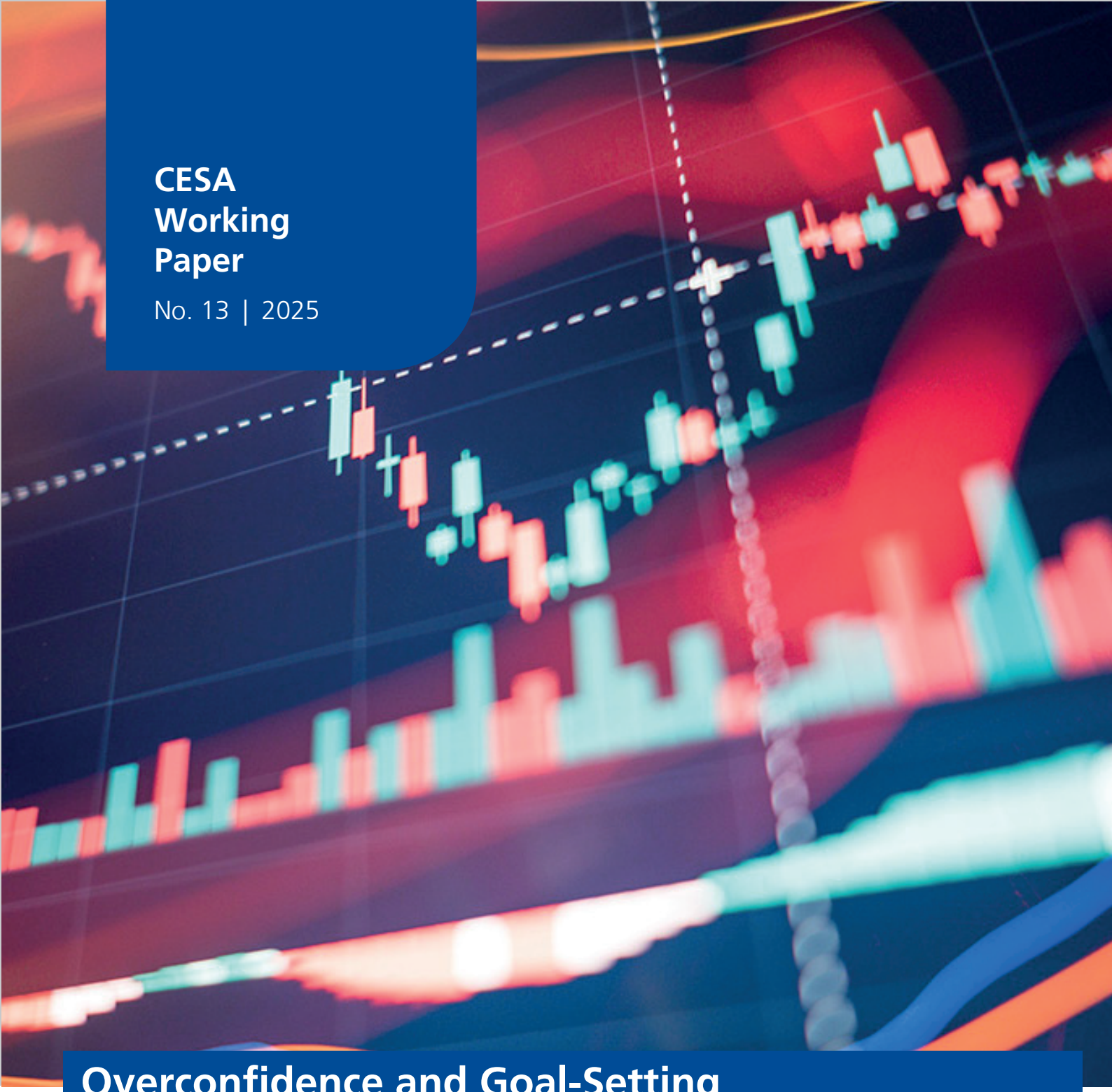


CESA
Working
Paper
No. 13 | 2025

# Overconfidence and Goal-Setting

**Sarah Rezaei[(a)], Hendrik Sonnabend[(b)] and Bastian Westbrock[(c)]**

*a Utrecht University*

*b University of Hagen (FernUniversität in Hagen)*

*c Swansea University*

FernUniversität in Hagen

# Overconfidence and Goal-Setting*

Sarah Rezaei† *Utrecht University*

Hendrik Sonnabend ‡ *Hagen University*

Bastian Westbrock§ *Swansea University*

**Abstract**

Goal-setting is widely used to overcome self-control problems but often yields underwhelming results. One likely explanation is that many individuals set overly ambitious goals to preserve their positive self-image. To counteract this tendency, we compare standard goal-setting with a goal-setting practice that includes a financial reward for goal achievement to offset the self-image costs of setting a lower, more realistic goal. In our field experiment with multiple cohorts of first-year undergraduates, the reward led to significantly more realistic goals, followed by higher study efforts and improved exam performance. These effects were particularly pronounced among the most confident students who, in the absence of rewards, actually performed worse than their peers in a control group without goal-setting. Overall, our findings stress the motivational value of goal-setting practices but also warrant caution against the formation of overly ambitious goals.

**JEL-Code**: C93, I23

**Keywords:** Goal-setting practices; Financial incentives; Field experiment; Overconfidence

## 1 Introduction

People set goals to achieve what they want, be it to get in shape for a marathon, control their desires, or write a PhD thesis. The concept of SMART (specific, measurable, attainable, relevant, and timed) goals, for example, originally popularized by management

*We are deeply grateful to Ansgar Wohlschlegel, Anke Gerber, John List, Julia Nafziger, and Elena Fumagali for their valuable and constructive comments. We also thank Annette van den Berg and Achim Zeileis for their continuous support in data collection, as well as participants of the 2021 EALE Conference in Barcelona, the 2023 ESA World Meeting in Lyon, the 2024 AFE Conference at LSE, and seminar attendees at Swansea University for their thought-provoking questions. All remaining errors are our own.

†Corresponding author. Utrecht University School of Economics, t.rezaei@uu.nl

‡Hagen University, Department of Economics and Business Administration, hendrik.sonnabend@fernuni-hagen.de

§Swansea University, School of Social Sciences, bastian.westbrock@swansea.ac.uk

practitioners (Doran, 1981), has gained widespread appeal across various domains, from sports and relationship building to academic pursuits. This popularity is not entirely surprising, as goal-setting is easily learned and maintained. Moreover, numerous laboratory studies have shown that personal goals can have an actual impact on the actions and outcomes they intend to promote, provided that a goal is sufficiently specific, challenging, and supported by the goal-setter's commitment (see Locke and Latham, 2019, for a review).[1]

However, when tested in the field, many goal-setting practices have failed to meet these high expectations (e.g., Dobronyi et al., 2019; Oreopoulos and Petronijevic, 2019; van Lent, 2019; van Lent and Souverijn, 2020; Islam et al., 2024). One likely explanation is that individuals often set too ambitious goals for themselves. Indeed, there is an abundance of research suggesting that the overestimation of one's skills and abilities is a common phenomenon among students, managers, drivers, and the population at large (e.g., Svenson, 1981; Kahneman and Tversky, 1982; Babcock and Loewenstein, 1997; Malmendier and Tate, 2015). This overconfidence poses a serious problem for goal-setting practices. The reason is that goals lose their meaning as a motivational device when they are overly ambitious. In the best case, such goals might be dismissed as something "never meant to be taken seriously". In the worst case, however, they may even have severe consequences for both the goal-setters and those around them, especially if the goal-setters remain committed to their unrealistic plans.

In this paper, we provide evidence of overambitious goals in a field setting and introduce a goal-setting practice designed to mitigate this problem. Our particular solution involves a reward for achieving one's self-set goals—an approach motivated by theories of self-image concerns. According to these theories, individuals maintain self-serving beliefs about their skills because they derive motivational benefits (Bénabou and Tirole, 2002; Brunnermeier and Parker, 2005) or direct ego-utility from them (Köszegi, 2006). Never-

[1]The underlying logic is rooted in the theory of motivation. Goals can serve as personal reference points that help to commit to self-chosen standards below which an individual feels dissatisfied with her performance (Locke and Latham, 1990; Koch and Nafziger, 2011; Hsiaw, 2013). In this way, they can help overcome potential self-control problems at the time when the individual needs to take action, but other impulses and desires become more salient (Ainslie, 1975; O'Donoghue and Rabin, 1999).

theless, these theories also assert that the benefits are consciously traded off against the costs of maintaining unrealistic beliefs, thereby creating room for incentives to promote more accurate self-assessments (Brunnermeier et al., 2017; Zimmermann, 2020).

We first present a model of goal-setting that incorporates this trade-off between self-image and realism and that extends on the framework of Koch and Nafziger (2011). The main prediction of our model is that a reward for a more thoughtful reflection on one's personal goal can offset the perceived self-image loss of a lower, more realistic goal. This, in turn, fosters greater effort and improved performance as the realistic goal becomes more worthwhile to pursue.

We then tested this prediction in a large-scale, pre-registered field experiment at two major European universities, where we introduced our goal-setting practice to three cohorts of first-year students. Out of the 2,729 participants, one-third were randomly assigned to a rewarded goal-setting intervention. In this treatment, students were asked to set a personal goal for an upcoming exam and received a financial reward for achieving that goal. The remaining two-thirds were divided between the control group, where students were not asked to set goals, and a non-rewarded goal-setting group, in which goal achievement was not incentivized.

Our findings indicate that, compared to both the control and non-reward groups, students in the reward group significantly increased their class attendance and achieved higher exam scores. In particular, they scored 1.8 points higher (out of 100) than the control group and 1.2 points higher than the non-reward group. This not only supports the key prediction of our theory that personal goals become more effective when combined with rewards, but it also places our proposed practice at the upper end of earlier goal-setting interventions in education, which reported either null effects (e.g. Dobronyi et al., 2019) or, at best, exam score improvements below a 1.8-point effect (e.g. Clark et al., 2020; van Lent and Souverijn, 2020).

Our findings also support two key mechanisms in our theory that drive this result. First, all behavioral and outcome effects in our experiment stem from a small financial reward compared to the original grade-point incentives of the exams—the unconditional

expected reward was just 2 Euros per participant. While effects from such a small payoff are difficult to reconcile with standard incentive theories, they align well with theories of self-image concerns. According to these theories, the self-image value of an action derives from the information it conveys about the actor's true ability (e.g., Bénabou and Tirole, 2004; Köszegi, 2006). However, as little can arguably be learned from briefly deliberating on one's personal goal, there are also little self-image costs in lowering this goal. Consistent with this, we find that even the small reward in our experiment can lead to significant goal reductions, amounting to over 8.2% of the goals in the non-reward group.

Second, our model predicts that the reward's positive effects should manifest among the most confident individuals while less confident individuals remain unaffected. The argument is again rooted in theories of self-image concerns. Individuals are more likely to recall and invest in planned activities that provide ego-enhancing information, such as pursuing a realistic goal, while ego-threatening information is more likely suppressed (see Eil and Rao, 2011; Zimmermann, 2020). Hence, as confident individuals set the most ambitious goals, they also have the strongest incentives to later on avoid working towards their execution. In other words, confident individuals are expected to benefit the most from mechanisms that help restrain their overambitiousness. Our findings strongly support this.

In fact, the entire performance gap between the reward and non-reward groups stems from the fact that the most confident students in the non-reward group performed even worse than their counterparts in the control group. Two possible explanations are choking under pressure (Baumeister, 1984) or effort avoidance to protect one's self-image—an extreme form of self-handicapping (Bénabou and Tirole, 2002). In either case, this suggests that goal-setters may remain committed to their ambitious plans even when such commitment becomes dysfunctional. As a practical implication, our findings thus highlight the need for instructors to exercise caution when implementing goal-setting practices without complementary measures.

Overall, thus, the key contribution of our study is a novel goal-setting practice to

mitigate the risk of excessive ambition. Traditionally, educational systems have relied on teachers to address this problem. However, years of experience with this paternalistic approach have shown that students often lack the desired commitment to these externally imposed goals, leading to calls for more self-regulation in education (e.g. Zimmerman, 1990; Lavecchia et al., 2016). Two previously tested alternatives are mentoring programs (e.g., Bursztyn and Jensen, 2015; Oreopoulos and Petronijevic, 2019) and goal-setting in front of academic mentors or peers (e.g., van Lent and Souverijn, 2020; Brandts et al., 2021; Islam et al., 2024). Yet, a key finding from these studies is that involving peers or mentors in this process might even prove counterproductive, as goals set in front of others tend to be even more ambitious than those set in solitude.

While prior research on the combination of self-set goals and rewards exists, most evidence comes from lab experiments (Corgnet et al., 2015; Dalton et al., 2016; Gonzalez et al., 2020). The only exception is Islam et al. (2024), who also examine a rewarded goal-setting intervention in an educational setting. However, their design made peer and teacher recognition so salient that the reward failed to promote more realistic goal-setting. In contrast to this study, we present a practice that succeeded in this regard.

The remainder of this paper is structured as follows: Section 2 introduces our theoretical model, Section 3 describes the experimental design, and Section 4 presents the results. Finally, Section 5 concludes.

# 2 Model

In our model, we combine personal goal-setting with self-image concerns. Our starting point is the theoretical model of Koch and Nafziger (2011), in which goals serve as personal reference points that help individuals to commit to a self-chosen standard below which they feel dissatisfied with their performance.

In the original framework, individuals have no choice but to engage with the task, and any effort below the targeted level inflicts a dissatisfaction cost on them, such as the loss of self-image associated with failing to perform. In our extension, the individuals have an additional option: they can choose to disengage entirely from the task, for instance, by

distracting themselves with alternative activities to avoid dissatisfaction. Moreover, we introduce an immediate ego-utility benefit from setting an ambitious goal, capturing the idea that individuals derive satisfaction from maintaining an optimistic view of themselves (Köszegi, 2006). We then study how this self-image benefit influences both the goal-setting stage and the choice of the outside option.[2]

To formalize our model, consider an individual who sets a goal in period one, implements the efforts in period two, and receives the returns in period three. The period-two self can select from four effort levels: $e \in \{N, L, M, H\}$, where $N$ stands for task avoidance, and $L$, $M$, and $H$ correspond to low, mediocre, and high effort, respectively. Effort comes at a cost $c_N = 0 < c_L < c_M < c_H$ and yields intrinsic benefits of $g_N = 0 < g_L < g_M < g_H$ in period three (e.g., grades in an exam). In addition, the individual may receive a (monetary) reward $R$ in period three from an external instructor for achieving her personal goal.

To capture the self-control problem, we assume that the individual has $(\beta, \delta)$-preferences with $\delta = 1$ and $0 < \beta < 1$. Specifically, we assume that, without a goal, period two self-prefers a low effort ($e = L$) and would never choose a high effort ($e = H$), even when offered a reward. Accordingly, the period-two self evaluates the future returns and present costs as follows:

$$\beta g_L - c_L > \beta g_M - c_M > \beta g_N - c_N = 0 > \beta g_H - c_H + \beta R\,. \tag{1}$$

In contrast, the period-one self ranks the costs and returns in the following way:[3]

$$\beta(g_M - c_M) > \beta(g_L - c_L) > 0 > \beta(g_H - c_H + R)\,. \tag{2}$$

[2]Note that in our model, the potential adverse effect of self-image concerns does not consist of the pursuit of an overly ambitious plan but rather of the opposite: individuals may withdraw effort entirely from the task.

[3]For consistency between (1) and (2), we require that present bias is strong but not excessive

$$\frac{c_M}{g_M} < \beta < \frac{c_M - c_L}{g_M - g_L} < 1\,.$$

Hence, these two relations establish a standard self-control problem: in the absence of a goal, the period-two self prefers the lower effort $e = L$, while the period-one self favors the higher effort $e = M$. Moreover, $e = H$ represents an unrealistically high effort level that is never optimal from either perspective, as the costs consistently outweigh the returns.

By setting a goal, the period-one self achieves two things within this framework. First, goal-setting can help overcome the self-control problem. Specifically, the period-one self can choose between the goal levels $t \in \{M, H\}$, where setting a goal introduces a dissatisfaction cost $D \geq 0$ that is incurred in period two if actual effort falls short of the target. By contrast, selecting the outside option $e = N$ avoids this cost. Second, goal-setting provides an immediate ego-utility benefit $b$ to the period-one self, which reflects the positive self-image conveyed by the goal. The more ambitious the goal, the greater this self-image benefit, such that $b_H \geq b_M$.

**Effort choices.** We begin with the optimal efforts in period two, considering three cases depending on the levels of $R$ and $D$. Consider first the case where the goal has a low commitment value, meaning that $D$ and $R$ are sufficiently small to satisfy

$$\beta g_L - c_L - D > \beta g_M - c_M + \beta R\,.$$

In this case, the future self ignores the goal, regardless of whether it is set at $t = H$ or $t = M$. As a result, she chooses the default effort level ($e^* = L$), implying that goal-setting has no effect in this case.

Thus, consider next a goal with a moderate commitment value:

$$\beta g_M - c_M + \beta R > \beta g_L - c_L - D > 0\,.$$

Here, the effectiveness of goal-setting depends on its ambition: If the goal is set at $t = H$, the future self still prefers $e^* = L$ because she fails to reach the goal regardless of whether she chooses $e = L$ or $e = M$. However, since $\beta g_H - c_H + \beta R < 0$, she finds the effort required to achieve $t = H$ too costly and instead defaults to $e^* = L$—as if the goal was

never meant to be serious. By contrast, when the goal is set at $t = M$, the dissatisfaction cost $D$, combined with the reward $R$, makes the goal worthwhile enough so that the future self follows through and chooses $e^* = M$. Thus, setting a realistic goal ($t = M$) is beneficial for effort.

Finally, consider the case of a goal with a high commitment value,

$$0 > \beta g_L - c_L - D\,.$$

Now, if the goal is set at $t = H$, then $e = L$ is no longer the best option because the self-image cost of trying and failing is too high. Instead, the future self opts for complete disengagement ($e^* = N$). In other words, when a goal has a high commitment value, setting an overly ambitious goal can even lead to lower effort than if no goal had been set at all. By contrast, when the goal is set at $t = M$, the condition $\beta g_M - c_M + \beta R > 0$ ensures that the future self still finds it worthwhile to exert effort, so $e^* = M$.

**Goal-setting.** We now turn to the goal-setting stage, where we disregard the case of a low-commitment goal for simplicity, as goal-setting has anyway no effect in this case. Thus, we assume

$$\beta g_M - c_M + \beta R > \beta g_L - c_L - D\,. \tag{3}$$

We analyze two cases in turn: first, the case of a naive goal-setter who is unaware of his self-control problem and the value of goal-setting, meaning he believes that $\hat{\beta} = 1$ and $\hat{D} = 0$ in period two. Second, the case of a sophisticated goal-setter who correctly anticipates that $\hat{\beta} = \beta$ and $\hat{D} = D$ but who has sufficiently strong self-image concerns.

We begin with the naive goal-setter. Such an individual assumes that his future effort choice is unaffected by goal-setting. Consequently, he expects, based on (2), to exert moderate effort ($e = M$) regardless of the present goal level. The discounted utility of a

moderate, respectively high, goal is thus given by

$$u(t, \hat{\beta}, \hat{D}) = \begin{cases} b_H + \beta(g_M - c_M), & \text{if } t = H \\ b_M + \beta(g_M - c_M + R), & \text{if } t = M \end{cases}.$$

Therefore, without a reward ($R = 0$), the present self prefers the ambitious goal ($t^* = H$) whenever $b_H - b_M > 0 \equiv \Delta_N(0)$.[4] By contrast, with a reward ($R > 0$), the present self prefers the more realistic goal $t^* = M$ if and only if

$$b_H - b_M < \beta R \equiv \Delta_N(R)\,. \tag{4}$$

In other words, for a naive goal-setter, the choice between an overambitious and a realistic goal depends on the trade-off between the self-image gain from setting the more ambitious goal and the foregone reward from failing to achieve it. From our period-two analysis, it then follows that rewards are clearly beneficial for effort and performance when they encourage the adoption of a realistic goal $t^* = M$.

Consider now the case of a sophisticated goal-setter who correctly anticipates the behavioral consequences ($\hat{\beta} = \beta$ and $\hat{D} = D$). Based on the expected effort in period two, the discounted utility of setting a moderate or high goal is given by

$$u(t, \beta, D) = \begin{cases} b_M + \beta(g_M - c_M + R) & \text{if } t = M \\ b_H + \beta(g_L - c_L - D) & \text{if } t = H \quad \text{and} \quad \beta g_L - c_L - D > 0 \\ b_H & \text{otherwise} \end{cases}.$$

Hence, there exists a threshold $\Delta_S(R)$ such that if the self-image gain from a high goal satisfies $b_H - b_M > \Delta_S(R)$, the present self always prefers setting an overly ambitious goal

[4]When a naive goal-setter has no self-image concerns ($b_H = b_M$) and receives no reward ($R = 0$), goal-setting has no utility consequences. We assume that such an individual sets a goal at the expected effort level $e = M$.

($t^* = H$), regardless of whether a reward is involved or not.[5] However, there also exists a second threshold, $0 < \Delta_S(0) < \Delta_S(R)$, such that if

$$\Delta_S(R) > b_H - b_M > \Delta_S(0)\,, \tag{5}$$

The reward shifts the goal-setter's preferences from an overambitious goal to a realistic one ($t^* = M$). Thus, even for a sophisticated individual, a reward can improve effort and performance by encouraging more realistic goal-setting.

All this leads to our first prediction for the experiment, in which a randomly selected share of participants is asked to set a rewarded performance goal.

**Prediction 1** *Offering a financial reward for goal achievement leads to weakly lower goal-setting, with goals being strictly lower for participants with moderate self-image concerns (i.e., when $\Delta(R) > b_H - b_M > \Delta(0)$). As a result, the reward increases effort and performance compared to a non-rewarded goal. Furthermore, the reward also enhances effort and performance compared to no goal-setting when $\Delta(R) > b_H - b_M$.*

# 3 Experimental Design

To test this prediction, we conducted a field experiment at two large, publicly funded universities in Europe. The experiment, along with its expected findings, was ethically approved and pre-registered.[6]. Our goal-setting intervention targeted several cohorts of students enrolled in the compulsory first-year courses Microeconomics and Mathematics. In the following, we provide a more detailed description of these courses and the position of our experiment within them.

[5]The threshold is given by

$$\Delta_S(R) = \begin{cases} \beta(g_M - c_M - (g_L - c_L) + D + R) & \text{moderate commitment} \\ \beta(g_M - c_M + R) & \text{high commitment} \end{cases}$$

[6]The two preregistrations can be found under AsPredicted-82808 and AEARCTR-0011737

## 3.1 Experiment in Microeconomics

Microeconomics is a first-semester undergraduate course at a university in the Netherlands. We ran our experiment over the academic years 2021–2023, during which the course was consistently taught by the same experienced instruction team. The course consists of weekly lectures delivered to all 500–650 course participants and two weekly tutorials held in smaller groups of 20–30 students. Attendance at these activities is not mandatory; however, students must demonstrate a minimum level of effort to qualify for a retake of the course assessment. This assessment includes a sit-in midterm exam conducted midway through the course (contributing 25% to the final grade) and a sit-in final exam at the end (contributing 75%). While all course elements are typically held in person, the lectures and tutorials were delivered online during the nationwide COVID-19 lockdown in 2021. We discuss the implications of this format change in Section 4.3.2.

Our experiment began right at the end of the midterm exam. It included three surveys distributed at different points in time: a pen-and-paper questionnaire distributed immediately after the midterm exam to all students present in the room (Survey 1), an online questionnaire issued one day after the midterm grades were published (Survey 2), and a final pen-and-paper questionnaire distributed to all students present at the final exam (Survey 3). The key questionnaire in our experiment is Survey 2. Here, we asked a random subset of students to set a personal goal for their final exam. Since this survey was similar in both Microeconomics and Mathematics, we provide its details below.

In the other two surveys, we collected additional information on the socio-demographic backgrounds (gender, nationality, GPA, parental education), learning efforts (lecture and tutorial attendance, self-study hours), and students' expected grades for the exams they had just completed. To encourage participation, students could earn up to 5 bonus points on their final exam scores. All they had to do was to write their names and student IDs on the survey sheets; neither consent nor completion was mandatory. Nevertheless, both the consent and completion rates exceeded 99%. Moreover, we were able to recruit 962 students to complete Survey 2 (56.6% of the course population), 795 students to complete both Surveys 1 and 2 (46.7%), and 749 students to answer all three surveys.

## 3.2 Experiment in Mathematics

Mathematics is a compulsory first-semester course in the Economics program of a university in Austria. We conducted our experiment during the academic years 2021–2024, with the same experienced instructional team delivering the course throughout. The format and content follow the standard curriculum for introductory mathematics courses. It includes weekly lectures in a flipped classroom format, delivered to the entire cohort of 500–700 students, alongside weekly tutorials in smaller groups of 30–40 students. Like the Microeconomics course, assessment is based on sit-in midterm exams (two exams, jointly accounting for 60% of the final grade) and a sit-in final exam (40%). Similarly, all lectures and tutorials were delivered online during the COVID-19 lockdown in 2021.

The Mathematics experiment closely mirrored the one used in the Microeconomics course. However, unlike in the latter, we only administered Survey 2, timed at a similarly critical point in the course: shortly after the midterm grades were released and two weeks before the final exam. Participation was incentivized with three bonus points, which were awarded independent of consent or completion. Nevertheless, both consent and completion rates remained again above 99%, and we recruited 1,767 students to participate in our experiment (ca. 75% of the course population).

Appendix A.2 presents summary statistics for all students who participated in Survey 2 of Mathematics and Surveys 1 and 2 of Microeconomics. Additionally, it provides evidence on the balance of our sample across the control and two treatment groups.

## 3.3 Goal-setting Intervention

We administered our goal-setting intervention to a randomly selected subset of students completing Survey 2. The remaining two-thirds of participants were divided between the control and the other treatment groups, where goal achievement was not incentivized. All three survey versions began with a set of questions about students' satisfaction with their midterm exams, followed by five self-control questions validated with the Maas

et al. (2017) self-control cognition questionnaire.[7] Additionally, students in Mathematics answered the sociodemographic questions included in Survey 1 of Microeconomics.

The goal-setting question appeared at the end of Survey 2. In this question, we asked students to set a personal goal for their final exams, offering them a choice between seven pre-specified target scores, ranging from 60% to 90% of the maximum score. The key difference between our two treatments lies in the incentivization: students in the reward treatment earned experimental tickets for achieving their self-set goals. These tickets increased their chances of winning a monetary prize of up to 100 Euros at the end of the course.

To discourage systematic underambitious goal-setting in the reward treatment, we implemented a progressive reward scheme in which higher target scores yielded more experimental tickets upon achievement. Nevertheless, the overall reward was deliberately kept at a modest level, so the unconditional expected payout was just 2 Euros per participant. Further details on the reward scheme and the exact wording of the goal-setting question can be found in Appendix A.1.

# 4 Main Results

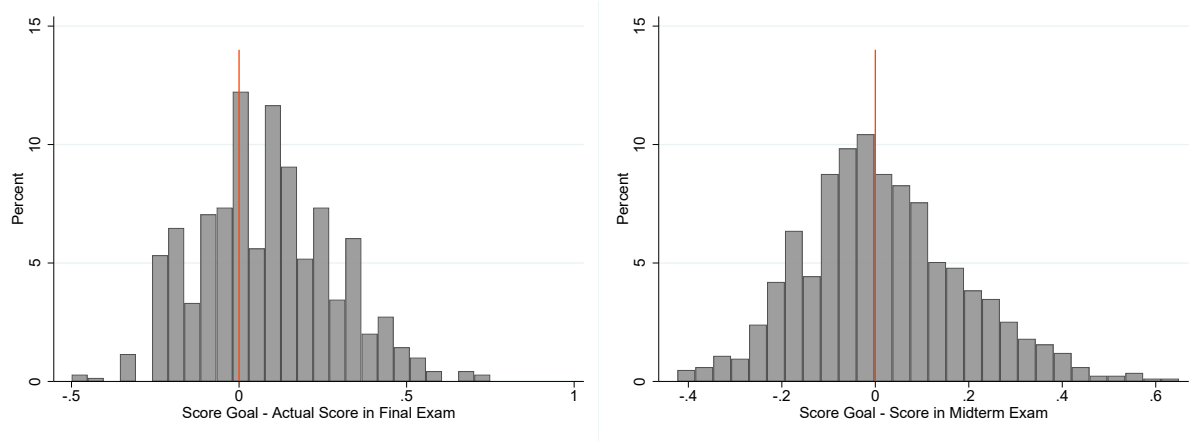
We begin with an analysis of the goal-setting behavior in the non-reward treatment before we turn to the effects of the reward.

## 4.1 Descriptive Statistics on Non-rewarded Goals

Motivated by the general observation that individuals tend to set overly ambitious goals when left unguided, Figure 1 presents two measures of overconfident goal-setting among the participants in our non-reward treatment. The left panel compares participants' self-set goals with their actual performance in the end-term exam, the right panel compares them with their performance in the preceding midterm exam.

[7] Although prior research suggested that goal-setting interventions might be more effective for individuals with lower self-control (e.g., Koch and Nafziger, 2011; Clark et al., 2020), our analysis did not reveal any moderating effect of the self-control measures elicited.

Figure 1: Goals in the Non-reward Treatment




*Notes:* Figures illustrate the frequencies of goals in the non-reward treatment that exceeded participants' actual exam performance. The left panel compares goals with final exam scores, and the right panel compares them with midterm scores.

As suggested by the left panel, more than 60% of participants failed to achieve their target scores, a figure that rises to even 67% when including those in the treatment group who did not attend the final exam. Similarly, the right panel shows, though somewhat less extreme, that more than 49% set goals that exceeded their preceding midterm exam performance. Together, these patterns point to a consistent tendency toward overambitious goal-setting.

## 4.2 Effects of Goal Incentives

We now examine the effects of introducing a financial reward for goal achievement. Our findings are summarized in Table 1. The table presents results from five OLS regressions that analyze the impact of participating in one of our treatment groups on students' self-set performance goals, post-midterm learning efforts (class attendance and self-study hours), final exam performance, and the likelihood of achieving their goals. Accordingly, each regression includes two indicator variables capturing a student's random assignment to one of the treatment groups.

To improve the fit of these models, we additionally controlled for students' sociodemographic characteristics (gender, nationality, midterm exam score, parental education), pre-treatment self-study efforts, course-year fixed effects, and interactions between these

student characteristics and the course-year effects. Moreover, we clustered the standard errors of these models at the tutorial group level to account for potential spillovers between students within the same group. Nevertheless, all our results remain robust—though estimated with lower precision—when we do not control for these factors. Table A3 in Appendix A.3 provides further details.

Table 1: Treatment Effects on Goals, Efforts, and Performance

| | Goal Level | Class Attendance | Self-study | Exam Score |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Effect of Asking Students to Set... | | | | |
| Non-rewarded Goal | Baseline | .001 | .001 | .006 |
| | | (.022) | (.012) | (.010) |
| Rewarded Goal | -.054*** | .049** | .004 | .018** |
| | (.005) | (.022) | (.011) | (.011) |
| Δ Treatment Effect | -.054*** | .048** | .003 | .012 |
| (F-statistic) | (106.02) | (4.86) | (0.06) | (1.83) |
| Mean in Reference Group | .661 | .607 | .147 | .580 |
| Observations | 1,697 | 722 | 725 | 2,128 |

*Notes:* The table reports the results from five OLS regressions, where the main regressors are two indicator variables for participation in one of the treatment groups. All dependent variables are normalized on a scale from 0 (lowest) to 1 (highest). The sample varies across columns: Column (1) includes all experimental participants in one of the treatment groups, columns (2) and (3) include all Microeconomics participants, column (4) includes all participants who took the final exam, and column (5) all treated students who took the final exam. The reference group differs accordingly: in Columns (1) and (5), it consists of participants in the non-reward treatment, while in Columns (2)–(4), it comprises participants in the control group. Standard errors in parentheses clustered at the tutorial group level. Significance levels * p<0.1 ** p<0.05, *** p<0.01.

### 4.2.1 Effect on Goal-Setting

Column (1) examines the effect of our treatments on students' behavior during the goal-setting stage. Accordingly, the analysis is restricted to the 1,345 students who participated in one of the two goal-setting treatments. The results clearly confirm one of the key predictions of our theory: the reward led to a significant reduction in goal levels. On average, participants in the reward group set final exam targets that were 5.4 points lower than those in the non-reward group (significant in a two-sided test at $p < .001$). In

relative terms, this corresponds to an 8.2% decrease from the non-rewarded group's mean target score of 66.1 points, suggesting a substantial behavioral adjustment.

Figure 2 examines this effect more closely. As shown in the top panel, students consistently lowered their target scores along the non-reward group's entire grade goal distribution. Moreover, as shown in the bottom panel, students in the reward group did not just lower their target scores to the lowest possible level. Rather, they spent significantly more time on their goal-setting question: on average, 22 seconds longer than their non-rewarded peers, supporting our interpretation that the reward encouraged more careful consideration of a realistic and achievable goal.

#### 4.2.2 Effect on Study Efforts

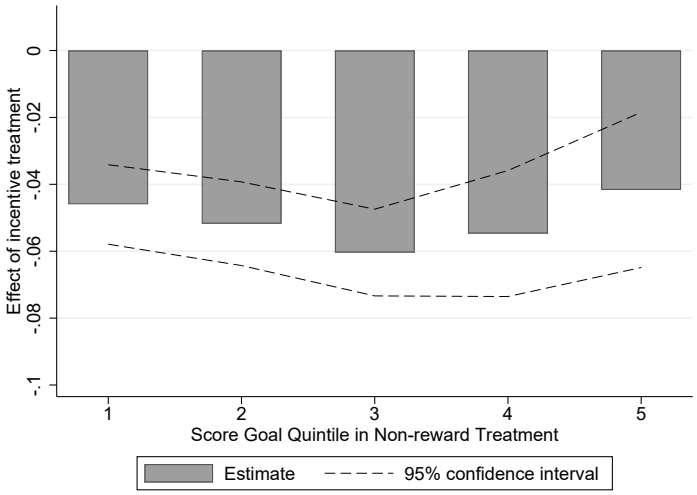
Columns (2) and (3) of Table 1 examine another key mechanism behind our theoretical predictions: the impact of the reward treatment on students' post-treatment learning efforts, measured by class attendance and self-study time in the remaining weeks before the exam. While these measures are self-reported and available only for students in Microeconomics, they nonetheless provide valuable insights because any performance improvement from our intervention would need to be driven by increased study efforts.

Indeed, column (2) shows that students in the reward treatment significantly increased their class attendance in lectures and tutorials by 4.9 percentage points relative to the control group ($p = .029$) and by 4.8 points compared to the non-reward group ($p = .031$). Moreover, even though the treatment's effect on self-study effort, reported in column (3), is not statistically significant, the coefficient is still in line with our expectations. By contrast, the non-reward treatment had no measurable impact on either attendance or self-study effort.

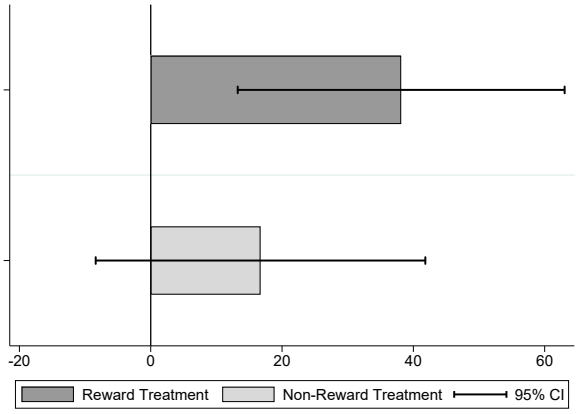
#### 4.2.3 Effect on Exam Performance

Column (4) turns to the core results of our analysis: the effects of our two treatments on students' final exam performance. This analysis includes all 2,128 experimental partici-

Figure 2: Effect of a Reward for Goal-Achievement on Goal-Setting Behavior




(a) Effect on Goal Quintiles of the Non-reward Group




(b) Effect on Survey Completion Time

*Note:* Panel (a) shows the estimated effect of the offered reward on the goal levels across different quintiles of the non-reward group. Panel (b), in turn, displays the estimated effects on the median response time for Survey 2. All estimates are derived from quantile regressions using the same set of independent variables as in Table 1.

pants who also participated in that exam.[8] In line with our expectations, the reward in our experiment—despite its modest size—had a significant impact on exam performance. Students in the reward group scored, on average, 1.8 points higher (out of 100) than those in the control group (significant at p=.04 in a two-sided test) and 1.2 points higher than those in the non-reward group (not significant: $p = .18$). By contrast, students in the

[8] As shown in Table A4 in Appendix A.3, students' dropouts from the final exam, and hence any remaining attrition in the courses after our treatments, were unrelated to them. The very same table also shows that none of our treatments had any discernible impact on the average student's likelihood of passing the exam.

non-reward group scored only an insignificant 0.6 points more than those in the control group, a difference that is not statistically significant ($p = .44$).

Hence, while not all differences in Table 1 are statistically significant, the overall pattern supports the key notion of our theory that a reward for goal achievement can enhance the motivational value of a goal-setting practice by encouraging individuals to set more realistic goals.[9]

## 4.3 The Role of Self-Image Concerns

So far, we have established that students were, on average, quite overconfident in their goal-setting choices, a tendency that was (partially) remedied by the promise of a financial reward. In this section, we present additional evidence suggesting that self-image concerns play a key role in this behavior.

The following prediction guides our first analysis:

**Prediction 2** *When participants have small self-image concerns (i.e., $b_H - b_M \leq \Delta(0)$), both rewarded and non-rewarded goal-setting yield identical effort and performance improvements compared to no goal-setting.*

The intuition is the following: As individuals without self-image concerns do not derive any benefit from making overly ambitious plans, they set the same realistic goals ($t* = M$), regardless of whether they are rewarded for their achievements. As a result, even their non-rewarded goals become inherently worthwhile to pursue for them. By contrast, individuals with self-image concerns are still subject to Prediction 1 and, consequently, benefit from the opportunity to temper their ambitions.

In the following, we test these discriminatory predictions using individual confidence measures as proxies for participants' self-image concerns.

[9]While Table 1 presents test results for single outcome variables, our findings are also partially robust to joint significance tests. Specifically, compared to students in the control group, our reward treatment's effort and performance effects remain jointly significant after applying a Holm-Bonferroni correction for multiple hypotheses ($\alpha = .058$). In contrast, when compared to the non-reward group, joint significance can only be confirmed for goal level and class attendance effects, but not for students' exam performance. Moreover, further tests confirmed that our experiment had no unintended side effects on the outcomes of parallel courses.

Table 2: Self-Confidence among Different Demographic Groups of Students

| | High edu. Male | Low edu. Male | High edu. Female | Low edu. Female |
|---|---|---|---|---|
| Expected - Actual Midterm Score | .03 | -.01** | .03 | .02 |
| Non-Rewarded Goal - Midterm Score | .03 | .03 | .02 | -.02*** |
| Non-Rewarded Goal - Endterm Score | .11 | .08 | .04*** | .08 |

*Notes:* The table presents mean values of three confidence measures for different groups of experimental participants, distinguished by gender and parental education. In particular, low versus high parental education distinguishes between students whose parents did (not) go to the university themselves. Numbers of observations vary across measures: Expected midterm scores are only available for Microeconomics students, while non-rewarded goals are only available for students who participated in the non-reward treatment. Asterisks indicate significant mean differences compared to the group of high-education males. Significance levels: * p<0.1 ** p<0.05, *** p<0.01 in two-sided tests.

### 4.3.1 Moderating Effect of Self-Confidence

To investigate the role of individual differences in self-image concerns, we draw on insights from prior literature documenting inherent confidence gaps between different demographic groups. Notably, there is extensive research suggesting that men tend to hold more optimistic self-views than women (e.g., Niederle and Vesterlund, 2011; Buser et al., 2017; Hoyer et al., 2020), while other studies suggest that family background plays an important role as well (Sacerdote, 2007; Waldfogel and Washbrook, 2011). Synthesizing these insights, Almås et al. (2016) find that, in a nationally representative sample of adolescents, the group of male students from high-socioeconomic status families exhibits the highest level of self-confidence. In contrast, their study does not reveal a discernible confidence gap between girls and boys from lower-status backgrounds.

This pattern is also evident within our student populations. As shown in Table 2, rows two and three, male students from families with a university background set the most ambitious goals for their final exams, both relative to their midterm exam performance and their actual final score. Moreover, as indicated in row one, these students also had the highest grade expectations when asked at the end of their midterm exams to predict their midterm result.

Building on this measured confidence gap between high-education male students and the rest, Table 3 revisits the exam performance effect of our reward treatment again,

this time interacted with an indicator variable for a student from this group. Column (1) reproduces the average treatment effects from Table 1. Column (2) introduces the interaction term. As anticipated, the results indicate that the treatment effect is concentrated within the high-education male group: students in this group experienced a significant additional increase of 3.1 exam points ($p = .043$) compared to the non-reward group. At the same time, the reward had no discernible impact on the other three groups, who benefited equally much from both treatments, improving their exam scores by 3.2 points compared to the control group ($p = .005$). Thus, as expected as well, less confident individuals do not need a reward for their goal-setting to be effective.[10]

### 4.3.2 Moderating Effect of Peer Exposure

Table 3 reveals another intriguing result: The reason male students from high-education families benefited from the reward treatment is that their exam performance actually deteriorated in the absence of rewards. As shown in column (2), their exam scores dropped by a notable 3.5 points compared to no goal-setting ($p = .013$).

This finding is not surprising from the viewpoint of our theory, which predicts:

**Prediction 3** *When participants have strong self-image concerns (i.e., $b_H - b_M > \beta(g_M - c_M)$), effort and performance deteriorate in the commitment value of a goal, so that the goal-setter exerts even less effort than without a goal.*

The intuition is that committed goal-setters do not simply dismiss their overly ambitious plans as something "never meant seriously" and subsequently adjust their goals or revert to default effort. Instead, they may avoid engaging with the task altogether as a means of protecting their self-image.

How, then, can we explain students' strong, self-defeating commitment to their final exam goals? One way to understand this is that our goal-setting intervention did not occur

[10]The results in Table 3 are robust with regard to various alternative regression specifications: (1) without student controls, (2) split regressions for students of differing gender and/or educational background, (3) other sample splits along the parental education dimension. Moreover, the differential exam score effects shown in this table also affect students' likelihood of passing the final exam. Table A5 in Appendix A.3 provides more evidence.

Table 3: Performance Effects for Students of Low versus High Baseline Confidence

| | Final Exam Score | | | |
|---|---|---|---|---|
| | 2021–24 | | COVID year | post COVID |
| | (1) | (2) | (3) | (4) |
| Non-rewarded Goal | .006 | .032*** | .059*** | .023 |
| | (.010) | (.012) | (.020) | (.013) |
| Non-rewarded Goal × High-education Male | | -.067*** | -.069** | -.067*** |
| | | (.019) | (.028) | (.022) |
| Rewarded Goal | .018** | .032*** | .044** | .027** |
| | (.009) | (.011) | (.021) | (.011) |
| Rewarded Goal × High-education Male | | -.035** | -.033 | -.038* |
| | | (.018) | (.028) | (.023) |
| Δ Treatment Effect | .012 | 0 | -.015 | .004 |
| (F-statistic) | (1.83) | (.01) | (.43) | (.13) |
| Δ Treatment Effect for High-education Male | | .031** | .021 | .033** |
| (F-statistic) | | (5.06) | (.99) | (4.20) |
| Non-reward Treatment Effect for High-education Male | | -.035** | -.010 | -.044** |
| (F-statistic) | | (6.38) | (.23) | (6.83) |
| Mean in Reference Group | .580 | .580 | .510 | .607 |
| Observations | 2,128 | 2,128 | 602 | 1,526 |

*Notes:* Columns report OLS regressions of final exam scores on a similar set of independent variables as in Table 1. Standard errors in parentheses clustered at the level of tutorial groups. Significance levels * p<0.1 ** p<0.05, *** p<0.01.

in isolation. Instead, students had the opportunity to discuss their goals in hallways or classrooms afterward, thereby transforming their goals into binding commitments. To examine the potential impact of this peer commitment channel, we divided our sample into two periods for our final analysis: the COVID-19 lockdown year (2021) and the post-COVID period (2022–2024). We then re-assessed the treatments' effects in each period separately. Columns (3) and (4) of Table 3 present the findings. As expected, the negative impact of the non-reward treatment on males with highly educated families disappeared in 2021, aligning with the idea that students could dismiss their ambitious goals more easily under the unique conditions of the lockdown. In contrast, as shown in column (4), the negative effect resurfaced in the post-COVID period, with high-education male students losing an even larger 4.4 exam points compared to the control group ($p = 0.011$).

# 5 Conclusion

This study investigates a critical but often overlooked aspect of personal goal-setting practices: the tendency to set overly ambitious goals. Using a large-scale field experiment, we replicate this phenomenon and provide evidence for a potential remedy—rewarding individuals for setting more realistic goals. This intervention is motivated by a theoretical model suggesting that overambitious goals arise from ego-enhancing utility benefits associated with them. Our empirical findings indicate that even a small financial reward can counteract this self-image effect, significantly improving the behaviors and outcomes that the goal aims to promote—in our case, university students' class attendance and exam performance.

Beyond the immediate relevance of this finding for other goal-setting practices, our study also contributes to the broader literature on self-regulation tools in education (e.g. Zimmerman, 1990; Paris and Paris, 2003). Many of these tools have underperformed relative to expectations (Oreopoulos and Petronijevic, 2019), with inaccurate self-reflections and self-assessments——an integral component of many such tools——likely playing a role. Offering rewards may thus also enhance the effectiveness of these tools.

A second key insight from our study concerns the mechanisms behind the positive effects of rewards. Our model suggested the possibility, and our experiment confirmed, that a significant part of this effect stems from mitigating a potential backfiring of goal-setting without rewards. We found that highly self-confident students performed even worse than their counterparts who were not asked to set goals. Hence, our findings highlight a potentially serious risk of implementing self-regulation tools without appropriate safeguards and underscore the importance of accounting for individual differences when evaluating their impact.

# References

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82(4):463.

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., and Tungodden, B. (2016). Willingness to compete: Family matters. *Management Science*, 62(8):2149–2162.

Babcock, L. and Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives*, 11(1):109–126.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, 46(3):610.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Bénabou, R. and Tirole, J. (2004). Willpower and personal rules. *Journal of Political Economy*, 112(4):848–886.

Brandts, J., El Baroudi, S., Huber, S. J., and Rott, C. (2021). Gender differences in private and public goal setting. *Journal of Economic Behavior & Organization*, 192:222–247.

Brunnermeier, M. K., Papakonstantinou, F., and Parker, J. A. (2017). Optimal time-inconsistent beliefs: Misplanning, procrastination, and commitment. *Management Science*, 63(5):1318–1340.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.

Bursztyn, L. and Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics*, 130(3):1329–1367.

Buser, T., Peter, N., and Wolter, S. C. (2017). Gender, competitiveness, and study choices in high school: Evidence from switzerland. *American Economic Review*, 107(5):125–130.

Clark, D., Gill, D., Prowse, V., and Rush, M. (2020). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics*, 102(4):648–663.

Corgnet, B., Gómez-Miñambres, J., and Hernán-Gonzalez, R. (2015). Goal setting and monetary incentives: When large stakes are not enough. *Management Science*, 61(12):2926–2944.

Dalton, P. S., Gonzalez, V., and Noussair, C. N. (2016). Self-chosen goals: Incentives and gender differences. *CentER Discussion Paper Series No. 2016-036*.

Dobronyi, C. R., Oreopoulos, P., and Petronijevic, U. (2019). Goal setting, academic reminders, and college success: A large-scale field experiment. *Journal of Research on Educational Effectiveness*, 12(1):38–66.

Doran, G. T. (1981). There's a smart way to write managements's goals and objectives. *Management Review*, 70(11).

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38.

Gonzalez, V., Dalton, P. S., and Noussair, C. (2020). The dark side of monetary bonuses: Theory and experimental evidence. *CentER Discussion Paper Series No. 2020-001*.

Hoyer, B., van Huizen, T., Keijzer, L., Rezaei, S., Rosenkranz, S., and Westbrock, B. (2020). Gender, competitiveness, and task difficulty: Evidence from the field. *Labour Economics*, 64:101815.

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

Islam, A., Kwon, S., Masood, E., Prakash, N., Sabarwal, S., and Saraswat, D. (2024). All pain and no gain: When goal setting leads to more effort but no gains in test scores. *Economics of Education Review*, 103:102594.

Kahneman, D. and Tversky, A. (1982). Intuituve prediction: Biases and corrective procedures. In D, D. K., Slovic, P., and Tversky, A., editors, *Judgment Under Uncertainty: Heuristics and Biases*, pages 414–421. Cambridge University Press, Cambridge, UK.

Koch, A. K. and Nafziger, J. (2011). Self-regulation through goal setting. *Scandinavian Journal of Economics*, 113(1):212–227.

Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4):673–707.

Lavecchia, A. M., Liu, H., and Oreopoulos, P. (2016). Behavioral economics of education: Progress and possibilities. In *Handbook of the Economics of Education*, volume 5, pages 1–74. Elsevier.

Locke, E. A. and Latham, G. P. (1990). *A theory of goal setting & task performance.* Prentice-Hall, Inc.

Locke, E. A. and Latham, G. P. (2019). The development of goal setting theory: A half century retrospective. *Motivation Science*, 5(2):93.

Maas, J., Keijsers, G. P. J., Cangliosi, C. M., van der Veld, W., Tanis-Jacobs, J., and van Minnen, A. (2017). The self-control cognition questionnaire: Cognitions in the maintenance of unwanted habits. *European Journal of Psychological Assessment*, 33(5):328–335.

Malmendier, U. and Tate, G. (2015). Behavioral CEOs: The role of managerial overconfidence. *Journal of Economic Perspectives*, 29(4):37–60.

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

O'Donoghue, T. and Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1):103–124.

Oreopoulos, P. and Petronijevic, U. (2019). The remarkable unresponsiveness of college students to nudging and what we can learn from it. *NBER Working Paper No. 26059*.

Paris, S. G. and Paris, A. H. (2003). Classroom applications of research on self-regulated learning. *Educational Psychology*, 36(2):89–101.

Sacerdote, B. (2007). How large are the effects from changes in family environment? a study of korean american adoptees. *The Quarterly Journal of Economics*, 122(1):119–157.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2):143–148.

van Lent, M. (2019). Goal setting, information, and goal revision: A field experiment. *German Economic Review*, 20(4):949–972.

van Lent, M. and Souverijn, M. (2020). Goal setting and raising the bar: A field experiment. *Journal of Behavioral and Experimental Economics*, 87:101570.

Waldfogel, J. and Washbrook, E. (2011). Early years policy. *Child Development Research*, 2011(1):343016.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1):3–17.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–363.

# A Appendix

## A.1 Goal-setting Questions

**Non-rewarded goal-setting:** What are your ambitions for the final exam? My ambition is to achieve a grade of at least ... in the exam. Please specify a grade goal below, but think carefully before you set this goal!

◯ My ambition grade is 6 or higher.

◯ My ambition grade is 6.5 or higher.

◯ My ambition grade is 7 or higher.

◯ My ambition grade is 7.5 or higher.

◯ My ambition grade is 8 or higher.

◯ My ambition grade is 8.5 or higher.

◯ My ambition grade is 9 or higher.

**Rewarded goal-setting:** What are your ambitions for the final exam? My ambition is to achieve a grade of at least ... in the exam. Please specify a grade goal below. But think carefully, as you earn experimental tickets when your actual grade is at least as high as your stated grade goal. In this case, you will earn X experimental tickets as stated below. Otherwise, you earn 0 tickets.

◯ My ambition grade is 6 or higher. You earn 14 tickets if you achieve at least this grade in the final exam, nothing otherwise.

◯ My ambition grade is 6.5 or higher. You earn 16 tickets if you achieve at least this grade in the final exam, nothing otherwise.

◯ My ambition grade is 7 or higher. You earn 18 tickets if you achieve at least this grade in the final exam, nothing otherwise.

◯ My ambition grade is 7.5 or higher. You earn 20 tickets if you achieve at least this grade in the final exam, nothing otherwise.

◯ My ambition grade is 8 or higher. You earn 23 tickets if you achieve at least this grade in the final exam, nothing otherwise.

◯ My ambition grade is 8.5 or higher. You earn 26 tickets if you achieve at least this grade in the final exam, nothing otherwise.

◯ My ambition grade is 9 or higher. You earn 30 tickets if you achieve at least this grade in the final exam, nothing otherwise.

**Incentives for Participation and Goal-setting:** Participation in the experiment was incentivized with course grade points, which were directly added to students' final exam scores. Students earned these points simply by providing their names and IDs on the surveys; neither consent nor survey completion was required. Students could additionally earn up to 100 experimental tickets by completing the surveys and accurately answering several verifiable items. Importantly, as already indicated in the sample question above, participants in the reward treatment could earn up to 30 experimental tickets for meeting their personal performance goals. To encourage some level of ambition, we implemented a progressive reward scheme where higher target scores yielded disproportionally more tickets upon achievement. Each ticket earned in this way increased the student's chances of winning one of several monetary prizes of 100 Euros, which were awarded through a lottery at the end of the course.

## A.2 Pre-treatment descriptives and balance tests

In this appendix, we present the results of balance tests to confirm that students assigned to the reward and non-reward treatments did not differ in any measurable pre-treatment characteristic from those in the control group. The following tables present test results separately for the student cohorts in the Netherlands (Microeconomics 2021–2023) and Austria (Mathematics 2021–2024).

Table A1: Comparison of Students in the Microeconomics experiment

| | *Control-Reward comparison* | | | | |
|---|---|---|---|---|---|
| | Control | Rewarded | Difference | SE | p-value |
| Male | .65 | .68 | -.036 | .041 | .378 |
| Parental education | 8.42 | 8.44 | -.018 | .156 | .906 |
| High-education Male | .80 | .84 | -.036 | .040 | .378 |
| Foreigner | .49 | .47 | .025 | .043 | .560 |
| Highschool GPA | 7.56 | 7.52 | .041 | .094 | .665 |
| Midterm exam score | .56 | .59 | -.024 | .160 | .127 |
| | *Control-Non-reward comparison* | | | | |
| | Control | Non-rewarded | Difference | SE | p-value |
| Male | .65 | .60 | .051 | .043 | .237 |
| Parental education | 8.43 | 8.25 | .180 | .170 | .290 |
| High-education Male | .80 | .77 | .030 | .046 | .510 |
| Foreigner | .49 | .51 | -.012 | .044 | .792 |
| High school GPA | 7.56 | 7.58 | -.024 | .094 | .795 |
| Midterm exam score | .56 | .58 | -.020 | .016 | .217 |

Table A2: Comparison of Students in the Mathematics experiment

| | *Control-Reward difference* | | | | |
|---|---|---|---|---|---|
| | Control | Rewarded | Difference | SE | p-value |
| Male | .57 | .55 | .019 | .029 | .514 |
| Parental education | 7.52 | 7.36 | .158 | .160 | .319 |
| High-education Male | .57 | .51 | .060 | .039 | .118 |
| Foreigner | .57 | .54 | .026 | .029 | .379 |
| High school GPA | 8.06 | 8.17 | -.107 | .076 | .162 |
| Midterm exam score | .67 | .67 | .008 | .010 | .389 |
| | *Control-Nonreward difference* | | | | |
| | Control | Non-rewarded | Difference | SE | p-value |
| Male | .57 | .56 | .012 | .029 | .692 |
| Parental education | 7.52 | 7.46 | .062 | .162 | .701 |
| High-education Male | .57 | .58 | -.014 | .038 | .718 |
| Foreigner | .57 | .54 | .028 | .029 | .327 |
| High school GPA | 8.07 | 8.0 | .067 | .080 | .404 |
| Midterm exam score | .67 | .67 | .003 | .010 | .762 |

## A.3 Robustness checks

**Raw Treatment Effects:** Table A3 reports the raw treatment effects derived from regression estimations analogous to those presented in the main text but without including student characteristics as control variables. As shown, the coefficients are qualitatively similar to those reported in Table 1, albeit estimated with lower precision.

Table A3: Raw Treatment Effects on Goals, Efforts, and Performances

| | Goal Level | Class Attendance | Self-study | Exam Score |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Non-rewarded Goal | Baseline | -.001 | -.004 | .010 |
| | | (.022) | (.013) | (.012) |
| Rewarded Goal | -.059*** | .042* | -.006 | .015 |
| | (.006) | (.022) | (.011) | (.011) |
| $\Delta$ Treatment Effect | -.059*** | .043* | -.002 | .005 |
| (F-statistic) | (82.65) | (3.83) | (0.03) | (0.14) |
| Mean in Reference Group | .652 | .607 | .147 | .592 |
| Observations | 1,421 | 722 | 725 | 2,128 |

*Notes:* Columns report OLS regressions with similar specifications as in Table 1 but without any student characteristics as controls. Standard errors in parentheses clustered at the level of tutorial groups. Significance levels * p<0.1 ** p<0.05, *** p<0.01.

**Treatment Effects on Exam Participation and Passing Rate.** Table A4 reports the effects of our two treatments on the likelihood of participating in the final exam and the chance of passing it. As it becomes clear, none of our treatments had any systematic impact on the average student's behavior in these dimensions.

**Treatment Effects Differing by Gender and Parental Background.** Table A5 reports the treatment effects on students' final exam performance separately for students of differing genders and/or educational backgrounds. Clearly, the effects of our reward treatment only become significant and sizable for the group of high-education males.

Table A4: Treatment Effects on Likelihoods of Exam Participation and Passing Rate

| | Exam Participation (1) | Passing the Exam (2) |
|---|---|---|
| Non-rewarded Goal | -.014 | .002 |
| | (.015) | (.023) |
| Rewarded Goal | -.006 | .009 |
| | (.017) | (.022) |
| Δ Treatment Effect | .008 | .007 |
| (F-statistic) | (.25) | (.13) |
| Mean in Reference Group | .852 | .665 |
| Observations | 2,533 | 2,128 |

*Notes:* Columns report OLS regressions with similar specifications as in Table 1. Standard errors in parentheses clustered at the level of tutorial groups. Significance levels * p<0.1 ** p<0.05, *** p<0.01.

Table A5: Treatment Effect Differences by Gender and Parental Background

| | Final Exam Score | | | |
|---|---|---|---|---|
| | Low edu. | High edu. | Female | Male |
| Non-rewarded goal | .042*** | -.008 | .040*** | -.018 |
| | (.013) | (.011) | (.013) | (.011) |
| Rewarded Goal | .041*** | .009 | .049*** | -.002 |
| | (.015) | (.011) | (.013) | (.013) |
| Δ Treatment | -.000 | .017* | .009 | .016 |
| (F-statistic) | (.00) | (3.07) | (.045) | (1.93) |
| Observations | 794 | 1,334 | 885 | 1,243 |
| | Low edu. Male | High edu. Male | Low edu. Female | High edu. Female |
| Non-rewarded goal | .018 | -.035** | .060*** | .026 |
| | (.019) | (.014) | (.021) | (.017) |
| Rewarded Goal | .001 | -.004 | .087*** | .030* |
| | (.022) | (.015) | (.021) | (.017) |
| Δ Treatment | -.018 | .031** | .026 | .004 |
| (F-statistic) | (.74) | (5.48) | (1.93) | (.06) |
| Observations | 413 | 810 | 361 | 524 |

*Notes:* Columns report findings from eight independent OLS regressions of final exam scores on the two indicator variables for our treatment groups and the same array of control variables already used in Table 1. Standard errors in parentheses clustered at the tutorial group level. Significance levels * p<0.1 ** p<0.05, *** p<0.01.

# CESA Working Paper Series

**CESA Working Paper No. 1 | 2023**
Andreas Hauptmann, Benjamin Schwanebeck and Hans-Jörg Schmerer
Plant-level adjustments to imports and exports at the extensive margin

**CESA Working Paper No. 2 | 2023**
Joscha Beckmann, Marco Kerkemeier and Robinson Kruse-Becher
Regime-specific exchange rate predictability

**CESA Working Paper No. 3 | 2023**
Joscha Beckmann, Timo Heinrich and Jennifer Rogmann
Inflation expectations and cognitive uncertainty

**CESA Working Paper No. 4 | 2024**
Daniel Monsees and Matthias Westphal
Disruptions in Primary Care: Can Resigning GPs Cause Persistently Negative Health Effects?

**CESA Working Paper No. 5 | 2024**
Giulio Callegaro, Mario Lackner and Hendrik Sonnabend
The Napoleon complex revisited: New evidence from professional soccer

**CESA Working Paper No. 6 | 2024**
Mario Lackner and Hendrik Sonnabend
When performance melts away: Heat causes mental errors in high-stakes competitions

**CESA Working Paper No. 7 | 2024**
Johannes Hollenbach, Hendrik Schmitz and Matthias Westphal
Gene-environment interactions with essential heterogeneity

**CESA Working Paper No. 8 | 2024**
Kamila Cygan-Rehm and Matthias Westphal
School Starting Age and the Gender Pay Gap over the Life Cycle

# CESA Working Paper

Series

**CESA Working Paper No. 9 | 2024**
Marco Kerkemeier
(Co-)Explosiveness of corporate credit spreads

**CESA Working Paper No. 10 | 2024**
Christoph Wegener, Robinson Kruse-Becher and Tony Klein
EU ETS Market Expectations and Rational Bubbles

**CESA Working Paper No. 11 | 2025**
Anna Krumme and Matthias Westphal
Monetary returns to upper secondary schooling, the evolution of unobserved heterogeneity, and implications for employer learning

**CESA Working Paper No. 12 | 2025**
Robinson Kruse-Becher and Philip Letixerant
Oil price expectations in explosive phases

**CESA Working Paper No. 13 | 2025**
Sarah Rezaei, Hendrik Sonnabend and Bastian Westbrock
Overconfidence and Goal-Setting

FernUniversität in Hagen