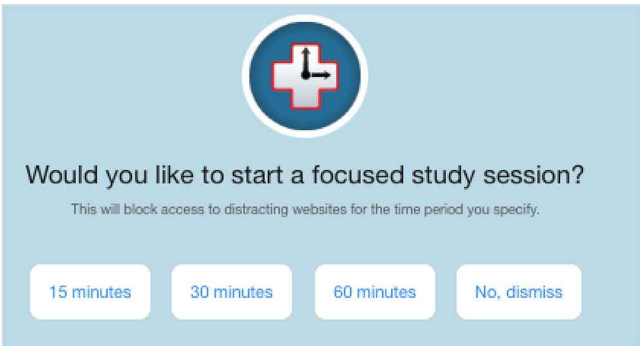

Fig. A.6. Distraction Blocking Screen.

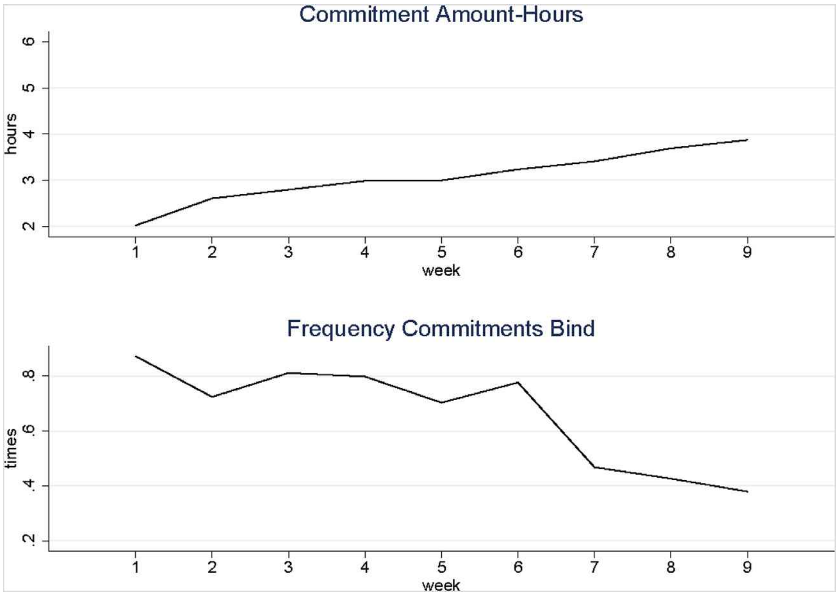

Fig. A.7. Commitment Device Patterns.

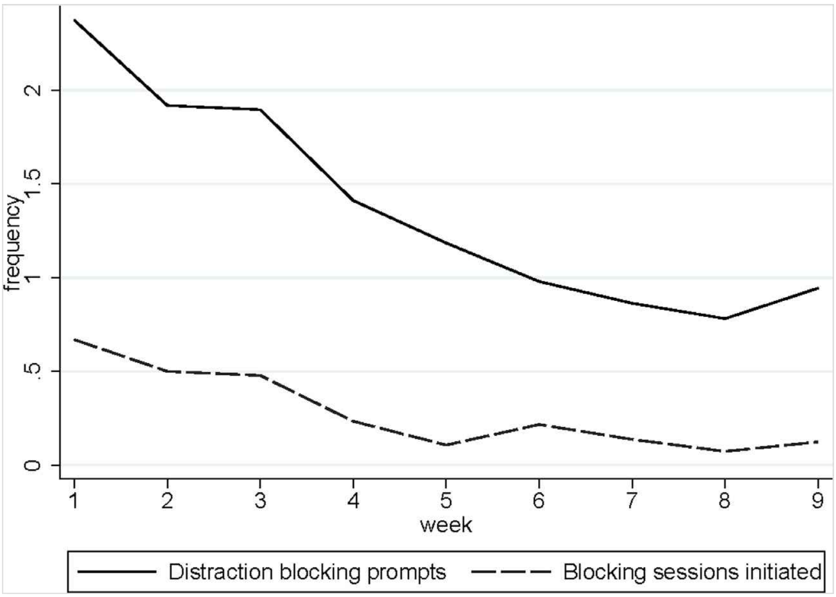

Fig. A.8. Distraction Blocking Study Patterns.

Appendix B. Model of course completion

To generate predictions for how students respond to the commitment, alert, and distraction blocking treatments, I develop a simple three-period model of online course completion that allows for a student to be impatient, forget about the course, and to be distracted away from working on the course.[48] The predictions of this model generalize to online course environments where students are enrolled in a course with multiple sections and must complete one course task per section. The primary predictions of this model are that the alert treatment typically increases the probability of course completion for students who exhibit limited memory, the distraction blcoking treatment typically increases the probability of course completion for students who experience willpower depletion related to Internet distractions, and the commitment device increases course completion for students who have present-biased preferences, exhibit limited memory, or experience willpower depletion, but there are certain conditions under which the introduction of behavioral tools leads to unexpected outcomes.

In this three-period model, a student is enrolled in a two period course that requires work in one period to complete. In periods 1 and 2, the student chooses whether or not to work on the course, with the choice in each period indicated by $x_t \in \{0, 1\}$. For reasons that will become clear, a student may choose to work on the course but not follow through to complete the coursework. Whether a student completes a task in period t is indexed by $y_t \in \{0, 1\}$. If the student completes the coursework ($y_1 = 1$ or $y_2 = 1$), she receives a benefit b in period 3. Working on the course has an immediate cost c_t that is allowed vary.[49] I assume that c_t is drawn from continuous distribution $f(c)$. I also assume that the choice not to work ($x_t = 0$) yields a constant flow utility ($u(x_t)$) equal to 0. In this model, students may procrastinate coursework (exhibit present-bias preferences), forget about the option of working on the course (exhibit limited memory), or succumb to distractions after starting work on course (exhibit limited willpower).

B1. Present biased preferences

When deciding whether to work on the course, students may procrastinate coursework due to present-bias preferences. I model the possibility that students exhibit present-bias preferences with a simplified quasi-hypberbolic discounting model (Laibson, 1997; O'Donoghue and Rabin, 1999). In this model, a student's discounted utility in period t is represented by $U_t = u_t + \beta \sum_{\tau=t+1}^{3} u_\tau$, where u_t is the flow utility in period t and β is a present-biased discount factor.[50] I also assume that a student may be sophisticated or naïve about her present bias. A student is sophisticated about her present bias if she is aware that she will behave more impatiently than she would like to in the future and is naïve if she does not anticipate her future impatience. Formally, a student has beliefs about her future discount factor $\hat{\beta} \in \{\beta, 1\}$ where an student is sophisticated if $\hat{\beta} = \beta$ and naïve if $\hat{\beta} = 1$.

The model, which is formally solved below, provides several predictions about the behavior of students who exhibit present–present biased preferences, but not limited memory or limited willpower. In particular, a student who exhibits present-biased preferences may fail to complete the course, even when completion is utility maximizing from a long-run perspective. The smaller the β (or larger degree of present-bias), the more likely an student is to procrastinate coursework that maximizes long-run utility. Additionally, naïvete about present bias makes a student less likely to complete the course in the first period. This is because a naïve student anticipates that she will behave patiently in the future and is more willing to delay coursework than is utility maximizing.

While students who exhibit present-bias preferences are likely to behave more impatiently than they would like to from a long-run perspective, commitment devices can increase the probability that present-bias students finish the course. In this setting, I introduce a commitment device technology that increases the future cost of spending time on distracting Internet activities. The change in the relative cost of coursework induced by the commitment device in period t is represented by κ_t. If a student in period 1 is given the option to set a commitment for period 2, she will to choose a commitment level that increases the probability she completes the course ($\kappa_2 < 0$) if she is (1) present biased, (2) sophisticated about her present bias, (3) the expected benefit of choosing κ_2 exceeds the cost.[51] When students are present biased and other behavioral factors are absent, commitment devices unambiguously increase a sophisticated student's expected utility.

B2. Limited memory

In addition to being present biased, students may forget about coursework. To incorporate the possibility that students may forget about the choice to work on the course, I allow for the probability of considering the course (ρ_t) to be less

[48] This model is closely related to Ericson's (2014) model of limited memory and present biased preferences and Taubinsky's (2014) model of inattention.

[49] While the cost c_t represents the general opportunity cost of working on the course in period t, I assume recreational/distracting Internet activity is a significant contributor to this opportunity cost.

[50] Quasi-hyperbolic discounted utility models often include an exponential discount factor δ such that $U_t = u_t + \beta \sum_{\tau=t+1}^{3} \delta^t u_\tau$. I make the simplifying assumption that $\delta = 1$. Additionally, I assume that if no other behavioral factors are present, a student follows through on any decision to work on course such that $y_t = x_t$.

[51] Specifically, a student will choose a commitment if there exists a κ^* such that $(P[x_2|\kappa_2 = \kappa*] - P[x_2|\kappa_2 = 0)b - (1 - P[x_2|\kappa_2 = \kappa*])\kappa_2 > 0$. When other behavior factors are absent, a commitment device set in period 1 for period 2 increases the overall probability that a student completes the course, but decreases the probability the probability that a student completes the course in period 1.

than 1, such that $\rho_t \in (0, 1]$. I also let $\alpha_t \in \{0, 1\}$ be an indicator for whether an individual is attentive in period t. I assume that forgetting is a transitory "slipping of the mind" (Ericson, 2014) that is independent of whether or not a task was remembered previously.[52] Just as a student may not anticipate their future present-biased tendencies, a student may or may not be aware of her tendency to forget about coursework $\hat{\rho}_t \in \{\rho_t, 1\}$, and is sophisticated about limited memory if $\hat{\rho}_t = \rho_t$ and naïve if $\hat{\rho}_t = 1$.

When other behavioral factors are absent, an increase in the probability that students remember the course (ρ_t) unambiguously increases the probability that students complete the course. Although an increase in ρ_t in either period 1 or 2 increase the overall probability that a student completes the course, awareness of a coming alert in period 2 decreases the probability that a sophisticated student ($\hat{\rho}_{t+1} = \rho_{t+1}$) completes the course in period 1.

B3. Limited willpower

Finally, I assume that after a student chooses to work on the course, she may become distracted and fail to complete the coursework she chose to do. I let $\pi_t \in (0, 1]$ represent the probability that a student has sufficient willpower to complete the coursework she chooses and $\sigma \in \{0, 1\}$ be an indicator for whether or not the student finishes the coursework that she starts such that $y_t = \sigma x_t$. I assume that where π_t is decreasing in the level of distractions to which she is exposed. I also assume that students who are distracted from completing the task do not incur the cost of work c_t.[53] Just as with present-biased preferences, I allow for naïvete and sophistication about limited willpower $\hat{\pi}_t \in \{\pi_t, 1\}$. An increase π_t in period 1 or 2 increases the overall probability that a student completes the course, but awareness of a coming increase in the probability of following through in period 2 reduces the probability that a sophisticated student completes the course in period 1.

B4. Combining behavioral factors

When present-bias, limited memory, and limited willpower are isolated, the impact of commitment devices, and factors that increase the probability of remembering the course (alerts) and following through on a decision to work on the course (distraction blocking tools) have straightforward impacts on utility and the probability that students complete the course. However, when behaviors are combined, behavioral tools may have unanticipated impacts on student outcomes. First, a student who is sophisticated about her time-inconsistent preferences, but naïve about her limited memory and limited willpower may choose to utilize a commitment device that reduces her overall utility. Naïvete leads a student to overestimate the probability that she will complete the course with the help of a commitment device, and may lead a student to choose a commitment device that reduces her overall well-being.

Furthermore, a student who is sophisticated about her limited memory but naïve about her limited willpower or present bias may actually be less likely to complete a course when she knows that she will get an alert in the following period. Naïvete about limited willpower or present bias leads a student to overestimate the impact of an alert in the following period. An increase in the anticipated probability of remembering coursework in the future decreases the probability that a student chooses to work on the course in the current period. If a student sufficiently overestimates the impact of an alert in the future then the alert may decrease the overall probability that she completes the course. Symmetrically, awareness of the availability a future distraction blocking tool may make a student who is sophisticated about her limited willpower but naïve about her present bias or limited memory less likely to complete the course.

B5. Model solution

Below is a full solution to the model which incorporates present bias and beliefs over present bias ($\beta, \hat{\beta}$), limited memory ($\rho_t, \hat{\rho}_t$), and limited willpower ($\pi_t, \hat{\pi}_t$). Students solve for their utility maximizing choice by backwards induction.

In the period 2, the final decision period, a student will choose to work on the course if she considers the choice to work on the course, and the discounted benefit exceeds the cost. Formally, the choice to work on the course can be characterized by the following:

$$x_2 = \begin{cases} 1, & \text{if } \alpha_2 = 1, y_1 = 0, \text{and } \hat{\pi}_2 c_2 + \hat{\pi}_2 \beta b > 0 \\ 0, & \text{otherwise} \end{cases}$$

where α_2 is an indicator of whether the student considers the choice of coursework, $\hat{\pi}_2$ is the belief of the probability that the student will complete a task she begins, β is the present-bias discount factor, c_2 the cost of completing the course, and b the benefit of completing the course. Note that a student will never work on the course if $y_1 = 1$, because there is no benefit to working on the course if the course was already completed in period 1.

---

[52] Some previous work has assumed that limited memory follows a dynamic process where the probability of forgetting to be increasing in previous memory failure (Ericson, 2014; Mullainathan, 2002; Taubinsky, 2014). In the most extreme case Ericson (2014), assumes that once something is forgotten it can never be remembered again.

[53] This assumption implies that students are distracted from completing the task shortly after deciding to work on the task. While this pattern matches observed behavior, the model generates similar predictions when this assumption is relaxed.

The probability that a student completes the course in the second period ($Pr[y_2 = 1]$) given that $y_1 = 0$, depends on the distribution of costs $f(c)$, the discounted benefit of action βb, and the probability of considering the course ρ_2 and the probability of following through with a decision to work on the course π_2, such that:

$$\begin{aligned} Pr[y_2 = 1 | y_1 = 0] &= \pi_2 \rho_2 \int_{-\beta b}^{\infty} f(c) dc \\ &= \pi_2 \rho_2 (1 - F(-\beta b)) \end{aligned} \tag{1}$$

where $F(\cdot)$ is the cumulative distribution function (CDF) of $f(c)$ and other variables are as previously specified.

In period 1, a student will choose to work on the course if the net value of working on the course in period 1 ($c_1 + \beta b$) exceeds $\hat{v}_{(2,0)}$—the expected value of the choice to work on course in period 2. Formally:

$$x_1 = \begin{cases} 1, & \text{if } \alpha_1 = 1, \text{ and } \hat{\pi}_1 c + \beta \hat{\pi}_1 b + (1 - \hat{\pi}_1) \hat{v}_{2,0} > \hat{v}_{2,0} \\ 0, & \text{otherwise} \end{cases}$$

where:

$$\hat{v}_{2,0} = \beta \hat{\rho}_2 \hat{\pi}_2 \int_{-\hat{\beta} b}^{\infty} (c + b) f(c) dc \tag{2}$$

and other variables are as previously specified. Note that students may anticipate the possibility of choosing to work on the course but not follow through, and that part of the value of choosing $x_1 = 1$ in period 1 is $(1 - \hat{\pi}_1)\hat{v}_{2,0}$, or the expected probability of failing to complete coursework multiplied by the expected value of value of the choice of x in period 2. Given the choice above, the probability that a student completes the course in the first period is:

$$\begin{aligned} Pr[y_1 = 1] &= \rho_1 \pi_1 \int_{-\beta b + \hat{v}_{2,0}}^{\infty} f(c) dc \\ &= \pi_1 \rho_1 (1 - F(-\beta b + \hat{v}_{2,0})) \end{aligned} \tag{3}$$

Having calculated the conditional probability of completing the course in the second period, $Pr[y_2 = 1 | y_1 = 0]$ and the unconditional probability of completing the course in the first period, the total probability of completing the course can be expressed by the following equation:

$$\begin{aligned} Pr[\mathbf{y} = 1] &= \rho_1 \pi_1 \int_{-\beta b + \hat{v}_{2,0}}^{\infty} f(c) dc + [1 - \rho_1 \pi_1 \int_{-\beta b + \hat{v}_{2,0}}^{\infty} f(c) dc] \rho_2 \pi_2 \int_{-\beta b}^{\infty} f(c) dc \\ &= \pi_1 \rho_1 (1 - F(-\beta b + \hat{v}_{2,0})) + [1 - \pi_1 \rho_1 (1 - F(-\beta b + \hat{v}_{2,0}))] \pi_2 \rho_2 (1 - F(-\beta b)) \end{aligned} \tag{4}$$

B6. Comparative statics

In this Section 1 examine how changes in the probability that students remember the choice to work on the course (ρ_t), the probability that students have sufficient willpower to complete coursework (π_t), and the relative cost of coursework (κ_t) impact course completion.

B6.1. Changes in the probability in remembering

Increasing the probability that students consider the course tends to increase the probability that students complete coursework. However, the magnitude and direction of the impact of changing ρ_t depends on the period t and value of other parameters in the model.

The impact of increasing ρ_1 on the probability that a student completes the course is:

$$\frac{\partial Pr[y = 1]}{\partial \rho_1} = [1 - \rho_2 \pi_2 \int_{-\beta b}^{\infty} f(c) dc] \pi_1 \int_{-\beta b + \hat{v}_{2,0}}^{\infty} f(c) dc \tag{5}$$

and the impact of increasing ρ_2 on the probability of course completion is:

$$\begin{aligned} \frac{\partial Pr[y = 1]}{\partial \rho_2} &= -\rho_1 \pi_1 [1 - \rho_2 \pi_2 \int_{-\beta b}^{c_l} f(c) dc] f(-\beta b + \hat{v}_{2,0}) \frac{\partial \hat{v}_{2,0}}{\partial \rho_2} \\ &\quad + [1 - \rho_1 \pi_1 \int_{-\beta b + \hat{v}_{2,0}}^{\infty} f(c) dc] \pi_2 \int_{-\beta b}^{\infty} f(c) dc \end{aligned} \tag{6}$$

where

$$\frac{\partial \hat{v}_{2,0}}{\partial \rho_2} = \begin{cases} \beta \hat{\pi}_2 \int_{-\hat{\beta} b}^{c_l} (c + b) f(c) dc, & \text{if } \hat{\rho}_2 = \rho_2 \\ 0, & \text{if } \hat{\rho}_2 = 1 \end{cases}$$

The equations above highlight several properties of increasing the probability of remembering the course. Eq. (5) shows that increasing the probability of ρ_1 unambiguously increases the probability that a student completes her coursework. Furthermore the impact of an increase in ρ_1 on completion is increasing in π_1 and decreasing in ρ_2, π_2, $\hat{\rho}_2$, $\hat{\pi}_2$, and $\hat{\beta}$.

Intuitively, an increase in the probability that a student remembers the choice to work on the course on the first period has the largest impact for students who are likely to follow through on their choice to work and who are unlikely (and aware that they are unlikely) to complete the coursework in the second period.

Eq. (6) shows the impact of increasing the probability that a student will remember the choice to work (ρ_2) in period 2. If a student is naïve about her tendency to forget, then the first line of Eq. (6) is equal to zero, and the increase in the probability in unambiguously positive. However, if a student is aware that she of the increase in ρ_2 in period 1, the first line of Eq. (6) shows the impact of the increase in ρ_2 is diminished, and may even be negative. Sophistication about ρ_2 further diminishes the probability of course completion if individuals are naïve about limited willpower and present-bias preferences.

B6.2. Changes in the probability that students follow through on coursework

Increasing the probability that students follow through on a decision they make to work (π_t) has symmetric implications to increasing the probability that students remember to consider the course. The impact of increasing the probability that a student will follow through with a choice in period 1 (π_1) on the probability that students complete the course is:

$$\frac{\partial Pr[y=1]}{\partial \pi_1} = [1-\rho_2\pi_2\int_{-\beta b}^{\infty} f(c)dc]\rho_1\int_{-\beta b+\hat{v}_{2,0}}^{\infty} f(c)dc \tag{7}$$

And the impact of increasing the probability that a student follows through with a choice in period 2 (π_2) on course completion is:

$$\frac{\partial Pr[y=1]}{\partial \pi_2} = -\rho_1\pi_1[1-\rho_2\pi_2\int_{-\beta b}^{c_l} f(c)dc]f(-\beta b+\hat{v}_{2,0})\frac{\partial \hat{v}_{2,0}}{\partial \pi_2} + [1-\rho_1\pi_1\int_{-\beta b+\hat{v}_{2,0}}^{\infty} f(c)dc]\rho_2\int_{-\beta b}^{\infty} f(c)dc \tag{8}$$

Eq. (7) shows that tie impact of increasing π_1 increases the probability that a student completes the course, and that the impact of increasing π_1 is increasing in ρ_1 and decreasing in ρ_2, π_2, $\hat{\rho}_2$, $\hat{\pi}_2$, and $\hat{\beta}$. Eq. (8) shows that increasing π_2 has an unambiguously positive impact on completion for who do not believe they may fail to follow through on their decision to complete the course, but the impact is reduced and may even be negative for students who are sophisticated about their limited willpower and naïve about their limited memory or present bias.

B6.3. Changes in the relative costs in coursework

A student may have access to a technology that changes the cost of the alternate choice to coursework in period t. An change of in the cost of the alternate choice leads to a corresponding shift in the relative cost of coursework, such that the new cost of the coursework is $c_t+\kappa_t$. If $c_t+\kappa_t$ is substituted for c_t in the student's utility maximization problem, the resulting probability of completing the course is:

$$Pr[\mathbf{y}=1] = \rho_1\pi_1\int_{-\beta b+\kappa_1+\hat{v}_{2,0}}^{\infty} f(c)dc + [1-\rho_1\pi_1\int_{-\beta b+\kappa_1+\hat{v}_{2,0}}^{\infty} f(c)dc]\rho_2\pi_2\int_{-\beta b+\kappa_2}^{\infty} f(c)dc$$

where:

$$\hat{v}_{2,0} = \beta\hat{\rho}_2\hat{\pi}_2\int_{-\hat{\beta}b+\kappa_2}^{\infty}(c+b)f(c)dc \tag{9}$$

Increasing the relative cost of coursework in period 1 (κ_1) has the following impact on course completion:

$$\frac{\partial Pr[y=1]}{\partial \kappa_1} = -\rho_1\pi_1[1-\rho_2\pi_2\int_{-\beta b+\kappa_2}^{\infty} f(c)dc]f(-\beta b+\kappa_1+\hat{v}_{2,0}) \tag{10}$$

while increasing the relative cost of coursework in period 2 (κ_2) had the following impact on course completion:

$$\frac{\partial Pr[y=1]}{\partial \kappa_2} = -\rho_1\pi_1[1-\rho_2\pi_2\int_{-\beta b+\kappa_2}^{\infty} f(c)dc]f(-\beta b+\kappa_1+\hat{v}_{2,0})\frac{\partial \hat{v}_{2,0}}{\partial \kappa_2} - \rho_2\pi_2[1-\rho_1\pi_1\int_{-\beta b+\kappa_1+\hat{v}_{2,0}}^{\infty} f(c)dc]f(-\beta b+\kappa_2) \tag{11}$$

where:

$$\frac{\partial \hat{v}_{2,0}}{\partial \kappa_2} = -\beta\hat{\rho}_2\hat{\pi}_2((1-\hat{\beta})b+\kappa_2)f(-\hat{\beta}b+\kappa_2) \tag{12}$$

Eq. (10) shows the unsurprising result that an increase in κ_1, the relative cost of course completion, leads to a decrease in the probability of course completion. The first line of Eq. (11) shows an anticipated increase the cost of course completion in period 2 increases the likelihood a student completes the course in period 1 line 2 of Eq. (11) indicates that higher period 2 costs decrease the probability that a student completes the course in period 2. A comparison of terms in the first and second line of Eq. (11) indicates that the overall impact of increasing the relative cost of the course decreases the total probability that students complete the course.

B7. Impact of treatments on model parameters

Each of the commitment device, alert, and distraction blocking treatments are designed to target different aspects of time-management problems. Below I describe how the treatments in this study are likely to impact different model parameters.

B7.1. Commitment device

The commitment device treatment prompts students via a daily email set to limits for the amount of distracting time they spend on their computer. This treatment is likely to impact several parameters in the model. First, if students experience limited memory, then daily emails and blocked websites are likely to increase ρ_t—the probability that students consider the course. Second, if students have limited willpower, the commitment device may also reduce π_t—the probability that a student follows through on a decision to work on the course. When a commitment device binds, then distractions are removed, which may increase the probability that students follow through on a decision to work on the course (π_t). Finally a pre-commitment to a restrictive time-limit decreases the relative cost of working on the course in a future period by κ_{t+1}, which increases the probability of completing the course in the following period. Given the predictions of the model, the commitment device may increase course completion by addressing limited memory (through increasing ρ_t), limited willpower (through increasing π_t), or present bias preferences (through decreasing κ_{t+1}).

B7.2. Alert

The alert treatment provides students with a reminder after each half hour of distracting time which includes a link to the course website. If students exhibit limited memory, this alert is likely to increase the probability that students remember the decision to work on the course ρ_t. It is also possible that the alert increases the cost of distracted Internet browsing, either by annoyance or guilt. If the reminder decreases the relative cost of coursework (κ_t) by making distracted Internet time less enjoyable, this may lead to increase the probability of course completion. This, however, is not evidence of present-bias, limited memory, or limited willpower—a decrease in the relative cost of coursework is likely to increase completion for all students, including those who experience no behavioral issues. Therefore, a positive impact of the alert treatment is consistent with a model limited memory, but does not rule out other models of behavior.

B7.3. Distraction blocking tool

The distraction blocking tool allows students to block out distracting websites for up to 60 min upon going to the course website. If a student has limited willpower and the presence of Internet distractions reduce the probability that she completes coursework she decides to do, then blocking distractions may increase the probability that she follows through with her decision to work on the course (π_t). Because a student must go to the course in order to interact with the distraction blocking tool, the distraction blocking tool is unlikely to impact on the probability that students consider the course ρ_t. Also, it is difficult for students to use the distraction blocking tool as a commitment device because choosing to utilize the distraction blocking tool impacts the relative costs of coursework for no more than an hour, and the distraction blocking tool takes immediate effect. As a result, a response to the distraction blocking tool is unlikely to be explained by present-bias[54] or limited memory, but is consistent with a model of limited willpower.

References

Abadie, A., Chingos, M.M., West, M.R., 2013. Endogenous stratification in randomized experiments. Rev. Econ. Stat. Forthcoming.

Allen, I.E., Seaman, J., 2013. Changing Course: Ten Years of Tracking Online Education in the United States. Babson Survey Research Group, Babson Park, MA.

Ariely, D., Wertenbroch, K., 2002. Procrastination, deadlines, and performance: self-control by precommitment. Psychol. Sci. 13 (3), 219–224.

Ashraf, N., Karlan, D., Yin, W., 2006. Tying odysseus to the mast: evidence from a commitment savings product in the philippines. Q. J. Econ. 121 (2), 635–672. doi:10.1162/qjec.2006.121.2.635.

Austin, S.M., Balas, E.A., Mitchell, J.A., Ewigman, B.G., 1994. Effect of physician reminders on preventive care: meta-analysis of randomized clinical trials. In: Proceedings of the Annual Symposium on Computer Application in Medical Care. American Medical Informatics Association, p. 121.

Banerjee, A.V., Duflo, E., 2014. (Dis)organization and success in an economics MOOC. AEA Pap. Proc 104 (5), 514–518.

Baumeister, R.F., Bratslavsky, E., Muraven, M., Tice, D.M., 1998. Ego depletion: is the active self a limited resource? J. Pers. Soc. Psychol. 74 (5), 1252–1265.

Baumeister, R.F., Vohs, K.D., 2003. Willpower, choice, and self-control. In: Loewenstein, G., Read, D., Baumeister, R. (Eds.). In: Time and decision: Economic and psychological perspectives on intertemporal choice. US: Russell Sage Foundation, New York, NY, pp. 201–216.

Breslow, L.B., Pritchard, D.E., DeBoer, J., Stump, G.S., Ho, A.D., Seaton, D.T., 2013. Studying learning in the worldwide classroom: research into edX's first MOOC. Res. Pract. Assess. 8, 13–25.

Bryan, G., Karlan, D., Nelson, S., 2010. Commitment devices. Annu. Rev. Econ. 2 (1), 671–698.

Cadena, X., Schoar, A., 2011. Remembering to pay? Reminders vs. financial incentives for loan payments (No. w17020). National Bureau of Economic Research.

Calzolari, G., Nardotto, M., 2011. Nudging with information: a randomized field experiment on reminders and feedback (No. 8571). CEPR Discussion Papers.

Chetty, R., Looney, A., Kroft, K., 2009. Salience and taxation: Theory and evidence. Am. Econ. Rev. 99 (4), 1145–1177.

Doherty, W., 2006. An analysis of multiple factors affecting retention in web-based community college courses. Internet Higher Educ. 9 (4), 245–255.

Ericson, K.M.M., 2011. Forgetting we forget: overconfidence and memory. J. Eur. Econ. Assoc. 9 (1), 43–60.

Ericson, K.M.M., 2014. On the interaction of memory and procrastination: implications for reminders.

Fudenberg, D., Levine, D.K., 2006. A dual-self model of impulse control. Am. Econ. Rev. 96 (5), 1449–1476.

Fudenberg, D., Levine, D.K., 2012. Timing and self-control. Econometrica 80 (1), 1–42. doi:10.3982/ECTA9655.

[54] To make this assertion, I assume that students treat time within an hour as being within the present period.

Ginder, S., Stearns, C.. Enrollment in distance education courses, by state: Fall 2012. National Center for Education Statistics (NCES 2014023). https://nces.ed.gov/Pubsearch/pubsinfo.asp?pubid=2014023
Giné, X., Karlan, D., Zinman, J., 2010. Put your money where your butt is: a commitment contract for smoking cessation. Am. Econ. J.: Appl. Econ. 2 (4), 213–235.
Houser, D., Schunk, D., Winter, J., Xiao, E., 2018. Temptation and commitment in the laboratory. Games Econ. Behav. 107, 329–344. doi:10.1016/j.geb.2017.10.025.
Karlan, D., McConnell, M., Mullainathan, S., Zinman, J., 2010. Getting to the top of mind: how reminders increase saving.
Kast, F., Meier, S., Pomeranz, D., 2012. Under-savers anonymous: evidence on self-help groups and peer pressure as a savings commitment device.
Kaur, S., Kremer, M., Mullainathan, S., 2011. Self-Control at Work.
Krishna, S., Boren, S.A., Balas, E.A., 2009. Healthcare via cell phones: a systematic review. Telemed. e-Health 15 (3), 231–240.
Laibson, D., 1997. Golden eggs and hyperbolic discounting. Q. J. Econ. 112 (2), 443–478.
Levitt, S. D., List, J. A., Neckermann, S., Sadoff, S., 2012. The behavioralist goes to school: leveraging behavioral economics to improve educational performance.
Madrian, B.C., Shea, D.F., 2001. The power of suggestion: inertia in 401 (k) participation and savings behavior. Q. J. Econ. 116 (4), 1149–1187.
McDaniel, M.A., Einstein, G.O., 2007. Prospective Memory: An Overview and Synthesis of an Emerging Field. Sage Publications.
Mullainathan, S., 2002. A memory-based model of bounded rationality. Q. J. Econ. 117 (3), 735–774.
O'Donoghue, T., Rabin, M., 1999. Doing it now or later. Am. Econ. Rev. 151 (3712), 867–868.
Ozdenoren, E., Salant, S.W., Silverman, D., 2012. Willpower and the optimal control of visceral urges. J. Eur. Econ. Assoc. 10 (2), 342–368. doi:10.1111/j.1542-4774.2010.01050.x.
Perna, L., Ruby, A., Boruch, R., Wang, N., Scull, J., Evans, C., Ahmad, S., 2013. The life cycle of a million MOOC users.
Romano, J.P., Wolf, M., 2016. Efficient computation of adjusted -values for resampling-based stepdown multiple testing. Stat. Probab. Lett. 113, 38–40. doi:10.1016/j.spl.2016.02.012.
Taubinsky, D., 2014. From intentions to actions: a model and experimental evidence of inattentive choice. Working paper.
Thaler, R.H., Benartzi, S., 2004. Save more tomorrow: using behavioral economics to increase employee saving. J. Polit. Econ. 112 (S1), S164–S187.
Waldrop, M.M., 2013. Massive open online courses, aka MOOCs, transform higher education and science. Nat. Mag. 495, 160–163.
Wilkowski, J., Deutsch, A., Russell, D.M., 2014. Student skill and goal achievement in the mapping with google MOOC. In: Proceedings of the First ACM Conference on Learning at Scale Conference. ACM, pp. 3–10.
Winters, F.I., Greene, J.a., Costich, C.M., 2008. Self-regulation of learning within computer-based learning environments: a critical analysis. Educ. Psychol. Rev. 20 (4), 429–444.
Xu, D., Jaggars, S., 2011. Online and hybrid course enrollment and performance in washington state community and technical colleges.
Zizzo, D.J., 2010. Experimenter demand effects in economic experiments. Exp. Econ. 13 (1), 75–98.
Zogg, J.B., Woods, S.P., Sauceda, J.A., Wiebe, J.S., Simoni, J.M., 2012. The role of prospective memory in medication adherence: a review of an emerging literature. J. Behav. Med. 35 (1), 47–62.