TI 2017-001/VII
Tinbergen Institute Discussion Paper

Goal Setting and Raising the Bar: A Field Experiment

Max van Lent[1]
Michiel Souverijn[1]

[1]Department of Economics, Erasmus University Rotterdam, and Tinbergen Institute

Goal Setting and Raising the Bar: A Field Experiment*

Max van Lent† and Michiel Souverijn‡

December 19, 2016

Abstract

We study goal setting using a randomized field experiment involving 1092 first-year undergraduate students. Students have private mentor-student meetings during the year. We instructed a random subset of mentors to encourage students to set a course-specific grade goal during one of the mentor-student meetings (goal treatment). A random subset of those mentors was further instructed to challenge students to set more ambitious goals if deemed appropriate (raise treatment). We find that students in the goal treatment perform significantly better as compared to students in the control group, and more so when they performed poorly prior to the experiment. Next, we find that students in the raise treatment do not perform significantly different from the control group. Finally, students who set a goal and are challenged to set a more ambitious goal perform significantly worse than comparable students in the goal treatment.

Keywords: Goal setting, motivation, education, field experiments.
JEL: C93, I23

*We would like to express our gratitude to Robert Dur and Josse Delfgaauw for their invaluable support and advice throughout the entire research project. We would like to thank Sacha Kapoor, Jan Stoop, as well as seminar participants at EUR, LSE, and Tinbergen Institute, and conference participants at 10th NCBEE, LPEx 2015, Field Days 2016 at WZB, and 2nd SBRE Workshop on EPEO for helpful comments.

†Department of Economics, Erasmus University Rotterdam, and Tinbergen Institute. E-mail: vanlent@ese.eur.nl

‡Department of Economics, Erasmus University Rotterdam, and Tinbergen Institute. E-mail: souverijn@ese.eur.nl

1 Introduction

People often set goals. For example dieters commonly set a target weight, runners aim for a certain time, and managers set goals for employees in the form of targets. Using a series of field experiments, the psychologists Locke and Latham (1979) were the first to provide evidence that goals help to increase performance.[1] More recently, goal setting has also been studied by management scientists and economists. Economic theory papers have shown how goals can be used as reference points in order to increase performance for loss averse agents or hyperbolic discounters (see e.g. Suvorov and Van de Ven 2008, Hsiaw 2013, Koch and Nafziger 2011, and Koch et al. 2014), and that meeting goals can lead to a sense of self-achievement that makes pursuing goals worthwhile (Gomez-Minambres 2012). A rapidly growing empirical literature tests the effects of goal setting on performance in the laboratory and in the field.

This paper examines whether goal setting can help to increase student performance and to decrease drop out in an academic course. Furthermore, we are interested to learn whether challenging students to be more ambitious by increasing the goal's difficulty can increase performance further. This is relevant given the widely held belief that many students should be more ambitious, and the recently increased focus on student performance in higher education.

We start by developing a simple theory which explains how and when setting a goal and increasing a goal's difficulty can increase performance. We derive the following predictions. In line with the literature, people are willing to set a goal since setting a goal increases both performance and utility. Having an outsider propose a more ambitious goal can, but need not, increase performance further. To be precise, a proposal to raise the

[1] Locke and Latham found that goals set by an outsider (a peer or a manager), goals set in cooperation, and self-set goals can all lead to a better performance as compared to not setting goals.

goal increases performance when there is a cost of rejecting the proposal to raise and if the raise is not too large.[2] Performance will be the same as without the proposal to raise if rejecting the raise is not so costly, and the alternative is the goal initially set. Finally, increasing the goal too much might lead students to give up on their goal, in which case performance will be similar to that of a student who did not set a goal.

We test our predictions by means of a field experiment among 1092 first-year economics students. Each of these first-year students regularly has individual meetings with a mentor (who is a senior student). Mentors help students to get used to studying at a university, teach them study skills, help them with their (study) motivation, monitor their performance, and give suggestions in order to increase their study performance. We ran our experiment during the second of three individual meetings between students and their mentor. In one treatment *(goal treatment)* we instructed mentors to ask their students whether they had a specific grade goal in mind for the main course they participated in at that moment, and if not, whether they wanted to set a grade goal. In another treatment *(raise treatment)* mentors received identical instructions as in the goal treatment, and were in addition instructed to encourage students to raise their goal if deemed appropriate. We subsequently measured performance using the grades the student obtained for the course.

We find that students whose mentor was instructed to motivate students to set a goal perform 0.16 better on a 10-point scale (which is 9.3% of a standard deviation) than students in the control group. This effect is driven by students in the goal treatment dropping out less often than students in the control group. Students whose mentor was instructed to also ask students to raise their goal do not perform significantly different from the control group. Finally, being asked to raise the goal in the raise treatment leads to a significant drop in performance as compared to similar students in the goal treatment.

Setting goals can also have adverse effects such as a narrow focus and

[2] The cost of rejecting the proposal to raise the goal can for example be a psychological cost or a loss in reputation towards the mentor.

ignorance of non-goal tasks or even unethical behavior (see Ordonez et al. 2009). In our setting a concern is that students increase effort and performance on the course for which they set a goal at the expense of the other course they take at the same time. We estimate the effect of the treatment on performance in the other course, and do not find such a negative effect. This implies that motivating students to set a goal is actually good for study performance overall.

Next, we look at heterogeneous treatment effects. We test whether there are heterogeneous effects of the treatments dependent on the student's prior study results, the mentor's experience, mentor's gender, and a match between mentor's and student's gender. We find that motivating students to set goals increases performance mainly for students who were initially performing poorly. We do not find a significant difference in treatment effect for students with more experienced mentors, or for students that have the same gender as their mentor.

There is a rich literature in psychology studying goal setting and its effects on performance (see Locke 1996, Locke and Latham 2002, and Locke and Latham 2006 for literature reviews). Research in psychology groups goals in roughly three categories: goals set by an outsider, cooperatively set goals, and self-set goals. Our *goal treatment* and *raise treatment* come closest to self-set goals and cooperatively set goals, respectively. Further, the literature shows that other factors such as goal commitment, goal specificity, and how challenging the goal is are important predictors for the success of goals (see for example Hollenbeck et al. 1989, Locke 1996, and Seijts et al. 2004). Our finding that the attempt to raise goals decreases students' performance as compared to goal setting by the student may be explained by a change in commitment to the goal, leading to a decrease in (study) motivation and hence performance.

Our paper is related to a rapidly increasing number of experiments in economics that study the effects of different types of goal setting on performance in various contexts. Experiments range from self-set goals to goals set by others. In some papers goals are combined with monetary incentives (see e.g. Goerg and Kube 2012, Dalton et al. 2015, and Corgnet et al. 2015,

2016) and in other papers goals are set without monetary incentives (see e.g. Goerg and Kube 2012, Sackett et al. 2014, and Clark et al. 2016). These studies typically find that when ambitious but attainable goals are set, goals increase performance, and more so when they are combined with monetary incentives. Our main contribution to this literature is that we investigate the effects of raising goals, by increasing its difficulty in a cooperative manner.

Also closely related to our research is the literature on (non-monetary) incentives for students in education. This literature considers a number of ways besides setting goals, in which students performance can be increased.[3] Lavecchia et al. (2015) review studies of interventions in education designed to improve students' performance. The interventions target a wide range of behaviors, varying from a too little focus on the future, overreliance on routines, student self-confidence and the information on and number of choices in education. Further, Sanders and Chonaire (2015) show that in education usually (very) small effect sizes are found. The effect we find from goal setting is around the median effect size found in the sample of Sanders and Chonaire.[4]

Goal setting by students has received a lot of attention from psychologists, see e.g. Ames and Archer (1988) and Schunk (1990). Many of these papers in the psychology literature have tested whether goal setting can increase students' performance (see also Linnenbrink 2005, Morisano et al. 2010, Bettinger and Baker 2013, Schippers et al. 2015, and Travers et al. 2015). Students not subjected to goal setting are typically subjected to other activities in these studies. As a consequence these papers are unable to estimate the causal effect of motivating students to set goals. In our experiment the only difference between the control and treatment groups is that in the treatment groups mentors encourage students to set a goal. Hence we are able to estimate the causal effect of motivating students to set goals on study performance. In addition we are the first to consider chal-

[3] For example changes in the class size (see Angrist and Lavy 1999 and Bandiera et al. 2010), providing feedback to students (see Bandiera et al. 2015), and several financial and non-financial incentives (see Levitt et al. 2016).

[4] While the median effect size in Sanders and Chonaire (2015) is 10% of a standard deviation, our (almost) costless intervention has an effect of 9.3% of a standard deviation.

lenging the goals that students set by asking them to increase their goal's difficulty.

Besides the contribution of our paper to the literature on goal setting, incentives and performance in education, our paper can also be useful to management practitioners. There is a large and growing literature on designing the optimal contract (see Gibbons 2005 for a review), and recent work on the use of goals as an incentive device (see e.g. Gomez-Minambres 2012). Our result that an encouragement to increase a (self-set) goal in order to motivate a students results in lower performance is of particular interest. In a workplace where a manager evaluates his workers, it can be common practice to set goals or targets. Our findings are a first indication that challenging workers to increase their goal's difficulty might be detrimental for performance.

This paper is organized as follows. In the next section we explain the experimental context and describe the data. In section 3 we present a simple theoretical framework and derive our hypotheses. In section 4 we explain the empirical strategy. Section 5 presents the descriptive statistics, section 6 the results followed by a discussion and conclusion in the final section.

2 Experimental Set-up and Data Description

2.1 Experimental Context

The experiment involved 1092 first-year students enrolled in several undergraduate programmes at Erasmus School of Economics in Rotterdam, The Netherlands during the 2014-2015 academic year. The year is divided into five blocks of eight weeks. In each block students take 12 study credits (ECTS) worth of courses. All courses that students take at this point are obligatory, hence all students within a study programme take the same courses. Our experimental treatments take place during the second block when students have their second individual meeting with their mentor.

Each first-year student has a mentor. Mentors are senior students and are randomly assigned to students enrolled in the same programme at the

start of the academic year. All mentors are employed by the university and are paid a flat wage. Our study involves all 84 mentors, and each mentor has 10 to 15 students. Mentors regularly meet with their students, both in groups and individually. The mentor-student meetings are intended to teach students study skills, monitor their motivation, and more generally to provide a point of contact within the university. Motivation and individual prospects are the primary subjects of the three individual mentor-student meetings held over the course of the academic year. The first individual mentor-student meeting takes place arround the start of the academic year in September, while the second and third take place in November and January, after the results of respectively the first and the second block of courses have been released. Our treatments are administered during the second individual mentor-student meeting.

While the first meeting at the start of the academic year primarily serves to discuss the student's motivation and to detect possible issues, the second and third meetings serve to evaluate results and prospects of the students. Due to university rules and national legislation, students with a weak performance record may be better off dropping out before February, which is in the third block of courses. Dropping out on time results in minimal grant loss and additionally allows students to re-enroll in the same programme the following academic year, which students that otherwise fail to meet first year requirements are not allowed to do. Thus, the second meeting is a natural moment to look forward towards the rest of the academic year and to discuss what results are necessary in order to make it sensible for the student to continue their current study programme. The last individual meeting after the release of the results for the second block serves mostly to determine whether it is better for the student to drop out given her motivation and study results.

Students take two courses in the second block, an introductory course in microeconomics worth 8 ECTS and a programme specific 4 ECTS course.[5]

[5] Students enrolled in the Economics and Business Economics, Fiscal Economics, and Law and Economics programmes take besides microeonomics an ICT course, while Econometrics students take a Calculus course.

Our treatment is focused on the microeconomics course. The course is taught in Dutch (824 enrolled students) and English (268 enrolled students). The Dutch and English version are identical in all respects except for the lecturers and language spoken. The course follows a standard setup of three non-compulsory plenary lectures each week complemented by two compulsory tutorials taught by teaching assistants. The tutorials serve to review the course material, practise and discuss exercises, and in general to provide students an accessible way to obtain further explanation and clarification of the material. Tutorials are taught in 42 tutorial groups. One tutorial group consists of the students of two mentor groups. Examination of the course follows a standard format with two midterms counting 15% each and a written exam for the remaining 70%. For both midterms and the final exam students receive a grade on a 10 point scale, ranging from 1 to 10 with 10 being the best grade. In addition students could obtain a bonus, which was equal to at most half a point of the final grade, by participating in weekly online tests.

2.2 Experimental Design

Our experiment revolves around the second individual mentor-student meeting. We instructed a random subset of 54 of the 84 mentors to motivate their students to set a course specific grade goal during this mentor-student meeting. As discussed before, this second meeting is an excellent opportunity for such a discussion as its purpose is to reflect on past performance and consider what results for the current courses are necessary. This means that discussion of the progress of the current courses is natural, and a focus on microeconomics is expected since it is the most important course in the second block due to its weight in ECTS. Our treatment builds on this discussion.

During meetings with all mentors in the period between 22 and 31 October 2014, we informed the mentors that some of them would be expected to take a somewhat different approach to the second individual meeting. Selected mentors were sent instructions by e-mail about how to complement

the discussion regarding the current courses one and a half week before the meetings. The instructions were accompanied by a simple flow diagram (see Appendices 1 and 2). All 54 selected mentors confirmed that they understood the instructions.

Randomly selected mentors were instructed to ask students whether they have a specific goal in terms of a grade in mind for the main course, microeconomics, and if so to elicit that grade goal. If the student did not have a grade goal in mind, the student was asked whether she wants to set one on the spot, again eliciting the goal set. Students were free not to set a goal. Mentors were asked to write down their evaluation of the goal of the student, evaluating the student's goal as either "too easy", "doable" or "too hard". The description of the treatment so far describes the *goal treatment*. Thus in this treatment, mentors are asked to induce their students to set themselves a specific grade goal for the main course in the second block.

A second group of mentors were randomly selected to perform the *raise treatment*. In the raise treatment mentors implement the goal treatment but are additionaly requested to attempt to raise the goal (if any) set by the student when deemed appropriate. If the mentor described the goal as "doable" or "too easy" the mentor was instructed to challenge the student by asking whether the student shouldn't be more ambitious and aim for a higher grade, specifically the student's self-set goal + 1 (e.g. if the student's goal was to get a 6 the mentor suggested aiming for a 7). The raise treatment serves to determine whether raising self-set goals can (further) improve study performance. Figure 1 illustrates the similarities and differences between the goal and raise treatment using a flowchart.

We chose to elicit a grade goal instead of other course related goals for multiple reasons. First, the final grade is (one of) the most important motivations to study for many students, hence students might find it more useful to set grade goals compared to other goals. Second, choosing an output goal (the final grade) instead of an input goal (e.g. study hours) leads to lower measurement error because we cannot perfectly measure study hours. Finally, a grade goal is specific and measurable, which are important factors that influence the success of a goal (Locke and Latham 2002).

As our measure of performance we do not take the final microeconomics grade. Instead we use a normalized version of the microeconomics grade without the first midterm result as our treatment is administered in the week of the first midterm. Hence, not all students have received treatment prior to the first midterm, while those that do have very limited time to respond.

Mentors were asked to record the outcome of the meetings on a form. We specifically asked mentors to note whether or not they brought up goals in order to identify treated subjects. Mentors record whether students set a goal, what the goal is, and their estimate of the difficulty of the goal. In the raise treatment mentors further record whether they asked students to raise their goal and whether or not the student accepted this higher goal. The mentor's estimate of the difficulty of the initial goal allows us to compare students in the raise treatment whose goal was challenged with similar students in the goal treatment whose goal was not challenged but would have been challenged if they were in the raise treatment.

Besides the forms filled in by the mentors selected to implement the treatments we obtain information on all the students from administrative data from the microeconomics course and the central administrative office. This gives us information on the student's performance in other courses, attendance of microeconomics tutorial sessions, gender, age, study programme, and mentor.[6] From the administration office we further obtained the mentor's gender and whether the mentor had experience in mentoring in previous years.

Only the mentors and lecturers were aware an experiment was being implemented, although mentors were not explicitly told so. Our introduction to all mentors in a general mentor instruction meeting necessitated that we informed all mentors that some of them would be asked to implement a small change in the upcoming individual mentor-student meetings. However, those not sent specific instructions were not aware of the exact change implemented. We specifically instructed the mentors who were selected for

[6] From students in Dutch study programmes who attended a Dutch high school, we also have highschool grades.

a treatment not to talk to anyone regarding our request. Selected mentors may deduce the purpose of the research but were not informed beyond their own instructions provided in Appendices 1 and 2. Finally, both authors of this paper were involved in the microeconomics course as teaching assistants. Because of this we took precautionary measures to prevent ourselves from learning the treatment assignment.[7]

2.3 Assignment Procedure

The assignment of students to both treatments and the control group is randomized at the mentor level. Assignment at the mentor level was chosen in order to increase compliance and prevent contamination. With assignment at the student level, a given mentor would be charged with treating her students differently, in a random order over the talks (students select a timeslot), likely leading to mistakes. In addition to accidental non-compliance, student level assignment might also result in more selective non-compliance by mentors selecting the treatment for their student(s) that they think is most appropriate.

The assignment of mentors to treatment was randomized in a stratified manner as follows. First, given that the tutorial group has a large impact on student performance as it is the main instruction method for many students, we ensure that a tutorial group is always of mixed composition in terms of treatments and control. This serves to create similar conditions for students in all treatments, but comes with the risk of contamination because students from treatment and control are in the same tutorial group. Randomization takes place within the various study programmes offered by the school as the effect of treatment can differ by programme due to the selection of students in a programme and the difficulty of the other course offered. Finally, several teaching assistants teach multiple tutorial groups. We therefore enforce that

[7] The randomization was programmed by one of the authors who received the list of mentors linked to tutorial groups. A researcher from the department was asked to perform the randomization and send only the list of mentor contacts and treatment assignment to the other author. Since the author receiving this information was unaware which mentors belong to which group it was impossible for either of the authors to relate mentors to (half of) a tutorial group.

classes taught by the same teaching assistant have an (even) mix of control and treatment groups.[8] In doing so we ensure that teaching assistants who teach classes in two different study programmes have a mix of control and treatment groups.

Randomization takes place by taking one random draw for each teaching assistant. Draws were compared between teaching assistants teaching the same number of mentor groups. The first mentor group is assigned to the control group if the draw belongs to the highest third of the draws. The middle and lowest third of the draws were assigned to the goal and raise treatment respectively. The assignment of the other groups taught by a teaching assistant then follows from the assignment of the first group by cycling through the list of possible assignments in order. The procedure is illustrated in Figure 2. We prioritized first the control and then the goal treatment. The final result of this randomization is that 30 mentors are in the control group, 28 in the goal treatment and 26 are in the raise treatment. This corresponds to 389 students in control, 367 in the goal treatment and 336 students in the raise treatment.

3 Theoretical Framework and Predictions

We are interested in the effects of goal setting and attempts to raise goals on students' study performance. To fix ideas, let us consider the following stylized framework.

Consider a student who values obtaining a high grade but dislikes to study. Let her utility be given by:

$$U = e - \frac{1}{2\theta}e^2$$

where e is her study effort which results in grade e. The student's ability is given by θ such that more able students have a lower cost of study effort. In this scenario the student optimally sets $e = \theta$ yielding utility $\frac{1}{2}\theta$.

[8] For example, if a teaching assistant teaches two groups he teaches four mentor groups of which at least one group is assigned to each treatment and at least one group is a control group.

Now let the student set a goal to motivate herself. Assume that the student values meeting her goal and that her utility from reaching this goal increases in goal difficulty. Meeting a goal may be intrinsically rewarding (see e.g. Gomez-Minambres 2012) or there may be some external motivation, for instance reputational concerns towards someone who is aware of the goal. Specifically, let the student's utility function in case she sets a goal g be given by:

$$U = e + I\left(e \geq g\right) g - \frac{1}{2\theta} e^2,$$

where $I\left(e \geq g\right)$ is an indicator function that equals 1 if the goal is met (i.e. if $e \geq g$) and 0 otherwise.[9] Since the student already exerts $e = \theta$ without a goal, setting a non-challenging goal $g \leq \theta$ does not affect her study performance. In that case she would be best off setting a goal $g^{NC} = \theta$, which yields her utility $U^{NC} = \frac{3}{2}\theta$.

Next consider the student setting a goal that challenges her to exert more effort, $g > \theta$. The student optimally meets such a goal by exerting $e = g$.[10] Given this, the student sets her goal to maximize her utility resulting in the optimal challenging goal $g^C = 2\theta$. The student obtains utility $U^C = 2\theta$ from setting herself the challenging goal, exceeding the utility $U^{NC} = \frac{3}{2}\theta$ derived from setting the non-challenging goal. Thus the student is best off setting a challenging goal for herself, boosting her study performance. This demonstrates our first prediction:

Prediction 1: Setting goals increases student performance.

Now consider what happens when the goal is raised above g^C by an outsider. Given that the student is better off under her optimal goal g^C as compared to either not setting a goal or setting an unchallenging goal she

[9] Most economic theory papers on goal setting model the agent's utility function assuming loss aversion. Agents get utility if they reach their goal and a disutility from not reaching this goal. Since in our simplified model effort maps directly into a grade, i.e. without any noise or uncertainty, agents never end up in the loss domain. Hence, these richer models would yield the same predictions as our simplified model. In case there is noise or uncertainty and agents are loss averse, the results marginally change. Some agents may no longer be willing to set goals, goals become less ambitious, but there is still some room to raise goals.

[10] Note that the student will never choose $e > g$, because then she would have a strictly higher utility if she would set a goal g', such that $e = g' > g$.

will still be better off under goals that deviate from g^C slightly as compared to not setting a goal ($U = \frac{1}{2}\theta$) or setting a non-challenging goal ($U^{NC} = \frac{3}{2}\theta$). Thus changing the goal from g^C to a higher goal can improve performance.[11] This leads us to our second prediction:

Prediction 2: Raising goals can increase student performance.

A raised goal can lead to higher performance if the student accepts the proposal to raise the goal. A student will accept this proposal to raise if there is a cost of rejecting the goal that is proposed by the outsider, and if the proposed raise is not too high.[12] If the proposed raise is too high, the student will not accept the proposed goal. This leads to a similar performance as when the student was not asked to raise the goal (this happens when the student's outside option is the initial goal), or this leads to performance that is similar to setting non-challening or even no goals (this happens when the student's outside option is no goal or a non-challenging goal). Alternatively, since the goal is raised by an outsider, the student may not derive as much utility from meeting the goal as from self set goals.[13] This leads the student to perform worse because she is less motivated to reach the goal.

4 Empirical Strategy

We estimate the effects of motivating students to set goals and attempts to raise students' self-set goals in two ways. First we estimate an intention to treat effect, comparing the results of students of mentors assigned to treatments to the results of students of mentors assigned to the control group. Random assignment of mentors to treatments coupled with random assignment of students to mentors should result in ex ante similarity between students. We thus attribute differences between students in the control and treatments after our intervention to the intervention. We estimate the

[11] Note that although performance (i.e. e) increases when a goal higher than g^C is achieved, utility will be lower compared to the student setting and reaching g^C.

[12] Costs of rejecting the proposal to raise the goal can for example be psychological or reputational costs.

[13] See e.g. Hollenbeck et al. (1989) for evidence on commitment to self-set goals versus assigned goals.

intention-to-treat (ITT) effect by:

$$P_i = \beta_0 \ + \beta_1 G_i + \beta_2 R_i + \beta_3 X_i + \varepsilon_i$$

where P_i is student i's study performance, X_i a vector of control variables and ε_i the error term. G_i and R_i are treatment dummies indicating whether a student's mentor was assigned to the goal or raise treatment respectively. To be more precise on student performance, P_i is not the final grade of a student. The final grade for the course is composed of two midterm exams (both with weight 15%) and a final exam (with weight 70%). Since the mentor-student meeting is in the same week as the first midterm, students hardly change their study behavior for the first midterm, and so we expect the treatment to only affect the later exams of the course. Hence we take as student performance a normalized combination of the second midterm and the final exam. Our performance measure is hence calculated as (0.15*midterm2+0.7*final)/0.85.[14] The coefficients β_1 and β_2 are the intention-to-treat estimates of the effect of having a mentor who was assigned to treatment. The intention-to-treat effect is an imperfect measure of the effect of a student setting a goal or of attempts to raise that goal as there is bound to be some non-compliance. Not all students who are intended to get treated will get treated, for instance due to more pressing concerns in the meeting such as personal circumstances of the student. Likewise, although mentors not assigned to treatment are unaware of the nature of the treatment, some students of mentors that were assigned to the control group might self-treat by setting a goal and discuss this with the mentor. Thus while this estimate does not isolate the effect of setting and attempting to raise goals per se, it does provide an unbiased estimate of the intention to treat. Further, as the intention to treat effect relies on assigned, not actual, treatment non-compliance is likely to result in an underestimation of the actual effect.

Given that we ask treated mentors to report which students set them-

[14] The weights assigned to the second midterm and the final exam in our performance measure, 0.15 and 0.70 respectively, are the same as the weights used in the compostion of the final grade for students.

selves a goal and which students they asked to raise their goal, it may be tempting to directly compare students setting goals to those that do not. However, this would yield a misleading estimate of the effect of treatment if selection into or out of treatment is not random as it would compare individuals that are not ex ante identical. Instead we estimate the effect of treatment on those students whose treatment status is changed as a result of the experiment, also known as the local average treatment effect (LATE). In a first stage we regress actual treatment status on student characteristics and treatment assignment. Here treatment assignment serves as an instrument for actual treatment. Predicted treatment status then takes into account the observable characteristics of those in treatment and can be used in a second stage regression to explain study performance.

Thus the effect of setting a goal on those induced to set a goal may be estimated using as a first stage:

$$T_i = \varphi_0 + \varphi_1 G_i + \varphi_2 R_i + \varphi_3 X_i + \eta_i$$

where T_i indicates whether student i actually sets a goal, η_i is the error term and G_i and R_i are the dummy variables indicating whether the student's mentor was assigned to the goal or raise treatment respectively. This first stage is then followed by estimating:

$$P_i = \beta_0 + \beta_1 \widehat{T}_i + \beta_2 X_i + \varepsilon_i$$

Here again P_i is student i 's study performance, X_i a vector of control variables and ε_i the error term. The effect of interest is the coefficient β_1 of predicted treatment status $\widehat{T}_i$ following from the first stage regression.

There are two main reasons to include covariates in our regressions. First, since we assign treatment randomly conditional on the student's programme and teaching assistant we include dummies for the tutorial groups which subsume both these categories. Second, we include statistics on past study performance. Past study performance is highly predictive of present study performance and hence including measures of past performance re-

duces noise in the data, allowing for more precise estimates. We additionally include the student's gender, the mentor's gender and a dummy for the mentor's experience since, as will be discussed later, actual treatment depends on these variables to some extent. Since treatment is assigned at the mentor level and students' performance can be affected by something that is mentor specific (e.g. mentors' social skills), there is the possibility of confusing treatment effects with unobservable mentor level effects. To deal with this the best we can, we cluster the errors of the regressions at the mentor level.

We assign students that do not complete the course a failing grade for those grade components that they do not complete. By giving the highest and lowest possible failing grade we derive lower and upper bounds of the total effect of treatment.[15] The highest failing grade is a 4.4 and the lowest a 1.0 at a 1 to 10 scale. Students who score a 4.5 or higher can still pass the course by scoring well in other courses. In our results we will focus on the lowest grade as the lowest possible failing grade is the grade that is actually given to students who do not pass the course. Further, for context consider that those who do not take the final exam but do take the second midterm score a 1.5 on average compared to the overall average of 5.7. The total effect of our treatments that we measure in this manner is composed of an effect on the intensive margin and an effect on the extensive margin. We also provide separate estimates for the effect on study performance for those students who complete the course, and for the effects on course participation (demonstrating selection effects).

5 Descriptive Statistics

Our dataset contains information on 1092 students, 824 of whom are enrolled in a Dutch language programme with the remaining 268 students enrolled in an English language programme. Given that students are randomly as-

[15] By assigning the highest and lowest possible failing grade we get a lower and upper bound respectively because (as we will show later) the positive treatment effect conditional on completing the course is combined with a lower drop out rate.

signed to mentor groups at the start of the academic year and the mentors are randomly assigned to treatment and control we do not expect to find any ex ante differences between the two treatment groups and the control group. Table 1 gives the descriptive statistics for the control (C), goal (G) and raise (R) group, as well as giving the p-value for two-sided comparisons of the means of these groups. Although the control and treatment groups appear to be comparable, there are some differences between the groups. Specifically, the characteristics of mentors of students in the treatment and control groups differ.[16] Students in the control group are significantly more likely to have a female mentor whereas students in the raise treatment are more likely to have an experienced mentor.[17] Furthermore treatment students in a Dutch language economics track (as opposed to students in an English language economics track) scored lower for the 8 credits accounting course in the first block than students in the control group, but there is no such difference regarding the mathematics course, which is more important for microeconomics.[18] In the analysis we control for differences in observables.

At first sight Table 1 suggests that the treatments had no effect as there is no significant difference between the various groups in terms of the final grade received for the microeconomics course. However this simple direct comparison does not take into account the characteristics of students.

Selection into or out of treatment is an issue affecting the generalizability of the results to the whole population. In our experiment there are three sources of selection out of the treatment. First, despite our best efforts to get all mentors to cooperate and ensure their understanding of the instructions, not all mentors assigned to treatment applied the treatment or took notes when administering the treatment. There are seven mentors for whom we do not have data about what happened during the individual student-mentor meetings. Anecdotal evidence suggests that some mentors

[16] At the time of the randomization the information on mentor characteristics was not available to us. Hence we could not stratify our randomization on mentor characteristics.

[17] We define an experienced mentor as a mentor who mentored students in earlier years.

[18] In Table 1 we tested for differences between control and treatment groups using t-tests. We obtain similar results if we use nonparametric tests.

have administered the treatment but not recorded the results while others did not administer the treatment at all. Thus this missing data forms a combination of measurement error and selection out of treatment. On average these mentors have less experience (5 out of 7 have no experience) than other mentors assigned to treatment (25 out of 47). Six of these seven mentors were assigned to the raise treatment, which was somewhat more demanding for mentors as mentors were asked to challenge their students goals. Mentors may also feel more apprehension to administer treatment when their students had weak prior performance. However, compared to other students assigned to treatment, students of non-complying mentors do not differ in terms of prior performance.

The missing data on treatment administration has diverse effects on our estimated treatment effects. Estimation of an intention-to-treat effect requires knowledge of assigned treatment only. Thus, the missing data on treatment administration has no effect on our estimates of the intention to treat effect. However, the estimates of the treatment effect on the treated (LATE) are affected as those estimates require knowledge of treatment administration to students. The possibly non-random missing data, caused by mentors who did not administer the treatment, may lead to biased estimates of the LATE.

Second, there is some treatment dilution as mentors do not administer the treatment to all students. Mentors assigned to treatment ask students for their grade goal in 93% of the cases although they were instructed to administer the treatment to all. Moreover, mentors are selective in which students they target for treatment. Specifically, students who performed poorly in previous courses are less likely to be asked about their goals as is shown in Table 2. In cases in which mentors did not ask students about their goals they often noted a lack of time due to the necessity to discuss other issues. Also, conditional on receiving the data from the mentor, we find that more experienced mentors are less likely to administer treatment.[19] In the raise treatment, mentors were instructed to attempt to raise the student's

[19] Experienced mentors ask 90.5% of their students to set a goal, while non-experienced mentors ask 95% of their students to set a goal.

goal when he/she deemed the goal to be either too easy or doable. Of the 193 students setting a goal in the raise treatment 163 set a goal that met this requirement. However, mentors attempt to raise the goal in only 95 of these cases (58%), including all 47 cases where the goal is deemed too easy.[20] Overall students who are asked to raise their goal have slightly higher grades than those not asked, although differences are largely insignificant.[21] See Table 3 for more descriptives of the comparison between students asked and not asked to raise their goal.

Also 12% (59) of the 492 students asked for a goal do not set a goal. These are all students who previously did not have a goal in mind. While students are more likely to set a goal if they have a female mentor there are no significant differences between those setting and not setting a goal in terms of past performance as shown in Table 4. Of all students asked to raise their goal half accept a higher goal. Again there is no significant difference in terms of past study results, but students are less likely to accept a raise from more experienced mentors (p-value 0.03), see Table 5. Furthermore, the level of the initial goal set has no influence on the acceptance of a suggested raise of the goal.

It is of interest to note that 270 of the 492 students (55%) asked to set a goal already had a grade they wanted to achieve in mind. The fact that many students already have a grade goal in mind implies that any effect of our treatment comes either from those who previously did not have a goal, or from the fact that students make the goal known to their mentor. The average initial goal set by the student is 6.9, a histogram of the goals set is shown in Figure 3. As expected higher (lower) goals are more likely to be deemed too hard (too easy) to achieve for the student by their mentor. Mentors appear to be able to gauge goal difficulty, as a regression of the difference between the final grade achieved and the initial goal set on the estimate of the mentor shows in Table 6. Goals that were expected to be too

[20] In addition, there are 9 instances where the mentor asks a student to raise the goal even though she estimated the goal to be too difficult.

[21] The low number of observations for the Dutch econometrics courses is due to the fact that three of the four Dutch econometrics mentors assigned to the raise treatment failed to provide data.

difficult were not achieved on average whereas goals that were too easy are indeed beaten by a significant margin. Furthermore all point estimates of the judgment categories differ significantly from each other, indicating that mentors differentiate well between the three categories. On average students failed to meet their goal by 0.4 of a point. There may be a concern that mentors assigned to the raise treatment are more likely to report that they expect the student's goal te be too difficult, in order to avoid challenging the students to raise their goal. We test whether the distribution of the mentor's estimates of the students' goal differs across treatments. We do not find evidence of such an effect.

Of the students who were asked to raise their goal we see that 50 percent (52) rejects the goal the mentor proposed. The average goal proposed does not differ (two-sided p-value of 0.62) between students who accept and reject the goal.

Finally, there are some differences between students participating in Dutch language and international (English) programmes. The programmes are identical in all respects, featuring the same courses, study materials, and examinations, but have a somewhat different application procedure. Students applying for a Dutch language programme need to have only a high school diploma meeting the requirements, whereas those applying for an international programme are additionally selected based on a motivation letter. This suggests that students in the international language programmes might be better motivated. Consistent with this, Table 7 shows that, with the exception of an accounting course, students enrolled in international programmes scored significantly higher in the courses completed prior to the experiment. In our analysis we control for programme enrolment and additionally provide separate estimates of the effects within the Dutch and international programmes.

6 Results

We first provide the total effect of the treatments, imputing a failing grade for students who did not complete the course. We then provide the results

for students that complete the course before turning our attention to the results for students who did not complete the course.

6.1 Total effect

We estimate the total effect of the treatments by imputing the highest and lowest possible grade that would result in failing to pass the course for those graded aspects of the course that the student did not complete. As discussed in section 4, we focus on the case in which we impute a missing grade as 1.0, as this appears to be the most relevant case.

Table 8 gives the result of the intention to treat estimations. We find a weakly significant ($p < 0.1$) positive effect of 0.16 of a gradepoint (i.e. 9.3% of a standard deviation) for students in the goal treatment and an insignificant negative effect of the raise treatment. The positive effect of assignment to the goal treatment is in line with our hypothesis that setting goals improves student performance. The insignificant negative effect in the raise treatment shows that attempts to raise a goal backfire, resulting in performance similar to students in the control group. The results are estimated separately for men and women in columns 3 and 4 of Table 8. We see that men hardly respond to the treatments, and the effects are (mainly) driven by women. Women in the goal treatment show a substantial positive effect of 0.36, amounting to approximately 20% of a standard deviation of the grade. There is also a significant negative effect of -0.26 of a gradepoint of the raise treatment. Columns 5 and 6 of Table 8 provide separate estimates for students in Dutch and international programmes. These estimates show a widely divergent response to the treatments in these two groups, with Dutch programme students responding positively to the goal treatment while students in international programmes respond to the raise treatment in a strongly negative manner. Finally, in the second column of Table 8 we show the results excluding the students of mentors who did not hand in a form. These students are less likely to have been subjected to treatment, therefore the second column may give an indication of the effect size when compliance rates are higher than they were in our experiment.

Recall from the descriptive statistics section that students of mentors not handing in forms do not differ from those of other mentors in the treatment groups in observable characteristics. The positive effect of the goal treatment is stronger in these estimates while the effect of the raise treatment is again negative but insignificant. The overall picture that emerges from the intention to treat estimates is that setting goals can help, but attempting to raise goals undoes any positive effect of setting goals and may even result in worse performance as compared to students that were assigned to the control group.

The effects of actually having set a goal during the individual student-mentor meeting and of having been asked to raise the goal is given in Table 9. In the first column we provide the overall effect of having set a goal in the meeting without differentiating between treatments. This gives the relevant effect if one does not expect raising the goal set to have any effect. However there are differences between the effect of the goal and the raise treatment as shown in the second column. Setting a goal by itself has a significantly positive effect on student performance under the goal treatment. Setting a goal under the raise treatment has no such positive effect, the point estimate of the effect is even negative. In column 3 we compare students who set a goal in the two treatments with each other. Specifically column 3 shows that the difference between the goal and the raise treatment is due to the fact that in the raise treatment some students are asked to raise their goal, which on average lowers performance by a substantial 0.87 gradepoint (i.e. more than half a standard deviation) as compared to similar students who set a goal but were not asked to raise their goal. These results are in line with the intention to treat estimates, showing that setting a goal during the individual student-mentor meetings improves performance, but attempts to raise those goals undo that positive effect and may result in even lower performance.[22]

[22] To make sure that we compare students who are asked to raise their goal in the raise treatment with their counterparts in the goal treatment we controlled for the mentor's estimate about the difficulty of the students goal. There might be a concern that mentors in the raise treatment report a biased estimate because they want to avoid asking students to raise their goal. We tested for such an effect by comparing treatment effects for those

The total effect of our treatments discussed above consists of two effects. First the treatments may have an impact on the students who complete the course inducing them to alter their efforts. This is the effect on the intensive margin. Second, our treatments may affect the extensive margin, the decision to participate in the course. These two effects cannot be interpreted separately as this risks the confusion of selection effects on the extensive margin for the effects of treatment on the intensive margin. We turn to these two effects now.

6.2 Intensive margin

Table 10 gives the intention-to-treat estimates for those students that complete the course. The results are largely in line with the overall estimates provided in Table 8. We find that the overall positive effect of being assigned to the goal treatment is no longer significant for the students that complete the course. But results in the female and Dutch programme subsamples are similar and significant. The results confirm the overall impression that setting goals can improve student performance, and that raising goals has an insignificant negative effect on performance.

The effects of having actually set a goal during the individual student-mentor meeting and of having been asked to raise the goal is given in Table 11. Also these results are largely in line with the overall treatment effect, and show the positive impact of actually setting a goal which is negated by being asked to raise the goal.

6.3 Extensive margin

The results above indicate a positive effect of the goal treatment on course performance for those who complete the course. Differences between the

students of which the mentor reports that the goal is too difficult. If mentors in the raise treatment bias their estimate in order to avoid asking students to raise their goal we would expect the treatment effect to be more positive in the raise treatment than in the goal treatment. We do not find such an effect. In addition we compare the distribution of mentors' estimates about the goals across the treatments. Because of randomization the distributions should be similar. We find no evidence that the distribution differs across treatments.

total effect estimates and the estimates on the intensive margin may be due to selection effects induced by the treatments. For instance, our treatments may affect course completion by creating greater commitment. To study the effects of our treatments on course completion we estimate a linear probability model in much the same way as above. Given that students have to attend at least 10 out of 13 tutorial sessions we limit our estimation to the sample of students who attended at least 3 sessions such that all students were still able to meet this requirement by the time the treatment took place. In the control group 6.2 percent (24) of the students having attended at least 3 sessions dropped out of the course. This dropout rate is lowered by 2 percentage points on average in the goal treatment as can be seen in Table 12 providing the intention to treat estimates on the dropout rate.[23] The results on the dropout rate are similar to those on course performance given course completion in that the goal treatment again has an effect whereas the raise treatment does not have an effect but has an oppositely signed coefficient. In contrast to the effect on the course grade however, the reduction in dropouts is concentrated among men rather than women. This is most likely due to the fact that women have a substantially lower baseline dropout rate than men (6.8% for men compared to 2.8% for women in the control group, two-sided p-value 0.12).

The pattern that the goal treatment has desirable effects while the raise treatment has undesirable effects continues to hold when considering the effect on those who actually set goals as shown in Table 13. Those in the goal treatment who actually set a goal show a 3.7 percentage point lower dropout rate than those in either the control or the raise treatment. Among those students who set a goal, those who are asked to raise their goal are 12 percentage points more likely to drop out than those who are not asked to raise their goal. Even taking into account that the comparison group consists of students with a lower dropout rate of 3.9 percentage point due to the effect of setting a goal itself on the dropout rate discussed above this is a sizeable effect. This result shows that attempts to raise a student's goal

[23] No estimates on subsamples for the Dutch and international programmes are provided as there are too few dropouts from the international programmes, resulting in collinearity.

in order to improve performance can backfire by leading to a substantially lower chance of course completion.

6.4 Further results

In this section we present a number of additional analyses that shed light on further questions. To start, we consider whether there is a heterogeneous effect of treatment due to differences in ability. We measure ability by taking the average of the grades achieved in the first block, and centering this grade average by subtracting the overall mean average score of 6.2 (std. dev. 1.65). We then interact the ability measure with students' treatment assignment. The intention-to-treat estimates in Table 14 show that students who performed better in previous courses respond less to the goal treatment. Thus our intervention had a stronger effect on weaker students than it did on top students. There is no such heterogeneous effect regarding the raise treatment.[24]

Second, as we have seen above, students set higher goals in front of a female mentor (two-sided p-value of 0.012). This suggests that the motivation to set a goal may differ depending on the gender of the mentor. It is thus natural to ask whether our treatments have heterogeneous effects depending on the mentor's gender. Overall there is no sign of a heterogeneous treatment effect depending on the gender of the administering mentor as shown in Table 15.

The desire to impress a member of the opposite gender may result in a heterogeneous treatment effect based on the gender combination of mentor and student. Estimates on male or female students reveal no such significant interaction effect, although the signs of the point estimates do conflict in the expected manner which is consistent with such an effect. In the subsample of Dutch students there is a clear interaction effect of gender on treatment response. The point estimates suggest that having a female mentor completely reverses the effect of the treatments such that the raise treatment

[24] For students who attended a Dutch high school we also have high school grades. If we use high school grades (as a measure of ability) for this subsample we find a qualitatively similar result but a decrease in power because of the smaller sample size.

has an insignificant negative impact when administered by a male mentor but a positive impact when administered by a female mentor (p-value of difference between male and female mentor administering the raise treatment is 0.002). Male and female mentors have the same effect on Dutch programme students in the goal treatment. Furthermore, no gender differences are found in the sample of international students, although this may be due to a lack of power.

Another possible channel through which treatment may be affected is the experience of the mentor. More experienced mentors may project more authority or may be better able to fit the treatment into the conversation in a natural manner. A heterogeneous treatment effect based on the mentor's experience may also speak to a possible channel through which setting a goal in the individual mentor-student meeting affects performance. If students care for their reputation in the eyes of their mentor they may value this reputation more if the mentor is more experienced. Taken combined, we expect more experienced mentors to strengthen the treatment effects. Recall however that we have already seen some evidence speaking against a better implementation of the treatments by more experienced mentors in the form of a weakly lower rate of inducement (two-sided p-value of 0.06) and goals that are approximately three tenths of a point lower.

The estimation results including a heterogeneous effect for experienced mentors is given in Table 16. We see that more experienced mentors do not have a significant positive impact on the raise treatment overall. Men in the goal treatment with a more experienced mentor show an increase in performance whereas there is no such increase in performance when they have an inexperienced mentor. Female students in the goal treatment seem to do worse with an experienced mentor compared to an unexperienced mentor, although this effect is not significant. This may suggest that men respond more to authority. Although there appears to be a significant negative interaction effect of mentor experience on assignment to the goal treatment for international students, the effect of having a more experienced mentor implementing the treatment is no different from that of an unexperienced mentor (p-value of 0.21). These results show that generally the students do

not appear to be more responsive to more experienced mentors. Although we suggested that students might be more responsive to a more experienced mentor due to reputational concerns, this finding does not discredit the reputation explanation for why goals can improve performance. For instance, students may not distinguish that much between different mentors, or value their reputation in the eyes of their mentor more based on other factors such as the mentor's own ability.

We are fully aware that there are two different effects of setting a goal in our goal treatment. The effect of thinking about a goal and setting a goal for oneself, and the effect of sharing this goal with the mentor. Thinking about a goal may create a reference point with which to compare one's performance as posited in the literature. This may result in higher performance by giving positive utility if the goal is met or exceeded and negative utility if performance falls short. Sharing the goal with the mentor may result in reputational concerns towards the mentor, in the sense that the mentor may evaluate the student based on her performance relative to her goal. Empirically our best way to learn whether sharing the goal with the mentor leads to the increased performance, is by considering when reputational concerns are likely to be stronger and comparing students in those situations with similar students who face weaker reputational concerns. We posit two such situations, facing a more experienced mentor who is likely better in evaluating students performance, and facing a mentor of the opposite gender who may be more valuable to impress. For neither of these situations we find a heterogeneous impact on the treatment effect, suggesting that goals primarily work through creating a reference point for oneself. Of course our proxies may not capture enough variation in the strength of reputational concerns.

Furthermore, asking students to raise their goal can have effects through two different channels. The challenge of the goal itself can have an effect, i.e. if students get utility from reaching a goal they may exert additional effort if the goal becomes more challenging. Another channel is that asking the student to raise her goal can give the student information about her ability. Depending on the substitutability or complementarity of ability and

effort this can lead to a decrease or increase in effort and performance. Here too, our best way to deal with this is comparing the treatment effects for students with experienced mentors with the treatment effects for students with unexperienced mentors. The reason is that we expect that students who have experienced mentors expect their mentor to be more able to know the student's ability, and hence the information aspect is stronger. We find that there is no significant difference in treatment effects for students who have experienced and unexperienced mentors. This implies that either the information component is small, or that there is not much heterogeneity in the extend to which the message contains information by experience of the mentor.

We find a consistent pattern in our data that female students respond stronger to our treatments than male students. At first sight this may seem surprising. There is a very rich literature on heterogeneous gender effects to monetary incentives and non-monetary incentives when there is a competition element. Many papers find that males respond stronger to competitive incentives or information about their (relative) ranking, see for example Gneezy et al. (2003), Barankay (2011), and Niederle and Vesterlund (2011), while others find that there is no gender difference, see Dreber et al. (2011) and Delfgaauw et al. (2013). An important difference between these incentives and the incentives in our treatments is that goal setting in this experiment does not have any competitive element, which might drive the gender effect in the current literature.

Finally, a potential concern of goal setting in multitasking environments is that goals lead subjects away from other (non-incentivized) tasks. Our students take one other course at the same time as microeconomics. Doing our previous analyses with the other course's grade as a dependent variable shows that there is no evidence of such a substitution effect (these results are available on request). Hence, the positive effect of goal setting in this study is a net increase in performance.

7 Discussion and Conclusion

We conducted a field experiment in order to test the effects of encouraging students to set goals and encouraging students to increase the ambitiousness of their goals during mentor-student meetings in a university study programme. We designed two treatments. In the *goal treatment* we instructed mentors to encourage students to set a grade specific goal. In the *raise treatment* we gave mentors the same instruction and in addition instructed them to raise this goal if deemed appropriate.

We find that students in the goal treatment perform better than students in the control group. Students whose mentor was assigned to the goal treatment score 0.16 gradepoints (i.e. 9.3% of a standard deviation) higher than students in the control group. Students in the raise treatment perform similarly to students in the control group, although there are some indications that their performance is even lower. This is true in terms of both the dropout rate and the grades achieved that are conditional on completing the course. The null effect of the raise treatment is in line with the goal becoming unacceptable due to the raise, indicating that the size of the raise was too high. Finally, being asked to raise the goal leads to a significant drop in performance as compared to similar students in the goal treatment.

An alternative explanation for the result that students in the raise treatment perform worse than students in the goal treatment is the nature of the goal. While in the goal treatment students set themselves a goal, a proposal to raise this self-set goal can be seen as a goal of a different kind, namely a cooperatively set goal (or even an assigned goal). Changing the nature of the goal can change the commitment of the student to the goal (see Hollenbeck et al. 1989), which implies that the intrinsic motivation (i.e. the utility gain when reaching the goal) changes between the two treatments. As a consequence students perform worse in the raise treatment than in the goal treatment. Further, if some students in the control group set themselves a goal then this could even lead to a lower performance of students in the raise treatment as compared to the control group.[25]

[25] Our finding that 55% of students that are asked about goals already have a goal in

Next, we looked at heterogeneous treatment effects. First, we find that students that performed poorly prior to the experiment benefit most from setting goals. Second, we expected stronger effects for students who are assigned to more experienced mentors. The reason for this is that we expect experienced mentors to be better able to incorporate our treatments in their meetings, and because these mentors might have more authority. We find that overall there is no heterogeneous effect of experience on the treatments. There is an effect on male students suggesting male students are more affected by authority. This however does not imply that overall reputational concerns of students towards their mentor do not play a role, as students may care more about (unobservable) characteristics other than experience of their mentor. Third, students' motivation to set a goal might differ by the gender of their mentor, as well as whether student and mentor are of the same gender. Overall we find that there is no effect of a mentor's gender on the treatment effects. More surprisingly, if we focus only on the students in Dutch education programmes, we see that the students assigned to the raise treatment perform better when their mentor is female, and worse when their mentor is male as compared to the control group. This might indicate that female mentors are more able to motivate the initially less motivated students by challenging them than male mentors.

To summarize we have shown that students setting and sharing goals with their mentors can help raise study performance, and more so for initially poorly performing students. We have furthermore shown that although it may be tempting to try to push students to raise the bar a bit higher, raising the bar can be more demotivating than inspiring. Hence, one should be cautious in attempting to push students beyond what they themselves aim to achieve, even if that bar appears to be low.

It is interesting to learn whether these results are generalizable to other settings, for example manager-worker settings. Our findings may have implications for the optimal design of appraisal interviews. In order to generalize our findings we need more (experimental) evidence in other settings.

Finally, our paper is (relatively) silent on the mechanisms that drive goal

mind supports this idea.

setting. It is interesting to learn to what extent present bias preferences and loss aversion, as is posited in economic theory papers as important drivers that make goals work, are predictors of the success of goal setting. For example our result that goal setting works mostly for initially poor performing students may be explained by poor performing students having stronger present bias preferences, are more loss averse, or by the fact that extra effort more easily increases performance when initial performance is low. One way to test how the effect of goal setting on performance interacts with people's present bias preferences and loss aversion is by running a laboratory experiment with a sample of the students that participated in this study.

References

[1] Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80(3), 260.

[2] Angrist, J.D., & Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics*, 114(2), 533-575.

[3] Bandiera, O., Larcinese, V., & Rasul, I. (2010). Heterogeneous class size effects: New evidence from a panel of university students. *The Economic Journal*, 120(549), 1365-1398.

[4] Bandiera, O., Larcinese, V., & Rasul, I. (2015). Blissful Ignorance? A Natural Experiment on the Effect of Feedback on Students.' Forthcoming in *Labour Economics*.

[5] Barankay, I. (2011). Gender differences in productivity responses to performance rankings: Evidence from a randomized workplace experiment. *Unpublished working paper, The Wharton School.*

[6] Bettinger, E.P., & Baker, R.B. (2013). The Effects of Student Coaching An Evaluation of a Randomized Experiment in Student Advising. *Educational Evaluation and Policy Analysis.*

[7] Clark, D., Gill, D., Prowse, V.L., & Rush, M. (2016). Using Goals to Motivate College Students: Theory and Evidence from Field Experiments.

[8] Corgnet, B., Gómez-Miñambres, J., & Hernán-Gonzalez, R. (2015). Goal Setting and Monetary Incentives: When Large Stakes Are Not Enough. *Management Science*, 61(12), 2926-2944.

[9] Corgnet, B., Gomez-Minambres, J., & Hernan-Gonzalez, R. (2016). Goal Setting in the Principal-Agent Model: Weak Incentives for Strong Performance. Working Paper.

[10] Dalton, P.S., Gonzalez Jimenez, V.H., & Noussair, C.N. (2015). Paying with Self-Chosen Goals: Incentives and Gender Differences. (CentER Discussion Paper; Vol. 2015-021). Tilburg: Economics.

[11] Delfgaauw, J., Dur, R., Sol, J., & Verbeke, W. (2013). Tournament Incentives in the Field: Gender Differences in the Workplace. *Journal of Labor Economics*, 31(2), 305-326.

[12] Dreber, A., Von Essen, E., & Ranehill, E. (2011). Outrunning the gender gap—boys and girls compete equally. *Experimental Economics*, 14(4), 567-582.

[13] Gibbons, R. (2005). Incentives between firms (and within). *Management Science*, 51(1), 2-17.

[14] Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics*, 118(3), 1049-1074.

[15] Goerg, S.J., & Kube, S. (2012). Goals (th) at Work–Goals, Monetary Incentives, and Workers' Performance. *MPI Collective Goods Preprint*, (2012/19).

[16] Gómez-Miñambres, J. (2012). Motivation through goal setting. *Journal of Economic Psychology*, 33(6), 1223-1239.

[17] Hollenbeck, J.R., Williams, C.R., & Klein, H.J. (1989). An empirical examination of the antecedents of commitment to difficult goals. *Journal of Applied Psychology*, 74(1), 18.

[18] Hsiaw, A. (2013). Goal-setting and self-control. *Journal of Economic Theory*, 148(2), 601-626.

[19] Koch, A.K., & Nafziger, J. (2011). Self-regulation through Goal Setting. *The Scandinavian Journal of Economics*, 113(1), 212-227.

[20] Koch, A., Nafziger, J., & Nielsen, H.S. (2014). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115, 3-17.

[21] Koch, A.K., Nafziger, J., Suvorov, A., & Van de Ven, J. (2014). Self-rewards and personal motivation. *European Economic Review*, 68, 151-167.

[22] Latham, G.P., & Locke, E.A. (1979). Goal setting—A motivational technique that works. *Organizational Dynamics*, 8(2), 68-80.

[23] Lavecchia, A.M., Liu, H., & Oreopoulos, P. (2015). Behavioral Economics of Education: Progress and Possibilities (No. 8853). Institute for the Study of Labor (IZA).

[24] Levitt, S.D., List, J.A., Neckermann, S., & Sadoff, S. The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *American Economic Journal: Economic Policy,* forthcoming.

[25] Linnenbrink, E.A. (2005). The Dilemma of Performance-Approach Goals: The Use of Multiple Goal Contexts to Promote Students' Motivation and Learning. *Journal of Educational Psychology*, 97(2), 197-213.

[26] Locke, E.A. (1996). Motivation through conscious goal setting. *Applied and Preventive Psychology*, 5(2), 117-124.

[27] Locke, E.A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705-717.

[28] Locke, E.A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science,* 15(5), 265-268.

[29] Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of Applied Psychology*, 95(2), 255-264.

[30] Niederle, M., & Vesterlund, L. (2011). Gender and Competition. *Annual Review of Economics,* 3(1), 601-630.

[31] Ordóñez, L.D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *The Academy of Management Perspectives*, 23(1), 6-16.

[32] Sackett, A.M., Wu, G., White, R.J., & Markle, A. (2014). Harnessing optimism: How eliciting goals improves performance. Working Paper.

[33] Sanders, M., & Chonaire, A.N. (2015). "Powered to Detect Small Effect Sizes": You keep saying that. I do not think it means what you think it means (No. 15/337). Department of Economics, University of Bristol, UK.

[34] Schippers, M.C., Scheepers, A.W., & Peterson, J.B. (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. *Palgrave Communications*, 1.

[35] Schunk, D.H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist,* 25(1), 71-86.

[36] Seijts, G.H., Latham, G.P., Tasa, K., & Latham, B.W. (2004). Goal setting and goal orientation: An integration of two different yet related literatures. *Academy of Management Journal*, 47(2), 227-239.

[37] Suvorov, A., & Van de Ven, J. (2008). Goal setting as a self-regulation mechanism. Working Paper.

[38] Travers, C.J., Morisano, D., & Locke, E.A. (2015). Self-reflection, growth goals, and academic outcomes: A qualitative study. *British Journal of Educational Psychology*, 85(2), 224-241.

Tables

Table 1: Descriptives by assigned treatment

	Control		Goal		Raise		C - G	C - R	G - R
	mean	sd	mean	sd	mean	sd	p	p	p
Age	18.72	1.71	18.63	1.39	18.75	1.27	0.57	0.85	0.39
Female	0.29	0.46	0.30	0.46	0.30	0.46	0.87	0.95	0.93
Female mentor	0.52	0.50	0.31	0.46	0.34	0.47	0.00***	0.00***	0.46
Experience mentor	0.31	0.46	0.38	0.49	0.50	0.50	0.03**	0.00***	0.00***
EC Accounting	6.45	1.61	6.13	1.51	5.96	1.73	0.06*	0.01**	0.31
ECX Accounting	5.82	1.72	5.95	1.91	5.74	1.92	0.70	0.79	0.56
EC Math	5.98	1.56	5.87	1.43	5.91	1.63	0.49	0.70	0.79
ECX Math	6.48	1.51	6.64	1.63	6.66	1.38	0.58	0.51	0.96
ET Matrix Alg.	6.48	1.94	6.13	1.88	6.02	2.38	0.32	0.29	0.79
ETX Matrix Alg.	7.54	2.11	6.92	1.81	7.58	2.18	0.28	0.95	0.31
ET Precalculus	6.46	1.50	6.04	1.78	5.72	2.20	0.17	0.05**	0.41
ETX Precalculus	7.42	1.79	6.54	2.29	7.25	1.98	0.13	0.74	0.31
ET Statistics	5.09	2.07	4.45	1.68	4.43	2.23	0.07*	0.13	0.96
ETX Statistics	6.35	2.51	5.50	2.31	5.29	2.17	0.23	0.13	0.78
Microeconomics	6.48	1.76	6.48	1.62	6.29	1.89	0.95	0.19	0.20
Midterm I	4.89	2.25	4.80	2.10	4.78	2.23	0.58	0.55	0.94
Attendance	10.66	2.98	10.66	3.15	10.71	2.85	0.98	0.81	0.84
Dropout	0.14	0.35	0.11	0.31	0.14	0.34	0.21	0.94	0.26
Max. Observations	389		367		336				

Students enroll in an economics (EC) track or an econometrics (ET) track in a Dutch or international (X) programme. Tracks in the Dutch and international programme are identical. Different tracks feature different courses, although some courses (e.g. Microeconomics) are common to all tracks.

Table 2: Students asked and not asked to set a goal within the treatment groups

	Not asked to set goal		Asked to set goal		
	Mean	Observations	Mean	Observations	p-value
Age	18.50	26	18.71	308	0.43
Female	0.30	37	0.27	483	0.69
Female mentor	0.22	37	0.34	492	0.14
Exp. mentor	0.62	37	0.46	492	0.06
EC Math	5.02	19	5.87	294	0.02
ECX Math	5.60	6	6.71	78	0.08
ET Precalculus	4.60	3	6.20	64	0.14
ETX Precalculus	5.50	2	7.43	22	0.24
EC Accounting	4.55	19	6.12	293	0.00
ECX Accounting	4.85	6	5.88	77	0.20
ET Matrix Alg.	4.47	3	6.35	63	0.10
ETX Matrix Alg.	6.70	2	7.85	22	0.46
ET Statistics	2.77	3	4.83	61	0.06
ETX Statistics	3.50	2	5.95	22	0.14

See Table 1 for an explanation of terms.

Table 3: Students asked and not asked to raise their goal in the raise treatment

	No raise proposed		Raise proposed		
	Mean	Observations	Mean	Observations	p-value
Age	18.79	84	18.59	75	0.29
Female	0.26	127	0.25	100	0.87
Female mentor	0.29	128	0.38	104	0.17
Exp. mentor	0.58	128	0.45	104	0.06
EC Math	5.63	70	6.04	74	0.13
ECX Math	6.32	21	7.24	17	0.03
ET Precalculus	6.26	7	6.97	3	0.73
ETX Precalculus	7.31	13	7.88	4	0.63
EC Accounting	5.69	69	6.24	75	0.06
ECX Accounting	5.25	21	6.31	17	0.10
ET Matrix Alg.	6.31	7	6.40	3	0.97
ETX Matrix Alg.	7.58	13	8.80	4	0.30
ET Statistics	5.36	7	7.40	3	0.22
ETX Statistics	5.44	13	6.47	4	0.37

See Table 1 for an explanation of terms.

Table 4: Students that set and do not set goals when asked to set a goal

	Set Goal		No goal set		
	Mean	Observations	Mean	Observations	p
Age	18.74	265	18.58	43	0.48
Female	0.27	424	0.27	59	0.94
Female Mentor	0.35	433	0.22	59	0.05
Exp. Mentor	0.47	433	0.41	59	0.39
EC Math	5.85	254	5.99	40	0.60
ECX Math	6.67	73	7.16	5	0.49
ET Precalculus	6.28	57	5.57	7	0.32
ETX Precalculus	7.62	20	5.55	2	0.19
EC Accounting	6.12	252	6.17	41	0.85
ECX Accounting	5.96	72	4.78	5	0.18
ET Matrix Alg.	6.28	57	7.00	6	0.39
ETX Matrix Alg.	7.89	20	7.40	2	0.74
ET Statistics	4.90	55	4.18	6	0.38
ETX Statistics	6.03	20	5.20	2	0.62
Midterm I	4.77	426	4.47	55	0.32

See Table 1 for an explanation of terms.

Table 5: Students that accept and reject raise when asked to raise the goal

	Reject raise		Accept raise		
	Mean	Observations	Mean	Observations	p-value
Age	18.53	38	18.65	37	0.60
Female	0.22	50	0.28	50	0.49
Female mentor	0.46	52	0.29	52	0.07
Exp. mentor	0.56	52	0.35	52	0.03
EC Math	5.93	37	6.15	37	0.57
ECX Math	6.78	9	7.75	8	0.10
ET Precalculus	4.70	1	8.10	2	.
ETX Precalculus	8.85	2	6.90	2	0.34
EC Accounting	6.52	38	5.95	37	0.13
ECX Accounting	6.12	9	6.51	8	0.70
ET Matrix Alg.	4.60	1	7.30	2	.
ETX Matrix Alg.	9.90	2	7.70	2	0.33
ET Statistics	6.10	1	8.05	2	.
ETX Statistics	7.25	2	5.70	2	0.54
Initial goal	6.60	52	6.53	52	0.62

See Table 1 for an explanation of terms.

Table 6: Mentor estimates are accurate

	grade - goal
Goal too hard	-1.507***
	(0.261)
Goal too easy	0.337***
	(0.412)
Baseline: doable	-0.446***
	(0.261)
Observations	406
R^2	0.111

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 7: Previous course results in Dutch and international programmes

	International		Dutch		
	Mean	n	Mean	n	p
EC Math	6.58	172	5.92	531	0.00
EC Accounting	5.84	170	6.18	529	0.02
ET Matrix Alg.	7.39	72	6.23	162	0.00
ET Statistics	5.84	71	4.67	158	0.00
ET Precalculus	7.14	71	6.10	164	0.00

See Table 1 for an explanation of terms.

Table 8: Intention to treat effect: total effect

Grade	Overall	Excl. 7 mentors	Male	Female	Dutch	International
Missing grade=1.0						
T: Goal	0.164*	0.256***	0.140	0.359**	0.333***	-0.102
	(0.090)	(0.088)	(0.110)	(0.172)	(0.115)	(0.119)
T: Raise	-0.156	-0.113	-0.0181	-0.262*	0.0179	-0.421**
	(0.101)	(0.102)	(0.126)	(0.148)	(0.117)	(0.154)
Missing grade=4.4						
T: Goal	0.099	0.167**	0.074	0.303*	0.251**	-0.102
	(0.079)	(0.079)	(0.092)	(0.159)	(0.095)	(0.119)
T: Raise	-0.138	-0.110	-0.017	-0.227	0.040	-0.421**
	(0.089)	(0.089)	(0.103)	(0.146)	(0.094)	(0.154)
Observations	955	868	678	277	719	236
Tutorgroups	84	77	84	81	60	24
R^2 (Upper)	0.619	0.630	0.621	0.691	0.615	0.641
R^2 (Lower)	0.616	0.624	0.621	0.695	0.610	0.641

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 9: Local average treatment effect: total effect

Grade	Overall	By Treatment	AskedToRaise
Missing grade=1.0			
Set Goal	0.027		
	(0.113)		
Set Goal T: Goal		0.257**	
		(0.118)	
Set Goal T: Raise		-0.263	
		(0.164)	
Asked to Raise			-0.870***
			(0.334)
Missing grade=4.4			
Set Goal	-0.014		
	(0.101)		
Set Goal T: Goal		0.164	
		(0.104)	
Set Goal T: Raise		-0.238	
		(0.146)	
Asked to Raise			-0.492
			(0.316)
Controls	yes	yes	yes
Instruments	T. assignment	T. assignment	R. assignment
Observations	966	966	411
Tutorgroups	84	84	47
R^2 (Upper)	0.613	0.618	0.639
R^2 (Lower)	0.610	0.613	0.634

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 10: Intention to treat effect: intensive margin

Grade	Overall	Excl. 7 mentors	Male	Female	Dutch	International
T: Goal	0.112	0.193**	0.069	0.361**	0.265***	0.011
	(0.078)	(0.077)	(0.093)	(0.178)	(0.091)	(0.125)
T: Raise	-0.147	-0.117	-0.007	-0.262*	0.054	-0.428**
	(0.092)	(0.095)	(0.103)	(0.151)	(0.091)	(0.177)
Controls	yes	yes	yes	yes	yes	yes
Observations	940	854	661	279	697	243
Mentorgroups	84	77	84	82	60	24
R^2	0.609	0.616	0.626	0.674	0.605	0.627

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 11: Local average treatment effect: intensive margin

Grade	Overall	By Treatment	Asked to Raise
Set Goal	0.072		
	(0.088)		
Set Goal T: Goal		0.251***	
		(0.096)	
Set Goal T: Raise		-0.146	
		(0.117)	
Asked to Raise			-0.409
			(0.292)
Controls	yes	yes	yes
Instruments	T. assignment	T. assignment	R. assignment
Observations	854	854	408
Tutorgroups	77	77	47
R^2	0.614	0.617	0.631

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 12: Intention to treat effect: dropout

Dropout	Overall	Excl. 7 mentors	Male	Female
T: Goal	-0.020**	-0.029***	-0.0215**	-0.016
	(0.008)	(0.008)	(0.010)	(0.014)
T: Raise	0.005	0.0005	0.00002	0.009
	(0.008)	(0.008)	(0.012)	(0.012)
Observations	955	868	678	277
Tutorgroups	84	77	84	81
R^2	0.169	0.183	0.175	0.269

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 13: Local average treatment effect: dropout

Dropout	Overall	By Treatment	Asked to Raise
Set Goal	-0.020**		
	(0.008)		
Set Goal T: Goal		-0.037***	
		(0.010)	
Set Goal T: Raise		0.0005	
		(0.010)	
Asked to Raise			0.120***
			(0.017)
Controls	yes	yes	yes
Instruments	T. assignment	T. assignment	R. assignment
Observations	868	868	416
Tutorgroups	77	77	47
R^2	0.176	0.183	0.214

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 14: ITT interaction treatment with GPA in preceding block

	Overall	Excl. 7 mentors	Male	Female	Dutch	International
T: Goal	0.154**	0.177**	0.046	0.598***	0.274***	0.196**
	(0.077)	(0.085)	(0.095)	(0.157)	(0.094)	(0.092)
T: Goal × GPA	-0.156**	-0.147**	-0.187*	-0.260**	-0.152	-0.182*
	(0.071)	(0.071)	(0.097)	(0.100)	(0.093)	(0.105)
T: Raise	-0.084	-0.041	-0.007	0.011	0.082	-0.314**
	(0.092)	(0.100)	(0.114)	(0.175)	(0.089)	(0.149)
T: Raise × GPA	0.003	0.039	0.034	-0.109	-0.031	0.089
	(0.075)	(0.091)	(0.105)	(0.106)	(0.092)	(0.118)
Block 1	0.748***	0.754***	0.748***	0.784***	0.754***	0.743***
	(0.043)	(0.043)	(0.058)	(0.065)	(0.056)	(0.072)
Observations	896	815	631	265	661	235
Tutorgroups	84	77	84	80	60	24
R^2	0.490	0.505	0.502	0.568	0.486	0.509

GPA is the mean centered GPA of students calculated over all courses they participated in prior to microeconomics. Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 15: ITT interaction treatment with gender of mentor

Grade	Overall	Excl. 7 mentors	Male	Female	Dutch	International
Female mentor	-0.0002	-0.010	-0.053	-0.146	0.159	-0.105
	(0.134)	(0.136)	(0.148)	(0.300)	(0.113)	(0.282)
T: Goal	0.171	0.218*	0.002	0.554**	0.387***	0.007
	(0.113)	(0.117)	(0.128)	(0.262)	(0.082)	(0.156)
T: Goal × Fem. mentor	-0.198	-0.083	0.085	-0.711	-0.520**	0.002
	(0.228)	(0.212)	(0.241)	(0.535)	(0.258)	(0.427)
T: Raise	-0.178	-0.131	-0.121	-0.430**	-0.091	-0.186
	(0.111)	(0.121)	(0.133)	(0.214)	(0.092)	(0.271)
T: Raise × Fem. mentor	0.070	0.033	0.262	0.336	0.449**	-0.487
	(0.178)	(0.169)	(0.212)	(0.395)	(0.205)	(0.292)
Observations	940	854	661	279	697	243
Tutorgroups	84	77	84	82	60	24
R^2	0.609	0.616	0.626	0.678	0.609	0.629

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 16: ITT interaction treatment with experience of mentor

Grade	Overall	Excl. 7 mentors	Male	Female	Dutch	International
T: Goal	0.067	0.202*	-0.142	0.416*	0.124	0.123
	(0.100)	(0.102)	(0.124)	(0.225)	(0.140)	(0.091)
Experienced Mentor	0.087	0.047	-0.172	0.544*	0.052	0.760**
	(0.137)	(0.125)	(0.133)	(0.280)	(0.124)	(0.303)
T: Goal × Exp. Mentor	0.129	-0.032	0.606***	-0.250	0.292	-1.217***
	(0.200)	(0.192)	(0.229)	(0.424)	(0.231)	(0.348)
T: Raise	-0.040	-0.088	0.124	-0.075	0.141	-0.254
	(0.132)	(0.137)	(0.144)	(0.264)	(0.137)	(0.268)
T: Raise × Exp. Mentor	-0.198	-0.073	-0.091	-0.502	-0.136	-0.862
	(0.167)	(0.159)	(0.183)	(0.433)	(0.188)	(0.557)
Observations	940	854	661	279	697	243
Tutorgroups	84	77	84	82	60	24
R^2	0.610	0.616	0.629	0.676	0.606	0.630

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Figures

Figure 1

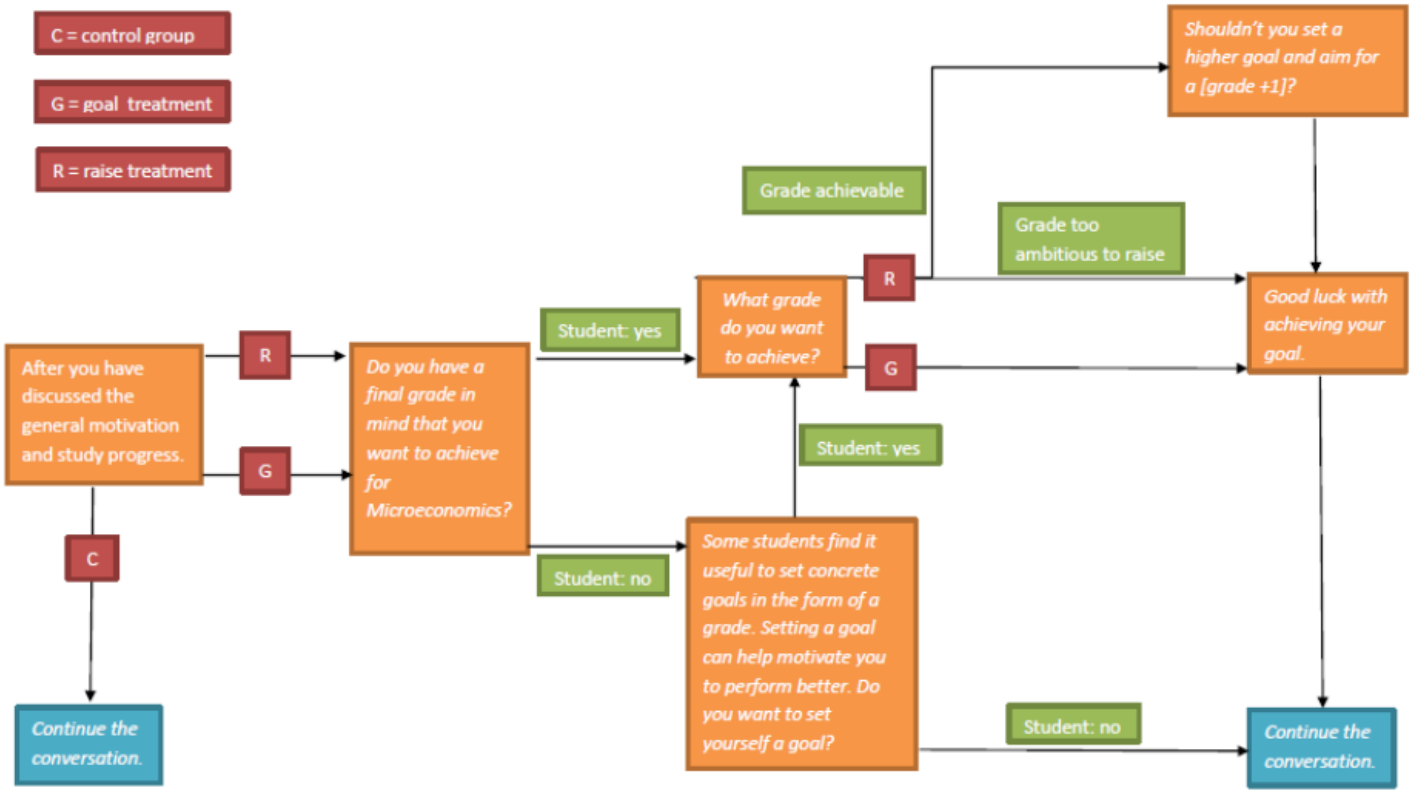


Figure 2

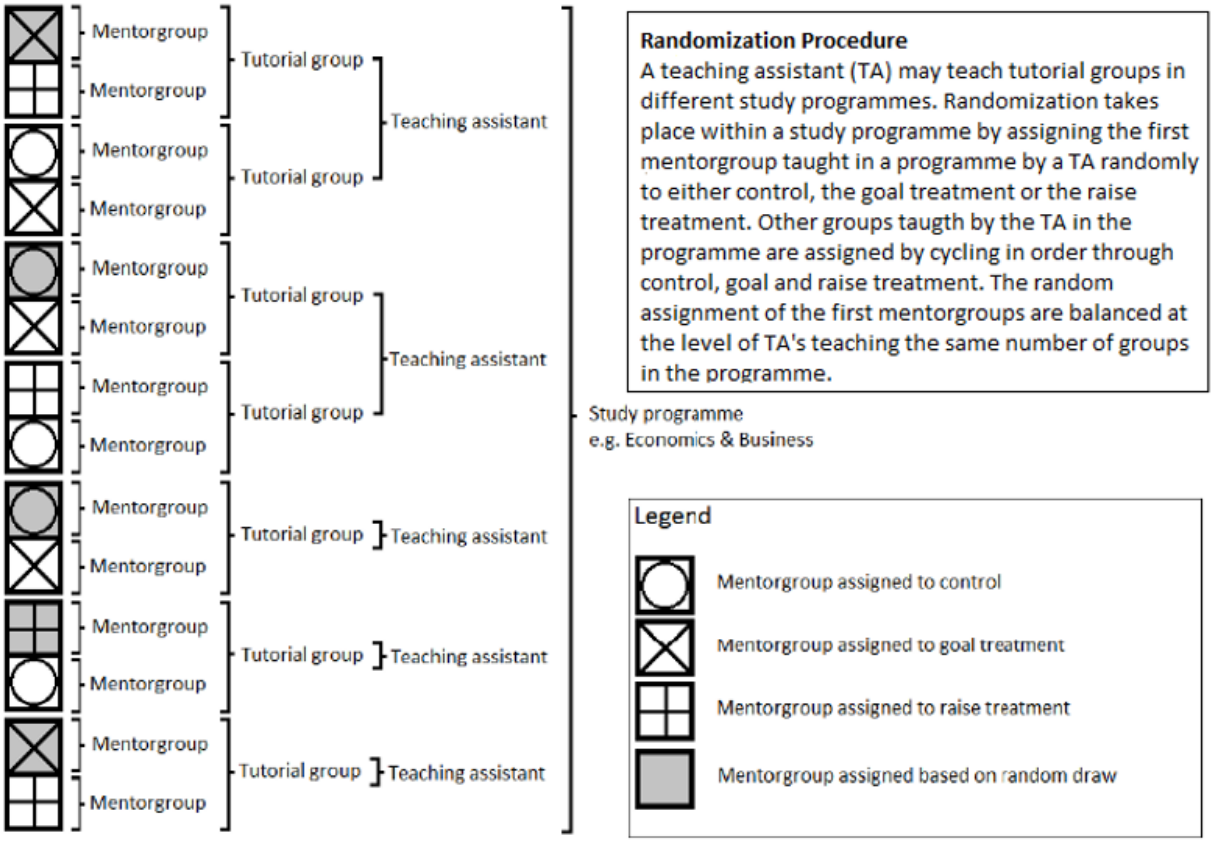
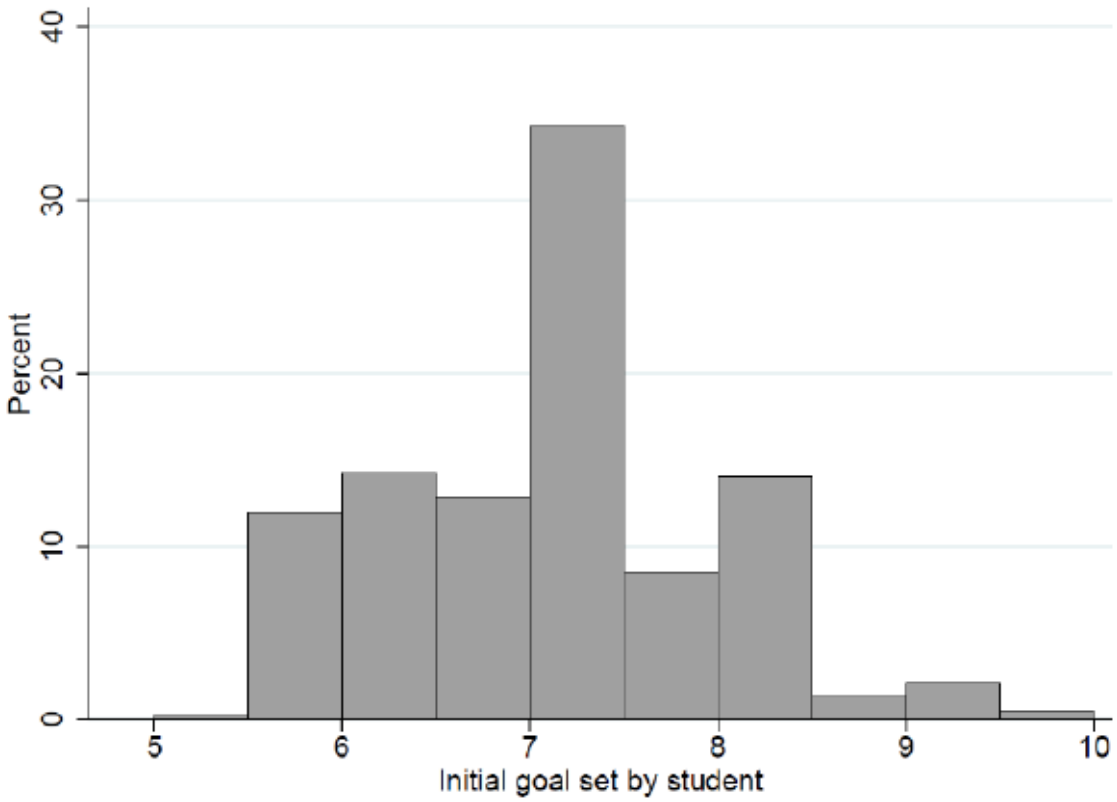

Mentorgroup
Mentorgroup
Tutorial group
Teaching assistant
Mentorgroup
Mentorgroup
Tutorial group
Mentorgroup
Mentorgroup
Tutorial group
Teaching assistant
Mentorgroup
Mentorgroup
Tutorial group
Mentorgroup
Mentorgroup
Tutorial group
Teaching assistant
Mentorgroup
Mentorgroup
Tutorial group
Teaching assistant
Mentorgroup
Mentorgroup
Tutorial group
Teaching assistant
Study programme
e.g. Economics & Business
Randomization Procedure
A teaching assistant (TA) may teach tutorial groups in different study programmes. Randomization takes place within a study programme by assigning the first mentorgroup taught in a programme by a TA randomly to either control, the goal treatment or the raise treatment. Other groups taugth by the TA in the programme are assigned by cycling in order through control, goal and raise treatment. The random assignment of the first mentorgroups are balanced at the level of TA's teaching the same number of groups in the programme.
Legend
Mentorgroup assigned to control
Mentorgroup assigned to goal treatment
Mentorgroup assigned to raise treatment
Mentorgroup assigned based on random draw

Figure 3



Histogram of goals initially set by students (prior to any attempts to raise goals).

Appendix

Appendix 1

Dear X,

Following our introduction during the tutor instruction session, we request you to adjust the progress meetings with your students. Your participation contributes to research regarding the possibilities to increase students' study success by improving the tutor meetings.

The instructions regarding the progress meetings that you conduct in the week of 17th to 21st November follow.

After you have discussed the general motivation and study progress of the student, you are expected to ask some additional questions while discussing the current courses the student follows. These questions relate to Microeconomics. The intention is to motivate the students to set a goal. Ask the questions in italics.

Do you have a final grade in mind that you want to achieve for Microeconomics?

If YES: *What grade do you want to achieve?*

(if the student answers: I want to pass the course, then try to specify this, for example: Are you aiming for 5,5 or a 6?)

If NO: *Some students find it useful to set concrete goals in the form of a grade. Setting a goal can help motivate you to perform better. Do you want to set yourself a goal?*

If YES: *What final grade do you want to achieve for Microeconomics?*

If NO: continue the conversation as usual.

If the student set a goal:

Good luck with achieving your goal.

It is important that you follow the instructions as much as possible. However, do try to incorporate the questions into the conversation naturally. Attached, you find a flowchart summarizing the script. You can use this flowchart to refresh your memory prior to the meeting.

We request you to complete the attached form after the meeting with each student. You can print this form yourself, or pick up a copy at H8-23 or H8-24. It can be useful to make notes. Please, read the form carefully before the meetings.

After the meetings we would like to receive the completed forms. The completed forms can be handed in at H8-23 or H8-24 or can be emailed to vanlent@ese.eur.nl or souverijn@ese.eur.nl. We request that you hand in the form at Friday 28th November at the latest.

For research purposes we request that you do not discuss these instructions with others.

If you have any questions, do not hesitate to contact us. You can find us in H8-23 or H8-24, and you can reach us at vanlent@ese.eur.nl, phone: 010 408 1793 or souverijn@ese.eur.nl, phone: 010 408 9038.

Max van Lent
Michiel Souverijn

P.S. Could you please confirm to us by email that you have received this email and that you have read the instructions.

Appendix 2

Dear X,

Following our introduction during the tutor instruction session, we request you to adjust the progress meetings with your students. Your participation contributes to research regarding the possibilities to increase students' study success by improving the tutor meetings.

The instructions regarding the progress meetings that you conduct in the week of 17th to 21st November follow.

After you have discussed the general motivation and study progress of the student, you are expected to ask some additional questions while discussing the current courses the student follows. These questions relate to Microeconomics. The intention is to motivate the students to set a goal. Ask the questions in italics.

Do you have a final grade in mind that you want to achieve for Microeconomics?

If YES: *What grade do you want to achieve?*

(if the student answers: I want to pass the course, then try to specify this, for example: Are you aiming for 5,5 or a 6?)

If NO: *Some students find it useful to set concrete goals in the form of a grade. Setting a goal can help motivate you to perform better. Do you want to set yourself a goal?*

If YES: *What final grade do you want to achieve for Microeconomics?*

If NO: continue the conversation as usual.

If you (as a tutor) think the goal (grade) set is achievable:

Shouldn't you set a higher goal and aim for a [grade +1]? [So if the student chooses a 6 as a goal and you think this is achievable, propose to aim for a 7.]

If the student set a goal:

Good luck with achieving your goal.

It is important that you follow the instructions as much as possible. However, do try to incorporate the questions into the conversation naturally.

Attached, you find a flowchart summarizing the script. You can use this flowchart to refresh your memory prior to the meeting.

We request you to complete the attached form after the meeting with each student. You can print this form yourself, or pick up a copy at H8-23 or H8-24. It can be useful to make notes. Please, read the form carefully before the meetings.

After the meetings we would like to receive the completed forms. The completed forms can be handed in at H8-23 or H8-24 or can be emailed to vanlent@ese.eur.nl or souverijn@ese.eur.nl. We request that you hand in the form at Friday 28th November at the latest.

For research purposes we request that you do not discuss these instructions with others.

If you have any questions, do not hesitate to contact us. You can find us in H8-23 or H8-24, and you can reach us at vanlent@ese.eur.nl, phone: 010 408 1793 or souverijn@ese.eur.nl, phone: 010 408 9038.

Max van Lent
Michiel Souverijn

P.S. Could you please confirm to us by email that you have received this email and that you have read the instructions.