Supporting Self-Control for Students in Online Courses

Final Report for the Amir Lopatin Fellowship

Daniel Greene

3/1/2015

Abstract

Online education is touted as a means to teach the world, but as it is currently offered, successful online learning depends on self-control capacities that are lacking for many people. In this study, I employ design-based research to test three tools to support student self-control in an open-enrollment nine-week online statistics course with approximately 10,000 students. The tools included an intelligent reminder of distracting-website time usage, a "commitment device" that allowed students to set a limit on the following day's distracting-website time, and a "focus-time" tool that allowed students to temporarily block distracting websites when they logged onto the course. In a randomized controlled trial, I found that the focus-time and commitment device tools had significant positive effects on student assignment submission for the first half of the course, but that effects largely disappeared for the second half. Implications and next steps for research are discussed.

Introduction

Online education and the need for self-control

The global demand for higher education is unprecedented. UNESCO's 2009 World Conference on Higher Education estimated that global participation in postsecondary education will grow from 165 million to 263 million by 2025 – a change that would require four major campus universities (with roughly 30,000 students each) to open every week for 15 years (Daniel, 2012). Bolstered by the recent worldwide prominence of MOOCs, online education has been hailed as a means of meeting this demand at a scale that is affordable to both educators and students (Bowen, 2013; Christensen, Horn, & Johnson, 2008; Vander Ark, 2011).

However, learning effectively online requires a substantial capacity for self-control (Bambara, Harbour, Davies, & Athey, 2009; Ehrman, 1990; Eisenberg & Dowsett, 1990). Self-control is "the voluntary regulation of behavioral, emotional, and attentional impulses in the presence of momentarily gratifying temptations or diversions to achieve long-term goals" (Angela L. Duckworth, 2011), and it is vitally important throughout our lives. When estimated reliably, the power of self-control for predicting a wide range of academic, health, and social outcomes rivals that of intelligence or socioeconomic status (Angela L. Duckworth, 2011; Moffitt et al., 2011). Online students have to schedule their own learning, stick to long-term goals, and resist a range of online distractions with little to no oversight or social pressure.

Unfortunately, there is much room for improvement in our own self-control abilities, both in academic work and in our time spent online. For example, approximately 75% of college students describe themselves as procrastinators, and almost 50% as procrastinating consistently and problematically (Day, Mensink, & O'Sullivan, 2000; Onwuegbuzie, 2000). Not surprisingly, procrastination is correlated with significantly lower academic performance (Beck, Koons, & Milgrim, 2000). Online, multitasking is widespread. A recent study of multitasking on personal computers estimated that users switched content every 19 seconds (Yeykelis, Cummings, & Reeves, 2014). However, research by Clifford Nass and colleagues on media multitasking makes it clear that in the majority of cases, multitasking harms performance on all tasks, and yet many people incorrectly report that they are able to multitask without problems (Ophir, Nass, & Wagner, 2009).

Not only do we struggle with self-control in academic and online settings, but self-control abilities do not appear to equally distributed across socioeconomic class lines, at least at a young age. Moffitt et al. (2011) found a correlation of .25 between children's self- control and their househould SES, and Evans and Rosenbaum (2008) found that family income predicted delay-of-gratification ability among a diverse national sample of elementary-schoolers, which itself predicted academic achievement in middle school. The longitudinal work of Mischel and colleagues (e.g., Mischel, Shoda, & Rodriguez, 1989) has also drawn links between childhood delay-of- gratification and later health and self-control outcomes. The causes of these relationships are still unclear (Heckman, Stixrud, & Urzua, 2006), but if people from lower-SES backgrounds tend to have less self-control, then, ironically, they may be less able to benefit from the online courses that are their only financial option. This issue poses a challenge to popular hopes that online education can close socioeconomic status gaps (Vander Ark, 2011).

Developing, supporting, and offloading self-control

How can we help students to stay focused in online courses? In this section, I outline four known general methods of supporting self-control: building self-control capacity directly, changing academic mindsets, practicing skills and strategies, and offloading self-control to the environment.

A decade of research by Roy Baumeister and colleagues suggests that self-control draws from a limited but expandable resource of energy. Like a muscle, it can be depleted in the short-term, leaving people vulnerable to temptations and distractions, and it can be built over time with regular practice (Baumeister, Vohs, & Tice, 2007). In one such study, experimental subjects who practiced effortful activities for a period of several weeks, such as using their non-dominant

hand to perform daily activities, showed improved abilities on unrelated self-control tasks in the laboratory (Baumeister & Tierney, 2011). Recent research on meditation also suggests that as little as five days' regular practice of a meditation program can improve subjects' performance on an attention task (Tang et al., 2007). Considering the importance of self-control for a range of life outcomes, this research is very exciting, but the time commitment and discipline required of subjects and the gradual rate of improvement observed makes self-control training inappropriate for an online learning setting.

Another set of solutions to the self-control problem focuses on influencing the academic mindsets that students hold, which in turn influence their motivation and their academic behaviors. For example, students' beliefs about the possibility of improving at an academic skill (Blackwell, Trzesniewski, & Dweck, 2007) or the relevance of course material to their future lives (Hulleman & Harackiewicz, 2009) have profound influences on their self-regulatory choices, because they change the meaning of momentary experiences like challenge or boredom. Researchers have developed a range of brief but psychologically-potent activities, typically involving reading an article and writing a response, that use persuasion techniques to lead students to adopt more adaptive mindsets. These activities can have surprisingly powerful effects for their cost and duration (Yeager & Walton, 2011).

Instead of focusing on mindsets, some researchers have studied mental processes related to planning and strategizing that are thought to be more psychologically proximal to the actual behaviors of self-control. One popular technique involves guiding subjects to set "implementation intentions" – simple if-then plans that specify an environmental trigger and a desired behavioral response (Gollwitzer, 1999). For example, a dieter might form the following plan: "If I am offered the dessert menu at a restaurant, then I will not even look at it and instead ask for the check."

In a meta-analysis of 94 studies with over 8,000 participants, implementation intentions had a remarkable effect size of .65 for helping subjects to achieve a variety of conscious goals, above and beyond the effect of the general intention of reaching the goal (Gollwitzer & Sheeran, 2006). In combination with another strategy called "mental contrasting", implementation intentions helped students studying for a standardized test to voluntarily complete 66% more practice problems over the summer than a control group (Duckworth, Grant, Loew, Oettingen, & Gollwitzer, 2011).

To summarize, research suggests that self-control can be supported by building it through repeated use, reframing it with a different mindset, or using the right strategies. In this project, I explored a fourth method of supporting self-control – providing tools to allow people to "offload" self-control into their environments.

Self-control in the environment

Examples of deliberately manipulating one's environment to maintain self-control can be found as far back as Odysseus tying himself to the mast of his ship to resist the tempting song of the Sirens (Fagles & Knox, 1997). But most empirical work on the subject has been in the study of addiction and health behavior change, where controlling one's environment is recognized as essential (Bernheim & Rangel, 2004; Dal Cin, MacDonald, Fong, Zanna, & Elton-Marshall, 2006). The recent popularity of self-control apps that monitor, reward, punish, and restrict one's behavior also offer promising signs that the right tool can aid self-control (Fowler & Ovide, 2013). However, to my knowledge, no research has investigated the use of digital tools that support student self-control. Below, I describe two classes of tools that I investigated in this research.

Reminders

Many theoretical models of self-control, such as those of Zimmerman (2002) and Carver and Scheier (Carver & Scheier, 2001), describe it in terms of a metacognitive loop of monitoring one's performance for problems. One must be vigilant for problems, detect them when they arise, interpret them correctly, search memory for the appropriate solutions, and then act on them (Carver & Scheier, 1981).

Reminders serve as a means of offloading part of this process to the environment (Levitin, 2015; Norman, 1988). At minimum, a reminder activates related concepts in memory, making them more easily accessible from memory and temporarily aiding vigilance (Ross, 2006). This process can even happen outside of conscious awareness (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001). However, mental accessibility has a short and variable time course (Althaus & Kim, 2006). Because of this, reminders can fail to activate related concepts at the time when they are needed. To avoid this problem, reminders are often designed to be seen at specific times when they are actionable or otherwise relevant. For example, a post-it note on one's keys can remind someone to take an action immediately before leaving the house. Implementation intentions (discussed in the previous section) use specific settings as a mental trigger to remind people to take pre-planned appropriate actions, thus automating the self-control loop.

In this study, I implemented a basic reminder system to interrupt students who may have forgotten to mentally monitor their distracting time usage. Every time a student accrued 30 minutes of time spent on distracting websites or programs in a day, he or she received a popup stating the current daily total of distracting time. This was intended to re-engage the student in monitoring their

distracting time usage and identifying problems with their current choice of activity.

Commitment devices

Much of the research on environment-based self-control is founded on research from behavioral economics on "hyperbolic discounting" of rewards over time (Ainslie & Haslam, 1992; Laibson, 1997). When offered choices between rewards of the same size at different times, people appear to discount more temporally distant rewards to a degree that is neither constant (e.g., minus 1 dollar per day) nor exponential (e.g., 99% of the previous day's amount). People strongly prefer present rewards to future ones, even in the very near future. This can lead people to hold time-inconsistent preferences – for example, they would prefer $10 today over $20 tomorrow, but $20 in 101 days over $10 in 100 days (Hoch & Loewenstein, 1991).

With this knowledge, it may be rational for an economic agent to use a "commitment device" to manipulate the feasibility and value of future choices in order to attain goals that the agent wants in the long term (Bryan, Karlan, & Nelson, 2010). In simple terms, commitment devices make temptations less accessible or enjoyable. Some everyday examples of commitment devices include not keeping junk food in the house, giving one's car keys to a friend, or undergoing bariatric surgery to reduce the size of one's stomach. In the domain of personal finance, the "Save More Tomorrow" financial plan has applied this principle to help people meet longer-term financial goals (Thaler & Benartzi, 2004). In the domain of academic performance, Ariely and Wertenbroch (2002) demonstrated that students given the option to pre-commit to optional early deadlines performed better in a college course.

In the research described here, I tested out two tools based on the principle of commitment devices. They were triggered differently and operated at different time scales. The first tool was a daily email to students that they saw whenever they chose to check their inbox. This email gave them the ability to set a limit on the amount of distracting website time that they could spend *tomorrow*. The second tool, referred to here as "focus-time", was triggered when students actually logged into their online course. It displayed a popup that gave students the ability to immediately block distracting websites for a desired amount of time, thus exercising controlling power over the immediate future. By comparing students' usage habits, survey feedback, and learning outcomes, I intend to take a first step toward developing a commitment device that students would voluntarily use and that would improve their learning outcomes.

Research Goals and Approach

This study is intended to be the start of a larger research project. My long-term goals for this project are both pragmatic and theoretical – to engineer tools that improve learning outcomes in online courses by supporting self-control, and to develop a better understanding of how self-control plays out in online learning environments.

To meet these goals, I chose to conduct design-based research (Collins, Joseph, & Bielaczyc, 2004; Reimann, 2011), an approach pioneered in the learning sciences in the 1980s and 1990s. Design-based research pursues theoretical and practical goals in parallel by iteratively developing and testing interventions in real educational settings. Design-based research projects typically rely on both qualitative and quantitative data in an attempt to capture the complexity of particular situations under study (Cobb et al., 2003).

Accordingly, the study described in this paper attempts to make incremental progress towards my larger goals highlighted above. My research questions for this first study are the following:

1. Will distracting-time reminders, a commitment device, or focus-time have any positive effect on student learning outcomes?
2. How do students' patterns of web browsing and tool usage differ across treatment conditions?
3. What do students report about the experience of using the tools?

For the first goal, I use a randomized controlled trial framework to assign students to use different tools, and I examine the effects of tool assignment on learning outcomes. I have no hypotheses about which tools will be most effective – my intent is to test a diverse range of tools, and to use the results to prioritize more focused research in follow-up studies on the most promising ones. As described in the analysis section, I use logistic regression to compare the proportions of groups while controlling for demographic and survey-response variables.

For the second and third goals, I use fine-grained snapshots of student computer usage and open-ended survey questions to develop a richer picture of how students actually use the tools. Analysis of these data is ongoing, so I do not report on them in this paper.

Experimental Methods and Data Sources

Overview

My colleague Richard Patterson and I conducted a randomized controlled trial in the MOOC "Statistics for Medical Professionals" (or "MedStats"), taught by Professor Kristin Sainani and hosted on Stanford's OpenEdX MOOC platform. The course started in late June 2014, lasted nine weeks, and covered a range of statistical concepts. Like many MOOCs, it had lecture videos with in-video quizzes, weekly auto-graded homework assignments, a class discussion forum, and a final exam. Students receiving a final grade of over 65% received a Certificate of Achievement. 10,681 students enrolled in the course, and 1,155 received certification – approximately a 10% certification rate, which is comparable to other MOOCs (Clow, 2013).

Students enrolling in MedStats saw a front-page announcement inviting them to complete in a voluntary paid study on time management. I also sent out invitation emails containing the same announcement immediately before and after the course started. The announcement contained a link to a Qualtrics survey containing a detailed consent form. Students were invited to participate in a study in which they would download software to their computers that would take continual records of the name of the program that was on the foreground of the screen, and, if the program is a web browser, the URL of the current website. The software would automatically be deactivated at the end of the course, and they would also be free to uninstall it at any time. In addition, the software would offer functions to help them stay focused and avoid procrastinating while working on the course. Finally, students would be asked a brief set of pre-survey questions immediately, and another set at the end of the course, whether or not the student had completed the course. For installing the software and completing the pre-survey, students would be paid $5 in an Amazon.com gift card sent to their email address, and for completing the post-survey and keeping the software installed for the duration of the course, they would be paid an additional $7.

Students who wished to participate completed the pre-survey, installed the software, and participated in the course however they desired. At the end of the course, I received student grade data from OpenEdX and combined it with the pre- and post-survey data and the data from the time-management software.

Pre-survey questions

<u>Reasons for enrolling:</u> I used the question developed by Schneider and Kizilcec (2014)to ask students about their reasons for enrolling, e.g. "relevant to job" and "for fun and challenge". Students were free to choose multiple reasons.

Prior experience: I asked students about the number of MOOCs that they had started and completed, number of statistics courses taken, and familiarity with the concepts covered in each week of the course (Likert scale 1-5).

Motivation, goals, and intended effort: I asked students to share their goal for the course. The options for goals were “No specific goals, I’m just exploring”, “Complete the entire course on time and receive a certificate of completion”, "Complete all of the assignments and watch all the videos, but at my own pace", "Complete a subset of the assignments at my own pace", "Watch a subset of the videos at my own pace", "I'm not sure", and "Other".

To measure aspects of student motivation, I asked for students’ interest in the course topic, their commitment to reaching their stated goal, their perceived likelihood of reaching their stated goal, and the perceived importance of their course goal (all Likert scales 1-5, from “not at all” to “extremely”). I also asked students for their intended number of hours/week to spend on the course.

Self-control and procrastination: I asked students to estimate the current amount of time per day that they spend on “distracting” websites, as well as the amount of time that they would *like* to spend, and the degree to which they would like to spend more or less time on distracting websites (five-point Likert scale, from “much more time on distracting websites” to “much less time on distracting websites”).

To assess self-control, I administered the Brief Self-Control Scale from Tangney et al. (2004), which consists of ten statements such as “I have a hard time breaking habits” and “People would say that I have very strong self-discipline”, assessed on a five-point Likert scale from “Not like me at all” to “Very much like me”.

Self-control beliefs: I assessed beliefs about whether willpower was depletable in the short term and improvable in the long term using questions adapted from Job et al. (2010) such as “When you have been working on a strenuous mental task, you feel energized and you are able to immediately start with another demanding activity". These were assessed on a six-point Likert scale from “Strongly Disagree” to “Strongly Agree” to polarize people into agreeing or disagreeing.

I was also interested in people’s beliefs about the degree to which self-control depends on internal properties of a person as compared to external properties of that person’s environment. I adapted the teacher/student responsibility scale

from Lewis (2001) to complete the sentence stem "I think that mental discipline…". Answers were on a seven-point Likert scale ranging from "Almost entirely depends on the person" to "Almost entirely depends on the environment".

Technology: I asked students about the type of computer on which they installed their software, and about the percentages of time they plan to spend on the course using different types of devices.

Demographics: I asked students for their highest educational qualification, gender, country of residence, race (only asked if they chose "USA" or "North America, outside USA" as their choice), and annual income in dollars.

Software and treatment conditions

All participating students installed some form of time-management software on their computers. This software was originally developed by RescueTime, a private company that sells subscriptions to software that offers a range of functions for tracking one's time. I worked with Richard Patterson and the RescueTime staff to develop simplified versions of their product. RescueTime leveraged the aggregated judgments of their users to create a database of programs and websites commonly considered unproductive, including social media, news, sports, and entertainment. It also maintains analogous lists of neutral and productive sites and programs, and a database of categories such as "news" or "sports" for sites and programs. All of the software in this study used these lists to categorize the current program or website every five seconds, and sends this data to a central server over the web. The software was set to run automatically at system startup and to not be easy to turn off. Students were also unable to change the categorization of any sites or programs.

Students were randomly assigned to one of four different software treatments. To ensure against self-selection problems, all students installed the same basic version of the software after completing the survey, the researchers set their treatment conditions remotely, and they were only told of their treatment condition after installing the software via an explanatory email sent to the address that they provided in the consent form. The software treatments were as follows:

Control: Students in the control condition were told that at any time, they could check a "dashboard" website with information about their patterns of distracting time usage. The dashboard displays graphs with information about their most and least productive days and times of day, their most popular categories, and

their recent history of productive time usage. Based on reports by RescueTime, I expected very few people to check this dashboard more than once, but the dashboard also provided a plausible cover story for why students should install the software. This makes it a strong control condition for purposes of comparison.

Figure 1: Control-condition tool. Students can access a "dashboard" of their computer usage habits via a drop-down menu. All other conditions also have this feature.

<u>Reminder:</u> Students in the reminder condition received cued reminders of their distracting time usage. Every time a student had accrued 30 minutes of time on distracting websites or programs in a day, a small popup would appear on their computer informing them of this. The clock for the reminders was reset every night at midnight in the time zone of the person using the software.

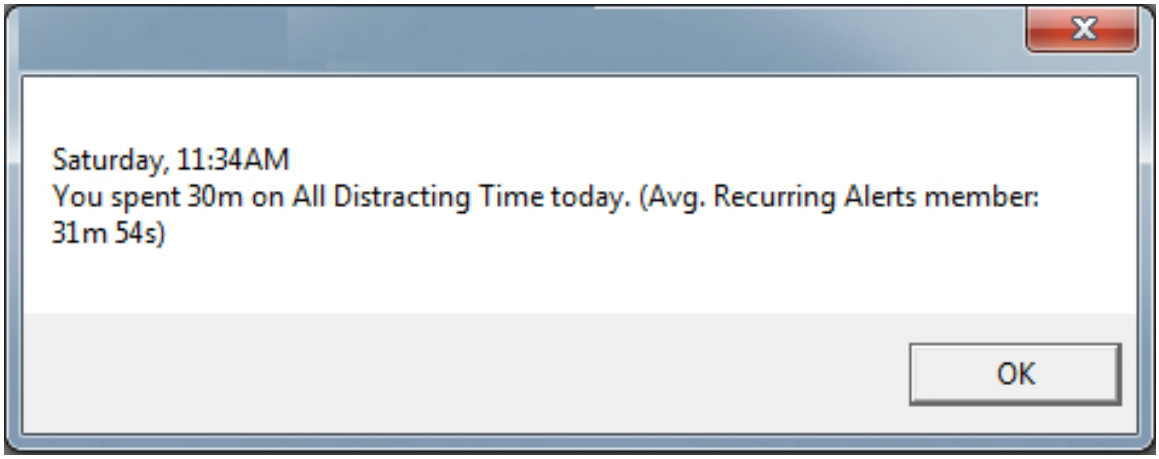

Figure 2: Reminder-condition tool. Students see a popup every time they have accumulated 30 minutes of time on distracting sites in a day.

Commitment device: Students in this condition were sent an email early every morning in their own time zone. This email invited them to visit a website where they could pre-commit to a maximum amount of time that they wanted to spend on distracting websites on the *following* day. (I did not say anything about distracting programs.) At midnight in their time zone, the value on this website would lock in. During the following day, once a student passed their pre-set time limit for distracting time, all distracting websites would be blocked. Students would see a message saying "This website has been blocked", and a button to "temporarily unblock this site". This button took students to a field asking them to provide a reason for unblocking the site. If they typed anything into the field and clicked Enter, they could access the site for the rest of the day.

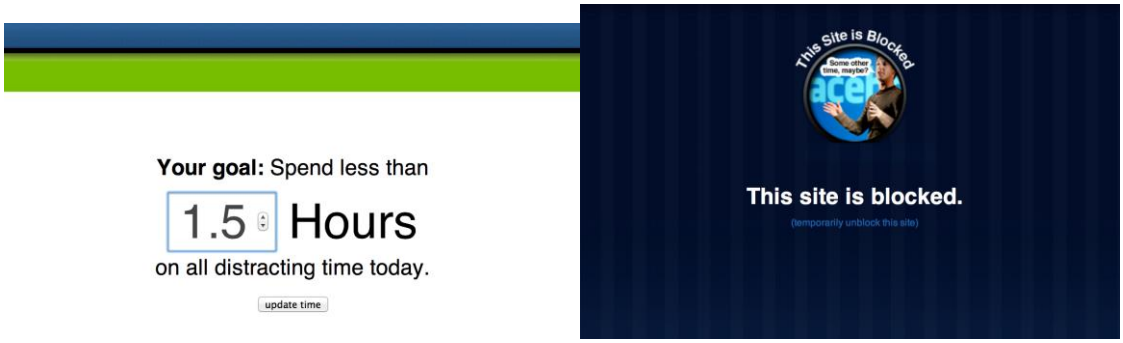

Figure 3: Commitment-device condition tool. Students set their distracting-time limit in the screen on the left, and they see the screen on the right if they pass their limit. (They can temporarily override the block by explaining why they chose to do so.)

Focus-time: Students in this condition received a popup within about 15 seconds of visiting the homepage for the MedStats course. The popup invited them to click on an option in the menu for their software that allowed them to immediately block distracting websites for any number of minutes (30 was the default). Websites were blocked in the same way as in the commitment device condition.

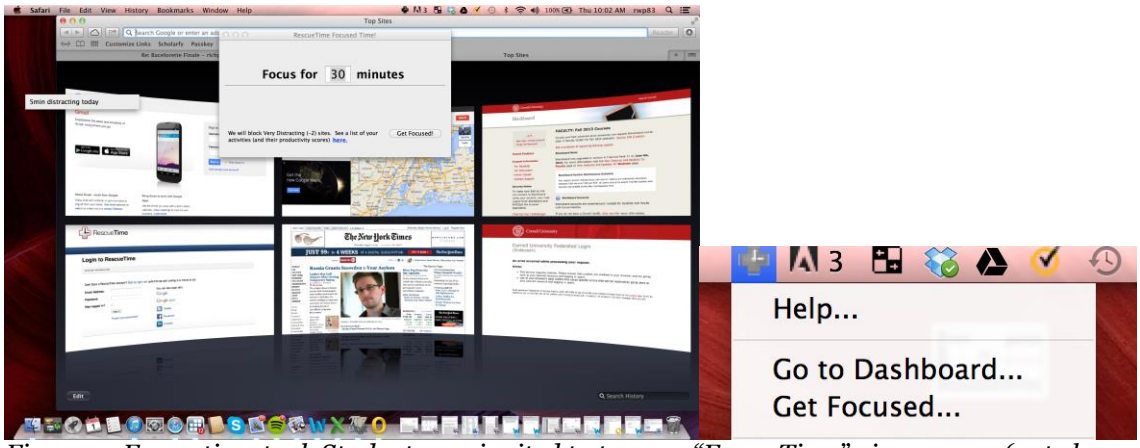

Figure 4: Focus-time tool. Students are invited to turn on "Focus Time" via a popup (not shown), and then they can set the desired time to focus. Distracting sites will be blocked for that time (see Commitment Device tool).

Post-survey questions

Self-control beliefs follow-up: All students were asked the self-control belief questions from the pre-survey a second time to assess pre-post changes.

Accomplishing goals: All students were presented with the goal that they set in the pre-survey, and asked about the degree to which they accomplished their goal (five-point Likert scale from "not at all" to "completely"). I also asked students about the importance of several factors in their success or failure at their course goals – "Forgetting to work on the course", "Procrastinating or delaying going to work on course website", "Difficulty staying on task after starting to work on the course", "Low interest in the course", "Low usefulness of the course", "Course difficulty", and "Work/School/Family responsibilities". Each of these was assessed on a five-point Likert scale from "not at all important" to "extremely important". I also asked the open-ended question "What (if any) other factors limited your ability to achieve your course goals?"

Time usage: In a similar format to the pre-survey questions, I asked students to report their estimated amount of time on the course per week, amount of distracting computer time per day, and whether they spent more or less distracting time than they would like (five-point Likert scale, from "much more distracting time than I would like" to "much less distracting time than I would like"). I also asked exactly how much more or less distracting time they wanted to spend.

User experience: I asked students to use a seven-point Likert scale to agree or disagree with eight statements about the value of the time-management software, such as “The time-management software helped reduce the time I spent on distracting websites.” I also asked students to rate how accurately the software categorized their distracting time (five-point Likert scale, from “not at all accurate” to “extremely accurate”).

I also asked the following open response questions:

- “Did you have any technical issues that made it difficult or impossible to use the software?”
- “How else could the time-management software have worked better for you?”
- “How did using the time-management software impact the way you spent your time?”
- “Any other comments?”

Questions only seen by students in the Reminder condition: As a simple confirmation that the experiment was working, I asked students if they had received at least one popup reminder about the course. I also asked students five questions about the value of the reminders, including “I reduced my unproductive time to avoid getting reminders” and “Seeing the reminder reminded me of my online coursework” (seven-point Likert scales, from “Strongly Disagree” to “Strongly Agree”).

I also asked students to estimate how often the alerts went off, how often the alerts reminded them that they were doing something that they didn’t want to be doing, and how often the alerts actually changed their behavior. Responses were on a 6-point Likert scale: “More than once per day”, “Once every 1-2 days”, “Once every 3-5 days”, “Once every 6-7 days”, Less than once every 7 days”, and “Never”.

Finally, I asked students “Which of these daily reminders about your course would have you liked best?” The available choices were "No reminders", "One reminder if I spend more than 15 minutes on distracting time", "One reminder if I spend more than 30 minutes on distracting time", "One reminder per day regardless of distracting time", and a reminder after every “15”, “30”, or “60” minutes of distracting time.

Questions only seen by students in the Commitment Device condition: As a simple confirmation that the experiment was working, I asked students if they had received at least one email from the software that allowed them to set limits

on their distracting website time. I also asked students six questions about the value of the reminders, including "I thought about my coursework when I thought about setting my distracting-time limit" and "Having a set limit helped me procrastinate less on coursework" (seven-point Likert scales, from "Strongly Disagree" to "Strongly Agree").

I was also interested in how students set their own distracting-time limits. I asked students how aggressive they were in setting their distracting time limit (five-point Likert scale, from "not at all aggressive" to "very aggressive"), how comfortable they were "with the idea of using the distracting-time limit to control yourself" (seven-point Likert scale, from "very uncomfortable" to "very comfortable"), and the open-ended question "How did you decide how to set your distracting-time limit?"

Grade data

I received a spreadsheet from OpenEdX containing the following information for each student enrolled in the course:

- student email address (used as an identifier)
- overall grade (from 0.00 to 1.00)
- grades for homework assignments 1 through 9
- grades for in-video quizzes 1 through 53
- grades for final exam questions 1 through 26

Results

Defining learning outcomes

In many MOOCs, variation in grades is mostly due to large amounts of dropout; typically only about 10% of students enrolled in a course actually receive a certificate of completion (Clow, 2013). Therefore, my primary outcomes of interest were based on both grades and assignment submissions.

A Shapiro-Wilk test for normality revealed that the distributions for both final grade and assignment-count violated assumptions of normality ($p < .0001$ in both cases), making traditional linear regression an inappropriate choice for modeling these outcomes. Instead, I chose to define three binary learning outcomes that reflected meaningful thresholds of course engagement – whether a student submitted *any* assignments, whether a student submitted *all nine* assignments, and whether a student received a certificate of completion by

scoring a final grade above 65%. I model these outcomes using logistic regression in order to avoid making false assumptions about normality.

Participation rates and outcomes

1,285 students finished the survey. Out of this sample, 605 were cut because they did not install the software in the first two weeks of the course, and 37 were cut because they reported goals that did not involve completing any assignments, making their assignment-completion outcomes meaningless. This left a final experimental sample of 643 students. This sample was 59% male, with a mean age of 27.

As shown in Figure 5, the experimental sample clearly outperformed the overall student population on all three outcome variables. This is likely due to the fact that the students in the experimental sample were engaged enough in the course to voluntarily participate in a related study, while the overall student population contained many students who enrolled but never or barely engaged with the course material. Implications for external validity are addressed in the Discussion section.

	Full student population	Study sub-population
Completed any assignments	0.275	0.633
Completed all assignments	0.071	0.213
Received certification	0.114	0.283

Figure 5: Table of outcome variables for subsample and total sample.

Model selection and predictive student characteristics

To select the variables to be used in my final models, I used a forward stepwise model-selection process to iteratively add variables to logistic regression models of each of my three outcome variables (completing any assignments, completing all assignments, and earning a certificate). Intermediate models in this process were evaluated using the Akaike Information Criterion to penalize models containing extraneous variables (Hu, 2007). After each of the models reached convergence, their chosen variables were pooled to make a final set of statistically-significant predictors. These were (all $p < .05$):

- course goal
- likelihood of achieving one's course goal
- importance of one's specific course goal
- importance of finishing the course

- self-reported belief that mental effort is energizing as opposed to draining
- self-reported belief that willpower can be increased with effort
- race
- education level
- gender
- number of previous statistics courses taken
- expected number of hours to commit to the course
- self-reported self-discipline
- interest in the course material
- current amount of time spent on distracting time (minutes)

Treatment effects

There were no significant differences between conditions in any of the categorical predictor variables of course goal, race, education level, and gender listed above (largest $X^2 = 18.103$, $p = 0.643$), indicating that random assignment was successful.

After controlling for all other predictor variables, logistic regression on the outcome of submitting at least one assignment found statistically significant effects of the focus-time condition ($\beta = .520$, $p = .046$) and marginal effects of the commitment-device condition ($\beta = .437$, $p = .090$). Students who were given access to the focus-time and commitment-device tools were approximately 69% and 65% likely to submit at least one assignment, respectively, compared to the reminder and control conditions, which were statistically indistinguishable at 58% and 61% (Figure 6, left panel).

However, logistic regression found no significant effects of any treatments on the outcome of submitting all nine assignments (for focus-time, $\beta = .216$, $p = .482$; for commitment-device, $\beta = .419$, $p = .169$; for reminders, $\beta = .279$, $p = .362$). The same was true for receiving a Certificate of Achievement, with the exception of the commitment device condition, which did have a marginal positive effect (for focus-time, $\beta = .041$, $p = .883$; for commitment-device, $\beta = .500$, $p = .065$; for reminders, $\beta = .104$, $p = .708$). Though 33.5% of commitment-device students received a certification, compared to roughly 26.5% of students in the other conditions, the large degree of unexplained variance in the model made it difficult to distinguish this effect from chance (Figure 6, middle and right panels).

Week-by-week analysis reveals that treatment effects are concentrated in the first few weeks of the course (Figure 7). The commitment-device and focus-time conditions appear to boost student average grade for the first two weeks, but the effects disappear with time, with the commitment device effects persisting for longer. Consistent with this interpretation, logistic regression finds varying

marginal- to significant positive effects of the commitment-device and focus-time conditions on likelihood of submission for the first five assignments, with no effects thereafter (Figure 8).

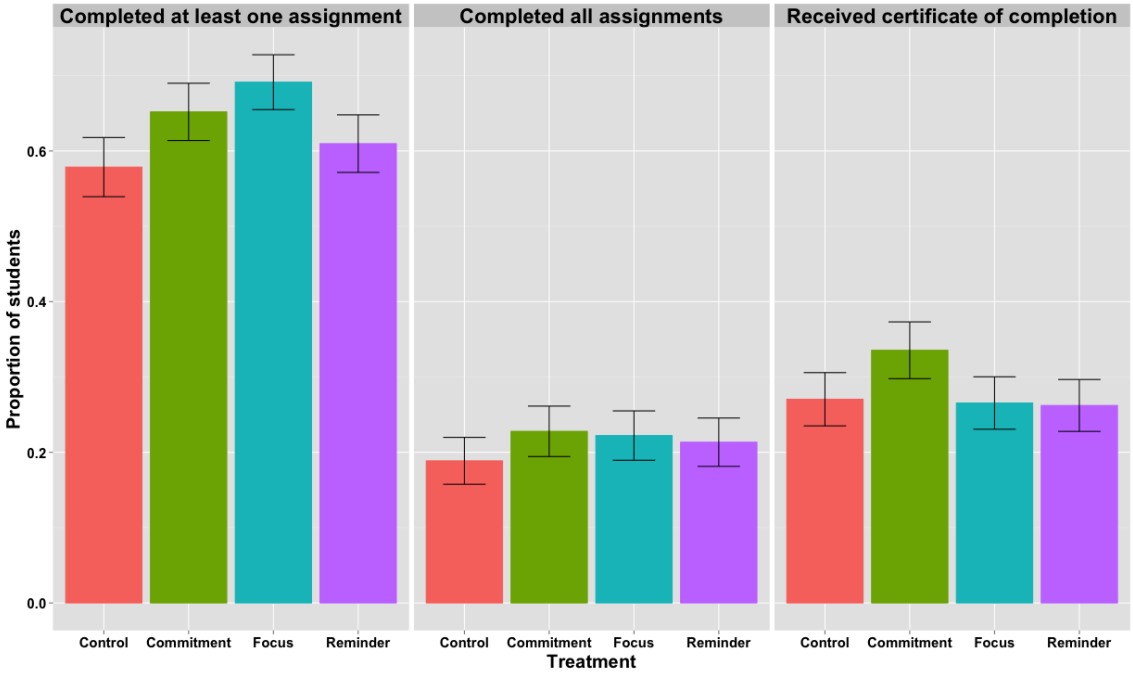

Figure 6: Treatment effects on the proportion of students achieving three learning outcomes.

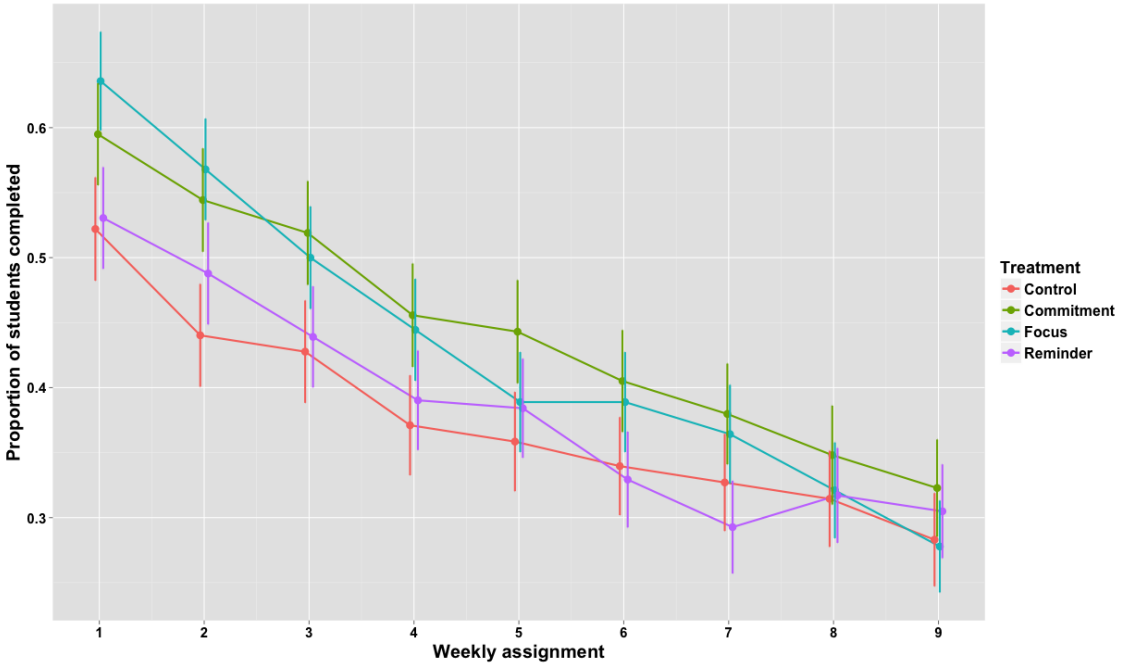

Figure 7: Week-by-week assignment completion rates by treatment. The commitment-device treatment has significant effects in the first half of the course, and the focus-time treatment has significant effects in week 2.

	Focus-time	Commitment-device	Reminders
β(Submitted HW 1)	0.496 *	0.445 '	0.155
β(Submitted HW 2)	0.537 *	0.523 *	0.289
β(Submitted HW 3)	0.302	0.532 *	0.147
β(Submitted HW 4)	0.406	0.549 *	0.238
β(Submitted HW 5)	0.168	0.544 *	0.269
β(Submitted HW 6)	0.238	0.408	0.046
β(Submitted HW 7)	0.212	0.39	-0.049
β(Submitted HW 8)	0.074	0.307	0.179
β(Submitted HW 9)	-0.006	0.331	0.247

Figure 8: Table of logistic regression coefficients (log-odds ratios) for treatments on probability of individual assignment submissions.
*Legend: * = [p < .05] ' = [.05 < p < .1]*

Discussion

Research questions

To recap, my research questions in this study were:

1. Will distracting-time reminders, a commitment device, or a temporary distraction-blocking tool have any positive effect on student learning outcomes?
2. How do students' patterns of web browsing differ across treatment conditions?
3. What do students report about the experience of using the tools?

Question 1 can be answered with "Yes – but only for the first half of the course." In a randomized controlled trial of three self-control tools in an online course, two tools designed to strategically block distracting websites significantly improved students' learning outcomes in the first half of the course. These tools outperformed both a matched control-condition tool and a tool that simply offered regular reminders of distracting website use. However, the effectiveness of the tools decreased over time, causing at best marginal improvements in course-completion. Questions 2 and 3 have relevant data collected, but have not yet been analyzed. In this section I discuss possible explanations for the observed effects, limitations of the current study, questions for continued investigation

with this data set, and questions for investigation with follow-up studies.

Further investigations with this data set

The tools developed for this study were prototypes intended to guide follow-up research, not carefully balanced conditions intended to prove or disprove a precise theoretical claim. For this reason, one should be hesitant about drawing any broad conclusions from this first study, particularly without a fuller examination of the data that are already available.

Qualitative student responses are one key data source that I collected but have not yet analyzed. Qualitative feedback is a vital part of most design-based research because it informs the researcher about whether an intervention was used as expected in a learning environment (Cobb et al., 2003). I also collected moment-to-moment process data from the software tools themselves. These data can also serve as powerful evidence of the ways that students actually use the tools.

These two data sources can enrich basic but unanswered questions about why some tools were more effective than others. For example, I hypothesized that the treatments would affect the degree to which students procrastinate online, which would lead them to have more time to devote to the course. Did students differ by condition in the overall amount of time that they spent on distracting websites and/or on the course website? And do these differences mediate treatment effects on learning outcomes in logistic regression?

An alternate explanation is that students in the more elaborate conditions simply gained a temporary boost in self-efficacy, and were thus more interested in participating in the course without necessarily spending less time on distracting websites. Or it could be the case that the commitment-device and focus-time tools are simply acting as well-timed reminders to participate, and students receive little added benefit from their site-blocking features. These explanations could be tested with a combination of qualitative self-report feedback, data on time spent on different websites, and data on the degree to which students use the features of the commitment-device and focus-time conditions.

The voluntary and customizable nature of the tools in this study also suggests that any treatment effects are likely to be moderated and mediated in complex ways by students' prior dispositions and by their cycles of successful or unsuccessful interaction with the tools. Some students may never try the tools because they are uncomfortable with the idea that they need to restrict their future choices in order to accomplish long-term goals. Others may start out skeptical and then have a positive experience early on that leads them to set stronger and stronger restrictions on themselves. Such recursive cycles are

hypothesized to drive impressive long-term effect sizes in some social-psychological interventions (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). In future work, I intend to explore statistical interaction effects of treatment condition with willpower beliefs. Promisingly, survey questions regarding students' beliefs about willpower as buildable and non-depleting were positive predictors of overall course-completion.

Another question that deserves further investigation is the decreasing of treatment effects over time. Some group of students who would otherwise drop out immediately are being retained for a few extra weeks as a result of the commitment-device and focus-time tools. One class of possible explanations for this diminishing effect involves changes in how students use the tools. For example, students could have been using the tools less and less frequently, setting them to be more and more permissive about allowing distracting time, or finding ways to work around the restrictions placed by the tools. We should be able to see these patterns clearly in the software data if they are present. Another class of explanations involves students becoming habituated to the idea of using a self-control tool, thus gradually losing some novelty effect. We should expect habituation to be independent of how and whether students actually use their tools, and we should also expect some reports on this in the qualitative data.

Limitations and future work

One important limitation of this study is that its sample consists of the subset of students who self-selected to participate in a survey. This subset performed dramatically better than the entire population (see Figure 1), raising potential concerns about the study's generalizability. However, I argue that this study should only be expected to generalize to students who pass a threshold value of commitment to the goal of completing the course. Theories of goal completion distinguish between a pre-commitment "deliberative" phase of weighing costs and benefits, and a post-commitment "pursuit" phase of defending one's effort from distractions and challenges (Oettingen & Gollwitzer, 2010). The tools in this study are intended to support self-control, which primarily plays a role in this second phase (Gollwitzer, 1999). MOOCs typically contain large numbers of students who enroll but never attend class once (Clow, 2013). These students can be reasonably assumed to never have been in the "pursuit" phase of a goal of course completion, so they should not be held as a goal for generalizability.

As the first iteration of a design-based research project, this study has revealed promising results for the commitment device and focus-time interventions, and no results for simple reminders. Without more careful analysis of the existing data, it is difficult to make concrete suggestions about future

iterations of this work. But in general, they should continue to pursue both theoretical and practical goals by introducing more fine-grained variation into the treatment design, collecting self-report data from students throughout the course in order to understand decreasing treatment effects, and looking more carefully at patterns of actual student tool usage to identify characteristic ways that students manage their own future choices.

Acknowledgements
This work is in memory of Amir Lopatin, and would not be possible without the generous donation of the Lopatin family via the Amir Lopatin Fellowship. Thank you for your support. I also thank Rich Patterson from Cornell University and the staff of RescueTime for their essential assistance in the design and implementation of this project.

References

Ainslie, G., & Haslam, N. (1992). Hyperbolic discounting. Retrieved from http://psycnet.apa.org/psycinfo/1993-97149-003

Althaus, S. L., & Kim, Y. M. (2006). Priming effects in complex information environments: Reassessing the impact of news discourse on presidential approval. *Journal of Politics*, *68*(4), 960–976.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, *13*(3), 219–224.

Bambara, C. S., Harbour, C. P., Davies, T. G., & Athey, S. (2009). Delicate Engagement The Lived Experience of Community College Students Enrolled in High-Risk Online Courses. *Community College Review*, *36*(3), 219–238.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*(6), 1014.

Baumeister, R. F., & Tierney, J. (2011). *Willpower: Rediscovering the greatest human strength*. Penguin Books. Retrieved from http://books.google.com.laneproxy.stanford.edu/books?hl=en&lr=&id=HpJ1bH3RLAsC&oi=fnd&pg=PT10&dq=baumeister+willpower&ots=u2crE9FJw3&sig=DFtQmU1VA60h83ntsEZ_5074LhU

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355.

Beck, B. L., Koons, S. R., & Milgrim, D. L. (2000). Correlates and consequences of behavioral procrastination: The effects of academic procrastination, self-consciousness, self-esteem and self-handicapping. *Journal of Social Behavior & Personality*. Retrieved from http://psycnet.apa.org/psycinfo/2002-10572-001

Bernheim, B. D., & Rangel, A. (2004). Addiction and cue-triggered decision processes. *The American Economic Review*, *94*(5), 1558–1590.

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

Bowen, W. G. (2013). *Higher Education in the Digital Age*. Princeton, N.J: Princeton University Press.

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

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation*. Springer. Retrieved from http://link.springer.com/book/10.1007/978-1-4612-5887-2/page/1

Carver, C. S., & Scheier, M. F. (2001). *On the self-regulation of behavior*. Cambridge University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=U9xi8wlfWccC&oi=fnd&pg=PR17&dq=carver+scheier&ots=gNgzbLZ_hu&sig=u077BWVU-7AHhp8mtyFqZhBHOz0

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disruptive innovation will change the way the world learns* (Vol. 98). McGraw-Hill New York. Retrieved from http://www.mcgraw-hill.com.au/professional/booksellers/resources/2008/June/PDF_FORMAT_BUSINESS.pdf

Clow, D. (2013). MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 185–189). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2460332

Cobb, P., Confrey, J., Lehrer, R., Schauble, L., & others. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, *324*(5925), 400–403.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15–42.

Dal Cin, S., MacDonald, T. K., Fong, G. T., Zanna, M. P., & Elton-Marshall, T. E. (2006). Remembering the message: the use of a reminder cue to increase condom use following a safer sex intervention. *Health Psychology*, *25*(3), 438.

Daniel, J. (2012). Dual-mode universities in higher education: way station or final destination? *Open Learning*, *27*(1), 89–95. http://doi.org/10.1080/02680513.2012.640791

Day, V., Mensink, D., & O'Sullivan, M. (2000). Patterns of academic procrastination. *Journal of College Reading and Learning*, *30*(2), 120–134.

Duckworth, A. L. (2011). The significance of self-control. *Proceedings of the National Academy of Sciences*, *108*(7), 2639–2640. http://doi.org/10.1073/pnas.1019725108

Duckworth, A. L., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. M. (2011). Self-regulation strategies improve self-discipline in adolescents: benefits

of mental contrasting and implementation intentions. *Educational Psychology*, *31*(1), 17–26.

Ehrman, M. (1990). Psychology: Psychological factors and distance education. *American Journal of Distance Education*, *4*(1), 10–24.

Eisenberg, E., & Dowsett, T. (1990). Student drop-out from a distance education project course: A new method of analysis. *Distance Education*, *11*(2), 231–253.

Evans, G. W., & Rosenbaum, J. (2008). Self-regulation and the income-achievement gap. *Early Childhood Research Quarterly*, *23*(4), 504–514.

Fagles, R., & Knox, B. (1997). *The Odyssey*. Penguin. Retrieved from http://books.google.com/books?hl=en&lr=&id=biBRxQmHPLIC&oi=fnd&pg=PA1&dq=the+odyssey+homer+fagles&ots=TEap8qHi6X&sig=DXeWwsa9wr5tIyjI7lus9OaMdAQ

Fowler, G. A., & Ovide, S. (2013, April 23). Sit Up Straight, and Other Advice From Big Mother. *Wall Street Journal*. Retrieved from http://online.wsj.com/news/articles/SB10001424127887324763404578432892976469664?mg=reno64-wsj&url=http%3A%2F%2Fonline.wsj.com%2Farticle%2FSB10001424127887324763404578432892976469664.html

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plans. *American Psychologist*, *54*(7), 493.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, *38*, 69–119.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). *The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w12006

Hoch, S. J., & Loewenstein, G. F. (1991). Time-inconsistent preferences and consumer self-control. *Journal of Consumer Research*, 492–507.

Hu, S. (2007). Akaike information criterion. *Center for Research in Scientific Computation*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.353.4237&rep=rep1&type=pdf

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410–1412.

Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion—Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science*, *21*(11), 1686–1693.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478.

Levitin, D. (2015). *The organized mind: Thinking straight in the age of information overload*. Penguin UK.

Lewis, R. (2001). Classroom discipline and student responsibility:: The students' view. *Teaching and Teacher Education*, *17*(3), 307–319.

Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science*, *244*(4907), 933–938.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Ross, S. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698.

Norman, D. A. (1988). *The psychology of everyday things*. Basic books. Retrieved from http://books.google.com/books?hl=en&lr=&id=TI5wnq1PimMC&oi=fnd&pg=PR7&dq=donald+norman&ots=xoahk8hJMl&sig=y6kLnhp8_3qcdQUJAwm_rnYMnuw

Oettingen, G., & Gollwitzer, P. M. (2010). *Strategies of setting and implementing goals*. Social Psychological Foundations of Clinical Psychology. New York: The Guilford Press. Retrieved from http://kops.ub.uni-konstanz.de/bitstream/handle/urn:nbn:de:bsz:352-opus-131945/oettingen_gollwitzer_strategies.pdf?sequence=1

Onwuegbuzie, A. J. (2000). Academic procrastinators and perfectionistic tendencies among graduate students. *Journal of Social Behavior & Personality*. Retrieved from http://doi.apa.org/psycinfo/2002-10572-008

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*(37), 15583–15587.

Reimann, P. (2011). Design-based research. In *Methodological Choice and Design* (pp. 37–50). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-90-481-8933-5_3

Ross, B. H. (2006). *The Psychology of Learning and Motivation: Advances in Research and Theory*. Academic Press.

Schneider, E., & Kizilcec, R. F. (2014). Why did you enroll in this course?: developing a standardized survey question for reasons to enroll. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 147–148). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2567852

Tang, Y.-Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., … others. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences*, *104*(43), 17152–17156.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271–324.

Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*(S1), S164–S187.

Vander Ark, T. (2011). *Getting smart: How digital learning is changing the world*. John Wiley & Sons. Retrieved from http://books.google.com/books?hl=en&lr=&id=jKjViK5jgq0C&oi=fnd&pg=PT11&dq=vander+ark+getting+smart&ots=zDxvwWS5cN&sig=caXxvNIvy3bGAEczbkXSS5padlE

Yeager, D., & Walton, G. (2011). Social-Psychological Interventions in Education They're Not Magic. *Review of Educational Research*, *81*(2), 267–301.

Yeykelis, L., Cummings, J. J., & Reeves, B. (2014). Multitasking on a single device: Arousal and the frequency, anticipation, and prediction of switching between media content on a computer. *Journal of Communication*, *64*(1), 167–192.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70.