Employee Recognition and Performance -
A Field Experiment

Christiane Bradler∗, Robert Dur†, Susanne Neckermann‡, Arjan Non§

February 2012

Very preliminary draft! Please, do not cite without permission!

Abstract
We conduct a natural field experiment to investigate the causal effect of formal recognition for a job well done on employee performance. We hired more than 400 students for a three-hour job. Participants worked individually on a data entry task in groups of eight. We randomized the unannounced provision of recognition (in the form of handing out a thank-you card) after two hours of work. We find that the provision of recognition to all workers in a group increases subsequent performance only weakly. By contrast, scarce recognition that is only provided to the best performers in a group raises subsequent performance substantially. Remarkably, workers who did not receive recognition are responsible for this performance increase. These results are consistent with workers having preferences for conformity and being reciprocal at the same time. All effects vanish completely when the reward is announced.

JEL Classification: C93, M52
Key words: employee motivation, recognition, reciprocity, conformity, field experiment

∗ZEW Centre for European Economic Research Mannheim
†Erasmus University Rotterdam
‡ZEW Centre for European Economic Research and University of Mannheim
§University of Maastricht
1 Introduction

"He who praises everybody praises nobody" - Samuel Johnson (1709-1784)

Recent years have seen a surge in popular business books on the importance of recognition for employee motivation. A prominent example is Bob Nelson’s book titled *1001 Ways to Reward Employees* (1995). He starts his book by stating that a number of surveys “confirm what almost every employee already knows: that recognition for a job well done is the top motivator of employee performance.” In his book he provides a massive amount of real-world examples to substantiate his view. Questionnaire studies reveal similar views among employees (Wiley, 1997; Kovach, 1995) and managers (Holton et al. 2009).

The vast amount of practitioner’s literature on employee recognition stands in strong contrast with the relatively modest amount of academic research on the topic. The academic literature provides relatively little direct evidence for a causal effect of employee recognition on job performance. Fred Luthans, together with different coauthors, conducted five field experiments on the effect of social recognition on employees’ performance, all reporting a positive effect. We should, however, be cautious in interpreting these findings, as these studies typically have a limited number of observations (a handful of plants) and lack a proper, randomized control group. More recently, Neckermann et al. (2010) analyze firm-level data to compare the behavior of employees receiving an award with observationally similar employees not receiving an award. They find that award recipients significantly increase performance as compared to non-recipients.

Several studies from a variety of settings provide indirect evidence on the motivational effects of praise and recognition. Markham et al. (2002) experimentally examine the effects of personal recognition on employee attendance, finding strong reductions in absenteeism. Smaller effects are reported by Fisher and Ackerman (1998), who investigate the effect on parents’ willingness to volunteer for their kids’ soccer club. Delin and Baumeister (1994) discuss several psychological mechanisms by which praise might operate to affect performance, both positive and negative. They conclude that for all of these mechanisms, little direct evidence is available. A noteworthy exception is presumably the effect of praise on intrinsic motivation. A

---

1 He continues with “yet most managers do not understand or use the potential power of recognition and rewards […] while money is important to employees, research shows that what motivates them really to perform […] is the thoughtful, personal kind of recognition that signifies true appreciation for a job well done.”

2 See Ottemann and Luthans (1975), Luthans et. al. (1983), Snyder and Luthans (1982), Luthans et. al. (1991), and Luthans et. al. (2008).

3 Some other studies rely on correlations between the provision of recognition and employee performance (e.g. Rhoades and Eisenberger 2002, Cropanzano and Mitchell 2005, and Wagner and Harter, 2006.) . Clearly, as these studies recognize, correlation does not prove causation. In fact, it seems natural to assume that good performance leads to recognition.
meta-analysis of 96 studies by Pierce and Cameron (1994) suggests that praise tends to increase intrinsic motivation. Finally, there is a small strand of literature in psychology that studies the effects of expressing gratitude. McCullough et al. (2001) start their review of the literature by noting that "empirical social scientists (..) have tended to neglect gratitude as a topic worthy of intensive and sustained empirical study" (p. 249). Nevertheless, a number of studies show that thanking individuals who contributed to the well-being of others promotes future helping behaviour, both towards the person expressing gratitude as well as towards unrelated strangers, see McCullough et al. (2001) and the references therein. Recently, Grant and Gino (2010) proposed that these effects are caused by a reduction in the feelings of uncertainty about whether the help will be appreciated.

Given the limited amount of academic research targeted at praise and recognition, it is not surprising that even less is known when it comes to the question how recognition should be provided. Should all employees receive recognition, even for performing daily routines or simply being on time, or is more differentiation desirable? In particular, does praising everybody demotivate the top-performers, thus creating a culture of mediocrity? Or does reserving praise for the best performers increase their performance at the cost of discouraging the less gifted employees? Likewise, it is an open question whether an element of surprise is essential for effective recognition.

In this paper we take up these questions by conducting a large-scale field experiment in a natural working environment. Over the course of November 2010 to May 2011, we hired more than 400 people (mainly students) for a three-hour data entry job. We created a work environment where eight workers shared the same room, but worked individually. Workers were paid a flat wage of 25 Euro and were not aware that they took part in an experiment. A random sample of workers received recognition after two hours of work. Recognition consisted of a thank-you card, personally signed by the head of the research institute and handed out by a research assistant. Thus, provision of recognition did not involve any material reward but clearly showed the management’s appreciation for workers’ effort. From the way the thank-you card was presented, workers could deduct that no further formal recognition would be provided in the remainder of the working period. Importantly (and most naturally), in most of the treatments possible provision of recognition was not announced at the start of the three-hour working period.

To address the questions posed above, we vary the scarcity of recognition over different treatments. In one treatment, all workers in a work group receive the thank-you card. In other treatments, the thank-you card is given only to the best performer or to the best three performers of a work group. When handing out the thank-you cards, it is publicly announced that their assignment is determined by performance in the first two working hours. Moreover,

---

4 Some early studies are Deci (1971), Anderson et. al. (1976), and Dollinger and Thelen (1978).
in one treatment, the research assistant announces at the start of the working period that a thank-you card will be given to those three employees who perform best, thereby mentioning the performance criterium as well.

Our main focus is on how the provision of the thank-you card affects subsequent performance of workers, in particular how the reaction of recipients differs from non-recipients. We argue that three behavioral mechanisms may be operational: signaling by the employer, reciprocity of employees, and conformity preferences of employees. As we shall see, our results are consistent with employees having preferences for conformity and being reciprocal at the same time.

The paper proceeds as follows. In the next section, we draw on existing behavioral theories to develop hypotheses on the effects of different forms of recognition. Next, we describe the experimental set-up in detail in Section 3. Section 4 lays out the empirical strategy and presents the results of our experiments. Section 5 concludes.

## 2 Theory

Why might the unexpected provision of a thank-you-card affect subsequent effort of workers? And does it matter whether all workers or only the best performing workers in a session receive a thank-you card? And what is the effect of announcement? We can think of three plausible behavioral mechanisms through which a thank-you card may affect subsequent effort.

First, workers may consider the thank-you card as a gift from the employer and feel inclined to reciprocate. A natural way to reciprocate is to increase effort in the remainder of the session. According to Blau (1964), the reason that gifts are reciprocated is that people "are interested in maintaining a balance between inputs and outputs and staying out of debt in their social transactions" (p 26). Following the seminal paper by Akerlof (1982) and the influential experimental work by Fehr et al. (1993), reciprocity has become a firmly established concept in economics. Most studies in economics have focused on wages as the employer’s means of exchange in reciprocal relationships with workers. An exception is Kube et al. (2011) who show that non-pecuniary gifts are particularly powerful in triggering subsequent work effort. Other social sciences have also considered socio-emotional gifts that address “social and esteem needs (and are often symbolic and particularistic)”(Cropanzano and Mitchell, 2005, p. 881). The thank-you card which is central to our treatments, clearly falls into this latter category.

To formalize this idea, suppose that workers have the following utility function (as in Dur et al. 2010):

$$u = \beta e + \gamma ae - \frac{1}{2}e^2,$$

where $e$ represents effort, $a$ the non-pecuniary rewards received from the employer, $\gamma$ indicates
the importance of reciprocal feelings, and $\beta$ captures other intrinsic motives to exert effort. Clearly, effort is increasing in the value of the rewards received (optimization of $u$ with respect to $e$ gives $e^* = \beta + \gamma a$). Workers who do not receive a reward do not exert any additional effort. Thus, in treatments where only a subset of workers receive a thank-you card, we only expect recipients to respond. Moreover, scarcity may increase the perceived value of the card (represented by $a$ in the model), implying a stronger response of recipients when the number of recipients decreases. Announcing the card in advance is likely to mute the effort responses to providing the card, as announcement changes the nature of the exchange. Announcing a reward can be viewed as an implicit demand from the employer for hard work, implying that receiving the reward is nothing but the employer’s reciprocation of the worker’s effort. Recipients consider the reward as earned by hard work, and therefore the reward cannot induce an obligation to reciprocate the ‘favor’.

A second reason for why the thank-you card may affect subsequent effort is that the card provides a signal to the workers about the employer’s kindness or care for the workers. The key assumption is that workers care more for an employer who cares for them. As a result, workers respond with increased effort when they perceive the thank-you card as a credible signal of the employer’s kindness (see Levine (1998), and more recently Dur (2009) and Non (forthcoming)). This tendency to improve the welfare of a kind person can be seen as a sort of reciprocity.\(^5\) The following game-theoretic model illustrates the basic mechanism.

Consider a group of workers who work for an employer whose altruism is private information. It is common knowledge, however, that the employer’s altruism, denoted $\alpha_j$, is a random draw from a uniform distribution over the interval 0 to 1. The utility of worker $i$ is described by

$$u_i = \beta e_i + \gamma \alpha_j \sum_{i=1} e_i - \frac{1}{2} e_i^2,$$

where the term $\gamma \alpha_j \sum_{i=1} e_i$ captures the idea that workers care more about the profits $\sum_{i=1} e_i$ of an altruistic employer. Representing our experiment, we assume that profits are simply the sum of all workers’ effort.\(^6\) The employer’s payoff $y_j$ is:

$$y_j = \sum_{i=1} e_i + \alpha_j \sum_{i=1} u_i - c(\alpha_j).$$

The employer can signal his type by creating thank-you cards, where $c(\alpha_j) > 0$ captures the effort, time, and money involved in creating a given amount of cards. When no cards

---

\(^5\)Some authors have stressed the importance of good intentions for reciprocity. The approach taken by Levine (1998) can be seen as a simple way to model intentions Gul and Pesendorfer (2010). Models that formalize intentions yield similar predictions, e.g. Falk and Fischbacher (2006)

\(^6\)Different profit functions yield the same qualitative predictions, as long as workers’ efforts are (weak) complements and the marginal benefits of individual effort are non-decreasing in effort.
are created, \( c(\alpha_j) = 0 \). In the spirit of Camerer (1988), Prendergast and Stole (2001), and Ellingsen and Johannesson (Forthcoming), we assume that the costs of creating thank-you cards are decreasing in the employer’s altruism, i.e. \( \frac{dc(\alpha_j)}{d\alpha_j} < 0 \).\footnote{The costs for a completely egoistic employer (\( \alpha_j = 0 \)) are possibly infinitely high, as he may not even think of thanking workers.} Assume further that there are two working periods of equal length and no time discounting. Costs associated with the cards are incurred at the start of the first working period, but the cards are always handed out between the two working periods. After the second period, the game ends. We deal with the possibility of announcing the card at the start of the first period below.

It is easily verified that profits are increasing in the workers’ beliefs on \( \alpha_j \): the optimal effort choice is \( e^* = \beta + \gamma \alpha_j \). The optimal effort choice in the first period reflects the prior belief \( E(\alpha_j) = \frac{1}{2} \). At the start of the second period, workers update their beliefs conditional on whether cards are handed out. As the cards are handed out in public, both recipients and non-recipients receive the same information, and should therefore hold the same beliefs on the employer’s type. Thus, we expect no difference in the response to the cards between recipients and non-recipients. As both effort and utility crucially depend on workers’ beliefs on \( \alpha_j \), all employer types have an incentive to be perceived as an altruist. However, this incentive is stronger for more altruistic employer types, as they genuinely care about the workers’ utility. More altruistic types can therefore distinguish themselves from lower types by sending a signal that is too costly for lower types to imitate. Thus, the cards signal altruism when the costs of providing the cards exceed their benefits for types below some threshold level of altruism \( \alpha^* \), while the cards are profitable for types above the threshold. The threshold level of altruism \( \alpha^* \) is defined formally in the appendix. An employer who hands out cards is believed to be a type from the interval \( \alpha^* \) to 1, while an employer who abstains from doing so is believed to be a type from the interval 0 to \( \alpha^* \). These beliefs are reflected in higher levels of effort and utility when cards are distributed. In particular, the equilibrium levels of effort with and without cards are \( \beta + \gamma \frac{1+\alpha^*}{2} \) and \( \beta + \gamma \frac{\alpha^*}{2} \), respectively.

Interestingly, an exogenous increase in the costs of providing cards strengthens the workers’ response to the cards. Higher costs imply that fewer employer types find it attractive signal, only the most altruistic types continue signaling. As a result, the difference between first and second-period equilibrium effort \( (\beta + \gamma \frac{1+\alpha^*}{2} - (\beta + \gamma \frac{1}{2}) = \gamma \frac{\alpha^*}{2}) \) becomes larger as \( \alpha^* \) increases. Thus, assuming the costs are increasing in the number of cards provided, an increase in the number of cards improves the strength of the signal. The strength of the effort response may therefore differ between treatments.

Now assume that, in case the employer has decided to provide three thank-you cards, he has the opportunity to announce the cards at the start of the first period. We assume that employers
are always truthful and do not announce cards when they don’t have cards. Is an announced card just as effective in signaling the employer’s care as an unannounced card? To answer this question, we compare the equilibria in which relatively altruistic employer types give announced cards rather than unannounced cards, while in both equilibria egoistic types do not give cards. The latter equilibrium is derived above, except that we assumed away the employer’s possibility to announce the cards in advance. When we relax this assumption, the equilibrium still exists as long as an announced reward is viewed negatively, see the appendix for details. Compare this to the equilibrium where altruistic types give an announced card, while the relatively egoistic types abstain from giving cards. Since the information revealed by the cards now becomes available in the first period already, the benefits of announced cards are twice as large as that of unannounced cards. As a result, distributing announced cards is also attractive for relatively low types: the threshold level $\alpha^*$ goes down. Announced cards are therefore a weaker signal of altruism than unannounced cards. Compared to effort in the control sessions (based on prior beliefs), we expect that announced cards induce only a small change in the worker’s beliefs and subsequent effort, smaller than when the card comes without prior notification. Announced cards are more likely to be perceived as an egoistic employer’s instrument to boost profits, and this effect would be exacerbated if we would assume that workers have an innate desire for receiving a thank-you card. In that case the announcement of the cards induces a competition between workers that stimulates first-period effort, hence further weakening the signaling value.

Closely related are other signaling models, for instance the model of Ellingsen and Johannesson (2008). The key idea in their model is that workers care more about social esteem when the principal shows that he is worth impressing. In the setting of our experiment, employers use the cards to show their care for the workers, while workers respond with increased effort to impress such a good employer. Alternatively, the thank you card may affect workers’ beliefs about how important their effort is to the employer. These theories yield similar predictions.

Third and last, in the treatments where only the best-performing workers receive a thank-you card, preferences for conformity may give rise to behavioral responses. Recipients in these treatments learn that their performance belongs to the top within their group. Likewise, non-recipients learn that they belong to the lower ends of the performance distribution. If workers have a preference for conformity (a formal model is described in the appendix, see also Bernheim, 1994, and Sliwka, 2007), we expect non-recipients to increase effort while recipients reduce effort. The effect on non-recipients should be less pronounced when the thank-you card is provided to very few workers, as less information is provided about one’s standing. Thus, we

---

8Doing so would in practice trigger a strong negative response, as it shows that the employer is untrustworthy. Alternatively, we could assume that there are employer types with negative $\alpha$. Failing to hand out announced cards is likely to be interpreted as a signal that the employer is an extremely low type, resulting in low or even negative effort choices in the second period. Thus, doing so would be a strictly dominated strategy.
expect smaller effects for non-recipients in the treatment where only a single thank-you-card is provided as compared to the treatment where the three best-performing workers receive a thank-you card. Conversely, recipients of the card should respond more strongly when fewer thank-you cards are provided. Lastly, as all effects come from the revealed information about one’s relative performance, we do not expect treatment effects when all workers are provided with a thank-you card, and we do not expect that announcement has any effect. Table 1 summarizes our theoretical predictions.

### 3 Design and Implementation

#### 3.1 Background

In order to test these hypotheses, we conducted a natural field experiment (according to the taxonomy laid out in Harrison and List, 2004). This means that we observed participants in a natural work environment and that subjects were not aware of the fact that they were part of an experiment.

---

9 The effect of relative performance feedback on performance has recently been studied in a variety of contexts, ranging from student test scores (Bandiera and Larcinese, 2009; Azmat and Iriberri, 2010; Tran and Zeckhauser, 2011), contributions to an online community (Chen et al., 2010), real-effort in the lab (Charness et al., 2010; Freeman and Gelber, 2010; Kuhnen and Tymula, 2011), to workers’ performance in a real working environment (Falk, 2006; Delfgaauw, 2009; Blanes i Vidal, 2011; Barankay, 2011a, 2011b). With the noteworthy exception of Barankay (2011a, 2011b), all of these papers find that relative performance feedback has on average a positive effect on performance. A subset of these papers also look at where in the performance distribution people are most responsive to feedback; results tend to be mixed in this respect.

10 In addition to the treatments reported in this paper, we conducted three additional treatments investigating the motivational production function. Results on these treatments are reported in Bradler and Neckermann (2012).
In the name of a German research institute – the Centre for European Economic Research (ZEW) – we hired students for a one-time data entry job between November 2010 and May 2011. The institute had just received several hundreds of surveys that could not be processed automatically, which made manual entry into a database necessary. For this job, we hired students from different universities in and around Mannheim via student job centers, flyers, and notes on bulletin boards. The announcement informed subjects that the research center needed short-term student helpers for a three-hour data entry job and that subjects would earn 25 Euro fixed wage in total. Students could sign up for the job online and were subsequently contacted via phone to arrange date and time. We minimized career concerns by informing participants that the job was one-time, i.e. a second participation was not possible due to tax reasons, and that no permanently employed research assistants were needed at the moment. We invited subjects in groups of eight. Sessions took place in the morning, at noon, and in the afternoon, in two locations. We implemented recognition by handing out thank-you cards (see Figure 2 in the Appendix) to participants. Depending on the treatment, these thank-you cards were either announced or came as a surprise after two hours of work; they were given to all participants, or only to the best three or the best one in the work session.

### 3.2 Implementation

Upon arrival, participants were seated in front of a workstation. The workspaces were arranged in a U-shape with sufficient space between individuals to ensure that subjects felt unobserved. Then, all formalities (privacy policy, wage agreement) were taken care of and the 25 Euro was paid in advance. This conveyed the fact that payment was independent of work performance. Subsequently, the participants received a short introduction to their employer (ZEW) and to the project the surveys belonged to. The questionnaires were part of an evaluation project in the educational context and involved surveys from several hundreds of students on their school and career preferences. Additionally, a short briefing was given on how to file the answers from the questionnaires into the database. The data were filed via a web interface; hence, access to the internet was visible and possible at any time. The interface was set up as an online version of the paper surveys that the participants had in front of them. Most survey questions were

---

11 The experiment has been conducted within the ethical guidelines of our home institutions.
12 The fixed wage was chosen according to hourly wage rates of student research assistants.
13 In case of a second participation, participants would exceed the threshold of earnings beyond which tax-based information needs to be collected. This would cause substantial administrative expenses for fiscal reporting.
14 The field experiment was carried out at the Centre of Economic Research in Mannheim and at the Ruprecht-Karls University of Heidelberg. In the regressions we control among others for time-of-day as well as location effects, but dropping these controls does not affect the results. All sessions were conducted by the same research assistant.
15 For an example of the screen see Figure 3 in the Appendix.
multiple choice, only a few questions required entering information from free text fields. The questionnaires were stacked in a high, non-transparent box placed in front of the participants, with an identical box next to it to deposit the completed ones. This prevented the participants from comparing their work with each other.

After the instructions were given, the research assistant left the room. She informed participants that she was working outside and was available for queries at any time, and that subjects could take breaks whenever necessary. These measures (individual breaks, payment in advance, internet access, absence of supervisor) were taken to give participants substantial leeway in the amount of time they spend on filing the questionnaires. Furthermore, a collective break was avoided to minimize possible group effects and communication between participants. The task was exhausting and monotonous, and we do not assume that there was much intrinsic motivation involved in completing the task. In fact, participants commented that the job was very monotonous and boring.

Eight subjects were assigned to each working session. However, since some subjects did not show up, the average number of participants per session was 7.02 (s.d. = 0.99). The timeline was as follows: the introduction lasted for about 20 minutes on average. Then, participants worked on the task for roughly 100 minutes. We refer to this first phase as working period 1. The working period 1 was longer than the subsequent working period 2 to allow for variation in the length of the introductory period and smoothing out learning effects which are supposed to be largest at the beginning of working period 1. Subsequent to working period 1 (i.e. two hours after participants’ official start of work), the research assistant entered the room and told subjects that the IT-department of the institute had just confirmed that the data was transmitted correctly from each computer to the central data base. Depending on the treatment, she handed out the thank-you-cards. Subsequently, subjects worked for approximately one additional hour – which we refer to as working period 2. After the second working period, feedback forms were handed out, which asked for improvement suggestions and comments on how we handled the short-term employment. These forms allowed us to gather some information on individual characteristics like gender and field of study.

---

16 The standard deviation of the introductory period is approximately 5 minutes. Differences in length are due to delays in arrivals of participants or different amount of questions which were raised during the introduction.

17 The standard deviation of the duration of working period 1 is 6.4 minutes. Differences in the length of working periods are due to delays in the introductory period. In our regressions we measure performance per minute and additionally control for the length of working period 1 to take care of potential learning or fatigue.

18 We chose this wording to ensure that in control as well as in the treatment sessions, subjects were aware of the fact that we could link data entered to the respective work station. This way, the perceived degree of monitoring was held constant across treatments as much as possible. The fact that in all treatments, the research assistant could easily monitor performance by looking at the pile of completed surveys also contributes to minimizing differences in perceived monitoring.

19 The mean of working period 2 is 52.6 minutes with a standard deviation of 3.92 minutes. We account for differences in the length of working periods in the statistical analysis by using productivity per minute of working time as the dependent variable.
3.3 Treatments

As described above, the treatment intervention took place after working period 1. The different scripts can be found in the Appendix. In all treatments, including the control treatment, the research assistant informed subjects that the data were transmitted correctly from all computers to the central server and that everything worked as planned. Nothing else happened in the Control group, so that we can measure the development of productivity without rewards.\(^{20}\)

In treatments involving thank-you cards, she continued that the research team together with the president of the institute had created thank-you cards for all workers as a symbol for the institute’s thankfulness and appreciation of the participants’ support in entering the data. To ensure that the card would be considered as a clear signal of appreciation, all cards were personally signed by the head of the institute, president Professor Franz.

We implemented the following treatments. In the treatment “Thank-you card for all”\((TC\ All)\), the research assistant handed a card to each participant. This treatment allows us to measure the response to general appreciation and recognition. In the treatment “Thank-you card for the best 3”\((TC\ Best3)\), subjects were informed that the research assistant had only three cards available, and had spontaneously decided to hand these out to those three persons who had performed best until then.\(^{21}\)

The treatment “Thank-you card for the best 3 announced”\((TC\ Best3\ announced)\) was identical to \((TC\ Best3)\) apart from the fact that there was only one recipient of a thank-you card. These treatments add scarcity and, hence, status and feedback to general recognition.

The scarcity of cards was explained by the fact that the president had not had enough time to sign a card for each participant. This excuse ensures that participants do not expect additional cards at the very end of the employment period. As there were fewer than eight workers in some sessions, we adjusted the number of cards in accordance to the actual group size to keep the percentage of rewarded subjects comparable.\(^{22}\)

Finally, in the treatment “Thank-you card for the best 3 announced”\((TC\ Best3\ announced)\), we announced at the beginning of working period 1 that thank-you cards would be handed out later on to the three participants who have performed best after roughly two hours. We showed the card to the subjects in advance to ensure common knowledge about what the card looked like.\(^{23}\)

---

\(^{20}\)By comparing treatment effects to the performance improvement in the control group we consider that participants might learn how to work on the task faster but might also get slower due to tiredness.

\(^{21}\)We measured performance in terms of the number of finished questionnaires in working period 1. Subjects were told that IT informed the research assistant of the three logins that had performed best. The wording was chosen in order to convey the notion that the research assistant could not perfectly monitor everyone’s performance.

\(^{22}\) For a group size of 7-8 persons we assigned three TCs, and for a group size of 5-6 persons we assigned two TCs. Obviously, the percentage of rewarded subjects is not kept constant by this variation. We control for group size in all regressions, although the results are not sensitive to this.
like, and to ensure that subjects understood that there was no material value attached to the reward. This treatment checks for incentive effects and for whether ex-post effects depend on the rewards being announced or handed out spontaneously.\textsuperscript{23}

4 Results

Table 1 reports some descriptive statistics by treatment group. The sample contains 414 observations, 81 in Control, 87 in TC All, 95 in TC Best3, 69 in TC Best3 announced, and 82 in TC Best1.\textsuperscript{24} There are no statistically significant differences between treatment groups with respect to gender and field of study, with the exception of a somewhat lower proportion of economics students in TC Best1. Sessions are spread out more or less evenly over the different times-of-day and locations, but there are some differences between treatments for organisational reasons. As shown in our main regression (see Table 3 in the Appendix), controlling for these characteristics does not alter our main results. Finally, treatments are balanced with respect to baseline performance, i.e. performance in the first working period. With the exception of TC Best3 announced, this is exactly what we would expect if our randomization is successful. Throughout the paper, performance is measured as the number of correct clicks per minute of working time.\textsuperscript{25}

We first examine the main treatment effects. A first impression is provided by Figure 4. It shows the average improvement in performance between working periods 1 and 2. The Control group shows almost no improvement in performance. Remarkably, after an announced reward is distributed in TC Best3 announced, participants seem to slack off. By contrast, all other treatments with recognition exhibit clear improvements in performance. This effect is most pronounced when the reward is scarce, as is the case in TC Best3 and TC Best1.

This picture is largely confirmed by a regression analysis. We estimate the following baseline

\textsuperscript{23} Even though we do not think that this is likely, we cannot rule out that word spread and that participants in later sessions expected certain interventions. Therefore, the more salient treatments (TC Best3, TC Best3 announced, TC Best1) were conducted subsequent to the less salient treatments (Control, TC All) to make sure that we perhaps positively surprised, but did not disappoint subjects with the intervention they received.

\textsuperscript{24} In total, 430 persons participated. Three sessions (two of Control and one of TC All) had to be dropped from the analysis due to severe server breakdowns during working time. Moreover, two participants had to be excluded from the analysis due to disabilities.

\textsuperscript{25} One click corresponds to the entry of an answer to a multiple choice question or to one word in the free-text fields. Correctness of an individual entry is determined as whether or not it corresponds to what the majority of participants clicked as answer for this particular question. This serves as a very reliable quality measure because each survey was entered on average 286 times (s.d.=67).
equation:

\[ y_{i,t=2} = \alpha + \beta_1 TCall + \beta_2 TCBest3 + \beta_3 TCBest1 + \beta_4 TCBest3ann \\
+ \delta_1 y_{i,t=1} + \delta_2 y_{i,t=1}^2 + \delta_3 y_{i,t=1}^3 + \epsilon_{i,s}, \]  

(1)

where \( y_{i,t} \) represents the performance of individual \( i \) in working period \( t \). We correct for initial performance differences between individuals by including performance in the first period, as well as squared and cubic terms. This ensures that we only compare individuals with similar baseline performance. Standard errors are clustered by session, denoted by the subscript \( s \).

The first column of Table 2 shows the results of estimating (1), the second column includes additional controls. The results of both specifications are similar. As already suggested by Figure 4, spontaneous, scarce rewards (\( TC Best3 \) and \( TC Best1 \)) have a positive and significant effect on performance. \( TC Best3 \) increases performance by 1.2 clicks per minute, which corresponds to a performance increase of 8% or a change of 0.3 standard deviations. The increase in performance in \( TC Best1 \) is smaller than in \( TC Best3 \), although the difference is not statistically significant (Wald test: \( p=0.49 \), one sided). Handing a reward to all participants also increases performance, albeit to a lesser extent than when the reward is contingent on good performance.\(^{26}\) The effect of \( TC All \) is only significant at the 10% level if all controls are included. Announced rewards have no overall effect on second period performance. The absence of any treatment effect of the announced treatment in working period 2 might be driven by relatively high effort in working period 1 when subjects respond positively to the announcement. However, we already saw in Table 1 that there is no significant difference between the performance of subjects in \( TC Best3 \) announced and \( Control \) in working period 1. Table 2 in the Appendix shows that this result is robust to the inclusion of the same control variables that are included in the main regressions presented above. Since the wording of the announcement made clear that the reward was assigned based on quantity, we check for differences in clicks per minute rather than the number of correct entries. As one can see, performance in working period 1 when the reward is announced is not statistically different from all other treatments that are identical in this first period. This result stands in remarkable contrast to the sizeable incentive effect found by Kosfeld and Neckermann (2011), but is in line with Kosfeld et al. (2012) who discuss reasons for the documented divergence in findings on social rewards.

In order to understand the underlying mechanisms better and to study the relevance of the different theories discussed above, we next look at recipients and non-recipients, separately. Figure 5 gives a first impression. The figure shows performance improvements by treatment.

\(^{26}\)Performance in working period 2 is significantly larger in \( TC Best3 \) and \( TC Best1 \) than in \( TC Best3 \) announced) (Wald test, \( p < 0.01 \)). The differences between the coefficients of \( TC Best3 \), \( TC Best1 \), and \( TC All \) are not statistically significant.
Interestingly, in all treatments where the research assistant was forced to differentiate between employees, the non-recipients improve by more than the recipients do. When the reward is unannounced, both recipients and non-recipients increase their performance, but non-recipients clearly show the largest improvement. When the reward is announced, the performance of the recipients strongly deteriorates, while the non-recipients show a negligible reduction in performance, comparable to the reduction in the control treatment.

The regression results for separate treatment effects of recipients and non-recipients are reported in columns III and IV of Table 5. Column III reports the results from estimating equation (1) with separate dummy variables for recipients and non-recipients of each treatment and column IV includes all controls. The results confirm the descriptive analysis in Figure 5 positive treatment effects are largely driven by non-recipients. In TC Best3 non-recipients increase their performance by 1.6 clicks per minute\(^{27}\) in TC Best 3 and by about 1.2 click per minute in TC Best1. The performance changes of non-recipients in TC Best3 and TC Best1 are statistically different from each other on the 1% level (p=0.003, two-sided). By contrast, receiving an unexpected scarce reward has no statistically significant impact on performance in any of the treatments, although all coefficients are positive.\(^{28}\) The difference between recipients and non-recipients is significant at the 10% level (p=0.08, two sided) in TC Best3. It is not significant in TC Best1. An announced award has no significant effect on performance, and the estimated coefficients are negative. Note, however, that also in this case, non-recipients do better than recipients, albeit not significantly so.

We also analyze treatment effects on quantity and quality separately, where quality is measured as the error rate (percentage of incorrect clicks). The results are reported in Table ?? in the Appendix. We find no significant treatment effects for quality levels. Moreover, quality levels are very high overall (only 1.6% of all entries were incorrect, s.d. = 2.7), so it is safe to conclude that all the treatment effects we have seen for correct entries are driven by changes in quantity, not quality.

As a robustness check, we split subjects into performance groups and run separate regressions for each group. In particular, we split the sample into the best 3 and worst 5 performers (corresponds to winners and nonrecipients in TC Best3) as well as the best 1 and worst 7 performers (corresponds to winners and nonrecipients in TC Best1). This allows a more direct comparison of individuals with similar baseline performance between treatments. Table 5 in the Appendix shows the results, which are in line with the results reported above. Receiving a spontaneous, scarce reward (recipients in TC Best1 and TC Best3) does not change performance significantly. In contrast, non-recipients increase performance, both in TC Best3 and TC Best1,\(^{27}\) This corresponds to a performance increase by 11 percent or a change of 0.45 standard deviations.\(^{28}\) We can rule out that top performers cannot increase their performance due to ceiling. In Bradler and Neckermann (2012) we use the same set-up and show that recipients do increase their performance subsequent to receiving a combination of reward and money.
and the coefficients are even larger than in our main regression 1. Column I of Table 5 shows that the performance increase of non-recipients in TC Best1 is mainly driven by those who performed very well but did not receive a reward. This also explains the significant positive effect of the three best but not the very best performer TC Best1.

As a further robustness check, we run specifications including interaction terms of treatments with gender. We find no evidence for gender effects. The only exception is TC Best3 announced where women perform significantly better than men in working period 2 (see column 6 of Table 5 in the Appendix).

Moreover, our results stay robust to the inclusion of additional controls for different time effects, such as the day-of-time, or whether the session took place during final exams period, during the semester, or during semester break. The results remain qualitatively similar and, depending on the specification, coefficients even get slightly larger.

### 5 Discussion

We have seen that an unannounced non-material reward has a positive impact on subsequent performance, in particular when provided exclusively to the best performers. This effect is primarily caused by the improvement in performance of those who do not receive the reward. When the reward is announced in advance, non-recipients also tend to do better than recipients, but the difference is not significant. Moreover, there is no incentive effect in the first period, and performance in the second period, if anything, tends to deteriorate.

How to reconcile these findings with the theories discussed earlier? A possible explanation is that individuals have a preference for conformity and are reciprocal at the same time. The performance improvement of non-recipients, which is largest in TC Best3 is well in line with conformity theory. As argued above, non-recipients should increase their performance when they learn that they perform below average, and this effect should be stronger when there are more rewards. Although the difference between TC Best3 and TC Best1 is only marginally significant, the finding that non-recipients improve by more than recipients is well in line with conformity theory.

However, conformity theory cannot explain the level-effects. It cannot explain the performance increase of recipients when the reward comes as a surprise, nor can it explain the weakly significant increase in performance in TC All. Moreover, conformity theory has difficulties explaining the differences between TC Best3 and TC Best3 announced. Hence, it seems likely that individuals are also reciprocal, which would explain the level-effects. When the reward comes as a surprise, reciprocity theory predicts that recipients increase their performance, which is exactly what we observe. Note that the point estimates of these effects are surprisingly similar,
see table 5. By contrast, an announced reward is likely to be perceived as instrumental in eliciting effort, and therefore does not lead to performance increases.

6 Conclusion

This paper presents a natural field experiment to test ex-post effects of recognition. We contribute to the literature by studying the effect of recognition in a work setting and by looking at ex-post effects of rewards. We find that materially worthless rewards have a statistically significant impact on performance. Specifically, non-recipients increase their subsequent performance in response to scarce rewards. In line with conformity models that predict that subjects adjust their performance to correspond to average performance, this effect is larger when there is a relatively large number of recipients. We conclude that it is important to take ex-post effects of rewards into account and that channels like conformity produce unexpected effects that should be investigated further. Limitations of this study are the student sample, the short-term employment (no established relationship between employer and employees) and the short duration of the work. Future work should address these issues.
References


Appendix A: Derivations of the Signaling Model

In this appendix, we formally characterize the threshold level of altruism $\alpha^*$ and show that an interior solution exists. Next, we show that the equilibrium in which a relatively altruistic employer offers an unannounced card, while a relatively egoistic employer offers an announced card, still exists when employers have opportunity to announce the card in advance. The threshold level of altruism $\alpha^*$ is defined as the employer type who is indifferent towards creating cards:

$$c(\alpha^*) = \sum_{i=1}^{n} e_i^* \left( \frac{1 + \alpha^*}{2} \right) - \sum_{i=1}^{n} e_i^* \left( \frac{\alpha^*}{2} \right) + \alpha^* \sum_{i=1}^{n} E (u_i | \alpha > \alpha^*) - \alpha^* \sum_{i=1}^{n} E (u_i | \alpha < \alpha^*),$$

where $e_i^* (\alpha)$ is a worker’s equilibrium effort when he believes that the employer is of type $\alpha$. By inserting $e_i = \beta + \gamma \alpha_j$ into the worker’s utility function, his utility can be written as

$$u_i = \frac{1}{2} (\beta + \gamma \alpha_j)^2 + \gamma \alpha_j (n - 1) (\beta + \gamma \alpha_j),$$

where $n$ is the total number of workers. Note that workers quite rationally assume that coworkers have the same beliefs as themselves, but nothing depends on this assumption. Existence of an interior solution is ensured by the fact that the marginal costs of signaling are decreasing in $\alpha^*$, while the marginal benefits are increasing in $\alpha^*$. The latter is true for two reasons. First, more altruistic employers put more weight on increased worker utility. Second, expected utility is quadratic in $\alpha_j$.

Now suppose the employer has the possibility to announce the cards in advance. Whether the separating equilibrium described above still exists depends on how deviation to an announced card is perceived by the workers. The equilibrium still exists as long as such a deviation is viewed negatively. A sufficient condition is that workers’ out-of-equilibrium beliefs are identical to the prior beliefs ($E(\alpha) = \frac{1}{2}$). To see this, note that in equilibrium the first-period beliefs equate the prior beliefs. Thus, offering announced cards is only attractive if it improves the second-period payoff. Employer types who in equilibrium hand out cards ($\alpha > \alpha^*$) cannot gain by deviating, as the second-period equilibrium belief $E(\alpha) = \frac{1 + \alpha^*}{2}$ is more optimistic than the beliefs when deviating ($E(\alpha) = \frac{1}{2}$). By contrast, deviation leads to more favorable second-period beliefs for employer types who do not hand out cards in equilibrium. This improvement in beliefs, however, is insufficient to recoup the costs of the cards, as in equilibrium condition (2) holds. This condition tells us that for all types $\alpha \leq \alpha^*$, creating cards is unprofitable when it leads to an update of beliefs from $\frac{\alpha^*}{2}$ to $\frac{1 + \alpha^*}{2}$. Hence, an improvement in beliefs from $\frac{\alpha^*}{2}$ to $\frac{1}{2}$ is clearly insufficient to justify the costs.
Appendix B: A model of Conformity

Consider 8 workers, who work for two periods. In both periods, workers choose an effort level $e$. Their utility function is described by

$$u = \beta e + ye - \frac{1}{2}e^2.$$ 

Workers condition their effort choice on what they consider an appropriate level of effort, denoted $y$. The level of $y$ cannot be observed, but it is common knowledge that $y$ is drawn from a uniform distribution over the interval 0 to 1. Although $y$ is unobservable, at the start of the first period workers receive a private signal $s_1$ about $y$. This signal can be either high ($s_1 = H$) or low ($s_1 = L$). The probability that the signal is high increases in $y$. In particular, we assume that $pr(s_1 = H \mid y) = y$ and $pr(s_1 = L \mid y) = 1 - y$. After the first period, the three best-performing workers unexpectedly receive a non-monetary reward. When workers have identical output, a random draw decides. Receiving and not-receiving a reward is denoted $s_2 = R$ and $s_2 = N$, respectively. This unexpected information gives workers the possibility to update their beliefs about $y$, denoted $\hat{y}$. We do not analyze the implications of announcing the cards in advance.\(^{29}\)

First, we analyze the effort choice in the first period. Straightforward maximization of expected utility reveals that optimal effort equates the expectation of $y$, i.e. $e^* = \hat{y}$. The worker’s belief depends on the signal received, as follows:

$$E(y \mid s_1 = H) = \frac{\int pr(s_1 = H \mid y) \cdot y dy}{\int pr(s_1 = H \mid y) dy} = \frac{\int y^2 dy}{\int y dy},$$

yielding $E(y \mid s_1 = H) = 2/3$, while similar calculations show that $E(y \mid s_1 = L) = 1/3$. Thus, workers who received a high signal work harder than those who received a low signal, and have a larger probability of receiving the reward in the second period as a result.

At the start of the second period additional information comes available via the public distribution of the reward to the best performers. After the reward is distributed, workers may have four possible beliefs, depending on the private signal received in the first period and whether the worker received a reward or not. First, we investigate the response of a recipient who received a high signal in the first round.

\(^{29}\)The reason is that doing so would be tedious. As all effects of the rewards come from the revealed information about one’s relative performance, we cannot generally rule out that workers have an incentive to manipulate their first-period effort in order to improve the informational content. Hence, there may not be an equilibrium in pure strategies.
\[ E(y \mid s_1 = H, s_2 = R) = \frac{\int pr(s_1 = H \cap s_2 = R \mid y) \cdot y \, dy}{\int pr(s_1 = H \cap s_2 = R \mid y) \cdot dy}, \]

which gives after evaluating the integral and rewriting \( \frac{25}{42} \), which is less than the first-period belief \( \frac{2}{3} \). So, the worker reduces effort upon receiving the reward, as it increases the likelihood that most other workers had a low signal, and hence \( y < \frac{2}{3} \). Similar calculations show that the reward also has a negative effect on a recipient who received a low signal initially: \( E(y \mid s_1 = L, s_2 = R) = \frac{1}{6} < \frac{1}{3} \). In fact, the reduction in effort is substantial, as receiving the reward despite low first-period effort reveals that there are at least 5 others with a low signal. Thus, receivers tend to reduce effort, and the lower effort in the first period, the stronger this reduction. Third, the effort of a non-recipient who received a high signal in the first round is given by \( E(y \mid s_1 = H, s_2 = N) = \frac{23}{30} \), which is more than \( \frac{2}{3} \). This means that he updates his beliefs positively: he learns that there are at least three others who also received the high signal, and hence \( y \) is probably high. Finally, the response of a non-recipient who received a low signal in the first round is described by \( E(y \mid s_1 = L, s_2 = N) = \frac{11}{30} \), implying a small positive update. As the probability of not receiving a prize after a low signal is high even if all others also received a low signal, there is little information revealed. To conclude, non-recipients exert more effort than in the first period, in particular when first period effort is high already.
Appendix C

Figure 2: Thank-you Card

Figure 3: Screenshot - Data Entry Form

Datenerfassung Projekt CBR-85010

1. Denkest Du heute schon daran, was Du später mal gerne im Beruf machen wirst?
   - ja, sehr häufig
   - ja, aber nicht so oft
   - nein, nie
   - unklar

2. Was möchtest Du nach der Schule gerne machen?
   - eine Berufsausbildung sofort nach der Schule
   - eine Berufsausbildung, aber erst später
   - eine weiterführende Schule besuchen
   - studieren
   - keine weitere Ausbildung
   - weiß ich noch nicht
   - unklar

3. Möchtest Du später einmal den gleichen Beruf wie Deine Eltern ausüben?
   - ja, wie mein Vater
   - ja, wie meine Mutter
   - nein
   - unklar

   - durch meine Eltern
   - durch Verwandte, Freunde oder Bekannte
   - durch meine Lehrer
   - durch die Berufsvorbereitung in der Schule
   - durch ein Praktikum
   - durch einen Nebenjob
   - durch die Arbeitsagentur/das Arbeitsamt, BIZ
   - durch Zeitschriften, Bücher, Fernsehen, Internet
   - unklar

3/12  OK  Fragebogen abbrechen
Figure 4: Difference in Performance by Treatment

![Graph showing performance improvement by treatment.

Figure 5: Difference in Performance by Treatment and Performance Group

![Graph showing performance improvement by treatment and performance group.]
Table 1: Summary Statistics by Treatment Group

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>TC All</th>
<th>TC Best3</th>
<th>TC Best3 ann.</th>
<th>TC Best1</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>81</td>
<td>87</td>
<td>95</td>
<td>69</td>
<td>82</td>
</tr>
<tr>
<td>Female</td>
<td>0.593</td>
<td>0.552</td>
<td>0.558</td>
<td>0.565</td>
<td>0.659</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.054)</td>
<td>(0.051)</td>
<td>(0.060)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Econ major</td>
<td>0.600</td>
<td>0.512</td>
<td>0.634</td>
<td>0.667*</td>
<td>0.457*</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.054)</td>
<td>(0.050)</td>
<td>(0.057)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Mannheim</td>
<td>0.556</td>
<td>0.460</td>
<td>0.779***</td>
<td>1.000***</td>
<td>0.744**</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.054)</td>
<td>(0.043)</td>
<td>(0.000)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Morning</td>
<td>0.407</td>
<td>0.299</td>
<td>0.389</td>
<td>0.406</td>
<td>0.268*</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.049)</td>
<td>(0.050)</td>
<td>(0.060)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Afternoon</td>
<td>0.235</td>
<td>0.437***</td>
<td>0.274</td>
<td>0.217</td>
<td>0.451***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.053)</td>
<td>(0.046)</td>
<td>(0.050)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Evening</td>
<td>0.358</td>
<td>0.264</td>
<td>0.337</td>
<td>0.377</td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.048)</td>
<td>(0.049)</td>
<td>(0.059)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>Baseline productivity(^1)</td>
<td>15.861</td>
<td>16.665</td>
<td>16.174</td>
<td>16.727</td>
<td>16.514</td>
</tr>
<tr>
<td></td>
<td>(0.342)</td>
<td>(0.415)</td>
<td>(0.431)</td>
<td>(0.421)</td>
<td>(0.414)</td>
</tr>
</tbody>
</table>

Note: The Table reports means for each group. Standard errors are displayed in parentheses. Significance levels indicate a difference of means (compared to the control group) and are denoted as follows: * p < 0.1, ** p < 0.05, *** p < 0.01. (1) Baseline productivity is measured as correct entries in working period 1.
Table 2: Productivity in Working Period 1

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC Best3 announced</td>
<td>0.452</td>
<td>-0.747</td>
</tr>
<tr>
<td></td>
<td>(0.521)</td>
<td>(0.484)</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>414</td>
<td>409</td>
</tr>
<tr>
<td>Sessions</td>
<td>59</td>
<td>59</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.002</td>
<td>0.123</td>
</tr>
</tbody>
</table>

Note: This table reports OLS coefficient estimates. Standard errors clustered by session are reported in parentheses. Productivity in working period 1 indicates performance in the first working period measured by total clicks per minute where one click refers to checking a box on the computer screen to transfer the choice of survey participants into the database. TC Best3 announced represents the treatment where three thank-you cards have been announced prior to the start of working period 1. All other treatments serve as the reference category. Column I shows the treatment effect without control variables. Column II includes controls for gender, university and economic field of study, location and time of day, groupsize, and length of working period 1. Significance levels are denoted as follows: * p < 0.1, ** p < 0.05, *** p < 0.01.
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC All</td>
<td>0.707</td>
<td>0.940*</td>
<td>0.677</td>
<td>0.914*</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.519)</td>
<td>(0.477)</td>
<td>(0.521)</td>
</tr>
<tr>
<td>TC Best3</td>
<td>1.145***</td>
<td>1.228***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.384)</td>
<td>(0.439)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best1</td>
<td>0.948**</td>
<td>1.084**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.397)</td>
<td>(0.410)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 announced</td>
<td>-0.133</td>
<td>-0.243</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.508)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 - Rec.</td>
<td></td>
<td>0.574</td>
<td>0.696</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.515)</td>
<td>(0.575)</td>
<td></td>
</tr>
<tr>
<td>TC Best3 - Non-Rec.</td>
<td>1.478***</td>
<td>1.586***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.399)</td>
<td>(0.458)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best1 - Rec.</td>
<td>0.394</td>
<td>0.559</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.791)</td>
<td>(0.840)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best1 - Non-rec.</td>
<td>1.004**</td>
<td>1.170***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.385)</td>
<td>(0.402)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 announced - Rec.</td>
<td>-0.322</td>
<td>-0.319</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.785)</td>
<td>(0.777)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 announced - Non-Rec.</td>
<td>-0.061</td>
<td>-0.148</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.484)</td>
<td>(0.523)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline performance</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other controls</td>
<td>Yes</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>414</td>
<td>409</td>
<td>414</td>
<td>409</td>
</tr>
<tr>
<td>Sessions</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>59</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.742</td>
<td>0.744</td>
<td>0.745</td>
<td>0.747</td>
</tr>
</tbody>
</table>

Note: This table reports OLS coefficient estimates (standard errors clustered by session are reported in parentheses). Correct indicates performance in the second working period measured by correct clicks per minute where one click refers to checking a box on the computer screen to transfer the choice of survey participants into the database. Correct clicks are all clicks entered minus all errors made by the participant. TC All represents a thank-you card for all subjects. TC Best3 refers to sessions where three thank-you cards were assigned to the best three subjects. TC Best1 indicates sessions where only one thank-you card was given to the best subject. TC Best3 announced represents sessions where three thank-you cards have been announced prior to working period 1. Treatment Control is omitted and serves as the reference category. Column I shows overall treatments effects with control for baseline performance where the latter is the productivity in working period 1. Further, we control for baseline performance squared and cubic. In Column II includes controls for gender, university and economic field of study, location and time of day, groupsize, and length of working period 1. Column III allows for a separate analysis for recipients and non-recipients of thank-you cards. Rec indicates recipients and Non-Rec indicate recipients and non-recipients of a thank-you card in the corresponding treatment. Column IV additionally includes all control variables. Significance levels are denoted as follows: * p < 0.1, ** p < 0.05, *** p < 0.01.
Table 4: Treatment effects on Total Productivity and Error Rates in Working Period 2

<table>
<thead>
<tr>
<th></th>
<th>Total Productivity</th>
<th>Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>TC All</td>
<td>0.592</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>(0.466)</td>
<td>(0.500)</td>
</tr>
<tr>
<td>TC Best3</td>
<td>1.231***</td>
<td>1.246***</td>
</tr>
<tr>
<td></td>
<td>(0.393)</td>
<td>(0.450)</td>
</tr>
<tr>
<td>TC Best1</td>
<td>0.974**</td>
<td>1.074**</td>
</tr>
<tr>
<td></td>
<td>(0.406)</td>
<td>(0.419)</td>
</tr>
<tr>
<td>TC Best3 - Rec.</td>
<td>-0.042</td>
<td>-0.155</td>
</tr>
<tr>
<td></td>
<td>(0.480)</td>
<td>(0.505)</td>
</tr>
<tr>
<td>TC Best3 - Non-Rec.</td>
<td>0.559</td>
<td>0.624</td>
</tr>
<tr>
<td></td>
<td>(0.528)</td>
<td>(0.590)</td>
</tr>
<tr>
<td>TC Best3 announced</td>
<td>1.632***</td>
<td>1.670***</td>
</tr>
<tr>
<td></td>
<td>(0.408)</td>
<td>(0.470)</td>
</tr>
<tr>
<td>TC Best1 - Rec.</td>
<td>0.458</td>
<td>0.584</td>
</tr>
<tr>
<td></td>
<td>(0.814)</td>
<td>(0.864)</td>
</tr>
<tr>
<td>TC Best1 - Non-rec.</td>
<td>1.026**</td>
<td>1.159***</td>
</tr>
<tr>
<td></td>
<td>(0.392)</td>
<td>(0.410)</td>
</tr>
<tr>
<td>TC Best3 announced - Rec.</td>
<td>-0.232</td>
<td>-0.239</td>
</tr>
<tr>
<td></td>
<td>(0.794)</td>
<td>(0.780)</td>
</tr>
<tr>
<td>TC Best3 announced - Non-Rec.</td>
<td>0.027</td>
<td>-0.043</td>
</tr>
<tr>
<td></td>
<td>(0.475)</td>
<td>(0.509)</td>
</tr>
</tbody>
</table>

Controls
Baseline performance | Yes | Yes | Yes | Yes | Yes | Yes
Other controls | Yes | Yes | Yes

Observations | 414 | 409 | 414 | 409 | 409 | 409
Sessions | 59 | 59 | 59 | 59 | 59 | 59
\(R^2\) | 0.742 | 0.744 | 0.746 | 0.747 | 0.084 | 0.084

Note: This table reports OLS coefficient estimates (standard errors clustered by session are reported in parentheses). In Columns I-IV, the dependent variable is the total number of clicks entered in the second working period where one click refers to checking a box on the computer screen to transfer the choice of survey participants into the database. In Columns V and VI, the dependent variable are error rates in the second working period. Error rate 2 is calculated by total errors divided by total clicks entered. Detailed explanations on treatment dummies TC All, TC Best3, TC Best3 announced, TC Best1 can be found in Table 5. Treatment Control is omitted and serves as the reference category. All columns show treatments effects with control for baseline performance where the latter is the total productivity in working period 1 (with control for squared and cubic baseline performance) in columns I-IV, and respectively error rate in working period 1 in columns V and VI. In Column I,II,IV and VI include controls for gender, economic field of study and university, location and time of day, groupsize, and length of working period 1. Column II, V and VI allow for a separate analysis for recipients and non-recipients of thank-you cards. _Rec_ indicates recipients and _Non-Rec_ indicate recipients and non-recipients of a thank-you card in the corresponding treatment. Significance levels are denoted as follows: * p < 0.1, ** p < 0.05, *** p < 0.01.
Table 5: Treatment Effects on Productivity in Working Period 2

<table>
<thead>
<tr>
<th></th>
<th>I Best 3</th>
<th>II Worst 5</th>
<th>III Best 1</th>
<th>IV Worst 7</th>
<th>V All</th>
<th>VI All</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC All</td>
<td>1.311*</td>
<td>0.796</td>
<td>0.857</td>
<td>0.963*</td>
<td>0.628</td>
<td>0.609</td>
</tr>
<tr>
<td></td>
<td>(0.697)</td>
<td>(0.598)</td>
<td>(1.029)</td>
<td>(0.515)</td>
<td>(0.550)</td>
<td>(0.555)</td>
</tr>
<tr>
<td>TC Best3</td>
<td>0.937</td>
<td>1.621***</td>
<td>0.199</td>
<td>1.434***</td>
<td>1.159**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.706)</td>
<td>(0.473)</td>
<td>(0.842)</td>
<td>(0.460)</td>
<td>(0.512)</td>
<td></td>
</tr>
<tr>
<td>TC Best1</td>
<td>1.603***</td>
<td>0.798*</td>
<td>0.515</td>
<td>1.242***</td>
<td>0.854</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.557)</td>
<td>(0.474)</td>
<td>(1.012)</td>
<td>(0.393)</td>
<td>(0.526)</td>
<td></td>
</tr>
<tr>
<td>TC Best3 announced</td>
<td>-0.372</td>
<td>0.053</td>
<td>-0.464</td>
<td>-0.158</td>
<td>-1.094</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.802)</td>
<td>(0.457)</td>
<td>(1.058)</td>
<td>(0.499)</td>
<td>(0.710)</td>
<td></td>
</tr>
<tr>
<td>Female*TC All</td>
<td>0.531</td>
<td>0.534</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.653)</td>
<td>(0.663)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female*TC Best 3</td>
<td>0.133</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.640)</td>
<td>(0.637)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female*Best 1</td>
<td>0.409</td>
<td>0.400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.655)</td>
<td>(0.662)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female*Best 3 ann.</td>
<td>1.493</td>
<td>1.470*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.906)</td>
<td>(0.874)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 - Rec.</td>
<td></td>
<td></td>
<td>0.659</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.663)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 - Non-Rec.</td>
<td></td>
<td></td>
<td>1.537***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.484)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best1 - Rec.</td>
<td></td>
<td></td>
<td>0.375</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.898)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best1 - Non-rec.</td>
<td></td>
<td></td>
<td>0.941*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.519)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 announced - Rec.</td>
<td></td>
<td></td>
<td>-1.053</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.948)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC Best3 announced - Non-Rec.</td>
<td></td>
<td></td>
<td>-1.049*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.611)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Controls

Baseline performance | Yes | Yes | Yes | Yes | Yes |
Other controls       | Yes | Yes | Yes | Yes | Yes |
Observations         | 163 | 246 | 58  | 351 | 409 |
Sessions             | 59  | 59  | 58  | 59  | 59  |
$R^2$                | 0.708 | 0.623 | 0.793 | 0.665 | 0.748 |

Note: The table reports OLS estimates with robust standard errors (clustered by session) reported in parentheses. For detailed explanations of variables see Table 5. The Control treatment is omitted and serves as the reference category. All columns include controls for baseline performance and other controls such as gender, university and economic field of study, location and time of day, groupsize and length of working period 1. Column I shows regression results for the subsample of only the best 3 performers within a session and Column II shows the subsample of only the worst 5 performers, accordingly. Additionally, Column III and IV are presenting estimates on the subsample of only the best 1 performer within a session and the worst 7 performers. Significance levels are denoted as follows: * p < 0.1, ** p < 0.05, *** p < 0.01.