Exercise Improves Academic Performance

Alexander W. Cappelen, Gary Charness, Mathias Ekström, Uri Gneezy, Bertil Tungodden

Exercise Improves Academic Performance

Alexander W. Cappelen, Gary Charness, Mathias Ekström, Uri Gneezy, Bertil Tungodden

August 24, 2017

Abstract: We report the results of a randomized controlled trial testing whether incentivizing physical exercise improves the academic performance of college students. As expected, the intervention increases physical activity. The main result is that it generates a strong and significant improvement in academic performance, particularly for students who struggled at the baseline in terms of lifestyle habits. We also provide evidence on the underlying mechanisms: Students who were incentivized to exercise have a healthier life style and improved self-control. Overall, the study demonstrates that incentivizing students to exercise can be an important tool in improving educational achievements.

1. Introduction

Is regular physical exercise important for educational achievements? This question relates to key debates in educational policy. In the United States, for example, there is a growing concern that physical education is no longer considered a critical element of schooling: In some areas, half of the students report having no physical education in an average week (New York Times, 2017). The lack of physical activity may contribute to bad health and obesity—but, importantly, it may also worsen students' academic performance (U.S. Department of Health and Human Services, 2010).

Physical activity may improve academic performance through enhanced psychological health, habit formation, and changes in the brain, supporting cognitive function and learning (Biddle and Asare, 2011; Booth et al., 2012; Reiner et al. 2013; Khan and Hillman, 2014; Mura et al., 2015). At the same time, it may also worsen academic performance by directing time and attention away from studying. While there is research suggesting a positive association between physical activity and academic performance, the evidence is mixed and, even more importantly, we know very little about the direction of the association (see Aaltonen et al., 2016, for a recent review). In particular, since this research is not experimental, we cannot impute causality; for example, one's level of self-control could affect *both* the ability to study and the tendency to exercise (Grissom, 2005).

In this paper, we report the results of a randomized-controlled trial that incentivized university students to go to the gym and then measured the effect of the intervention on their academic performance relative to an unincentivized control group. To the best of our knowledge, the present study is the first pre-specified randomized control trial investigating the causal effect of physical activity on academic performance in higher education. The study combines complete

and precise administrative data on both gym attendance and academic performance with subjective survey data (baseline and follow-up) on happiness, self-control and lifestyle habits. The administrative data on academic performance show zero attrition in the main variable of interest, and the study therefore offers a clean identification of the causal effect of incentivizing students to go the gym on academic performance. The administrative data on gym attendance and the detailed survey data (with less than 5% attrition) further allow us to provide evidence on the underlying mechanisms of the relationship between exercise and academic performance. Our analysis builds on a pre-analysis plan registered before the researchers had access to the data.[1]

As expected from previous work (beginning with Charness and Gneezy, 2009), the gym attendance data show that the treated students visited the gym more frequently than the students in the control group: 67% of the participants who had been incentivized attended during the period of our study, compared to 12% of the participants in the control group. Strikingly, we find that the incentives to exercise not only increased gym attendance, it also significantly increased academic performance (by, on average, about 0.15 standard deviations).

The study also provides evidence that sheds light on the underlying mechanisms driving the positive effect of exercise on academic performance. First, as hypothesized in the pre-analysis plan, we show that students who, before the intervention, struggled in terms of lifestyle habits (comprised of responses to three questions about sleep, tiredness, and health), study hours, happiness, and self-control drive the overall treatment effect. The intervention increases academic performance by more than 0.3 standard deviations for students who self-reported being

---

[1] The pre-analysis plan is available at the AEA RCT registry (number: AEARCTR-0001949). Prior to registration of the pre-analysis plan, the data was stored in a locked drawer at the Norwegian School of Economics, which none of the researchers had access to.

below median at the baseline on these four dimensions, while there is no treatment effect on academic performance for students who do not struggle on any of these dimensions. For the group that struggled at the baseline, we observe an increase both in the numbers of courses completed and grade average. Second, we show that the treatment had a positive effect on self-reported life style and self-control in the follow-up survey. Overall, these findings suggest that the increase in exercise induced a change in lifestyle that contributed to improved academic performance. Finally, we complement the main analysis with instrumental variable estimates, where assignment to treatment is used as an instrument for the level of exercise.

Our paper contributes to several different literatures. First, it offers causal evidence to the literature studying the relationship between physical education and academic achievements (see Rasberry et al., 2011, and Singh et al., 2012, for recent reviews), which so far has suffered from methodological challenges, most notably lacking tests showing causality. On the same note, we extend the small-but-growing non-experimental literature in economics that reports positive effects of participation in sports activities in adolescence on educational achievement (Lipscomb, 2007; Stevenson, 2010; Pfeifer and Cornelissen, 2010). Our experimental evidence on the impact of gym exercise substantiates these findings further and suggests that physical activity is important in driving the effects, as opposed to being entirely driven by the social environment associated with sport participation (see, Di Bartolomeo and Papa, 2017). Our results also relate to the economic literature connecting exercise to self-control (DellaVigna and Malmendier, 2006), as the effect on academic performance in the present study relates to improved self-control.

Turning to the experimental literature on behavioral interventions, there is strong evidence that financial incentives can affect behavior at least while these incentives are present

(e.g., Charness and Gneezy, 2009; Volpp et al., 2011; Acland and Levy, 2015; Babcock et al., 2015; Royer et al., 2015; Carrera et al., 2017). However, there has been less success in using financial incentives to increase academic performance (e.g., Fryer, 2011; Angrist et al., 2009; Levitt et al., 2016).[2] The present study indicates that targeting physical exercise is an effective approach to improving educational outcomes. We also show that the rate of return of behavioral interventions may be greater than a narrow perspective would indicate, in the sense that interventions not only affect the investments in the incentivized good (exercise in our case), but they also produce immediate spillover effects in other type of investment goods (education in our case).

Similar evidence on positive spillovers from exercise to academic performance is supported by two independent studies. In a follow up on Charness and Gneezy (2009), the authors sent the participating students in the original study an email two years after the intervention. In the email, students were asked about their exercising behavior as well as their academic performance. The results showed no impact on exercising, but a large effect on academic satisfaction as well as GPA (Charness and Gneezy, 2011). However, this study suffers from several shortcomings, most notably a low response rate (around 50%) and the use of a self-reported measure of academic performance. In a recent study, Fricke et al. (2017) incentivized university students to attend the student gym, and subsequently followed up on their academic performance. In contrast with us, they targeted students that already exercised regularly, and as a result have a smaller effect on exercise. They also report a positive effect of exercising on academic performance, even though the observed spillover effect appears to be less robust than

---

[2] See Gneezy et al., 2011, for a survey. Scott-Clayton (2011) and Garibaldi et al. (2012) report more positive results in the educational context using quasi-experimental approaches.

4

in the present study. They suggest that much of the spillover effect is due to exercise taking place close in time to studying, thereby enhancing the memory of students, but this must be considered explorative evidence since it relies on a follow-up survey with only a 50% response rate.

Finally, our paper contributes to the large literature on human capital formation (see, for example, Cunha et al., 2006), by providing evidence on the importance of non-cognitive skills, such as self-control, for educational outcomes. Since exercising seems to improve these non-cognitive skills, our results speak to the role of physical activity in a society's educational strategy. Our results suggest that physical activity may be a particularly important tool for targeting vulnerable groups, which implies that it may be an important dimension in policies aiming at reducing the socio-economic gradient of human-capital formation.

The paper proceeds as follows: Section 2 describes the experiment in more detail, Section 3 outlines the hypotheses and the empirical strategy, Section 4 presents the results, and Section 5 concludes.

2. The Experiment

We recruit university students and randomly assign them to a treatment or control group, where the treatment reduces the barrier to start exercising by providing (at minimum) free membership at the student gym SIB. We elaborate on this simple design below.

2.1 Recruitment

We recruited students at the university campuses of the University of Bergen (UIB) and the Bergen City College (HIB). We specifically target undergraduate students who are less likely to have a regular exercise regime, by inviting only students who did not already have a gym card.

Students who were interested in participating received a booklet containing an information sheet, a consent form, and a baseline survey. Each participant was required to sign the consent form and each person returning the booklet was paid 100 NOK (equivalent to $12.50 at the time of the experiment). On the information sheet we informed participants that everyone who fulfilled the stated requirements would have a 40% probability of winning a free SIB-card for the spring term 2016 (regular price: 1100 NOK), potentially topped with a 1,000 NOK bonus for attendance, or two hours with a personal trainer. Participants were also informed that we would conduct follow-up surveys via email.

Table 1: Timeline

Date	Event
January 11, 2016	Start of Spring semester
February 1, 2016	Deadline for course registration
February 1-8, 2016	**Recruitment of participants and baseline survey**
February 10-12, 2016	Screen out pre-existing SIB-members and randomize to treatment
February 15, 2016	Inform participants whether they won SIB-card or not
February 16, 2016	**Possible to pick-up SIB-card and begin exercise**
March 21-March 28, 2016	Easter Break
May 2-June 17, 2016	Exam period
May 9, 2016	**Follow-up survey sent via email and SMS to participants**
June 17, 2016	Academic semester ends
June 30, 2016	SIB-card expires

Note: Entries in bold are key events during the Experiment.

2.2 Timeline

Table 1 gives an overview of important dates and events in the study. Recruitment was done in the first week of February 2016. This choice had two advantages. First, it rules out most students who would have bought a gym card in the absence of our intervention, since the vast majority of those who buy a gym card do so early in the semester. Second, the deadline for course registration was February 1, which implies that participation (and potentially treatment

assignment) cannot influence the courses for which people chose to register.

The students who were assigned to a treatment could collect the gym card from February 16. Given that the academic semester ended June 17, the relevant exercise period thus encompasses roughly four months. The academic semester is divided in two parts: the teaching period (January-April) and the exam period (May-June). The follow-up survey was distributed to participants during the exam period in the middle of May.

2.3 The sample

In total, 823 students signed the consent form and returned the booklet. From this group we excluded 32 students who already had a SIB-card (as shown in the administrative records of SIB). Six students had filled in the booklet twice; in those cases we only included the first booklet turned in (i.e., we dropped six records). Another two students were registered at other universities than HIB and UIB (for which we did not have administrative data on academic performance), and one student withdrew from the study after handing in the booklet. After these exclusions, we were left with 782 participants who were included in the study.

We randomly assigned 400 participants out of the 782 to the treatment group. All participants in the treatment group received an email and a text message on February 15, in which we informed them that they had been randomly selected to receive a gym card that they could pick up at one of SIB's gyms. After the randomization had taken place, one participant (from the control group) withdrew from the study and three other participants (one control and two treated) turned out not to be registered as students at neither HIB nor UIB. We therefore base the analysis on the remaining 778 participants.

Table 2 presents mean values of the key background characteristics, measured at the baseline, for the overall sample (column 1), and separately for the control and treatment group (column 2 and 3). We also report *p*-values from a *t*-test of the difference in means between these groups.

Participants were on average 22 years old and were in their second year at the university; roughly 50 percent of the sample is female, and roughly 50 percent study at each of HIB and UIB. Lifestyle index, Study hours, Happiness and Self-control index are pre-specified self-reported background variables that we will use in heterogeneity analysis. Lifestyle index is the combination of the responses to three questions in the baseline survey, where each response has been standardized to have a zero mean and a standard deviation of one. The three questions are: 1) "How many days last week did you go to bed after midnight?" 2) "How many days last week did you feel tired/unrested?" 3) "How satisfied are you currently with your health?" (1-10; 1=Very unsatisfied, 10=Very satisfied). The index is the sum of the responses to these three questions with the values of questions 1) and 2) rescaled such that a higher number means going to bed earlier, feeling less tired, and being more satisfied with your current health; hence, a higher value on the index corresponds to what we define to be a proxy for a better lifestyle.

Study Hours is the self-reported number of hours studied, including attended lectures, the week prior to the baseline survey. We note that the average participant reported having spent 23 hours studying that week. Happiness is the answer to the following question: "How satisfied are you currently with your life in general?" (1-10; 1=Very unsatisfied, 10=Very satisfied). The average response is seven, which implies that the students on average appear to be somewhat satisfied with their life in general.

Table 2: Balance Test

	Overall	Control	Treatment	p-value (T-C)
Age	22.258	22.316	22.204	0.640
Female (1/0)	0.533	0.550	0.518	0.366
Year of study	2.008	2.034	1.982	0.561
HIB (1/0)	0.481	0.500	0.462	0.294
Lifestyle index	0.004	-0.025	0.032	0.703
Study hours	22.599	22.884	22.330	0.475
Happiness (1-10)	7.340	7.287	7.391	0.412
Self-control index (1-10)	5.814	5.751	5.875	0.249
N	778	380	398	

Notes: Age and Year of Study is in years. HIB = 1 if student at Bergen City College (0 if student at University of Bergen). Lifestyle index is a standardized measure of people's sleeping habits and satisfaction with health; a higher number means better lifestyle. Study hours is total hours spent studying (including lectures) the week prior to the baseline survey). Happiness is the answer to a question about general life satisfaction; a higher number means higher life satisfaction. Self-control index is comprised of four questions; a higher number means more self-control.

The Self-control index is constructed based on four questions: 1) "I am a person who often acts too hastily." 2) "I have difficulties resisting temptations." 3) "I have a tendency to procrastinate on things, even though it would be best to take care of them quickly." 4) "I am a person who follows my plans." Each of the four questions was answered on a scale 1-10, expressing the extent to which the participant agreed with the statement (1=completely disagree, 10=completely agree). The index is the rescaled average of these four questions, with a higher index corresponding to having more self-control. The last column of Table 2 shows that, as expected, all background characteristics balance between treatment and control groups.[3]

---

[3] See Table A1 in the appendix for an extended version of the balance table, also including variables that were not part of the pre-analysis plan. In particularly, we also show that academic performance in the fall semester prior to the intervention balance across treatments.

2.4 The treatment group

All students allocated to the treatment group received a free gym card. In addition, we randomly offered some of the individuals in the treatment group two hours with a personal trainer and some a bonus of 1,000 NOK if they fulfilled a specified attendance requirement.[4] The treatment group thus consisted of three sub-treatments: 100 participants received a free SIB-card (sub-treatment Card), 100 participants received a free SIB-card and were offered two hours with a personal trainer (sub-treatment PT), and 200 participants received a free SIB-card and a 1,000 NOK bonus if they fulfilled our specified attendance requirements (sub-treatment Bonus).

As specified in the pre-analysis plan, our main interest is in the effect of being allocated to the treatment group. We did not expect to have sufficient statistical power for a comparison of the sub-treatments, and therefore defer this analysis to the Appendix. Overall, as shown in Tables A6 and A7 in the Appendix, we find the same patterns for all three sub-treatments for the effects on both exercise and academic performance, with some differences to which we return in the analysis.

2.5 Data sources

We use four data sources in the analysis. First, from the baseline survey we have the background characteristics of the participants, including gender, age, and year of study.[5] In addition, to further explore the underlying mechanisms, we included a battery of questions related to lifestyle habits, hours spent studying, life satisfaction, and self-control.

---

[4] The requirement was that they would get a bonus of 1,000 NOK if they visited the student gym at least two days per week in at least seven out of eleven weeks between February 16 and May 1, 2016.
[5] The complete baseline survey and follow-up survey are provided in Appendix B.

Second, from SIB we have scanner data on each occasion a participant swiped their SIB-card at one of their gyms, which gives us the total number of gym visits during the semester. Third, from UIB and HIB we have the complete administrative data related to the academic performance of each student. This includes records for which exams a student registered, whether it is an ordinary exam or a re-take exam, the number of credits involved, and the grade received on each exam.

Fourth, from the follow-up survey we have self-reported data on relevant outcomes that are not captured in the administrative data. In particular, we included (as in the baseline) questions related to lifestyle habits, hours spent studying, life satisfaction, and self-control in order to understand whether treatment affected people along these dimensions.

3. Hypotheses and Empirical Strategy

We here outline our main hypotheses and the regression model specifications we employ in the analysis.

3.1 Gym attendance

Based on past research and in line with our pre-analysis plan, we expect that our treatment will increase gym attendance.

Hypothesis 1: Incentivizing physical exercise has a positive causal effect on gym attendance.

We test Hypothesis 1 by estimating the following OLS regression equation:

$$G_i = \alpha + \beta Treated_i + \delta \mathbf{X}_i + \varepsilon_i \qquad\qquad (1)$$

where G_i is the number of gym visits of individual i in the spring semester 2016, or a dummy variable equal to one if the student visited the gym at least once. *Treated* is an indicator taking the value one if individual i were assigned to the treatment group and the value zero if assigned to the control group, and **X** is a vector of pre-specified covariates included to increase precision (age, gender, institution, year of study, and dummies for being below the median in terms of the lifestyle index, study hours studied, happiness, or the self-control index). Since Hypothesis 1 is one-sided, the formal test is equivalent to:

$$H0: \beta \leq 0$$
$$H1: \beta > 0$$

3.2 Academic performance

Our main hypothesis, as stated in the pre-analysis plan, is that exercise improves academic performance, which means that our alternative hypothesis can be stated as follows:

Hypothesis 2: Incentivizing physical exercise has a positive causal effect on academic performance.

We test Hypothesis 2 by estimating the following OLS regression equation:

$$Y_i = \alpha + \beta Treated_i + \delta \mathbf{X}_i + \varepsilon_i \qquad (2)$$

where Y_i is a measure of academic performance. Since Hypothesis 2 also is one-sided, the formal test is equivalent to:

$$H0: \beta \leq 0$$
$$H1: \beta > 0$$

As stated in the pre-analysis plan, we consider the extent to which the treatment has

affected the two main dimensions of academic performance: the number of study points completed and grade average. The variable Completed Study Points is defined as the total number of study points a student *completed* in a given semester; the variable Grade Average is defined as the average grade per semester (weighted by study points), conditional on passing a course and excluding courses with a Pass/Fail system. In the calculation of the Grade Average we use a numeric scale (0-5) where A equals 5 and F equals 0.[6] To have an overall measure of academic performance, we introduce the variable Total Grade Points, which combines these two components multiplicatively.

We are also interested in whether the intervention has a stronger impact on academic performance for specific sub-groups. In particular, we pre-specified the hypothesis that people who score low on the lifestyle index, number of hours studied, general life satisfaction and degree of self-control would experience a larger treatment effect than those who score high on these measures.

Hypothesis 3: Incentivizing physical exercise improves academic performance more for people who score low on the lifestyle index, study hours, happiness or the self-control index at the baseline, than for those who score high on these dimensions.

We test Hypothesis 3 by estimating the following OLS regression:

---

[6] The majority of students take four courses per semester, each worth 7.5 study points, and most courses are graded on a scale A-F where A is best and F is worst. However, study points may differ between courses, and some courses are only graded Pass/Fail. It is important to note that our intervention could not influence which courses a student took in the spring of 2016, since the allocation to treatment only was revealed in mid-February, after the deadline for course registration.

$$Y_i = \alpha + \beta_1 Treated_i + \beta_2 Below\ median\ LIFESTYLE * Treated_i$$

$$+ \beta_3 Below\ median\ STUDYHOURS * Treated_i$$

$$+ \beta_4 Below\ median\ HAPPINESS * Treated_i$$

$$+ \beta_5 Below\ median\ SELFCONTROL * Treated_i + \delta\mathbf{X}_i + \varepsilon_i \qquad (3)$$

In (3), we interact the treatment dummy with dummy variables for being below the median in terms of the lifestyle index (LIFESTYLE), the number of hours studied (STUDYHOURS), general life satisfaction (HAPPINESS), and the self-control index (SELFCONTROL), where the subgroup dummy variables also appear independently in the vector $\mathbf{X}$. The estimated coefficients $\beta_2, \beta_3, \beta_4, \beta_5$ provide tests of whether the intervention worked differently for those who were below or above the median in the respective dimension at the baseline, while the estimates of β_1 and $(\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5)$ provide tests of how the intervention affected those who were above and below, respectively, the median at the baseline in all dimensions. In the analysis, we also consider a specification where we include each subgroup separately.

Finally, we report results from an instrumental variable analysis using two stage least squares (2SLS), which relies on the strong exclusion restriction that the effect of the intervention only works through the number of gym visits. The first stage corresponds directly to equation (1), where G_i is the number of gym visits. The second stage regresses academic performance (Y_i) on the predicted values from the first stage (G_i^*). Formally, the second stage is equivalent to:

$$Y_i = \alpha + \beta G_i^* + \delta\mathbf{X}_i + \varepsilon_i \qquad (4)$$

The estimated coefficient of β in (4) provides us with an estimate of the causal effect of increasing the number of gym visits on academic performance.

3.3 Other outcome variables

To shed further light on underlying mechanisms that may cause an increase in exercise to have spillovers to academic performance, we also study how the intervention affects dimensions that we hypothesize are important for succeeding as a student.

Hypothesis 4: Incentivizing physical exercise has a positive effect on lifestyle, study hours, happiness, and self-control.

We test Hypothesis 4 by estimating the following OLS regression equation:

$$Y_i = \alpha + \beta_1 Treated_i + \delta \mathbf{X}_i + \varepsilon_i \qquad (5)$$

In (5), Y_i represents standardized values of either: the lifestyle index, study hours, happiness, or the self-control index.

4. Results

We provide the main results in this section and relegate the supporting analysis to the appendix.

4.1 Effect on Exercise

We start out by studying whether the intervention, which targeted physical exercise, induced the students to attend the gym more frequently. Figure 1 provides a histogram showing the distribution of gym visits by treatment. We observe that there is a huge difference in gym

attendance between the treatment and control group: the great majority of students in the control group did not attend the gym at all, while the preponderance in the treatment group did exercise.[7]

Figure 1: Gym Visits by Treatment
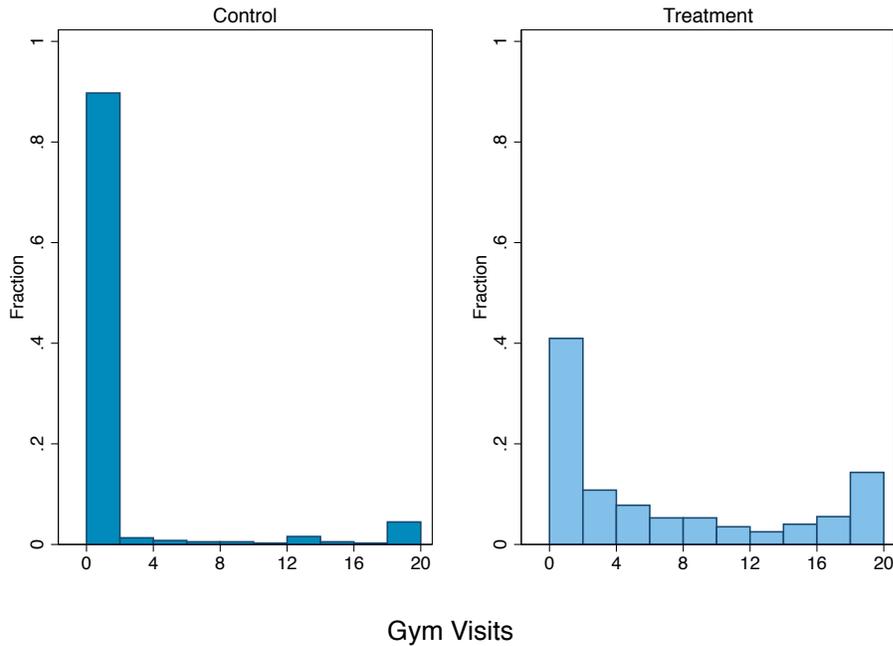
Notes: The figure provides the distribution of gym visits by treatment arm, censored at 20 gym visits.

Table 3 reports the effect of the treatment intervention on gym exercise using OLS regression analysis. Columns 1-2 report the effect on the (standardized) number of gym visits for the whole semester, where we find an average treatment effect of more than 0.6 standard deviations.[8] On average, treated students attend the gym 7.5 times in the semester, which is a more than 300 percent increase relative to the 1.8 visits in the control group.[9] We observe from

---

[7] Not surprisingly, as shown in Table A6, the treatment with the strongest effect on physical exercise is the sub-treatments with the strongest incentives (PT and Bonus).

[8] Throughout the analysis, we standardize our outcome measures by dividing the respective variable with its standard deviation, without subtracting the mean.

[9] Figure A1 in the appendix provides a disaggregated overview of the treatment effect on gym visits.

column 2 that the effect is robust to adding control variables. In columns 3-4, we report the effect of the intervention on the extensive margin, where we observe a highly significant increase in the share of students at least exercising once, from 11.6 percent in the control group to 68.6 percent in the treatment group.

Table 3: Regression Results on Gym Attendance

	Outcome Variable			
	Visits (std)	Visits (std)	Visited gym (1/0)	Visited gym (1/0)
Treated	0.629***	0.627***	0.570***	0.571***
	(0.0676)	(0.0678)	(0.0285)	(0.0288)
Age		-0.0145		-0.00438
		(0.00899)		(0.00436)
Female		-0.0176		0.0644**
		(0.0700)		(0.0297)
Year of study		-0.0607***		-0.00429
		(0.0214)		(0.0123)
HIB		0.0527		-0.0213
		(0.0698)		(0.0302)
Low_Lifestyle		0.123*		0.0412
		(0.0660)		(0.0300)
Low_Study		0.0135		-0.0171
		(0.0712)		(0.0303)
Low_Happiness		-0.0365		-0.00421
		(0.0791)		(0.0348)
Low_Selfcontrol		-0.0473		-0.00785
		(0.0686)		(0.0298)
Constant	0.199***	0.609***	0.116***	0.190*
	(0.0390)	(0.231)	(0.0164)	(0.102)
N	778	758	778	758

Notes: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2 Effect on Academic Performance

We now turn to the question of whether the intervention that targeted exercise also had an impact on the participants' academic performance. Our main focus is on overall academic performance, which takes into account both the number of completed courses and the grade in each course. We measure overall academic performance by total grade points (TGP), which is the product of the number of credits (study points) and the grade achieved in the course, summed over all courses in a semester. We supplement the analysis of TGP by also separately considering its two components: completed study points (CSP), which is the sum of course credits over all completed courses, and the grade average conditional on passing a course (GA).[10]

Figure 2 provides the distribution of TGP across conditions: 0 implies completing no courses and 150 implies completing courses worth 30 credits (which is equivalent of full-time study in Norway) and getting an average grade of A.[11] We observe that the share of students with less than 30 TGP is higher in the control group, while the share of students with more than 120 TGP is higher in the treatment group.

---

[10] Grades are given on a scale A-F, where A is the best grade and F means fail. We transform the grades on a 0-5 scale, where F=0 and A=5. The analysis focuses on ordinary exams and graded courses, but, as we show in Table A8 in the appendix, there are no significant differences between the treatment group and control group in retake exams and pass/fail courses.

[11] TGP is censored at 150 in Figure 2. It is possible to take more credits than 30 per semester, which a small fraction of students does, and thereby have a TGP above 150. In all tests, we use uncensored data.

Figure 2: Distribution of Total Grade Points (TGP) by Condition

Figure 3 summarizes the first set of main results of the intervention.[12] It shows the average TGP, CSP, and GA for both the control group and the treatment group, where outcomes are standardized so that any treatment effect can be interpreted in terms of standard deviations. As can be seen, there is a difference in TGP of about 0.15 standard deviations, which is almost entirely driven by a similar difference in CSP, while the grade average is unaffected.

---

[12] As shown in Table A7, we find very similar patterns in all sub-treatments, with no statistically significant differences across sub-treatments.

Figure 3: Differences in Outcomes Across Conditions
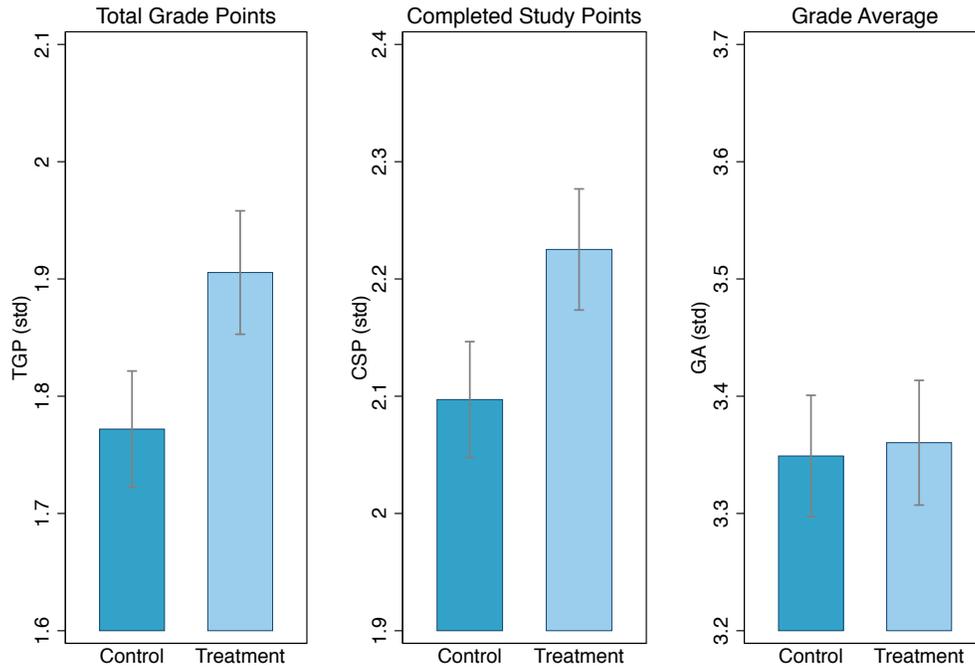
Notes: The figure shows the treatment effect on overall academic performance (TGP), and its two components (CSP and GA). Error bars indicate +/- one standard error.

Table 4 substantiates the graphical evidence by presenting the results from OLS regressions of the intention-to-treat effect on academic performance. The outcomes are again standardized so that the treatment effect can be interpreted in terms of standard deviations. In column 1, we regress TGP and only control for the treatment indicator variable (full sample). The coefficient is positive, economically relevant, and statistically significant. Simply being encouraged to exercise increases total grade points by 0.13 standard deviations. Column 2 expands the specification by adding the pre-specified control variables. The treatment effect is robust to the addition of control variables, and if anything increases in size, to 0.15 standard deviations ($p = 0.038$ in two-sided test).

Table 4: Regression Results on Academic Performance

| | Outcome Variable (std) | | | | | |
	TGP	TGP	CSP	CSP	GA	GA
Treated	0.134*	0.148**	0.128*	0.139*	0.0113	0.0346
	(0.0723)	(0.0714)	(0.0715)	(0.0714)	(0.0742)	(0.0728)
Age		-0.0323**		-0.0368**		-0.0299***
		(0.0142)		(0.0144)		(0.0113)
Female		0.128*		0.0651		0.117
		(0.0734)		(0.0727)		(0.0753)
Year of study		0.0425		0.00500		0.137***
		(0.0319)		(0.0330)		(0.0313)
HIB		0.267***		0.212***		0.290***
		(0.0781)		(0.0781)		(0.0765)
Low_Lifestyle		-0.0624		-0.0968		0.0235
		(0.0751)		(0.0718)		(0.0788)
Low_Study		-0.127		-0.163**		-0.0877
		(0.0790)		(0.0760)		(0.0778)
Low_Happiness		-0.0653		-0.102		-0.0404
		(0.0813)		(0.0798)		(0.0846)
Low_Selfcontrol		-0.181**		-0.126*		-0.170**
		(0.0747)		(0.0707)		(0.0788)
Constant	1.772***	2.438***	2.097***	3.043***	3.349***	3.669***
	(0.0496)	(0.313)	(0.0495)	(0.311)	(0.0517)	(0.253)
N	761	742	778	758	726	707

Notes: Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01

In columns 3-4 and 5-6, respectively, we break TGP into its two components: completed study points (CSP) and the grade average (GA). We note that treatment on average improves CSP by 0.14 standard deviations. Hence, treatment clearly makes people on average pass more courses. The estimated average effect on the grade average is also positive, but smaller, and not statistically significant. We return to a further discussion of this result in the heterogeneity analysis in the next section. Table A2 and A3 in the Appendix show that the results are robust to controlling for academic performance in the previous semester and to adding, as pre-specified,

alcohol consumption as a control variable.[13] Further, in Table A5 in the appendix, we show that the positive effect on academic performance as measured by administrative data is mirrored in a self-reported measure of academic performance, which improved, significantly, by 0.12 standard deviations for the treated group.

Table 5 complements the analysis by presenting the results from IV regressions, where we instrument the standardized number of gym visits by the treatment indicator. Hence, if one rules out all other potential channels, the coefficient on gym visits provides an estimate of the causal effect of a one standard deviation increase in gym visits on academic performance, which for TGP is 0.234 standard deviations.

Table 5: Instrumental Variable Regressions

| | Outcome Variable (std) | | | | | |
	TGP	TGP	CSP	CSP	GA	GA
Gym visits (std)	0.210*	0.234**	0.203*	0.222*	0.0175	0.0538
	(0.115)	(0.114)	(0.115)	(0.115)	(0.115)	(0.113)
Constant	1.729***	2.302***	2.057***	2.908***	3.345***	3.636***
	(0.0680)	(0.336)	(0.0677)	(0.329)	(0.0707)	(0.269)
Controls	No	Yes	No	Yes	No	Yes
N	761	742	778	758	726	707

Notes: The table reports results from 2SLS regressions where the endogenous variable Gym visits (std) is instrumented using the exogenous variable Treated. Controls are: Age, Female, Year of study, HIB, Low_Lifestyle, Low_Study, Low_Happiness and Low_Selfcontrol. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

4.3. Heterogeneous effects

Figure 4 considers heterogeneity along four pre-specified dimensions: lifestyle, study hours, happiness, and self-control. In particular, we hypothesized that the treatment effect would

---

[13] In line with the pre-analysis plan, we do not control for academic performance in the semester prior to the intervention in the main regressions because it reduces the sample size in a selective manner. Alcohol was part of the lifestyle index in the pre-analysis plan. However, many participants were unwilling to answer the questions on alcohol consumption. Including it therefore means that we need to reduce the sample size by 27 observations.

be stronger for those who were below the median on the respective measure. We see considerable support for this hypothesis. The difference between Control and Treatment is systematically greater for those subjects who were below the median in each of the respective dimensions, as well as in the aggregate.

Figure 4: Difference in effectiveness of treatment, by median split
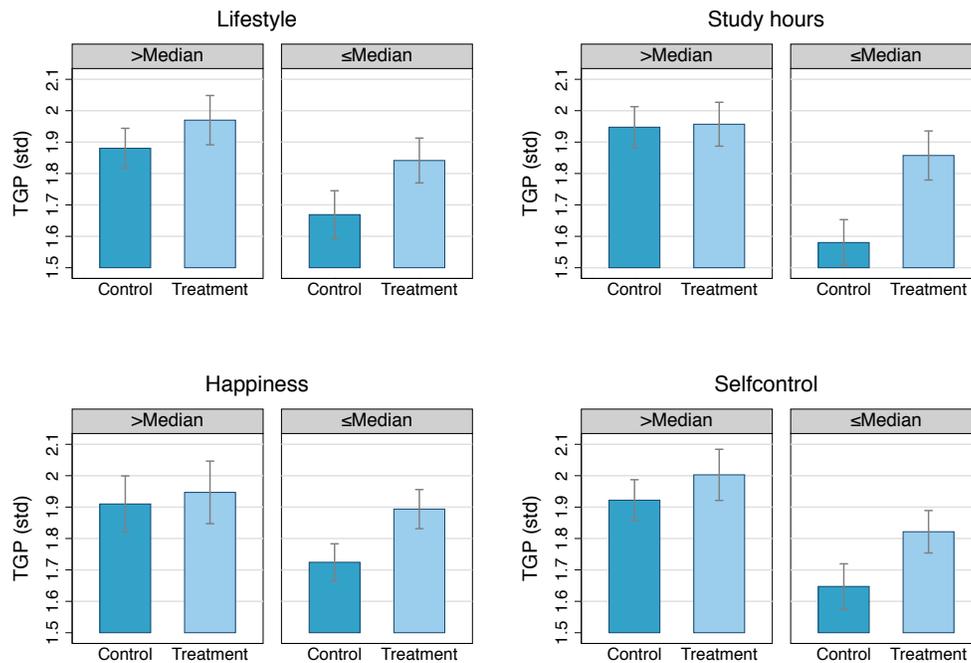
Notes: The figure shows the treatment effects on overall academic performance separately for students above and below the median in terms of the lifestyle index, hours studied, life satisfaction and the self-control index. Error bars indicate +/- one standard error.

Table 6 confirms the visual effects more formally with regression analysis: In the first four columns, we consider each dimension separately. We observe that the patterns are remarkably similar. For all four dimensions, we observe a positive (though not always significant) interaction effect, which is indicative of the intervention having strongest impact on the students who were struggling in the respective dimension.

Table 6: Academic Performance Heterogeneous Effects

	Outcome Variable (std)						
	TGP	TGP	TGP	TGP	TGP	CSP	GA
Treated	0.107	0.0329	0.0523	0.0803	-0.0684	0.0469	-0.181
	(0.0996)	(0.0944)	(0.130)	(0.103)	(0.157)	(0.168)	(0.162)
Low_Lifestyle	-0.193**				-0.0360	-0.132	0.0704
	(0.0968)				(0.101)	(0.0980)	(0.108)
Treated*Low_Lifestyle	0.0808				-0.0354	0.0739	-0.0810
	(0.143)				(0.151)	(0.144)	(0.156)
Low_Study		-0.297***			-0.238**	-0.208**	-0.225**
		(0.0975)			(0.102)	(0.100)	(0.106)
Treated*Low_Study		0.246*			0.217	0.0903	0.264*
		(0.142)			(0.151)	(0.146)	(0.151)
Low_Happiness			-0.182*		-0.121	-0.100	-0.0636
			(0.106)		(0.106)	(0.107)	(0.112)
Treated*Low_Happiness			0.135		0.106	-0.000637	0.0418
			(0.156)		(0.159)	(0.160)	(0.166)
Low_Selfcontrol				-0.278***	-0.226**	-0.135	-0.264**
				(0.0964)	(0.0997)	(0.0958)	(0.106)
Treated*Low_Selfcontrol				0.125	0.0859	0.0169	0.180
				(0.143)	(0.151)	(0.143)	(0.157)
Constant	2.321***	2.398***	2.354***	2.317***	2.522***	3.084***	3.769***
	(0.302)	(0.300)	(0.298)	(0.297)	(0.307)	(0.304)	(0.257)
Treated+Treated*Low	0.187*	0.279***	0.188**	0.205**	0.305**	0.227	0.224*
	(0.103)	(0.105)	(0.0847)	(0.0981)	(0.131)	(0.139)	(0.135)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	753	756	759	756	742	758	707

Notes: Controls are: Age, Female, Year of study, and HIB. Robust standard errors in parentheses.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

For all four dimensions, we also observe a large and highly significant treatment effect for the group that was below the median at the baseline (see Treated+Treated*Low near the bottom of the Table), while we see a smaller (and insignificant) effect for the group that was above the median. Importantly, we also observe for all four dimensions that a low score at the baseline is strongly correlated with poorer academic performance, which suggests these dimensions may be important for success as a student. It is therefore also interesting noticing

that the intervention appears to close the gap in academic performance between the students below and above the median in each of these dimensions.

Column 5 includes all four dimensions in a single regression model, which provides us with an estimate of the treatment effect on those that are below the median on all four dimensions (Treated+Treated_Low): students that struggle across all dimensions experience a 0.31 standard deviation improvement in academic performance ($p = 0.021$ in two-sided test). We also note that being treated has no effect on students who were above the median on all four dimensions. In fact, the point estimate for these students is negative (but not statistically significant), which may be seen as suggestive evidence of the possibility that exercise may worsen academic performance by directing time and attention away from studying.[14]

Columns 6-7 extend the analysis to consider the impact on completed study points (CSP) and grade average (GA), where we observe significant improvements (more than 0.2 standard deviations, again, highlighted by Treated+Treated_Low) in *both* CSP and GA for the students who struggled at the baseline (the *p*-values are 0.102 and 0.098, respectively, two-sided tests).

4.4. Effects on lifestyle, self-control and happiness

Table 7 shows the effects of the intervention on self-reported lifestyle, study hours, happiness, and self-control, shedding light on mechanisms for the effect on academic performance.[15]

---

[14] A comparison of the sub-treatments may also be suggestive of there being two opposing effects from physical exercise. In Table A6 and A7, we observe that the Bonus sub-treatment had the strongest effect on exercise, but the weakest effect on academic performance (even though this difference is not statistically significant).

[15] Table A4 in the appendix repeats the analysis for the sample that were below the median across all four dimensions at the baseline. The treatment effects for this subsample are roughly twice as large, except for self-control which is of similar size.

Table 7: Regressions Results Intermediary Outcomes

	Outcome Variable (std)							
	Lifestyle	Lifestyle	Study hours	Study hours	Happy	Happy	Self-control	Self-control
Treated	0.138*	0.132**	-0.00765	-0.00143	0.0506	0.0629	0.177**	0.151**
	(0.0735)	(0.0633)	(0.0736)	(0.0709)	(0.0737)	(0.0689)	(0.0732)	(0.0624)
Age		-0.00201		-0.00346		-0.0191		-0.00766
		(0.0108)		(0.0124)		(0.0128)		(0.0110)
Female		0.138**		0.0365		0.0521		0.0416
		(0.0636)		(0.0709)		(0.0694)		(0.0629)
Year of study		-0.00247		0.0291		0.0107		0.00430
		(0.0265)		(0.0291)		(0.0315)		(0.0283)
HIB		0.0949		0.0851		0.303***		-0.00114
		(0.0665)		(0.0690)		(0.0703)		(0.0639)
Low Lifestyle		-0.865***		0.00670		-0.247***		-0.211***
		(0.0686)		(0.0741)		(0.0741)		(0.0677)
Low_Study		-0.0172		-0.574***		0.0590		-0.150**
		(0.0667)		(0.0690)		(0.0706)		(0.0645)
Low_Happiness		-0.125*		-0.232**		-0.714***		-0.0521
		(0.0714)		(0.106)		(0.0755)		(0.0762)
Low_Selfcontrol		-0.314***		-0.0480		-0.0901		-0.980***
		(0.0682)		(0.0715)		(0.0753)		(0.0665)
Constant	-0.0705	0.558**	1.678***	2.119***	3.601***	4.503***	3.809***	4.711***
	(0.0538)	(0.238)	(0.0550)	(0.303)	(0.0550)	(0.266)	(0.0506)	(0.234)
N	740	723	741	724	740	723	740	723

Notes: The outcome variables are all measured in the follow-up survey (i.e., after the intervention), whereas all the controls are based on answers to questions in the baseline survey (i.e. before the intervention). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The variables that stand out in terms of positive treatment effects are lifestyle and self-control. Hence, it seems that the treatment *de facto* stimulated the development of good habits. In contrast, the effects on happiness and hours studied are less articulated. The null effect on the numbers of hours studied (last week) may come as a surprise, but may reflect that the follow-up study was conducted during the exam period. It is thus still possible that the treatment had a positive effect on the number of hours studied earlier in the semester. Alternatively, the absence of an effect on study hours may suggest that the positive effects on academic performance are obtained from more efficient studying. The null effect on general happiness may be less surprising, since it is a complex variable that encompasses many different aspects in the daily life. Table A5 in the Appendix goes behind general life satisfaction by also reporting the effect

on satisfaction regarding academic performance, health, social life, and one's economic situation, respectively. Interestingly, the largest, and only significant effect is for satisfaction with academic performance, which – as mentioned before – improves by 0.12 standard deviations, in line with the observed treatment effect on academic performance observed in the administrative data.

Although not of primary interest to us, one might wonder whether the effect on exercise improved students' health. Our data do not allow a direct test of this question, but we note that Charness and Gneezy (2009) report significant improvements on weight, waist size, BMI, body fat percentage and pulse rate, from a similar increase in exercise (0.6 additional weekly visits in the 13-week post-intervention period). This is also in line with the medical literature, where RCT's have found positive effects on these measures from exercising as little as 60 minutes per week in periods from 4-20 weeks (see Janssen and LeBlanc, 2010, for a review).

5. Conclusion

Exercising carries health benefits. Previous work has found that it is possible to incentivize exercise in the short run (while incentives are active), and to some degree also create new habits in the long run (after the incentives are removed). In this paper we test the hypothesis that incentivizing exercise may also carry benefits in other realms – in particular academic performance. The intuition is based on correlational evidence suggesting that 'a healthy body leads to a healthy mind' and a sense that exercise will improve academic performance. We test this intuition in the first pre-specified, large-scale randomized control trial, combining high quality administrative data on both gym exercise and academic performance and self-reported data with almost no attrition.

As expected, we find that incentivizing exercise leads to an increase in exercising. Our main contribution is that we also demonstrate that there is a strong causal effect of this intervention on academic performance. A measure of overall academic performance, taking into account both the number of credits taken and the grades in the courses taken shows a significant increase of 0.15 standard deviations. Importantly, the intervention primarily affects academic performance of students who struggle with personal habits at the baseline. The treatment causes an increase in total grade points by about 0.30 standard deviations in this group, where these students both complete more courses and improve their grade average. To put the size of the main effect in context, we note that it is comparable to the effect of: a one standard deviation improvement in teacher quality (Rockoff, 2004); a one standard deviation in university class size (Bandiera et al., 2010); and substantial individual incentives for academic performance in college (Angrist et al., 2009), which all have been shown to improve academic performance by about 0.15 standard deviations.

Our results suggest that physical exercise may be an important tool for improving academic performance, especially for students who lag in this respect. Given the enormous importance of education and human capital in our complex society, we believe these findings offer a very promising avenue for future policy interventions in the realms of health and education, where a number of open questions remain. For example, how persistent are the academic benefits from our intervention? Would providing a direct form of incentives for better academic performance be beneficial in combination with incentives for exercise? Would there be a synergy from such cross-incentivization or is an indirect incentive scheme best for improving academic performance? We hope that others will join us in the pursuit of better insights on these important questions.

References

Aaltonen, S., Latvala, A., Rose, R., Kujala, U., Kaprio, J., & Silventoinen, K. (2016). Leisure-Time Physical Activity and Academic Performance: Cross-Lagged Associations from Adolescence to Young Adulthood. *Scientific Reports, 6*, 39215.

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science, 61*(1), 146-160.

Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and Services for College Achievement: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics, 1*(1), 136-163.

Babcock, P., Bedard, K., Charness, G., Hartman, J., & Royer, H. (2015). Letting Down the Team? Social Effects of Team Incentives. *Journal of the European Economic Association, 13*(5), 841-870.

Bandiera, O., Larcinese, V., & Rasul, I. (2010). Heterogeneous Class Size Effects: New Evidence from a Panel of University Students. *Economic Journal, 120(549)*,1365-1398.

Biddle, S., & Asare, M. (2011). Physical activity and mental health in children and adolescents: a review of reviews. *British Journal of Sports Medicine, 45*, 886-895.

Booth, F., Roberts, C., & Laye, M. (2012). Lack of exercise is a major cause of chronic diseases. *Comprehensive Physiology, 2*, 1143-1211.

Carrera, M., Royer, H., Stehr, M., & Sydnor, J. (2017). *The Structure of Health Incentives: Evidence from a Field Experiment.* NBER Working Paper No. 23188.

Charness, G., & Gneezy, U. (2009). Incentives to Exercise. *Econometrica, 77*(3), 909-931.

Charness, G., & Gneezy, U. (2011). *On The Long Term Effect of Short Term Incentive to Exercise.* Unpublished manuscript.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education, 1*, 697-812.

DellaVigna, S., & Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review, 96*(3), 694-719.

Di Bartolomeo, G., & Papa, S. (2017). The Effects of Physical Activity on Social Interactions: The Case of Trust and Trustworthiness. *Journal of Sports Economics*, forthcoming.

Fricke, H., Lechner, M., & Steinmayr, A. (2017). *The Effect of Physical Activity on Student Performance in College: An Experimental Evaluation.* University of St Gallen Discussion Paper no. 2017-07.

Fryer, R. G. (2011). Financial Incentives and Student Achievement: Evidence from Randomized Trials. *Quarterly Jounal of Economics, 126*(4), 1755-1798.

Garibaldi, P., Giavazzi, F., Ichino, A., & Rettore, E. (2012). College Cost and Time to Complete a Degree: Evidence from Tuition Discontinuities. *Review of Economics and Statistics, 94*(3), 699-711.

Gneezy , U., Meier, S., & Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives, 25*(4), 191-210.

Grissom, J. (2005). *A Study of the Relationship Between Physical Fitness and Academic Achievement in California Using 2004 Test Results.* Sacramento: Califonia Department of Education.

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity, 7*(40).

Khan, N., & Hillman, C. (2014). The relation of childhood physical activity and aerobic fitness to brain function and cognition: a review. *Pediatric Exercise Science, 26*(2), 138-146.

Levitt, S. D., List, J. A., & Sadoff, S. (2016). *The Effect of Performance-Based Incentives on Educational Achievement: Evidence from a Randomized Experiment.* NBER Working Paper No. 22107.

Lipscomb, S. (2007). Secondary school extracurricular involvement and academic achievement: A fixed effects approach. *Economics of Education Review, 26*(4), 463-472.

Mura, G., Vellante, M., Egidio Nardi, A., Machado, S., & Giovanni Carta, M. (2015). Effects of school-based physical activity interventions on cognition and academic achievement: a systematic review. *CNS & Neurological Disorders-Drug Targets, 14*(9), 1194-1208.

New York Times. (2017). Retrieved May 15, 2017 from: https://www.nytimes.com/2017/03/21/well/family/why-kids-shouldnt-sit-still-in-class.html

Pfeifer, C., & Cornelissen, T. (2010). The impact of participation in sports on educational attainment— New evidence from Germany. *Economics of Education Review, 29*(1), 94-103.

Rasberry, C., Lee, S., Robin, L., Laris, B., Russel, L., Coyle, K., & Nishiser, A. (2011). The association between school-based physical activity, including physical education, and academic performance: a systematic review of the literature. *Preventive Medicine, 52*, S10-S20.

Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2011). Long-term health benefits of physical activity – a systematic review of longitudinal studies. *BMC Public Health, 45*, 886–895.

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review, 94*(2), 247-252.

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, Commitments, and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company. *American Economic Journal: Applied Economics, 7*(3), 51-84.

Scott-Clayton, J. (2011). On Money and Motivation: A Quasi-Experimental Analysis of Financial Incentives for College Achievement. *Journal of Human Resources, 46*(3), 614-646.

Singh, A., Uijtdewilligen, L., Twisk, J., van Mechelen, W., & Chinapaw, M. (2012). Physical activity and performance at school: a systematic review of the literature including a methodological quality assessment. *Archives of Pediatrics & Adolescent Medicine, 166*(1), 49-55.

Stevenson, B. (2010). Beyond the classroom: Using Title IX to measure the return to high school sports. *The Review of Economics and Statistics, 92*(2), 284-301.

U.S. Department of Health and Human Services. (2010). *The association between school-based physical activity, including physical education, and academic performance.* Atlanta, GA.

Volpp, K. G., Asch, D. A., Galvin, R., & Loewenstein, G. (2011). Redesigning Employee Health Incentives — Lessons from Behavioral Economics. *New England Journal of Medicine, 365*(5), 388-390.

Appendix A

Figure A1: Gym Visits per Week by Treatment
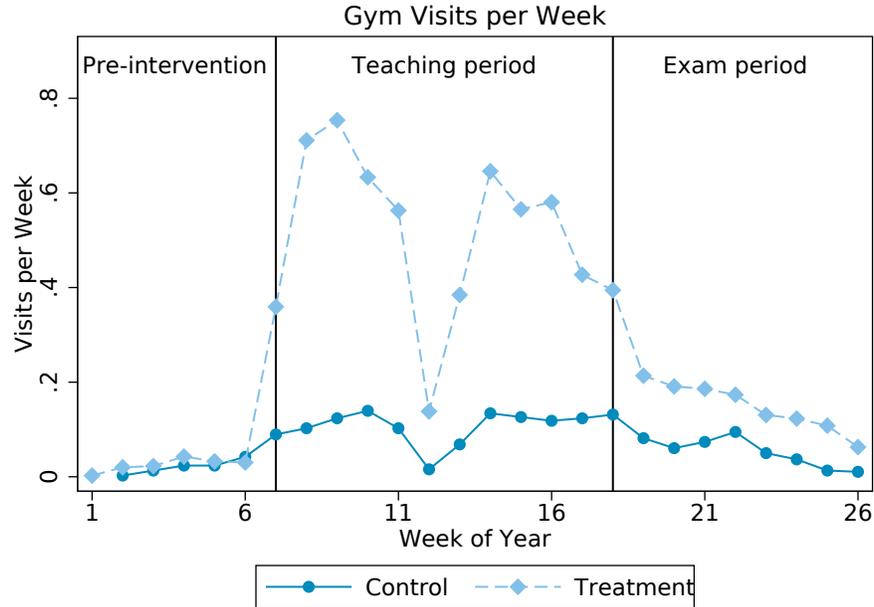
Notes: The figure plots the average number of gym visits per week in the treatment and control group. The solid vertical lines mark the week SIB-cards were available for pick-up at the gym and the beginning of the exam period. We observe that there was an immediate response to treatment. The week after distributing the free SIB-cards average attendance is 0.7 visits in the treatment group compared to 0.1 visits in the control group. Except for the Easter break these numbers are remarkably stable during the teaching period (i.e., until May 2 or week 18). From the figure we also observe that there are close to zero gym visits in our sample prior to distributing the SIB-cards in week 7 (February 16). That there are some visits in the pre-experiment period indicates that the screening of existing members by SIB was imperfect. Importantly, however, there is no difference in this regard between treatment and control.

Table A1: Extended Balance Test

	Overall	Control	Treatment	p-value (T-C)
Age	22.258	22.316	22.204	0.640
Female (1/0)	0.533	0.550	0.518	0.366
Year of study	2.008	2.034	1.982	0.561
HIB	0.481	0.500	0.462	0.294
Lifestyle index	0.004	-0.025	0.032	0.703
Study hours	22.599	22.884	22.330	0.475
Happiness (1-10)	7.340	7.287	7.391	0.412
Self-control index (1-10)	5.814	5.751	5.875	0.249
Work (1/0)	0.512	0.525	0.499	0.464
Hours worked	5.525	5.664	5.394	0.660
Member of other gym (1/0)	0.094	0.084	0.104	0.357
Previous gym member (1/0)	0.777	0.780	0.774	0.833
Times exercised	1.908	1.864	1.949	0.576
Gym visits Fall 2015	2.404	2.808	2.018	0.107
TGP Fall 2015	76.833	75.573	78.057	0.417
CSP Fall 2015	23.206	23.241	23.172	0.928
GA Fall 2015	3.177	3.140	3.212	0.310
N	778	380	398	

Notes: The variables Work-GA Fall 2015 were not specified as control variables in the pre-analysis plan. Work is a dummy variable equal to one if the student worked in addition to studying. Hours worked is the self-reported number of hours worked last week. Member of other gym is a dummy variable equal to one if the student was a member at another gym. Previous member of gym is a dummy variable equal to one if the student had been member at any gym before. Times exercised is the self-reported number of times exercised last week (all types of exercise, not only gym attendance). Gym visits Fall 2015 is the total number of gym visits at SIB in the semester prior to the intervention. TGP, CSP and GA Fall 2015 are the main academic outcome variables measured in the semester prior to the intervention.

Table A2: Regressions Results Including Alcohol in Lifestyle Index

| | Outcome Variable (std) | | |
	TGP	CSP	GA
Treated	0.136*	0.135*	0.0303
	(0.0718)	(0.0721)	(0.0742)
Constant	2.585***	3.128***	3.809***
	(0.287)	(0.304)	(0.256)
Controls	Yes	Yes	Yes
N	715	730	681

Notes: Controls are: Age, Female, Year of study, HIB, Low_Lifestyle, Low_Study, Low_Happiness and Low_Selfcontrol. The Lifestyle index used in these regressions includes total alcohol consumption the week before the baseline survey. Robust standard errors in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A3: Regressions Results Controlling for Academic Performance Fall 2015

| | Outcome Variable (std) | | |
	TGP	CSP	GA
Treated	0.127*	0.140**	-0.0476
	(0.0682)	(0.0699)	(0.0652)
TGP Fall 2015 (std)	0.334***		
	(0.0502)		
CSP Fall 2015 (std)		0.201***	
		(0.0565)	
GA Fall 2015 (std)			0.482***
			(0.0383)
Constant	1.392***	2.310***	1.697***
	(0.365)	(0.401)	(0.263)
Controls	Yes	Yes	Yes
N	727	758	651

Notes: Controls are: Age, Female, Year of study, HIB, Low_Lifestyle, Low_Study, Low_Happiness and Low_Selfcontrol. Robust standard errors in parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A4: Regressions Results Intermediary Outcomes for Subjects Below Median Across All Four Dimensions at Baseline

	Outcome Variable (std)			
	Lifestyle	Study	Happy	Selfcontrol
Treated	0.278*	0.0658	0.147	0.0930
	(0.165)	(0.143)	(0.177)	(0.142)
Constant	-0.821	0.655	3.848***	3.282***
	(0.615)	(0.472)	(0.574)	(0.411)
N	135	135	135	135

Notes: Controls are: Age, Female, Year of study and HIB. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A5: Different Happiness Measures

	Happiness Measure (std)				
	General	Academic	Health	Social	Economy
Treated	0.0629	0.122*	0.0836	-0.00246	0.0437
	(0.0689)	(0.0701)	(0.0676)	(0.0714)	(0.0726)
Constant	4.503***	3.659***	3.631***	4.652***	3.184***
	(0.266)	(0.265)	(0.241)	(0.269)	(0.301)
Controls	Yes	Yes	Yes	Yes	Yes
N	723	723	723	723	723

Notes: Controls are: Age, Female, Year of study, HIB, Low_Lifestyle, Low_Study, Low_Happiness and Low_Selfcontrol. Each outcome variable is coded 1-10. All is the average of the five different happiness measures. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A6: Gym Exercise by Sub-treatment

	Outcome Variable			
	Visits (std)	Visits (std)	Visited gym (1/0)	Visited gym (1/0)
Card	0.352***	0.359***	0.470***	0.474***
	(0.113)	(0.114)	(0.0523)	(0.0529)
PT	0.562***	0.562***	0.611***	0.615***
	(0.102)	(0.106)	(0.0478)	(0.0487)
Bonus	0.800***	0.792***	0.599***	0.598***
	(0.0913)	(0.0920)	(0.0360)	(0.0365)
Age		-0.0143		-0.00411
		(0.00902)		(0.00436)
Female		-0.0356		0.0614**
		(0.0696)		(0.0296)
Year of study		-0.0545**		-0.00246
		(0.0216)		(0.0124)
HIB		0.0539		-0.0212
		(0.0694)		(0.0301)
Low_Lifestyle		0.133**		0.0432
		(0.0657)		(0.0298)
Low_Study		0.0222		-0.0138
		(0.0706)		(0.0303)
Low_Happiness		-0.0405		-0.00578
		(0.0775)		(0.0343)
Low_Selfcontrol		-0.0390		-0.00558
		(0.0684)		(0.0298)
Constant	0.199***	0.592**	0.116***	0.179*
	(0.0390)	(0.231)	(0.0165)	(0.102)
N	778	758	778	758

Notes: Robust standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Academic Performance by Sub-treatment

	Outcome Variable (std)					
	TGP	TGP	CSP	CSP	GA	GA
Card	0.208*	0.249**	0.218*	0.259**	0.0167	0.0468
	(0.112)	(0.116)	(0.122)	(0.124)	(0.111)	(0.115)
PT	0.205*	0.211*	0.106	0.127	0.0705	0.0750
	(0.113)	(0.113)	(0.113)	(0.114)	(0.133)	(0.129)
Bonus	0.0650	0.0707	0.0946	0.0846	-0.0196	0.00959
	(0.0916)	(0.0898)	(0.0863)	(0.0856)	(0.0892)	(0.0874)
Age		-0.0322**		-0.0370**		-0.0298***
		(0.0142)		(0.0144)		(0.0114)
Female		0.137*		0.0710		0.120
		(0.0740)		(0.0731)		(0.0759)
Year of study		0.0400		0.00247		0.137***
		(0.0319)		(0.0330)		(0.0315)
HIB		0.265***		0.211***		0.290***
		(0.0782)		(0.0781)		(0.0766)
Low_Lifestyle		-0.0678		-0.100		0.0216
		(0.0750)		(0.0719)		(0.0789)
Low_Study		-0.130		-0.167**		-0.0882
		(0.0793)		(0.0764)		(0.0781)
Low_Happiness		-0.0628		-0.100		-0.0397
		(0.0811)		(0.0796)		(0.0847)
Low_Selfcontrol		-0.183**		-0.129*		-0.171**
		(0.0746)		(0.0705)		(0.0790)
Constant	1.772***	2.440***	2.097***	3.054***	3.349***	3.668***
	(0.0496)	(0.313)	(0.0495)	(0.311)	(0.0518)	(0.255)
N	761	742	778	758	726	707

Notes: Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A8: Academic Performance on Retake Exams and Pass/Fail Courses

	Outcome variable (std)			
	TGP: retake exams	TGP: retake exams	CSP: Pass/Fail courses	CSP: Pass/Fail courses
Treated	-0.0720	-0.0768	-0.0832	-0.0707
	(0.0722)	(0.0703)	(0.0718)	(0.0719)
Age		-0.0180**		0.0152
		(0.00764)		(0.0151)
Female		-0.0616		0.0197
		(0.0665)		(0.0736)
Year of study		0.113***		0.0782**
		(0.0384)		(0.0367)
HIB		0.0610		0.392***
		(0.0791)		(0.0847)
Low_Lifestyle		-0.00654		0.0224
		(0.0719)		(0.0740)
Low_Study		-0.0919		0.108
		(0.0830)		(0.0858)
Low_Happiness		-0.126		0.0792
		(0.0849)		(0.0740)
Low_Selfcontrol		0.147*		0.0313
		(0.0753)		(0.0815)
Constant	0.382***	0.615***	0.325***	-0.509
	(0.0569)	(0.183)	(0.0520)	(0.311)
N	776	756	778	758

Notes: Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Appendix B - Survey

B2: Baseline survey (in English)

(A)

1.   **How old are you (number of years)? _____**

2.   **What gender are you?**   **Female**   **Male**

3.   **At which educational institution are you currently studying? _____**

4.   **What profession do you study (Put a cross beside your answer)?**

Humanities   **Psychology**   **Mathematics and natural sciences**

Social sciences   **Law**   **Medicine and odontology**

Other

5.   **How long have you studied at college/university (Put a cross beside your answer)?**

  First year   **Second year**   **Third year**

  Fourth year   **Fifth year**   **More than five years**

6.   **How many hours did you study last week (including participation in lectures)? _____**

7.   **Do you have paid work in addition to your studies this semester? YES   NO**

8.   **How many hours did you spend doing paid work last week? _____**

9.   Are you a member of a gym other than SIB trening this semester?

                                                    YES   NO

10.   Have you ever been a member of a gym (including SIB trening)?

                                                    YES   NO

11.   If you answered "NO" on question 10, answer the following question (if you answered "YES" you can skip this question):

Have you ever considered becoming member of a gym?   YES   NO

12.   How do you prefer to work out?

On my own   With friends   In an organized group

13.   Have you ever used a personal trainer?   YES   NO

14.   If you answered "NO" on question 13, answer the following question(if you answered "YES" you can skip this question):

If you have not ever used a personal trainer, would you consider using one in the future?

                                                    YES   NO

15.   How many times did you exercise last week (including all exercise methods, not just in the gym)? _____

16.   Please enter your alcohol consumption last week:

a.   _____ glass/glasses of wine

b.   _____ Bottles/ cans of beer

c.   _____ units of other alcoholic beverages

17.   How many days last week did you go to bed after midnight (set ring)?

                                                    0 1 2 3 4 5 6 7

18.   How many days last week did you feel tired / not fully rested (set ring)?

                                                    0 1 2 3 4 5 6 7

(B)

For each of the questions below, please indicate on a scale of 1-10 to what extent you agree or disagree: 1 means you disagree completely while 10 means you completely agree.

                1 2 3 4 5 6 7 8 9 10

            (1= Disagree completely)   **(10= Agree completely)**

19.   During the last week I felt a lack of energy that made it difficult for me to get things done.

                                        1 2 3 4 5 6 7 8 9 10

20.   Why are you not a member of SiB trening?

a)   It is to expensive to become a member.   **1 2 3 4 5 6 7 8 9 10**

b)   I dont have time to exercise.   **1 2 3 4 5 6 7 8 9 10**

c)   I prefer other forms of exercise (for instance: running, team sports, etc).

                                        1 2 3 4 5 6 7 8 9 10

d)   I prefer another gym.   **1 2 3 4 5 6 7 8 9 10**

e)   Other reasons:
_____
_____
_____

21.   I feel comfortable with using the exercise equipment at the gym.
                                        1 2 3 4 5 6 7 8 9 10

22.   **I am a person who often makes hasty decisions.**   **1 2 3 4 5 6 7 8 9 10**

23.   **I find it hard to resist temptations.**   **1 2 3 4 5 6 7 8 9 10**

24.   **I tend to postpone things even if it would be best to take care of them as quickly as possible.**
     1 2 3 4 5 6 7 8 9 10

25.   **I am a person who follows my plans.**   **1 2 3 4 5 6 7 8 9 10**

(C)

For each of the questions below, please indicate on a scale of 1-10 to what extent you are dissatisfied or satisfied: 1 means you are very dissatisfied and 10 means you are very satisfied.

       1 2 3 4 5 6 7 8 9 10

     (1=very dissatisfied)   **(10=very satisfied)**

26.   **How satisfied are you with your life in general at the moment?**

     1 2 3 4 5 6 7 8 9 10

27.   **How satisfied are you with your academic achievements?**   **1 2 3 4 5 6 7 8 9 10**

28.   **How satisfied are you with your health at the moment?**   **1 2 3 4 5 6 7 8 9 10**

29.   **How satisfied are you with your social life?**   **1 2 3 4 5 6 7 8 9 10**

30.   **How satisfied are you with your financial situation?**   **1 2 3 4 5 6 7 8 9 10**

Thank you!

You are now done completing the form. If you deliver it to one of the research assistants from The Choice Lab, they will give you 100 kr in return.

We will send you the next questionnaire by e-mail at the end of May.

In 2-3 weeks, we will notify you by email if you have won a training card from SiB or not.

Thank you very much for your participation in the experiment!

Best regards,

The Choice Lab, NHH

B2: Follow-up survey (in English)

Note: This is a printed version of the electronic survey. In the actual survey, each question appeared on a separate page.

Thank you for participating in this survey.

Best regards,

The Choice Lab

Q1 How old are you (in number of years)?

Q2 What gender are you?

○　Male
○　Female

Q3 At which educational institution are you currently studying?

○　HIB
○　UIB
○　Other college/university
○　I am no longer a student

Q4 How many hours did you study last week (including participation in lectures)?

Q5 Do you have paid work in addition to your studies this semester?

○　YES
○　NO

Q6 How many hours did you spend doing paid work last week?

Q7 Are you a member of a gym other than SIB trening this semester?

○　YES
○　NO

Q8 How many times did you exercise last week (including all exercise methods, not just in the gym)?

Q9 How do you prefer to work out?

- ⭕ On my own
- ⭕ With friends
- ⭕ In an organized group

Q10 Have you ever used a personal trainer?

- ⭕ YES
- ⭕ NO

Q11 Would you consider using a personal trainer in the future?

- ⭕ YES
- ⭕ NO

Q12 Please enter your alcohol consumption last week:

Glass/Glasses of wine
Bottles/Cans of beer
Units of other alcoholic beverages

Q13 How many days last week did you go to bed after midnight?

- ⭕ 0
- ⭕ 1
- ⭕ 2
- ⭕ 3
- ⭕ 4
- ⭕ 5
- ⭕ 6
- ⭕ 7

Q14 How many days last week did you feel tired/not fully rested

○ 0
○ 1
○ 2
○ 3
○ 4
○ 5
○ 6
○ 7

Q15 For each of the questions below, please indicate on a scale of 1-10 to what extent you agree or disagree: 1 means you disagree completely while 10 means you completely agree.

Q16 During the last week I felt a lack of energy that made it difficult for me to get things done (1 = Disagree completely, 10 = Agree completely).

○ 1
○ 2
○ 3
○ 4
○ 5
○ 6
○ 7
○ 8
○ 9
○ 10

Q17 I feel comfortable with using the exercise equipment at the gym (1 = Disagree completely, 10 = Agree completely).

○ 1
○ 2
○ 3
○ 4
○ 5
○ 6
○ 7
○ 8
○ 9
○ 10

Q18 I am a person who often makes hasty decisions (1 = Disagree completely, 10 = Agree completely).

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9
- ○ 10

Q19 I find it hard to resist temptations (1 = Disagree completely, 10 = Agree completely).

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9
- ○ 10

Q20 I tend to postpone things even if it would be best to take care of them as quickly as possible (1 = Disagree completely, 10 = Agree completely).

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9
- ○ 10

Q21 I am a person who follows my plans (1 = Disagree completely, 10 = Agree completely).

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8
- ○ 9
- ○ 10

Q22 For each of the questions below, please indicate on a scale of 1-10 to what extent you are dissatisfied or satisfied: 1 means you are very dissatisfied and 10 means you are very satisfied.

Q23 How satisfied are you with:

	1	2	3	4	5	6	7	8	9	10
your life in general at the moment?	○	○	○	○	○	○	○	○	○	○
your academic achievements?	○	○	○	○	○	○	○	○	○	○
your health at the moment?	○	○	○	○	○	○	○	○	○	○
your social life?	○	○	○	○	○	○	○	○	○	○
your financial situation?	○	○	○	○	○	○	○	○	○	○

You have now answered all the questions. If you want the compensation of a 100 kr to be transferred to a different account than we mentioned in the email, you can enter a new account number below (otherwise you can leave it blank).

Please click the arrow to complete the survey!