Goal Setting, Information, and Goal Revision: A Field Experiment

Max van Lent (ID)

Leiden University

Abstract. *People typically set goals in settings where they cannot be sure of how they will perform, but where their performance is revealed to them in parts over time. When part of the uncertainty is resolved, initial goals may have turned out to be unrealistic and hence they no longer work as a motivation device. Revising goals may increase performance by making goals realistic, but may also adversely affect performance through reduced goal commitment. We study the effects of motivating university students to set goals and inviting them to revise their goals later, using a field experiment involving nearly 2,100 students. We use courses containing two midterms and a final exam, where midterms reduce uncertainty about students' potential performance. We find that motivating students to set goals does not affect performance on average. Students with midterm grades lower than their goal, decrease their performance. This effect is driven by students who were motivated to set goals without being made aware that they can revise their goals later. This finding may help explain why the evidence of the effectiveness of goals on study performance is mixed.*

Keywords: Goal setting; field experiments; economics of education.

1. INTRODUCTION

Goal setting by students, as a means of increasing study performance, is by now a widely observed phenomenon. However, often the goals students set are not reached.[1] One reason for this may be that at the time students set goals, the uncertainty about the productivity of their effort is typically high. When part of the uncertainty is resolved, goals may have turned out to be unrealistic, causing these goals to no longer work as a motivation device. So far there has not been empirical research on goal setting that takes the reduction of performance uncertainty into account. Studying this type of uncertainty can help explain the mixed evidence about the effectiveness of goal setting at university found in the literature.[2]

In this paper, we study goal setting for university students in courses where students' uncertainty about the productivity of their effort is reduced over time.[3]

---

1. For example, using data from university students in the Netherlands Van Lent and Souverijn (2017) find that only 38% of the students reach the goal they had set themselves. Similarly, Clark *et al.* (2018) find that 24% to 53% of their students reach their goal.
2. For example, Van Lent and Souverijn (2017) report positive effects, Clark *et al.* (2018) report no effects for grade goals and positive effects for number of practice exam goals, and Dobronyi et al. (2017) report no effects.
3. One reason for studying goal setting at university is that there is a widely held belief that many students do not exert enough effort. A goal-setting intervention could boost effort and thereby graduation prospects and eventually labor market outcomes.

Specifically, students learn about their ability and several aspects of the course through studying and participation in two midterm exams and thereby uncertainty reduces. We examine i) whether goal setting increases performance ii) whether emphasizing the possibility of goal revision decreases short-run performance, and iii) whether emphasizing the possibility to revise goals that are very distant from current performance increases subsequent performance. In order to do so, we run a large-scale field experiment involving almost 2,100 students from one large European university.

Students are randomly assigned to the control group or one of three treatment groups and receive a series of three surveys. In addition to questions that all students receive, students assigned to a treatment group are asked whether they want to set grade goals as well as other goals. Students assigned to treatment 1 (T1) are asked whether and what goals they want to set in survey 1 and are reminded about their goals in survey 2 and 3. In treatment 2 (T2), students are asked about goals in survey 1, reminded of and invited to revise their goals in survey 2 and reminded of their (newly set) goals in survey 3. In treatment 3 (T3), students are asked about their goals in survey 1, but are also informed that they will be invited to revise their goals in the next survey. Survey 2 and survey 3 are identical to the surveys that students in T2 receive. Hence, the only difference between T2 and T3 is whether the invitation to revise goals (in the second survey) comes as a surprise. The timing of the second and third survey is such that students receive a survey shortly after taking a midterm exam, and have thereby received additional information (e.g., about the difficulty of the course or the grading scheme) before deciding how much effort to exert for the next (midterm) exam. When deciding about the treatments, we had to trade-off the intensity of the treatment with the possibility of treatment contamination. The surveys between the control and treatment groups differ only by a couple of questions which comes at the risk that the treatment is not intensive enough. However, by making the treatments more different from the control group – for instance, by adding more questions or more reminders – we believe that students from both the control and treatment groups would learn during the treatment period that an experiment was going on which could induce an experimenter demand effect.

This experiment is motivated by the idea that goals that turned out to be too easy or too difficult do not motivate students, leading to lower performance than when goals would be moderately difficult.[4] Revising a goal solves the problem that the goal is too easy or too difficult, but may create other problems. First, students may be less motivated to reach revised goals as compared to the initial goal, for instance, because they are disappointed for having to revise their goal in the first place. We test this by asking students about goal motivation in all treatments. Second, if students know from the start that they will be invited to revise their goals later, they may initially be less motivated to exert effort, simply because they know that they can revise their goals later. We test this by comparing short-run performance of students assigned to T3 with students assigned to T1 and T2.

---

4. See Vroom (1964), and more recently Shah and Higgins (1997), Heath *et al.* (1999), and Wu *et al.* (2008) for evidence in line with this premise.

Our main findings are the following. We find small and marginally significant negative effects of the treatment on midterm grades. The treatment effect in the second midterm is negative and significant for those students who are motivated to set goals but are not told that they could revise goals from the start, i.e., T1 and T2. The size of the effect is an economically significant 12% to 15% of a standard deviation. We do not find such an effect for students in T3. For the final exam, we find no effect of the treatment. An explanation for this is that the final examination is already a high stakes test. The fact that we find mainly insignificant results for our treatments is unlikely to be a power issue. We are able to detect an effect size of 5.5% of a standard deviation, and the standard errors of the coefficients are small. Focusing on those students who set grade goals, we find that students who have a grade goal that is very different from their first midterm grade, subsequently decrease their performance. This effect is driven by those students who are not made aware that they can revise their goals.

There is a large literature on the effects of goal setting on performance, see Locke and Latham (1990, 2002) for a review of the goal-setting literature. Some of the key mechanisms through which goals affect performance are goal commitment and task complexity, see also Koestner *et al.* (2002). In the surveys, we ask students about their commitment toward their goals, and find overall a high goal commitment. Performance in a university education program can be considered a complex task in the sense that it is not obvious how increased effort can result in an increased performance. The goal-setting literature has found that when task difficulty is higher, the positive effects of goals are smaller, see also Wood *et al.* (1987).

Economic theory papers found that goal setting can increase performance for present biased and loss averse agents, see e.g. Hsiaw (2013), Koch and Nafziger (2011), Koch *et al.* (2014), and Suvorov and Van de Ven (2008). However, only very recently Koch and Nafziger (2016) showed how goal setting can impact performance under uncertainty in a dynamic setting.

In addition to the economic theory literature, there is an increasing number of empirical studies investigating the effects of goal setting on performance in the laboratory and in the field. See for instance Brookins *et al.* (2017), Corgnet *et al.* (2015, 2018), Dalton *et al.* (2016), Koch and Nafziger (2017), Markle *et al.* (2018), and Uetake and Yang (2018), and for experimental studies about goal setting with university students, see e.g., Clark *et al.* (2018), Dobronyi *et al.* (2017), and Van Lent and Souverijn (2017). In Van Lent and Souverijn (2017) goals are set during one-on-one interviews between students and their randomly assigned mentor. The goal-setting intervention takes at most five minutes and increases students' grades by 9% of a standard deviation. Dobronyi *et al.* (2017) implement an online goal-setting intervention in which students complete structured goal-setting exercises related to study performance but also related to other aspects of life. They find no impact of the goal-setting intervention on students' study performance. Clark *et al.* (2018) study goal setting of university students using a survey, and students are reminded about their goal each time they received a new grade. Clark *et al.* (2018) find that grade goals do not affect performance, but goals with respect to the number of practice exams increases performance for male students. Finally, psychologists and management scientists have also

studied goal setting by university students, see e.g., Ames and Archer (1988), Bettinger and Baker (2013), Linnenbrink (2005), Morisano *et al.* (2010), Schippers *et al.* (2015), Schunk (1990), and Travers *et al.* (2015). In most of these studies however, there is no control group, which makes it impossible to estimate a causal effect from motivating students to set goals.

In order to learn what drives the average null finding of being asked to set goals, we explore heterogeneity along several dimensions. We focus on i) self-reported differences in the amount of uncertainty; ii) time preferences; and iii) task/course motivation. We find no evidence that heterogeneity in uncertainty or differences in course motivation are important mechanisms.

One of the insights that this study provides is that one should not expect goal setting to always increase performance. Encouraging students to set goals using surveys may not be a strong enough intervention, and may even decrease performance for some students.

This paper is organized as follows. In the next section, we explain the experimental context and set up. Section 3 explains the empirical strategy. Followed by sections 4 and 5 in which we describe the data and explain the results. The last section discusses our findings and conclude.

2. EXPERIMENTAL SET-UP

2.1. Experimental context

Our experiment involved nearly 2,100 first-year and second-year students enrolled in several undergraduate programs at the Erasmus University Rotterdam during the 2016–17 academic year. The academic year is divided in five blocks of eight weeks (seven weeks of lectures and tutorials followed by one week of exams). In each block students take 12 study credits (ECTS) worth of courses. For first-year and second-year students all courses are mandatory. Hence, all students in our sample within a study program take the exact same courses.

Students typically take two courses in each block, one course worth 8 ECTS and a smaller 4 ECTS course. For first-year students, we implement our experiment during the 8 ECTS course Microeconomics, and for the second-year students we implement our experiment during the 8 ECTS course Applied Microeconomics. The 4 ECTS course that students take depends on the study program that the student is enrolled in.[5]

The final grade that students receive for the courses of interest are based on two midterm exams (worth 10% of the final grade each) followed by a final exam which counts for the remaining 80%. It is exactly the composition and timing of the components of the grades that makes this setting suitable to answer our research question, as we will further explain in the next subsection.

---

5. Students enrolled in the first year programs Economics & Business Economics and Econometrics take respectively a course in ICT and Calculus at the same time as the Microeconomics course. Students enrolled in the second year take the course History of Economic Thought at the same time as Applied Microeconomics.
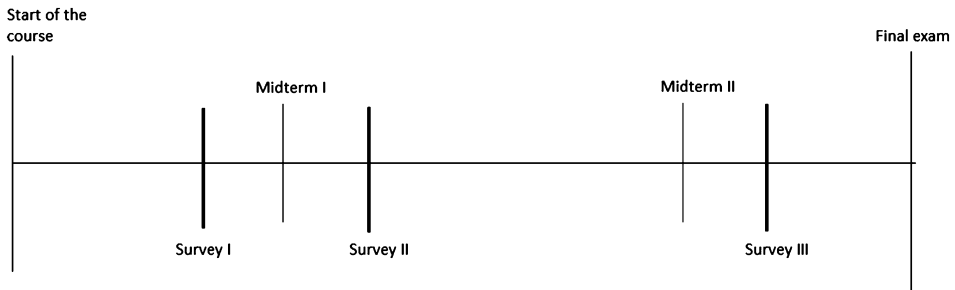
Figure 1 Planning of the course and surveys. The vertical axis represents time

2.2. Experimental design

Students are asked to complete the surveys on their mobile phones or laptops during tutorials, taught by tutorial teachers. Attending these tutorials is mandatory for first-year students, but not for second-year students. For the timing of the exams and surveys see Figure 1.

We randomly select a subset of the students and ask them whether they want to set goals. First, we ask whether students want to set grade goals (and what goal their grade goal is), and then ask whether they want to set other goals, such as an amount of study hours or preparing for each tutorial.[6] We ask the students about non-grade goals because some students may be more interested in setting these type of goals, and since non-grade goals may be more effective for some students. Note that we explicitly included the option for students to choose not to set any goals. Further, students are made aware that their survey responses will not be disclosed to any of their teachers, hence there is no reason for students to fill in these questions strategically.

In one treatment (T3), we inform students that in the second survey they will be invited to revise their goals, while in the other treatments (T1 and T2) we do not. In addition, we have a control group in which students are not motivated to set any goals, but are (just like in the treatment groups) asked questions about their preferences, behaviors, opinions, characteristics, and motivations. Since students fill in the surveys during their tutorials the probability of contamination is higher if students within a tutorial group would receive different surveys. We therefore randomize students into one of the treatments or the control group based on the tutorial group level.

After the first survey, students take the first midterm exam. Directly after the examination, students receive the second survey. Students in T1 are only reminded about their goals, while students in T2 and T3 are also invited to revise any of their goals. While students who are assigned to T3 knew that they would be granted the opportunity to revise their goals, this opportunity comes as a surprise for the students in T2. After the second midterm, students receive the third and final survey. In this last survey treated students are again reminded about their goals. Figure 2 shows the difference between the control and treatment

---

6. To be precise, the question we use in T1 and T2 is: "Do you want to set yourself a [grade/ non-grade] goal for [coursename]?" and the question in T3 is "Do you want to set yourself a [grade/ non-grade] goal for [coursename]? Note that you can decide to change your goal after the first midterm".

	Control	Treatment 1	Treatment 2	Treatment 3
Before Midterm I (5min)	Preferences, behaviour, opinions, characteristics, and motivation	Control + Motivated to set goals	Control + Motivated to set goals	Control + Motivated to set goals + Possibility to revise goals later
Post Midterm I (3min)	Behaviour, opinions, and motivation	Control + Reminded about goals	Control + Reminded about goals + Allow to revise goals	Control + Reminded about goals + Allow to revise goals
Post Midterm II (3min)	Behaviour, opinions, and motivation	Control + Reminded about goals	Control + Reminded about goals	Control + Reminded about goals

Figure 2 Differences in survey content between the control and treatment groups for each wave

groups for each survey. For more detail about the survey questions in each group and survey wave see the Supplementary Appendix S1.

We randomized treatment assignment based on tutorial group and stratify the randomization by teaching assistant and study program. Teaching assistants typically teach multiple tutorial groups. Since some of the shocks in student productivity will be teaching assistant specific, we always have a mix of treatment groups for teaching assistants who teach more than one group.[7] We also stratify our randomization on study program, since students in different programs take different courses, and since students in some study programs generally perform better than students who are selected into other programs as historical data showed. Finally, after students completed their survey, they were not able to review the questions.

3. EMPIRICAL STRATEGY

We estimate the effects of asking students to set goals, of inviting students to revise their goals, and the timing of this invitation, on various performance measures. We estimate an intention-to-treat effect because not all students filled in the survey. First, we estimate the effects of being asked to set goals, and then we estimate the effects separately per treatment:

$$y_i = \alpha_0 + \alpha_1 G_i + \alpha_1 X_i + \varepsilon_i$$

$$y_i = \beta_0 + \beta_1 T1i + \beta_2 T2_i + \beta_3 T3_i + \beta_4 X_i + \varepsilon_i$$

where y_i is student i's study performance. Our main variable of interest is students' grades. Since across courses (and instruction language) the grade distribution varies we standardize grades by taking student i's grade, subtract it by the mean grade of the course and then divide it by the standard deviation. The timing of the exams allows us to separate the responses to the treatment in the short, medium, and long run. Hence, as an outcome variable for grade, y_i, we use students' grades for the first midterm, the second midterm, the final exam, and the final grade (which is a weighted average of the midterm and final exam). In addition to grades, we study self-reported effort measures and tutorial attendance

7. Note that teaching assistants were only aware that students had to fill in surveys, not that there were different survey versions or that the students were part of an experiment.

Table 1 Descriptive statistics

Covariates	Mean	Treatment 1		Treatment 2		Treatment 3		F-test
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a) Gender and grades								
Prior GPA	6.67	0.01	0.93	−0.03	0.69	−0.06	0.39	0.79
Male	0.71	−0.02	0.40	0.01	0.70	−0.07	0.01	0.02
R^2		0.001		0.001		0.004		
p-value on F-statistic		0.84		0.33		0.04		
(b) Survey participation								
Survey 1	0.55	−0.07	0.02	−0.03	0.37	0.06	0.04	0.00
Survey 1 & 2	0.33	−0.09	0.00	−0.10	0.00	−0.02	0.48	0.00
Survey 1, 2, and 3	0.24	−0.04	0.11	−0.06	0.02	−0.02	0.54	0.11
R^2		0.006		0.004		0.008		
p-value on F-statistic		0.00		0.03		0.00		

Columns (2)–(7) report the coefficients and p-values from OLS regressions of each covariate on three indicators (one for each treatment). Column (8) reports p-values for joint test of significance of the three treatment indicators.

as proxies of study effort. G is a dummy that equals one if the student is assigned to a treatment, and $T1$, $T2$, $T3$ are dummies for respectively treatment 1, 2, and 3. Further, X denotes a rich set of control variables, including students' gender and GPA prior to the experiment, and characteristics of the tutorial teachers (i.e., gender and previous teaching experience).

4. DESCRIPTIVE STATISTICS

We targeted nearly 2,100 students from two cohorts to participate in a series of surveys. Of this group, 55% of the students (i.e., 1,134 students) filled in the first survey. Table 1a shows a test of balance between treatments for the variables of interest. Based on observables, treatment groups do not differ in terms of GPA prior to the experiment, but there are differences in the fraction of male students assigned to treatment and control groups. The fraction of male students is lower in treatment 3 as compared to the other groups. We control for these covariates in all the regressions.

Since the first survey differs between treatment groups and the control group we cannot rule out that students in some groups are more likely to stop participating in subsequent surveys than in others. Selective drop out of the surveys leads to challenges in the estimation of within student differences measured by the surveys (e.g., differences in student motivation and perceived uncertainty). Table 1b tests whether there is selective drop out from survey participation. The covariates Survey 1, Survey 1 and 2, and Survey 1, 2, and 3 are dummies that equal 1 if a student filled in the first survey only, the first and second survey only, and all surveys, respectively. Students are more likely to fill in the survey

in the control group and in T3 as compared to the other groups. One reason for the differences is likely to be that some teachers gave more time and opportunity to students to fill out the survey during class. Further analysis on the survey participation shows that there is a positive correlation between participating in the survey and tutorial attendance (which makes sense given that the students were supposed to fill in the surveys during class), but no differences in terms of GPA prior to the course, nor for treatment group students, the distance between the grade goal and midterm exam performance.

Table 2 provides data on goal setting separated by treatment. It is ambiguous whether students should be more willing to set goals when they know they have the opportunity to revise their goals later. On the one hand, students may be more willing to set a goal since they are able to revise goals that seem too easy or too difficult early on. On the other hand, students may be less motivated to reach (and therefore set) goals if they allow themselves to deviate from them later. From Table 2, we see that a large majority of students is willing to set goals, and that offering the opportunity to revise goals later does not affect students' willingness to set a goal. Initially, students who are informed that they can revise their goals later, set lower grade goals. We elicited goal motivation by asking how realistic students think their goal is and how important it is for them to reach this goal. We see that there are no differences based on students' responses to these questions. Further, students' willingness to set goals is independent of their time preferences. Finally, students are more willing to set grade goals as compared to other goals. Among the most frequently set non-grade goals are goals regarding: study hours, deeper understanding of the study materials, and making homework exercises. Hence, although the question eliciting non-grade goals did not impose much structure on the goals students can set, we observe that students set mostly specific and measurable goals, as the psychology literature on goal setting recommends.

Using data from the second survey, we see that 10% of the students who set a goal in treatment 2 or 3 revises their goals when asked. This fraction does not

Table 2 Descriptive statistics of goal setting by treatment

	Treatment 1	Treatment 2	Treatment 3
Fraction of students who sets a grade goal	0.70	0.70	0.70
Fraction of student who sets a non-grade goal	0.45	0.45	0.48
Fraction of students who sets a goal	0.83	0.84	0.82
Mean grade goal	8.00	7.92	7.88
Standard deviation grade goal	1.00	0.99	0.85
My grade goal is realistic	3.86	4.00	4.00
How important … to reach this grade goal	3.19	3.19	3.21
Observations	251	265	306

The sample here consists of those students who filled in (at least) the first survey. The answer to the statement 'My grade goal is realistic' is measured on a 5-point scale with 1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree. The anwer to the question 'How important do you find it to: reach this goal' is measured on a 4-point scale with 1 = Not at all important, 2 = Not that important, 3 = Important, 4 = Very important.

differ between treatments. Goal revision is unrelated to the absolute distance between the first midterm exam grade and the grade goal the student sets. We expected students to revise their grade goal when a goal turned out to be too easy or too difficult. Under the assumption that students feel that their goal is realistic is at least partly based on the distance between grade goal and performance, this finding is surprising. Finally we do see that students whose perceived uncertainty has decreased are more likely to revise their goal, see Table A1 ($p = 0.131$).

One key assumption throughout this paper is that over time the uncertainty of students decreases. Table 3 shows that indeed students' perceived uncertainty regarding what they should do in order to obtain a certain grade has decreased after the first midterm exam. Further, we see as expected that perceived uncertainty is not affected by the treatment.

5. RESULTS

5.1. Main results

We are interested in the effects of asking students to set goals, and the effects of inviting students to revise their goals on study performance. We start by analyzing the short-run effects by examining students' grades for the first midterm exam. The first two columns of Table 4 show the effect of the treatment on student performance in the first midterm. In the short run, motivating students to set goals has a small and marginally significant negative effect ($p = 0.092$). The treatment 3 coefficient is not significantly different from the other coefficients. This implies that making salient that students can revise their goals later, does not affect performance negatively in the short run. This is in line with the earlier finding that offering the opportunity to revise goals later did not change initial goal motivation (see also Table 2).

Table 3 The perceived uncertainty decreases over time, but no difference by treatment

	Control group	Treatment 1	Treatment 2	Treatment 3
Sample: participants survey 1				
Uncertainty S1: mean	1.39	1.44	1.38	1.31
Standard deviation	1.08	1.12	1.05	1.13
Observations	312	251	265	304
Sample: participants survey 1 & 2				
Uncertainty S1: mean	1.33	1.38	1.44	1.33
Standard deviation	1.07	1.08	1.02	1.14
Uncertainty S2: mean	1.58	1.63	1.60	1.56
Standard deviation	1.08	1.14	1.12	1.16
Observations	210	176	151	205

The value of uncertainty is between 0 and 3, with a higher score implying less uncertainty. A t-test shows that for the control group Uncertainty S2 is higher than Uncertainty S1 ($p < 0.01$). This implies that students' perceived uncertainty reduces over time.

Table 4 Intention to treat effects on grades in the short, medium, and long run

Dependent variable:	Midterm I		Midterm II		Final exam		Final grade	
Goal treatments combined	−0.12* (0.07)		−0.10* (0.05)		−0.02 (0.05)		−0.05 (0.04)	
T1		−0.09 (0.07)		−0.12 (0.07)		−0.04 (0.08)		−0.06 (0.07)
T2		−0.13* (0.07)		−0.15** (0.07)		−0.01 (0.07)		−0.06 (0.06)
T3		−0.14 (0.13)		−0.04 (0.07)		−0.02 (0.07)		−0.03 (0.06)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2.076	2.076	2.076	2.076	2.076	2.076	2.076	2.076

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and dividing by the st.dev. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

Columns three and four display the effects on students' performance in midterm 2. Goal setting decreases performance on average with 10% of a standard deviation ($p = 0.054$). The effect is driven by students in T1 and T2. Performance for these groups decreases by 12% to 15% of a standard deviation. Students in T3 do not perform significantly different from the control group. This indicates that in the medium run, students who set a goal and were not informed at the start that they can revise their goals later, perform worse than the control group, while explicitly offering the opportunity to revise goals seem to have neither a positive nor a negative effect.

The final exam has the largest impact on students' final grade. Columns five and six show that the treatments have no effect on performance in the final exam, the coefficients are close to zero and precisely estimated. One interpretation for this finding is that the final exam is already a high stake environment, since students can only take a limited number of resits at the end of the academic year, and hence goal setting cannot push students any further (or hold them back). First-year students who do not pass all their courses after resits in the summer cannot continue their education program and are forced to stop their studies and are not allowed to resubscribe for one year. Columns seven and eight show the effects on the final grades, which is a weighted average of the midterms and final exam. Motivating students to set goals does not have an overall effect on students' grades. We also measured the effects of our treatments on self-reported effort and tutorial attendance which are alternative effort measures, see Table A2. We find small and insignificant results that are in line with the findings displayed in Table 4.

Next we consider heterogeneity of the treatment effect, based on students' GPA before the experiment, and gender. We find that GPA prior to the

experiment is positively correlated with performance during the experiment, and that goal setting increases performance for below average GPA students (and decreases performance for above average GPA students), see Table 5, columns 1, 3, 5, and 7.[8] Hence, the goal-setting interventions decrease the grade variation across students. This heterogeneous treatment effect is mostly driven by students who were not explicitly offered the opportunity to revise their goals (T1), see columns 2, 4, 6, and 8. One explanation for this finding may be that students who have a higher goal than their midterm grade, give up when the opportunity to revise goals was not made salient, but less so when the opportunity to revise goals was emphasized. We further explore this possibility later.

In Table 6, we study whether the treatment effect is heterogeneous based on students' gender. We find some indication that goals have a negative effect for mainly male students in the short run, but no differential effect in the long run. This is opposite to the findings of Clark *et al.* (2018). Clark *et al.* (2018) find positive effects of goal setting for males and attribute this to students' time preferences. As we will show in the next section, there is in our study no heterogeneous treatment effect based on students' time preferences.

One reason for finding an average null effect may be that for some students goals have turned out to be too easy or too difficult and hence no longer work as a motivation device, while for students who set a realistic goal, the effect is positive. In order to analyze this, we estimate the effect of the distance between the grade goal and the first midterm grade on later performance, see Table 7.

We find in line with our prediction that those students who have a lower midterm grade than their grade goal subsequently decrease their performance. Students who have a first midterm grade that is 1 grade point lower than their goal decrease their second midterm exam grade by 6% of a standard deviation and their final grade by 7%. Both these effects are strongly significant (column 1 and 3). If we distinguish between those students who were told that they could revise their goals in survey 1 (T3) and the other treatments, we find that T3 students who have a grade that is very different from their goal do not decrease their performance, while those who were not offered the opportunity to revise at the start decrease performance by 9% to 12%. This suggests that goals that turned out to be either too easy or too difficult harm future performance, but only when students are not made aware from the beginning that they can revise their goals.

5.2. Mechanisms

In the previous section, we established that motivating students to set goals using surveys has on average no effect on study performance. However, there may be interesting heterogeneity underlying this finding. Therefore, we now shed light on some of the mechanisms that may drive the null finding. The mechanisms we study are: students' perceived uncertainty about their potential performance, intrinsic motivation, and time preferences.

---

8. The finding that motivating students to set goals increases performance for below average GPA students is in line with Morisano *et al.* (2010) who find positive effects of a goal setting intervention among poor performing students.

Table 5 Heterogeneous effects: GPA prior to the experiment

Dependent variable:	Midterm I		Midterm II		Final exam		Final grade	
Goal treatments combined	−0.10 (0.08)		−0.07 (0.06)		0.03 (0.05)		−0.01 (0.05)	
GPA	0.14*** (0.03)	0.14*** (0.03)	0.18*** (0.02)	0.18*** (0.02)	0.20*** (0.02)	0.20*** (0.02)	0.20*** (0.02)	0.20*** (0.02)
Goal × GPA	−0.03 (0.03)		−0.05* (0.03)		−0.08*** (0.03)		−0.08** (0.03)	
T1		−0.08 (0.08)		−0.09 (0.07)		0.01 (0.07)		−0.02 (0.07)
T2		−0.11 (0.08)		−0.12* (0.07)		0.03 (0.07)		−0.01 (0.06)
T3		−0.12 (0.14)		−0.01 (0.07)		0.035 (0.07)		0.01 (0.07)
T1 × GPA		−0.04 (0.03)		−0.06* (0.03)		−0.11*** (0.04)		−0.10** (0.04)
T2 × GPA		−0.02 (0.03)		−0.04 (0.03)		−0.06 (0.04)		−0.05 (0.04)
T3 × GPA		−0.05 (0.05)		−0.05 (0.04)		−0.08* (0.05)		−0.07 (0.05)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2.076	2.076	2.076	2.076	2.076	2.076	2.076	2.076

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and dividing by the st.dev. Control variables: student's gender, tutorial teacher characteristics, and study program.

Table 6 Heterogeneous effects: Gender

Dependent variable:	Midterm I	Midterm II	Final exam	Final grade
Goal treatments combined	−0.10	−0.17**	0.00	−0.06
	(0.09)	(0.07)	(0.09)	(0.08)
Male	−0.06	−0.24***	−0.03	−0.09
	(0.06)	(0.08)	(0.08)	(0.08)
Goal × Male	−0.19**	−0.32***	−0.06	−0.14*
	(0.08)	(0.07)	(0.09)	(0.08)
T1	−0.06	−0.17*	−0.01	−0.08
	(0.08)	(0.10)	(0.11)	(0.10)
T2	−0.14	−0.28***	−0.09	−0.16
	(0.11)	(0.09)	(0.12)	(0.11)
T3	−0.09	−0.06	0.09	0.06
	(0.15)	(0.10)	(0.12)	(0.12)
T1 × Male	−0.16*	−0.34***	−0.07	−0.15
	(0.10)	(0.09)	(0.11)	(0.11)
T2 × Male	−0.19**	−0.33***	−0.01	−0.11
	(0.08)	(0.08)	(0.10)	(0.09)
T3 × Male	−0.22*	−0.27***	−0.10	−0.17*
	(0.13)	(0.08)	(0.10)	(0.09)
Controls	Yes	Yes	Yes	Yes
Observations	2.076	2.076	2.076	2.076

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and dividing by the st.dev. Control variables: tutorial teacher characteristics, study program, and GPA prior to the experiment.

Table 7 Distance from the goal and subsequent performance

Dependent variable:	Midterm II		Final Grade	
Distance from the goal	−0.06***	−0.12***	−0.07***	−0.09***
	(0.19)	(0.03)	(0.02)	(0.03)
Revise treatment		0.05		0.05
		(0.07)		(0.07)
Revise treatment × Distance		0.07**		0.03
		(0.04)		(0.04)
Controls	Yes	Yes	Yes	Yes
Observations	569	569	569	569

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and dividing by the st.dev. 'Distance from the goal' is measured as the absolute difference between the goal and the first midterm grade, and 'Revise treatment' is defined as a dummy that equals 1 if the student has set a goal in T2 or T3, and equal to zero if the student has set a goal in T1. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

5.2.1. Uncertainty

Students' perceived uncertainty about how they will perform (conditional on effort) may be a reason why goal setting on average does not affect performance. We study whether there is a heterogeneous treatment effect by students' perceived uncertainty. In each of the survey waves, we ask students about this type of uncertainty. To be precise we ask: '*Do you find it difficult to estimate how hard you should work in order to pass the Microeconomics course with a [Grade X]?*'. Where we use the most frequent passing grades: 6.0, 7.0, and 8.0. Answer categories are limited to 'yes' and 'no'.

Students who are uncertain about what they need to do in order to receive certain grades may set a low grade goal in order to shield themselves from overestimating their ability and thereby incurring a loss of not reaching their goal. If subsequently their first midterm grade is high relative to their goal, then they may adjust their effort downwards, as compared to similar students in the control group. On the other hand, students who set a goal that is too high and learn that they will not reach their goal give up on their goal and thereby decrease their performance. In Table 8, we report the results of estimating our specification separately for the students who report that they do not know how hard they should work in order to obtain a certain grade, with those students who answer all the questions with 'yes'.

Students who perceive high uncertainty decrease performance when being asked to set goals and students perceiving little uncertainty increase performance, although this finding lacks statistical significance. Qualitatively, this finding is in line with the idea that goal setting hurts students who find it difficult to estimate what they need to do in order to perform a certain way, while students who do not perceive such uncertainty benefit from goal setting.

5.2.2. Intrinsic motivation

Students who set themselves goals are typically also intrinsically motivated to reach their goals. Triggering this type of intrinsic motivation may, however,

Table 8 Heterogeneous treatment effects: perceived uncertainty

Dependent variable: Final grade	Sample: no uncertainty		Sample: uncertainty	
T1		−0.04		−0.03
		(0.15)		(0.05)
T2		0.06		−0.05
		(0.13)		(0.06)
T3		0.12		−0.05
		(0.13)		(0.05)
Goal treatments combined	0.05		−0.04	
	(0.12)		(0.04)	
Controls	Yes	Yes	Yes	Yes
Observations	211	211	921	921

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and dividing by the st.dev. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment. The sample in column one and two consists of those students who responded '*No*' to the question: 'Do you find it difficult to estimate how hard you should work in order to pass the Microeconomics course with a [Grade X]?' for X = 6, 7, and 8. The sample in column three and four consists of students who answered '*Yes*' to at least one of the questions.

crowd out (or crowd in) intrinsic motivation for the course itself. There is by now a large literature showing that extrinsic motivation can crowd out intrinsic motivation, see for instance Benabou and Tirole (2003), and for field experimental evidence see Huffman and Bognanno (2017). However, the evidence for crowding out of one intrinsic motivation type (course motivation) for another (goal motivation) is by our knowledge scarce.

Each survey wave, we ask students to rate their course motivation using two questions, we ask: '*To what extent do you agree with the following statements: I find the course interesting*' and '*I think the course is useful*'. Answers are measured on a 5-point scale ranging from strongly disagree to strongly agree. Students in treatment groups fill in these questions before they are motivated to set goals (and are unable to go back to adjust their answers), hence we can compare course motivation within a student across surveys. In Table 9, we display the effect of goals on student motivation for the course.

We see that the treatments have no effect on course motivation, i.e., there is no crowding out or crowding in of motivation for the course as a consequence of the goal treatments. We see that this is the case for the aggregate course motivation (columns 1 and 2), as well as for students' response to the two questions separately, see columns 3 to 6.[9]

5.2.3. Time preferences
Economic theory has shown that the success of goal setting depends on peoples' time preferences. Koch and Nafziger (2011) show that present biased people can

---

9. However, since students selectively drop out of surveys 2, one should be cautious when interpreting these findings.

Table 9 Impact of the treatments on the difference in course motivation

	Course motivation		Course interesting		Course useful	
T1		−0.04		−0.04		0.01
		(0.14)		(0.08)		(0.09)
T2		0.30**		0.11		0.18**
		(0.14)		(0.07)		(0.08)
T3		0.07		−0.01		0.07
		(0.12)		(0.08)		(0.07)
Goal treatments combined	0.10		0.02		0.09	
	(0.11)		(0.06)		(0.06)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	675	675	675	675	675	675

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The sample consists of all students who completed at least the first and second survey. The dependent variable is measured by the respondent's answer to the questions: To what extend do you agree with the following statements: 'I find the course interesting' and 'I think the course is useful'. Answers are measured on a 5-point scale from 1 = Strongly disagree, to 5 = Strongly agree. In the first two columns we took the average of the response to both questions. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

increase performance by setting goals.[10] We attempt to learn whether students who are present biased react stronger to the goal-setting treatment, and whether there are differences between treatments. We measure time inconsistency using the same hypothetical questions as used in Ashraf *et al.* (2006). We ask: '*Would you prefer to receive 15 euros guaranteed today, or 20 euros guaranteed in one month?*' and '*Would you prefer to receive 15 euros guaranteed in 6 months, or 20 euros guaranteed in 7 months?*'. Those students who prefer receiving the 15 euros now, and prefer the 20 euros in 7 months are labeled time inconsistent (and are therefore expected to be more often present biased). According to this definition 15.6% (i.e., 180 students) of the sample is time inconsistent. We regress our main specification for both samples separately, and find no significant differences between the samples, see Table A3.

Students have to (re)subscribe to their education program each year before a certain deadline, and the university keeps track if and when students (re)subscribe. We use the timing at which students subscribe in order to learn whether students are procrastinators.[11] Students neither have an incentive to (re)subscribe early, nor have an incentive to wait until the deadline. Hence, procrastinators will subscribe late, while others will either subscribe early or late. Therefore, looking at the sample of students who subscribe late gives us a noisy estimate of the treatment effect for procrastinating students.[12] In order to do this, we

10. Present biased individuals are defined as individuals who place a higher value to current utility than to future utility, see Stotz (1956). Present bias is a particular form of time inconsistency.
11. Procrastination is a good predictor of time-inconsistent behavior, see Reuben *et al.* (2015).
12. For a similar approach see Himmler *et al.* (2019).

Table 10 Heterogenous treatment effect: time preferences

Dependent variable: Final Grade	Sample: early subscription		Sample: late subscription	
T1		0.03		−0.08
		(0.6)		(0.10)
T2		−0.11**		−0.01
		(0.5)		(0.09)
T3		−0.11*		−0.01
		(0.06)		(0.09)
Goal treatments combined	−0.06		−0.03	
	(0.05)		(0.07)	
Controls	Yes	Yes	Yes	Yes
Observations	1.047	1.047	1.029	1.029

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and and dividing by the st.dev. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

separate the sample and estimate our main specifications for both the group of students who subscribe to the education program relatively early (i.e., below the median date) and for those who subscribe relatively late (i.e., above the median). We see that there are no significant differences between the groups, see Table 10.

5.3. Further results

5.3.1. Substitution and long-run effects

Besides the course used for the experiment, students take one other program specific course at the same time. If the treatments increase students' effort in the course for which they set the goal, it may be that they substitute effort away from the other course. Using the grades of these program-specific courses, we find no effect of the treatments on the study performance of this other course, i.e., we find no substitution effect, see Table 11 column 1 and 2.

We also study whether there are effects of our treatments on student performance in later courses. This may happen when students change their study behaviors in a way that benefits performance only later. We find, using the first set of courses after the experiment that there is no difference between students assigned to the control group and any of the treatment groups, see Table 11 column 3 and 4.

5.3.2. Satisfaction

While goal setting did not affect study performance much, it may have increased students' utility. First, because goal setting may allow students to not work too hard, and thereby avoid stress which plays an important role in burnouts. Second, by motivating students to set non-grade goals, students may have focused on tasks that were not productive in increasing their grades. For instance, some students choose to set non-grade goals such as: 'I would like to become a better

Table 11 The substitution effect and future performance

Dependent variable:	Substitute		Future	
Goal treatments combined	−0.00		−0.03	
	(0.01)		(0.07)	
T1		−0.00		−0.01
		(0.02)		(0.08)
T2		0.00		−0.06
		(0.01)		(0.08)
T3		−0.00		−0.02
		(0.03)		(0.09)
Controls	Yes	Yes	Yes	Yes
Observations	2.076	2.076	1.649	1.649

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. Dependent variable is standardized by subtracting the mean and dividing by the st.dev. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

decision maker' or 'learn more about economics'. Hence it may be that while goal setting decreased study performance for some students it increased their utility. Third, by making progress toward their goal (grade and/or non-grade goals), students my increase their satisfaction, see e.g., Diener *et al.* (1999) and Koestner *et al.* (2002).

In each survey wave, we asked students how satisfied they are with their performance in the course and with their life in general. The questions we asked are as follows: '*How satisfied are you with: your study performance in [course name] so far*' and '*How satisfied are you with: your life in general*'. Students answer on a 10-point scale, ranging from *1 = Not at all satisfied*, to *10 = Fully satisfied*. We regress the difference in satisfaction between surveys for the same student, on treatment assignment, see Table 12.

We find no effect of the treatment on satisfaction with study performance, although the coefficients are negative which is in line with the coefficients for actual performance. Life satisfaction increases for students assigned to treatment as compared to the control group. However, the size of the effect is marginally significant and rather small.

6. DISCUSSION AND CONCLUSION

We studied goal setting during university courses where students perceive high levels of uncertainty initially about their potential performance, and where through midterm exam, participation uncertainty is reduced over time. We asked students to set goals using surveys and varied whether students were invited to revise their goals, and the timing of this invitation. We find that on average goal setting had no effect on study performance. Our findings are similar to Dobronyi *et al.* (2017) who found also no effects of goal setting on performance using surveys, but different from among others Morisano *et al.* (2010), Van Lent and

Table 12 Impact of the treatments on the difference in course and life satisfaction

	Course satisfaction		Life satisfaction	
T1		−0.13		0.29**
		(0.24)		(0.13)
T2		0.08		0.25*
		(0.27)		(0.14)
T3		0.01		0.14
		(0.26)		(0.13)
Goal treatments combined	−0.02		0.22*	
	(0.22)		(0.12)	
Controls	Yes	Yes	Yes	Yes
Observations	745	745	745	745

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is measured by the respondent's answer to the question: How satisfied are you with... 'your study performance in Microeconomics so far' and 'your life in general'. Answers are measured on a 10-point scale from 1 = Not at all satisfied, to 10 = Fully satisfied. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

Souverijn (2017), and Clark *et al.* (2018) who find positive effects from goal setting on study performance. While statistical power is not an issue, there are several other potential reasons for why we find that goal setting has no effect on study performance on average.[13]

The treatment may not have been salient enough. We choose to make the surveys in the control and treatment groups similar in order to avoid that students and teachers felt like they were in an experiment. Hence, while both groups received several questions regarding study behaviors and preferences, the only difference between the treatment and control group was a couple of additional questions regarding goal setting. It may therefore be that – although we reminded students of their goals in all later surveys – the goal-setting intervention was not salient enough. Comparing our intervention with other goal-setting interventions at university, we see that the intervention in most of the studies seems stronger.[14]

Students were not incentivized to take the survey seriously. Students were asked to fill in the survey by their teacher, and in addition those who filled in all surveys participated automatically in a lottery where they could win an Ipad. These incentives were set in order to stimulate students to participate, however, this may not have been sufficient for students to take the survey seriously. However, this is likely to be an issue in all goal-setting-related surveys or essays as long as reaching the goal is not related to, for instance, a monetary incentive. Hence, the absence of incentives seems not sufficient to explain the null result in this study.

13. We are able to detect an effect size of 5.5% of a standard deviation.
14. In Dobronyi et al. (2017) and Schippers *et al.* (2015), for instance, students complete structured goal-setting exercises.

Another reason for results that deviate from most of the literature may be that the student population in our study is different from those used in other goal setting experiments at university. For instance, institutional, program, or recruiting differences may have led to different students, and hence different results. Morisano *et al.* (2010) target struggling students at a top Canadian university, and Clark *et al.* (2018) target students at a public university in the United States. However Van Lent and Souverijn (2017) do find significant effects of a goal-setting intervention with participants of the same study program and university as in this paper. One key difference between this paper and Van Lent and Souverijn (2017) is that the latter paper elicits goals during a one-on-one conversation between students and their mentor.

Finally, there is the possibility for treatment contamination. Students complete the surveys in the classroom during tutorials taught by tutorial teachers. Every student who is assigned to the same tutorial session received the same survey (i.e., within a tutorial group, students all received the control or one of the treatment surveys). In addition to tutorials, students also have the opportunity to attend lectures, taught by professors. Hence, although students mainly interact with others in their tutorial group, we cannot rule out that (some) control group students may have learned about goal setting from treated students. This type of interaction could lead to treatment contamination. As a consequence, we may be unable to detect treatment effects (since in fact also the control group is treatment). While we cannot formally rule out that there was (some) contamination, there is also no evidence that students or teachers learned about the experiment ex post.

Further inspection of the data shows that students who are behind on their goals decrease their performance, but mainly so when students were not explicitly offered the opportunity to revise their goals. This suggests that if students set goals that are too high, this hurts their performance to the extent that on average students setting goals are worse off. This finding may explain the mixed evidence of the effects of goals on student performance in the literature. In addition, our findings can be interpreted as a cautionary note for people who would like to set long-run goals in settings containing a lot of uncertainty, such as study related goals and career plans.

APPENDIX

Table A1 Perceived uncertainty and goal revision

Dependent variable:	Revise goal
Uncertainty decreased	−0.06
	(0.04)
Controls	Yes
Observations	254

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is a dummy that equals one if a student who set a goal in T2 or T3 revised a goal, and equals zero if that person did not revise a goal. The variable: Uncertainty decreased is a dummy that equals one if the student's perceived uncertainty has decreased and zero if it has not. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

Table A2 Treatment effect on attendance and study hours

Dep. var: Attendance	Attendance	Attendance	Study hours S2	Study hours S2	Study hours S3	Study hours S3
T1		−0.11		0.10		0.07
		(0.42)		(0.83)		(0.83)
T2		−0.02		−0.12		0.21
		(0.37)		(1.09)		(1.03)
T3		0.12		−0.31		0.67
		(0.50)		(0.74)		(0.89)
Goal treatments combined	0.00		−0.10		0.30	
	(0.28)		(0.70)		(0.73)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1.994	1.994	860	860	670	670

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is the number of times that students attended tutorials and the self reported amount of hours studied on average per week. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment. The mean (st.dev) of the dependent variables are respectively, 7.76 (5.20), 15.25 (8.09), and 15.97 (8.91).

Table A3 Heterogeneous effects: Time inconsistency

Dependent variable:	Midterm I		Midterm II		Final exam		Final grade	
Goal treatments combined	−0.14 (0.11)		−0.12* (0.06)		0.01 (0.04)		−0.04 (0.04)	
Time inconsistent	−0.19 (0.12)	−0.19 (0.12)	−0.17 (0.15)	−0.17 (0.15)	−0.18** (0.08)	−0.18** (0.08)	−0.19** (0.08)	−0.19** (0.08)
Goal × TI	0.12 (0.15)		0.10 (0.17)		0.12 (0.11)		0.14 (0.11)	
T1		−0.05 (0.09)		−0.19** (0.08)		−0.03 (0.06)		−0.08 (0.06)
T2		−0.14 (0.10)		−0.11 (0.08)		0.06 (0.07)		−0.02 (0.06)
T3		−0.21 (0.19)		−0.08 (0.07)		−0.02 (0.06)		−0.03 (0.05)
T1 × TI		0.09 (0.16)		0.30 (0.21)		0.29** (0.14)		0.32** (0.13)
T2 × TI		−0.07 (0.20)		−0.18 (0.20)		−0.05 (0.16)		−0.02 (0.15)
T3 × TI		0.33 (0.21)		0.20 (0.17)		0.13 (0.14)		0.14 (0.13)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1.133	1.133	1.133	1.133	1.133	1.133	1.133	1.133

OLS regressions with standard errors clustered at the tutorial group level and in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$. The dependent variable is standardized by subtracting the mean and dividing by the st.dev. Control variables: student's gender, tutorial teacher characteristics, study program, and GPA prior to the experiment.

Address for correspondence: Max van Lent, Department of Economics, Leiden University, 2311 EZ Leiden, The Netherlands. Tel.: +31 715277756; e-mail: m.van.lent@law.leidenuniv.nl

REFERENCES

Ames, C. and J. Archer (1988), 'Achievement Goals in the Classroom: Students' Learning Strategies and Motivation Processes', *Journal of Educational Psychology* 80, 260–267.

Ashraf, N., D. Karlan and W. Yin (2006), 'Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines', *Quarterly Journal of Economics* 121, 635–672.

Benabou, R. and J. Tirole (2003), 'Intrinsic and Extrinsic Motivation', *Review of Economic Studies* 70, 489–520.

Bettinger, E. P. and R. B. Baker (2013), 'The Effects of Student Coaching An Evaluation of a Randomized Expeirment in Student Advising', *Educational Evaluation and Policy Analysis* 36, 3–19.

Brookins, P., S. Goerg and S. Kube (2017), 'Self-chosen Goals, Incentives, and Effort', Mimeo.

Clark, D., D. Gill, V. Prowse and M. Rush (2018), 'Using Goals to Motivate College Students: Theory and Evidence from Field Experiments', Natural Field Experiments 00647.

Corgnet, B., J. Gomez-Minambres and R. Hernan-Gonzales (2015), 'Goal Setting and Monetary Incentives: When Large Stakes Are Not Enough', *Management Science* 61, 2926–2944.

Corgnet, B., J. Gomez-Minambres and R. Hernan-Gonzalez (2018), 'Goal Setting in the Principal-Agent Model: Weak Incentives for Strong Performance', *Games and Economic Behavior* 109, 311–326.

Dalton, P. S., V. Gonzalez and C. N. Noussair (2016), 'Self-Chosen Goals: Incentives and Gender Differences', CentER Discussion Paper Series No. 2016-036.

Diener, E., E. M. Suh, R. E. Lucas and H. L. Smith (1999), 'Subjective Well-Being: Three Decades of Progress', *Psychological Bulletin* 125, 276–302.

Dobronyi, C. R., P. Oreopoulos and U. Petronijevic. (2017). 'Goal Setting, Academic Reminders, and College Success: A Large-Scale Field Experiment (No. w23738)', National Bureau of Economic Research.

Heath, C., R. P. Larrick and G. Wu (1999), 'Goals as Reference Points', *Cognitive Psychology* 38, 79–109.

Himmler, O., R. Jäckle and P. Weinschenk (2019), 'Soft Commitments, Reminders and Academic Performance', *American Economic Journal: Applied Economics* 11, 114–142

Hsiaw, A. (2013), 'Goal-Setting and Self-Control', *Journal of Economic Theory* 148, 601–626.

Huffman, D. and M. Bognanno (2017), 'High-Powered Performance Pay and Crowding Out of Nonmonetary Motives', *Management Science* 64, 4669–4680.

Koch, A. and J. Nafziger (2011), 'Self-Regulation through Goal Setting', *Scandinavian Journal of Economics* 113, 212–227.

Koch, A. K. and J. Nafziger (2016), 'Goals and Bracketing under Mental Accounting', *Journal of Economic Theory* 162, 305–351.

Koch, A. K. and J. Nafziger (2017), 'Motivational Goal Bracketing: An Experiment', IZA Discussion Paper No. 10955.

Koch, A., J. Nafziger and H. S. Nielsen (2014), 'Behavioral Economics of Education', *Journal of Economic Behavior & Organization* 115, 3–17.

Koestner, R., N. Lekes, T. A. Powers and E. Chicoine (2002), 'Attaining Personal Goals: Self-concordance Plus Implementation Intentions Equals Success', *Journal of Personality and Social Psychology* 83, 231–244.

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, 197–213.

Locke, E. A. and G. P. Latham (1990), *A Theory of Goal Setting & Task Performance*, Prentice-Hall Inc., Englewood Cliffs, NJ.

Locke, E. A. and G. P. Latham (2002), 'Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey', *American Psychologist* 57, 705–717.

Markle, A., G. Wu, R. White and A. Sackett (2018), 'Goals as Reference Points in Marathon Running: A Novel Test of Reference Dependence', *Journal of Risk and Uncertainty* 56, 19–50.

Morisano, D., J. B. Hirsh, J. B. Peterson, R. O. Pihl and B. M. Shore (2010), 'Setting, Elaborating, and Reflecting on Personal Goals Improves Academic Performance', *Journal of Applied Psychology* 95(2), 255–264.

Reuben, E., P. Sapienza and L. Zingales (2015), 'Procrastination and Impatience', *Journal of Behavioral and Experimental Economics* 58, 63–76.

Schippers, M. C., A. W. Scheepers and J. B. Peterson (2015), 'A Scalable Goal-Setting Intervention Closes Both the Gender and Ethnic Minority Achievement Gap', *Palgrave Communications* 1, 15014.

Schunk, D. H. (1990), 'Goal Setting and Self-Efficacy during Self-Regulated Learning', *Educational Psychologist* 25(1), 71–86.

Shah, J. and T. E. Higgins (1997), 'Expectancy × Value Effects: Regulatory Focus as Determinant of Magnitude and Direction', *Journal of Personality and Social Psychology* 73, 447–458.

Suvorov, A. and J. Van de Ven (2008), 'Goal Setting as a Self-Regulation Mechanism', Center for Economic and Financial Research (CEFIR) Working Paper no. w0122, Moscow.

Travers, C. J., D. Morisano and E. A. Locke (2015), 'Self-Reflection, Growth Goals, and Academic Outcomes. A Qualitative Study', *British Journal of Educational Psychology* 85, 224–241.

Uetake, K. and N. Yang (2018), 'Harnessing the Small Victories: Empirical Evidence from a Calorie and Weight Loss Tracking Application', Working Paper.

Van Lent, M. and M. Souverijn (2017), 'Goal Setting and Raising the Bar: A Field Experiment', Tinbergen Institute Discussion Paper TI 2017-001/VII.

Vroom, V. H. (1964), *Work and motivation*, Wiley, Oxford.

Wood, R. E., A. J. Mento and E. A. Locke (1987), 'Task Complexity as a Moderator of Goal Effects: A Meta-Analysis', *Journal of Applied Psychology* 72, 416.

Wu, G., C. Heath and R. Larrick (2008), 'A Prospect Theory Model of Goal Behavior', Unpublished manuscript.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Online Appendix.