Employee Recognition and Performance:
A Field Experiment

Christiane Bradler[1]
Robert Dur[2]
Susanne Neckermann[2]
Arjan Non[3]

[1] ZEW Centre for European Economic Research Mannheim;
[2] Erasmus School of Economics, Erasmus University Rotterdam, ZEW, and Tinbergen Institute;
[3] Maastricht University.

Employee Recognition and Performance:
A Field Experiment[*]

Christiane Bradler[†] Robert Dur[‡] Susanne Neckermann[§] Arjan Non[¶]

March 2013

Abstract

This paper reports the results from a controlled field experiment designed to investigate the causal effect of public recognition on employee performance. We hired more than 300 employees to work on a three-hour data-entry task. In a random sample of work groups, workers unexpectedly received recognition after two hours of work. We find that recognition increases subsequent performance substantially, and particularly so when recognition is exclusively provided to the best performers. Remarkably, workers who did not receive recognition are mainly responsible for this performance increase. This result is consistent with workers having a preference for conformity.

[†] ZEW Centre for European Economic Research Mannheim. E-mail: bradler@zew.de.

[‡] Erasmus University Rotterdam and Tinbergen Institute. E-mail: dur@ese.eur.nl.

[§] Erasmus University Rotterdam, ZEW and Tinbergen Institute. E-mail: neckermann@ese.eur.nl.

[¶] Maastricht University. E-mail: j.non@maastrichtuniversity.nl.

1   Introduction

Recent years have seen a surge in popular business books on the importance of recognition for employee motivation. A prominent example is the book by Nelson (2005) entitled *1001 Ways to Reward Employees*. 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."[1] Other questionnaire studies reveal similar views among employees (Kovach, 1995; Wiley, 1997) and managers (Holton et al., 2009).

The vast amount of practitioner literature on employee recognition is supported by a body of academic research in organizational and personnel psychology. Stajkovic and Luthans (2003) give an overview of field-experimental studies and report strong positive effects of recognition on employees' performance in a variety of workplace contexts. More recently, Grant and Gino (2010) experimentally study how a manager's verbal expression of gratitude affects employees' effort and find strong positive effects.[2]

While the existing literature provides a fairly consistent picture that recognition improves employee performance, much less is known about how exclusive or inclusive recognition should be. Should all employees receive recognition? Or is more differentiation desirable? In particular, does providing recognition to all workers demoralize the top-performers, thus creating a culture of mediocrity? Or does reserving praise for the best performers come at the cost of discouraging the others?

In this paper we take up these questions by conducting a large-scale field experiment in a controlled working environment. Over the course of November 2010 to May 2011, we hired 363 people (mainly students) for a three-hour data-entry job.[3] 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 were taking part in an experiment. In a random sample of work groups, workers received

---

[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] Effects of praise and recognition have also been studied in other contexts. Fisher and Ackerman (1998) investigate the effect of recognition on parents' willingness to volunteer for their kids' soccer club. An early lab experiment is Deci and Ryan (1971) showing that provision of praise increases students' willingness to work on a puzzle. Cameron and Pierce (1994) provide a survey of the subsequent lab-experimental literature. Finally, some studies in management provide correlational evidence, e.g., Rhoades and Eisenberger (2002), Cropanzano and Mitchell (2005), and Wagner and Harter (2006).

[3] Short-term data-entry jobs like ours have been frequently employed in recent field experiments studying the impact of rewards on performance, because they allow for a precise measurement of performance and provide for a high degree of control (see among others Gneezy and List, 2006, Hennig-Schmitt et al., 2010, Kosfeld and Neckermann, 2011, and Kube et al., 2012, forthcoming).

public recognition after two hours of work. Following Kosfeld and Neckermann (2011), recognition consisted of a thank-you card, personally signed by the head of the research institute and handed out by a research assistant.[4] Thus, the 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 deduce that no further formal recognition would be provided in the remainder of the working period. A defining feature of the type of recognition that we have in mind is that it is not expected and hence comes as a surprise to employees.[5] Therefore, and in contrast to Kosfeld and Neckermann (2011), the possible provision of recognition was not announced at the start of the three-hour working period.

To address the research questions posed above, we vary the scarcity of recognition between different treatments. In one treatment, all workers in a group receive the thank-you card. In the other treatments, the thank-you card is given only to the best performer or to the best three performers of a group of eight. When handing out the thank-you cards in these latter treatments, it is mentioned that their assignment is determined by performance during the first two working hours. Further, we make clear that the management recognizes the work input from all workers, but can only recognize a subset of the employees explicitly due to constraints. This resembles that, in practice, managers are often constrained in the number of people they can praise. Lastly, we take care that perceived monitoring is constant across all treatments.

Our focus is on how the provision of recognition affects subsequent performance of workers, how this depends on the scarcity of recognition, and how the response of recipients differs from that of non-recipients. We argue that three behavioral mechanisms may be operational: conditional altruism, reciprocity, and conformity preferences of employees. As we shall see, our results are consistent with a combination of conformity and reciprocity preferences.

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 setup in Section 3. Section 4 lays out the empirical strategy and presents the results of our experiment. Finally, Section 5 concludes.

---

[4] The use of such thank-you cards is not uncommon in business. Postcards producer Hallmark devotes a separate website to employee thank-you cards, see http://hallmark.businessgreetings.com/employeerecognitioncards-employeethankyoucards-3-0.

[5] This is often true in practice, even in organizations with a culture for praise. While recognition in general might be expected in such organizations, the particular situation, timing, and kind of recognition remain a surprise.

2   Theory

Why might the unannounced provision of recognition in the form of a thank-you card affect subsequent performance of workers? And does it matter whether all workers or only the best performing workers in a group receive a thank-you card? We can think of three plausible behavioral mechanisms through which provision of public recognition may affect subsequent performance.

2.1   Reciprocity of employees

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. (2012), 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 in the simplest possible way, suppose that workers have the following utility function (as in Dur et al., 2010):

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

where w is the base salary, e represents the worker's effort, a the amount or value of recognition received from the employer, γ indicates the importance of reciprocal feelings, β captures other intrinsic motives to exert effort, and the last term of the equation represents the worker's costs of effort. Clearly, a worker's optimal effort is increasing in the amount of recognition received (maximization of u with respect to e gives $e^* = \beta + \gamma a$). Workers who do not receive a thank-you card do not respond to the intervention. Thus, in treatments where only a subset of workers receives 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.

2.2 Conditional altruism of employees

A second reason for why the provision of a thank-you card may affect subsequent effort is that the card provides a signal to the workers about the employer's kindness or his care for the workers. The key assumption is that workers are conditionally altruistic, that is, they 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, 2012). This tendency to contribute to the welfare of a kind person can be seen as a form of reciprocity, but it has implications that are different from those of the simple reciprocity model considered above, particularly for the response of non-recipients.[6] The following game-theoretic model illustrates the basic mechanism.

Consider a group of n workers who work for an employer whose altruism towards them is private information. It is common knowledge, however, that the employer's altruism, denoted α_j, is randomly drawn from a uniform distribution over the interval 0 to 1. The utility of worker i is described by

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

where we use the same notation as above. The term $\gamma \alpha_j \sum_{i=1}^{n} (e_i - w_i)$ captures the idea that workers care more about the profits of the employer ($\sum_{i=1}^{n} (e_i - w_i)$) if the employer is more altruistic towards them. As α_j is unobservable, workers form beliefs over α_j, denoted by $\widehat{\alpha_j}$. We assume that profits are simply the sum of all workers' effort minus wage costs. The employer's payoff y_j is:

$$y_j = \sum_{i=1}^{n} (e_i - w_i) + \alpha_j \sum_{i=1}^{n} u_i - c(\alpha_j),$$

where the second term captures his altruistic utility. The employer can signal his type by creating thank-you cards, where $c(\alpha_j)$ captures the costs (of effort, time, and money) involved in creating a given number of cards. When no cards are created, $c(\alpha_j) = 0$ for any α_j. In the spirit of Camerer (1988), Prendergast and Stole (2001), and Ellingsen and Johannesson (2011), we assume that the costs of creating thank-you cards decrease in the

---

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

employer's altruism, i.e., $c'(\alpha_j) < 0$.[7] Further, to keep the model as simple as possible, assume that there are two working periods of equal length and no time discounting. Costs associated with the cards are incurred before the second working period starts. The cards are handed out in between the two working periods.

Applying backward induction, it is easily verified that workers' effort increases in their beliefs about the employer's altruism $\widehat{\alpha}_j$: the optimal effort choice is $e^* = \beta + \gamma\widehat{\alpha}_j$. The optimal effort choice in the first working period reflects the prior belief $\widehat{\alpha}_j = \frac{1}{2}$. At the start of the second period, the employer hands out cards or not, and workers update their beliefs using Bayes' rule. As the cards are handed out publicly, both recipients and non-recipients receive the same information, and therefore hold the same beliefs about the employer's type. Thus, on the basis of the signaling model, we expect no difference in the response to the cards between recipients and non-recipients.

Since workers' effort and utility increase in workers' beliefs $\widehat{\alpha}_j$, all employer types would like to be perceived as altruistic as possible. However, this incentive is stronger for more altruistic employer types, as they genuinely care about the workers' utility. Moreover, as $c'(\alpha_j) < 0$, providing recognition is less costly for more altruistic types. They can therefore distinguish themselves from less altruistic types by sending a signal that is too costly for less altruistic types to imitate. Thus, handing out thank-you cards provides a signal about the employer's altruism when the costs of providing the cards exceed their benefits for types below some threshold level of altruism α^*, while handing out cards is profitable for types above the threshold. The threshold level of altruism α^* is formally derived in Appendix A. An employer who hands out cards is believed to be a type from the interval α^* to 1, while an employer who abstains from doing so is believed to be a type from the interval 0 to α^*. Equilibrium effort in working period 2 with and without cards are $\beta + \gamma\frac{1+\alpha^*}{2}$ and $\beta + \gamma\frac{\alpha^*}{2}$, respectively. On the basis of signaling theory, we thus expect a positive effort response to provision of recognition for recipients and non-recipients alike.

An increase in the costs of providing cards strengthens the workers' response to the cards. The reason is that higher costs imply that fewer employer types find it attractive to signal, only the most altruistic types do so. Hence, assuming that costs increase in the number of cards, the signal is the stronger the more cards are provided. The effort response therefore differs between treatments, with stronger responses if more cards are distributed.[8] Note that this is opposite to the prediction of the simple reciprocity model in

---

[7] This assumption is not essential for the results, see Appendix A.

[8] Other signaling models are closely related, 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. Alternatively, the thank-you card may affect workers' beliefs about how important their effort is to the employer or how important the task is (task significance). These theories

the previous subsection.

2.3 Conformity preferences of employees

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 end of the performance distribution. If workers have a preference for conformity (for example, if they want to behave like others or to adhere to a group norm as in Bernheim, 1994, Sliwka, 2007, Fischer and Huddart, 2008, and Chen et al., 2010), this coarse rank information should increase effort of non-recipients and decrease effort of recipients. The effect on non-recipients should be less pronounced when the thank-you card is provided to fewer workers, as less information is provided about one's standing. Thus, we expect smaller effects for non-recipients in the treatment where only a single thank-you-card is provided to the best performer 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 distributed, as this provides a stronger signal of exceptionally high performance by the recipients. Lastly, on the basis of conformity preferences alone, we do not expect any treatment effects when all workers are provided with a thank-you card, since nothing is revealed about relative performance in this treatment.[9]

To illustrate this mechanism, consider 8 workers who work for two periods. In both periods, workers choose an effort level e. Their utility function is described by

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

Workers condition their effort choice on what they consider as the group norm, denoted

_____

yield similar predictions.

[9] The effect of relative performance feedback on performance has recently been studied in a variety of contexts, ranging from student test scores (Bandiera et al., 2009; Azmat and Iriberri, 2010; Tran and Zeckhauser, 2012), contributions to an online community (Chen et al., 2010), real-effort in the lab (Charness et al., 2012; Freeman and Gelber, 2010; Kuhnen and Tymula, 2012), to workers' performance in a real working environment (Falk and Ichino, 2006; Delfgaauw et al., forthcoming; Blanes i Vidal and Nossol, 2011; Barankay, 2011a,b). With the noteworthy exception of Barankay (2011a,b), all of these papers find that relative performance feedback has on average a positive effect on performance. A subset of these papers also looks at where in the performance distribution people are most responsive to feedback; results tend to be mixed in this respect. A distinguishing feature of our study is that we look at *ex post* effects of rank revelation that is not announced *ex ante* and in the *absence* of further rank concerns (as ranks will not be disclosed again). Most of the empirical papers as well as the theoretical literature center around the effect of rank revelation in ongoing tournaments.

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. 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 recognition (we deal with the case of recognition for the best employee below). When workers have identical output, a random draw decides. Receiving and not-receiving recognition is denoted by $s_2 = R$ and $s_2 = N$, respectively. This information gives workers the possibility to update their beliefs about y, denoted $\widehat{y}$.[10]

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

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

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 higher probability of receiving recognition at the start of the second period as a result.

At the start of the second period, additional information becomes available via the public distribution of recognition to the best performers. After recognition is provided, workers may have four possible beliefs, depending on the private signal received in the first period and whether the worker received recognition or not. First, we investigate the response of a recipient of recognition who received a high signal in the first round. His posterior belief about y is equal to

$$E(y \mid s_1 = H, s_2 = R) = \frac{\int pr(s_1 = H \cap s_2 = R \mid y) . y dy}{\int pr(s_1 = H \cap s_2 = R \mid y) . dy},$$

which is equal to $\frac{25}{42}$, and, hence, less than the first-period belief $2/3$. So, the worker reduces effort upon receiving recognition, as it increases the likelihood that many other workers had a low signal, and hence $y < 2/3$. Similar calculations show that receiving recognition also has a negative effect on a recipient who initially received a low signal: $E(y \mid s_1 = L, s_2 = R) = \frac{1}{6} < 1/3$. In fact, the reduction in effort is substantial as receiving recognition despite low first-period effort reveals that there are at least 5

---

[10]Alternative interpretations of y are task significance and whether the employer is kind and/or worth impressing as in Ellingsen and Johannesson (2007, 2008).

co-workers with a low signal. Thus, such a recipient tends to reduce effort. Third, the effort of a non-recipient who received a high signal in the first round is given by $E\left(y \mid s_1 = H, s_2 = N\right) = \frac{23}{30}$, which is higher than $2/3$. This means that he updates his belief positively: he learns that there are at least three co-workers who also received the high signal, and that, hence, y is probably high. Finally, the response of a non-recipient who received a low signal in the first round is described by $E\left(y \mid s_1 = L, s_2 = N\right) = \frac{11}{30}$, implying a small positive update. To conclude, the conformity model predicts that recipients exert less effort while non-recipients exert more effort than in the first period. It is straightforward to repeat the analysis for the case of recognition to the best employee only. Qualitatively, the results are the same, but stronger for the recipient and weaker for the non-recipients of recognition.

Table 1 summarizes our theoretical predictions.

3   Design and Implementation

3.1   Background

To estimate the causal effect of recognition on employee performance, we conducted a natural field experiment (Harrison and List, 2004).[11] This means that we observed participants in a controlled work environment and that subjects were not aware of the fact that they were part of an experiment. On behalf 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. At the time, 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 receive a fixed wage of 25 euro.[12] Students could sign up for the job online and were subsequently contacted by phone to arrange date and time of their assignment. We minimized career concerns by informing participants that the job was one-time, that a second participation

---

[11] In addition to the treatments reported in this paper, we conducted four additional treatments investigating other ways of providing recognition in the same setup. Results from these treatments are reported in Bradler and Neckermann (2013) and Bradler et al. (2013). All regressions reported below include observations from and dummies for these treatments to obtain more accurate estimates for the coefficients of the control variables.

[12] The fixed wage corresponds to a typical hourly wage of students for this kind of short-term job.

was not possible for tax reasons,[13] and that the research team was not looking for permanently employed research assistants at the moment. We invited subjects in groups of eight. Sessions took place at two different locations and three different times of the day (morning, noon, and afternoon).[14] Following Kosfeld and Neckermann (2011), we implemented recognition by handing out thank-you cards (see Figure 1) to participants. In contrast to Kosfeld and Neckermann (2011), these thank-you cards were not announced at the start of the working period. Instead, they were unexpectedly provided after two hours of work to either all or a subset of participants in the work session depending on the treatment. After this interruption, participants worked for another hour.

3.2 Implementation

Upon arrival, subjects chose a workstation and were assigned a login code for the data entry surface. The workspaces were arranged in a U-shape with sufficient space between individuals to ensure that subjects felt unobserved. After all formalities (privacy policy, wage agreement) were taken care of, the fixed wage of 25 euro was paid in advance to stress that payment was independent of work performance. Subsequently, the participants received a short introduction to their employer (ZEW) and to the project that the surveys stemmed from. The surveys were part of an evaluation project in the educational context and involved responses from several hundreds of students to questions about their school and career preferences. Additionally, a short briefing was given on how to file the answers from the questionnaires into the database. As the data were filed via a web interface, 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 multiple choice, only a few questions required entering information from free text fields. To minimize productivity spillovers between participants, questionnaires were stacked in a high, non-transparent box placed in front of each participant, with an identical box next to it to deposit the completed ones.

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, and absence of supervisor) were taken

---

[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 effort for fiscal reporting and was, hence, not desired by the research institute.

[14] The field experiment was carried out at the Centre for European Economic Research (ZEW) in Mannheim and the Ruprecht-Karls University of Heidelberg. In the regressions we control for time-of-day as well as for day-of-week and location effects, but dropping these controls does not affect the results. All sessions were conducted by the same research assistant.

to give participants substantial leeway in the amount of time they spend on filling the database. Furthermore, a collective break was avoided to minimize possible group effects and communication between participants.[15] When participants entered the room at the beginning, they freely chose a workstation. Each work station was provided with an individual login and a set of questionnaires. The research assistant publicly made note about which workstation was used by whom. Thus, subjects were aware of the fact that login codes of a workstation could be linked to their person, which holds perceived monitoring constant across treatments.

Eight subjects were assigned to each working session. However, since some subjects did not show up, the average number of participants per session was 7.21 (s.d. = 0.87). The timeline was as follows: the formalities and introduction lasted for about 18 minutes on average. Then, participants worked on the task for roughly 102 minutes. We refer to this phase as working period 1.[16] 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 were transmitted correctly from each computer to the central data base.[17] Depending on the treatment, she did or did not provide recognition by handing out thank-you cards, following a standard procedure described in the next subsection. Subsequently, subjects worked for approximately one additional hour – which we refer to as working period 2.[18] Working period 1 was longer than working period 2 to allow for variation in the length of the introductory period and to let learning effects smooth out which we believed to be largest at the beginning of working period 1. 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 information on field of study.

---

[15] Because the assistant was sitting outside, we have no perfect monitoring of the amount of communication in the room. However, the research assistant was sitting close to the door and could hear everything that was said loudly. According to her, the extent of communication was minimal in all sessions.

[16] The standard deviation of the introductory period is 5 minutes. Differences in length are due to delays in arrival of participants and differences in the number of questions that were asked during the introduction. The standard deviation of the duration of working period 1 is 6.6 min. Differences in the lengths of working period 1 are mainly due to delays in the introductory period.

[17] This statement ensures that in control as well as in the treatment sessions, subjects were aware of the fact that we could link data entered into the database to the login codes provided.

[18] The mean of working period 2 is 54.2 min with a standard deviation of 4.1 min. 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. Additionally, we control for the length of working period 1 in the regressions.

3.3 Treatments

As described above, the treatment interventions took place after working period 1. The different scripts can be found in Appendix B. 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. Nothing else happened in the *Control* treatment, so that we can measure the development of productivity without recognition but with disruption. In treatments involving thank-you cards, she continued by saying that the research team, together with the president of the institute, had created thank-you cards as a symbol of the institute's gratitude 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, see Figure 1.

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.[19] The treatment "Thank-you card for the best 1" (*TC Best1*) was identical to *TC Best3* apart from the fact that there was only one recipient of a thank-you card, the participant who performed best in the first working period.

The scarcity of cards was explained by saying that the president did not manage to sign a card for each participant. This excuse renders it unlikely that participants expect additional cards at the end of working period 2 and makes sure that subjects knew that there was a practical reason for making recognition scarce. 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 subjects who received cards comparable.[20]

Since sessions took place over the course of several months, we cannot rule out that participants in later sessions heard about treatments in earlier sessions. We took two measures to address this concern. First, we conducted the more salient treatments (*TC*

---

[19] We measured performance in terms of the number of finished questionnaires in working period 1. Subjects were told that the IT department had informed the research assistant which three logins had performed best. Hence, in all treatments, including the control, subjects knew that the research assistant herself could not directly observe the performance of subjects in real-time. She announced that login codes X, Y, and Z performed best and called the names of the participants who were assigned these logins.

[20] For a group size of 7-8 persons we assigned three thank-you cards, and for a group size of 5-6 persons we assigned two thank-you cards in treatment *TC Best3*. We control for group size in all regressions. The results are not sensitive to this.

Best3, TC Best1) subsequent to the less salient treatments (*Control, TC All*). Second, we elicited how people learned about the job in the feedback form at the end of the employment. Only a handful of subjects indicated that they had learned about the job from a friend. The exclusion of these subjects from the analysis does not change the results.

4 Results

Table 2 reports descriptive statistics by treatment. The sample from the described treatments contains 340 observations, 81 in *Control*, 82 in *TC All*, 95 in *TC Best3*, and 82 in *TC Best1*.[21] There are statistically significant differences between treatments in some of the observables. However, as shown below, controlling for these characteristics does not alter the results. Treatments are balanced with respect to baseline performance, i.e., performance in the first working period. Throughout the paper, performance is measured as the number of correct entries per minute of working time.[22]

We first examine the main treatment effects. Figure 2 gives a first impression. It shows the average improvement in performance between working periods 1 and 2 for the different treatments. Individuals in the *Control* treatment show almost no improvement in performance. By contrast, individuals in treatments with recognition clearly improve their performance. This effect is most pronounced in *TC Best3*. Also, performance improves slightly more in *TC Best1* as compared to *TC All* suggesting that recognition motivates and that exclusive recognition works better than praising everyone.

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

$$
\begin{aligned}
y_{i,t=2} &= \alpha + \beta_1 TCAll + \beta_2 TCBest3 + \beta_3 TCBest1 \\
&\quad + \delta_1 y_{i,t=1} + \delta_2 y_{i,t=1}^2 + \delta_3 y_{i,t=1}^3 + \sum_{j=2}^{8} \gamma_j r_{i,j} + \varepsilon_{i,s},
\end{aligned}
\tag{1}
$$

---

[21] In total, 363 persons participated in 48 sessions. We dropped four sessions (two of *Control* and two of *TC All*) from the analysis due to severe server breakdowns during working time. Moreover, we excluded two participants as they were hardly able to work on the task due to physical restrictions (one person was visually impaired and could not read the screen, the other person was over 60 years old and did not know how to operate the computer mouse).

[22] One entry corresponds to checking a box of a multiple choice question on the computer screen or typing one word in a free-text field of the data entry surface. Correctness of an individual entry is determined by whether or not it corresponds to what the majority of participants transferred as answer for this particular question. This serves as a very reliable quality measure because each survey was entered on average 21.6 times (s.d.=10.8).

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, linearly as well as squared and cubic.[23] Additionally, all regressions control for rank-specific effects that are independent of the treatment where $r_{i,j}$ is a dummy variable taking value one if the individual i was of rank j in working period 1, and taking value zero otherwise. Standard errors are clustered by session.

The first column of Table 3 shows the results of estimating (1), the second column includes gender, field of study, home university, location, time of day, group size, and length of working period 1 as additional controls. The results of both specifications are similar. As already suggested by Figure 2, 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 7.2% or a change of 0.28 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=0.19, one-sided). Giving recognition to all participants (*TC All*) also increases performance. The effect is only weakly statistically significant and smaller than when recognition is scarce.[24] Overall, these results clearly show that performance responds positively to the provision of recognition, suggesting that recognition might be a cost-effective tool to stimulate workers' effort.

In order to shed light on underlying mechanisms and to assess the relevance of the different theories discussed in Section 2, we next look at recipients and non-recipients of recognition separately. Recall that the theories have quite distinct predictions regarding the responses of recipients and non-recipients (see Table 1). Figure 3 gives a first impression. The figure shows performance improvements separately for recipients and non-recipients. Recipients increase their performance and this performance increase is slightly decreasing in the scarcity of the recognition. Standard errors are very high for recipients in *TC Best1* due to the small number of observations in this cell. Interestingly, in all treatments where the research assistant was forced to differentiate between employees, the non-recipients subsequently improve by more than the recipients do. The difference is most pronounced for the *TC Best3* treatment, where the improvement of non-recipients is more than twice as large as the improvement of recipients. Hence, the differences between the overall treatment effects that we saw above are mainly driven by the non-recipients rather than the recipients.

---

[23]This amounts to a standard diff-in-diff estimation when the coefficient for the linear term is one and the coefficients of the quadratic and cubic terms are zero. Our specification is more flexible and allows us to control for mean reversion, for which we find some indication in our data. However, treatment effects estimated with the more simple diff-in-diff are very close to the ones shown below. See Table 4.

[24] The difference between *TC All* and *TC Best3* is not statistically significant (Wald-test: p=0.22, one-sided).

An obvious point of criticism on these simple comparisons between recipients and non-recipients is that there is some randomness in employee performance; and thus we would expect reversion to the mean in the second working period even in the absence of any treatment. To account for this, we control for baseline performance (linear, squared, and cubic) and baseline rank in all our regressions, so that we compare 'like with like'.[25] Columns III and IV of Table 3 report the results from estimating equation (1) with separate dummy variables for recipients and non-recipients in each treatment. The results confirm the descriptive analysis in Figure 3: positive treatment effects are largely driven by non-recipients. The latter increase their performance by 1.59 clicks per minute[26] in *TC Best3* and by about 1.05 clicks per minute in *TC Best1*, which are both statistically significantly different from the change in performance of respective others in the control treatment. By contrast, recipients do not show a statistically significant increase in performance in *TC Best3*, and even a slight decrease in performance in *TC Best1*.[27] The difference between recipients and non-recipients is significant at the 5% level in *TC Best3* (Wald test: p=0.016, one-sided), and at the 10% level in *TC Best1* (Wald test: p=0.057, one-sided).[28]

We have seen that recognition has a strong positive impact on subsequent performance, in particular when recognition is scarce, but not too scarce. This effect is primarily caused by the improvement in performance of those who did not receive recognition. How can we reconcile these findings with the theories discussed in Section 2? Our results seem most consistent with individuals having a preference for conformity and being reciprocal at the same time. The strong performance improvement of non-recipients

---

[25] The estimated coefficients for baseline performance point to the presence of mean reversion, albeit small in magnitude. Additionally, the rank coefficients reveal that all ranks between 2-8 have a negative relation with performance in the second working period as compared to the base category of the first rank, and coefficients vary between -0.2 and -0.6. Most other control variables do not have a statistically significant effect apart from the university where participants study. Students from Mannheim and Ludwigshafen show a slightly better performance as compared to students from Heidelberg. Our results are robust to the inclusion of additional controls for different time effects, such as the day of the week or whether the session took place during final exam period, during the semester, or during semester break. The results remain qualitatively similar and, depending on the specification, the coefficients of interest even get slightly larger. Lastly, we ran specifications including interaction terms of treatments with gender and find no evidence for gender-specific treatment effects.

[26] This corresponds to a performance increase of 11.2 percent or a change of 0.46 standard deviations.

[27] We can rule out that recipients do not respond because they are at their maximum level of performance ('ceiling effects'). Bradler and Neckermann (2013) use the same setup and show that recipients (best 3 performers in the first working period) do increase their performance significantly and substantially subsequent to receiving a combination of recognition and money.

[28] The data allow us to analyze treatment effects separately for quantity and quality, where quality is measured as the error rate (percentage of incorrect entries). The results are reported in Table 5. We find no significant treatment effects for quality levels except for a smaller error rate in *TC All* (significant at the 10 percent level). One reason for small effects on the error rate is that quality levels are already at a very high level in working period 1 (only 1.5% of all entries were incorrect, s.d. = 2.4), leaving little room for improvement.

14

is well in line with conformity theory.[29] Moreover, the fact that the improvement in productivity of non-recipients is larger in *TC Best3* than in *TC Best1* nicely fits with the predictions of conformity preferences (Wald test: p=0.06, one-sided). At the same time, however, conformity theory cannot explain the increase in performance in *TC All*. Further, in isolation, conformity theory would predict that recipients of scarce recognition reduce their performance, while we find small increases in *TC Best3*. Hence, it seems likely that employee behavior is also partly driven by reciprocity, which explains these level-effects. This reciprocity may either stem from an inclination to return a gift (as discussed in subsection 2.1) or from signaling (as discussed in subsection 2.2). The data do not allow us to distinguish between these two forms of reciprocity in a meaningful way.

5   Conclusion

In a controlled workplace context, we have shown that provision of public recognition to employees causes a statistically and economically significant increase in performance. Further, we have shown that recognition works best when it is provided exclusively, but not too exclusively. In groups of eight workers, recognition to the best three performers in a group led to a stronger performance increase than either recognition to all employees and recognition to the best performer. The performance increases in response to exclusive recognition are mainly driven by strong positive responses of non-recipients. Conformity preferences are the most likely reason for these responses. Upon learning that one does not belong to the best three performers in a group of eight, non-recipients feel inclined to improve performance so as to adhere to the apparent group norm. In line with this interpretation, we find weaker (but still positive) responses of non-recipients when only the best performer in a group received recognition as this is a weaker signal of low relative performance than belonging to the bottom five out of eight workers. However, conformity preferences cannot explain all of our results. In particular, recipients of recognition do not always decrease performance, as conformity would suggest, but sometimes (slightly and insignificantly) increase it. Moreover, we find that recognition for all workers in a group increases performance. A natural candidate to reconcile these findings is that, in addition to conformity preferences, workers are also reciprocal and hence increase performance in

---

[29] Recall that in the treatments with scarce recognition the research assistant told the employees that she did not have any further thank-you cards to distribute, rendering it unlikely that the performance increase of the non-recipients that we observe here is caused by non-recipients' expectation of and striving for another round of cards at the end of working period 2. A further reason to be sceptical about this alternative explanation is that Bradler et al. (2013) do not find any evidence for incentive effects of *announced* thank-you cards to the three best performing workers in an identical setup.

response to recognition.

Our findings have clear management implications. On the one hand, the results suggest that recognition can be a cost-effective tool for increasing average effort, especially when recognition is limited to a substantial subset of high performers. On the other hand, managers need to take into account that the distribution of scarce recognition provides information to workers about the work norm prevalent in the group. Depending on the circumstances, this may or may not be beneficial for the company.

The current experiment raises a number of questions that remain open for future research. Since the performance increase was primarily driven by non-recipients, a similar effect might be reached at lower cost with pure relative performance information.[30] Another potentially very interesting follow-up study could look at whether the effects persist over time and when recognition is provided repeatedly. Moreover, it might be fruitful to replicate the design with experienced workers and in a setting where other incentives are present. Further, in our experiment we explained the lack of thank-you cards for everyone by stating that time constraints of the president prohibited providing cards to all. It would be interesting to see whether behavioral patterns change when it is stated that cards were deliberately made scarce. This may have important implications for the signaling value of the cards with respect to employer characteristics. These additions will also be helpful for managers who wonder to what extent the expression of one's gratefulness can be used strategically and repeatedly.

---

[30] Neckermann and Yang (2012) report the results of a lab experiment with a similar setup. In the experiment, five agents work for one principal on a real-effort task. The profit of the principal depends on the effort of the agents. In one treatment, participants learned about the identities of the top three performers in the group (pure information). In two other treatments, the principal could opt for a message of recognition or a financial bonus to the top three performers. They show that non-recipients increase their effort similarly in all treatments. This suggests that the information per se rather than the rewards drives the response of non-recipients. Interestingly, the authors do not find any performance response from the recipients in all of their treatments, either.

References

Akerlof, G. A., 1982. Labor contracts as partial gift exchange. Quarterly Journal of Economics 97 (4), 543–569.

Azmat, G., Iriberri, N., 2010. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. Journal of Public Economics 94 (7-8), 435–452.

Bandiera, O., Larcinese, V., Rasul, I., 2009. Blissful ignorance? Evidence from a natural experiment on the effect of individual feedback on performance, mimeo, London School of Economics and Political Science.

Barankay, I., 2011a. Rank incentives: Evidence from a randomized workplace experiment. Mimeo, University of Pennsylvania.

Barankay, I., 2011b. Rankings and social tournaments: Evidence from a crowd-sourcing experiment. Mimeo, University of Pennsylvania.

Bernheim, B. D., 1994. A theory of conformity. Journal of Political Economy 102 (5), 841–877.

Blanes i Vidal, J., Nossol, M., 2011. Tournaments without prizes: Evidence from personnel records. Management Science 57 (10), 1721–1736.

Blau, P. M., 1964. Exchange and Power in Social Life. John Wiley, New York.

Bradler, C., Dur, R., Neckermann, S., Non, A., 2013. Reward expectancy and performance : A field experiment. Mimeo.

Bradler, C., Neckermann, S., 2013. Money meets recognition: A field experiment on worker effort. Mimeo, ZEW.

Camerer, C., 1988. Gifts as economic signals and social symbols. American Journal of Sociology 94, 180–214.

Cameron, J., Pierce, W., 1994. Reinforcement, reward, and intrinsic motivation – a meta-analysis. Review of Educational Research 64 (3), 363–423.

Charness, G., Masclet, D., Villeval, M. C., 2012. The dark side of competition for status. Economics Working Paper 06-12, University of California at Santa Barbara.

Chen, Y., Harper, F. M., Konstan, J., Li, S. X., 2010. Social comparisons and contributions to online communities: A field experiment on movielens. American Economic Review 100 (4), 1358–98.

Cropanzano, R., Mitchell, M. S., 2005. Social exchange theory: An interdisciplinary review. Journal of Management 31 (6), 874–900.

Deci, E. L., Ryan, R. M., 1971. Effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology 18 (1), 105–115.

Delfgaauw, J., Dur, R., Sol, J., Verbeke, W., forthcoming. Tournament incentives in the field: Gender differences in the workplace. Journal of Labor Economics.

Dur, R., 2009. Gift exchange in the workplace: Money or attention? Journal of the European Economic Association 7 (2-3), 550–560.

Dur, R., Non, A., Roelfsema, H., 2010. Reciprocity and incentive pay in the workplace. Journal of Economic Psychology 31 (4), 676–686.

Ellingsen, T., Johannesson, M., 2007. Paying respect. Journal of Economic Perspectives 21 (4), 135–150.

Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: The human side of incentive theory. American Economic Review 98 (3), 990–1008.

Ellingsen, T., Johannesson, M., 2011. Conspicuous generosity. Journal of Public Economics 95 (9-10), 1131–1143.

Falk, A., Fischbacher, U., 2006. A theory of reciprocity. Games and Economic Behavior 54 (2), 293–314.

Falk, A., Ichino, A., 2006. Clean evidence on peer effects. Journal of Labor Economics 24 (1), 39–58.

Fehr, E., Kirchsteiger, G., Riedl, A., 1993. Does fairness prevent market clearing? An experimental investigation. Quarterly Journal of Economics 108 (2), 437–59.

Fischer, P., Huddart, S., 2008. Optimal contracting with endogenous social norms. American Economic Review 98 (4), 1459–1475.

Fisher, R. J., Ackerman, D., 1998. The effects of recognition and group need on volunteerism: A social norm perspective. Journal of Consumer Research 25 (3), 262–75.

Freeman, R. B., Gelber, A. M., 2010. Prize structure and information in tournaments: Experimental evidence. American Economic Journal: Applied Economics 2 (1), 149–64.

Gneezy, U., List, J., 2006. Putting behavioral economics to work: Field evidence on gift exchange. Econometrica 74 (5), 1365–1384.

Grant, A. M., Gino, F., 2010. A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior. Journal of Personality and Social Psychology 98 (6), 946–955.

Gul, F., Pesendorfer, W., 2010. Interdependent preference models as a theory of intentions. Princeton University Working Paper.

Harrison, G. W., List, J. A., 2004. Field experiments. Journal of Economic Literature 42 (4), 1009–1055.

Hennig-Schmitt, H., Rockenbach, B., Sadrieh, A., 2010. In search of worker's real effort reciprocity – a field and a laboratory experiment. Journal of the European Economic Association 8 (4), 817–837.

Holton, V., Dent, F., Rabbetts, J., 2009. Motivation and employee engagement in the 21st century: A survey of management views. An Ashridge Report.

Kosfeld, M., Neckermann, S., 2011. Getting more work for nothing? Symbolic awards and worker performance. American Economic Journal: Microeconomics 3 (1), 1–16.

Kovach, K. A., 1995. Employee motivation: Addressing a crucial factor in your organization's performance. Employment Relations Today 22 (2), 93–107.

Kube, S., Maréchal, M. A., Puppe, C., 2012. The currency of reciprocity – Gift-exchange in the workplace. American Economic Review 102 (4), 1644–1662.

Kube, S., Maréchal, M. A., Puppe, C., forthcoming. Do wage cuts damage work morale? Evidence from a natural field experiment. Journal of the European Economic Association.

Kuhnen, C. M., Tymula, A., 2012. Feedback, self-esteem and performance in organizations. Management Science 58 (1), 94–113.

Levine, D. K., 1998. Modeling altruism and spitefulness in experiments. Review of Economic Dynamics 1 (3), 593–622.

Neckermann, S., Yang, X., 2012. The impact of scarce rewards on worker effort: An experimental investigation. mimeo.

Nelson, B., 2005. 1001 Ways to Reward your Employees. Vol. 2. Workman Publishing Company, New York.

Non, A., 2012. Gift-exchange, incentives, and heterogeneous workers. Games and Economic Behavior 75 (1), 319–336.

Prendergast, C., Stole, L., 2001. The non-monetary nature of gifts. European Economic Review 45, 1793–1810.

Rhoades, L., Eisenberger, R., 2002. Perceived organizational support: A review of the literature. Journal of Applied Psychology 87 (4), 698–714.

Sliwka, D., 2007. Trust as a signal of a social norm and the hidden costs of incentive schemes. American Economic Review 97 (3), 999–1012.

Stajkovic, A. D., Luthans, F., 2003. Behavioral management and task performance in organizations: Conceptual background, meta-analysis, and test of alternative models. Personnel Psychology 56 (1), 155–194.

Tran, A., Zeckhauser, R., 2012. Rank as an inherent incentive: Evidence from a field experiment. Journal of Public Economics 96 (9–10), 645–650.

Wagner, R., Harter, J. K., 2006. 12: The Elements of Great Managing. New York: Gallup Press.

Wiley, C., 1997. What motivates employees what motivates employees according to over 40 years of motivation surveys. International Journal of Manpower 18 (3), 263 – 280.

Appendix A: Signaling in Equilibrium

In this section, we formally characterize the threshold level of altruism α^* and show under what conditions an interior solution exists. The threshold level of altruism α^* is defined as the employer type who is indifferent between creating and not creating cards:

$$
\begin{aligned}
c\left(\alpha^*\right) = & \sum_{i=1} e_i^* \left(\frac{1+\alpha^*}{2}\right) - \sum_{i=1} e_i^* \left(\frac{\alpha^*}{2}\right) + \\
& \alpha^* \sum_{i=1} E\left(u_i \mid \widehat{\alpha} = \frac{1+\alpha^*}{2}\right) - \alpha^* \sum_{i=1} E\left(u_i \mid \widehat{\alpha} = \frac{\alpha^*}{2}\right),
\end{aligned}
$$

where $e_i^*\left(\widehat{\alpha}\right)$ is a worker's equilibrium effort when his expected value of the employer's altruism equals $\widehat{\alpha}$. By inserting $e_i^* = \beta + \gamma\widehat{\alpha}_j$ into the worker's utility function, his equilibrium utility can be written as

$$
u_i = w_i + \frac{1}{2}\left(\beta + \gamma\widehat{\alpha}_j\right)^2 + \gamma\widehat{\alpha}_j\left(n-1\right)\left[\left(\beta + \gamma\widehat{\alpha}_j\right) - w_i\right],
$$

where n is the total number of workers employed by the employer. Note that workers rationally expect that coworkers hold the same beliefs as themselves (this is not crucial for the results). Existence of an interior solution is ensured if $c(0) > \frac{1}{2}\gamma n$ (so that selfish employers do not wish to create cards) and, following the condition above, that $c(1)$ is not too high (so that the most altruistic type finds it worthwhile to create cards). Note that the marginal costs of signaling are decreasing in α^*, while the marginal benefits are increasing in α^*. The latter is true for two reasons. First, more altruistic employers put more weight on the increase in workers' utility when workers' expectations about the employer's type become more favorable. Second, workers' equilibrium utility is quadratic in α^*.

Appendix B: Treatment Scripts

Control:

"May I bother you for a moment? I just talked to our IT-specialists. They did a brief check on our server. The data have been correctly transmitted from all computers so far. We were previously struggling with some software problems but now everything seems to work fine."

Thank-you card for all (TC All):

"May I bother you for a moment? I just talked to our IT-specialists. They did a brief check on our server. The data have been correctly transmitted from all computers so far. We were previously struggling with some software problems but now everything seems to work fine. We would also like to take the opportunity to thank you in advance for your help. In preparation for this data-entry job, our president, Wolfgang Franz, and we have decided to give everyone this thank-you card as a symbol of our appreciation and thankfulness for your support. [While assigning the cards:] Thank you very much for your commitment!"

Thank-you card for best 3 (TC Best3):

"May I bother you for a moment? I just talked to our IT-specialists. They did a brief check on our server. The data have been correctly transmitted from all computers so far. We were previously struggling with some software problems but now everything seems to work fine. We would also like to take the opportunity to thank you in advance for your help. In preparation for this data-entry job, our president, Wolfgang Franz, and we have decided to give everyone this thank-you card as a symbol of our appreciation and thankfulness for your support. However, Prof. Franz was only able to sign a small number of cards personally. Therefore, we have decided just now to hand these cards to those of you who entered most data so far. Our IT-specialists have told us that logins A, B, and C have transmitted most data so far. This should be Mr. /Mrs. X, Y, and Z, correct? [While assigning the cards:] A special thank to you, Mr./Ms. X, Y, and Z for your commitment."

Thank-you card for best 1 (TC Best1):

"May I bother you for a moment? I just talked to our IT-specialists. They did a brief check on our server. The data have been correctly transmitted from all computers so far. We were previously struggling with some software problems but now everything

seems to work fine. We would also like to take the opportunity to thank you in advance for your help. In preparation for this data-entry job, our president, Wolfgang Franz, and we have decided to give everyone this thank-you card as a symbol of our appreciation and thankfulness for your support. However, Prof. Franz was only able to sign a small number of cards personally. Therefore, we have decided just now to hand the card to that person who entered most data so far. Our IT-specialists have told us that login A has transmitted most data so far. This should be Mr./Mrs. X, correct? [While handing over the card:] A special thank to you, Mr./Ms. X for your commitment."

Figures and Tables

5.1 Figures

Figure 1: Thank-you Card

Figure 2: Improvement in correct entries per minute by treatment (+/- standard error)
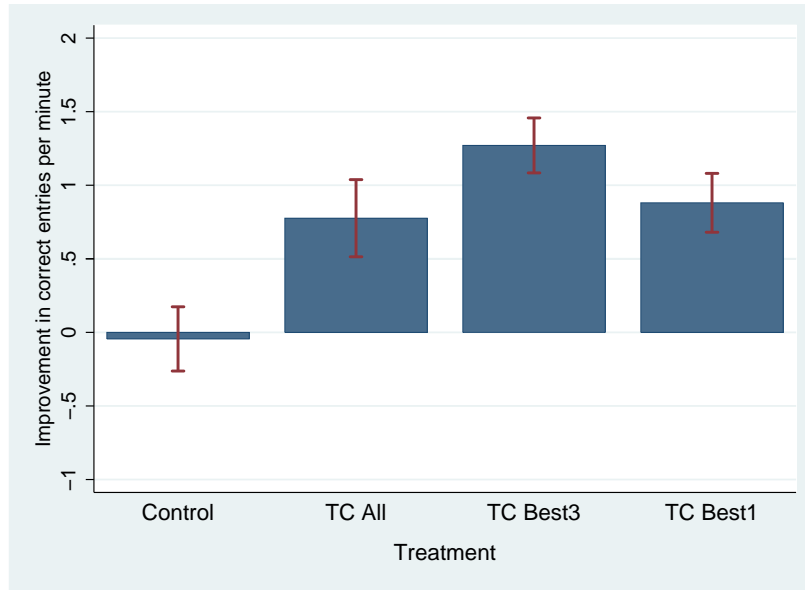
Figure 3: Improvement in correct entries per minute by treatment and performance group (+/- standard error)
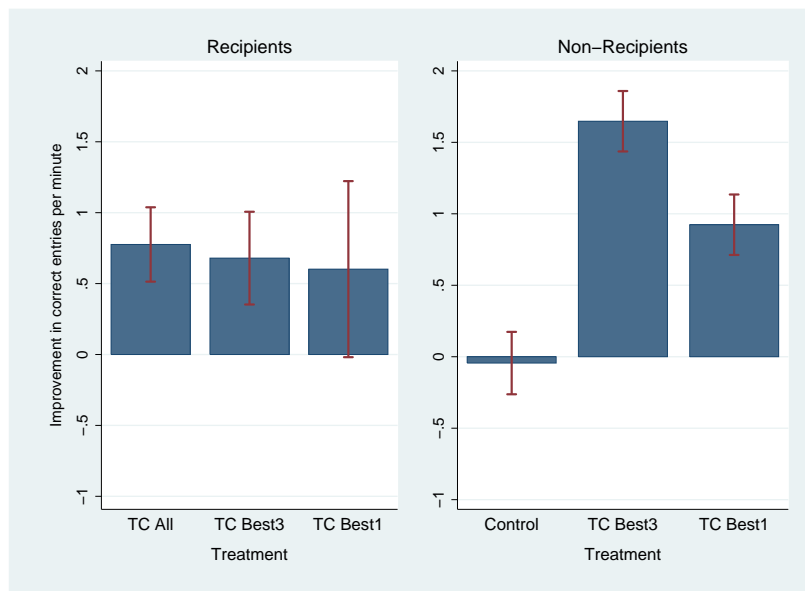
5.2 Tables

Table 1: Theoretical Predictions

	Card for all		Card for best 3		Card for best 1
Reciprocity	R: +	$<$	R: +	$<$	R: +
			N: 0		N: 0
Conditional altruism	+	$>$	+	$>$	+
Conformity	R: 0		R: $-$	$>$	R: $-$
			N: +	$>$	N: +

Note: R denotes recipients of recognition, *N* denotes non-recipients.

Table 2: Summary Statistics by Treatment

	Control	TC All	TC Best 3	TC Best 1
N	81	82	95	82
Female	0.593	0.549	0.558	0.659
	(0.494)	(0.501)	(0.499)	(0.477)
Economics major	0.383	0.317	0.368	0.512*
	(0.489)	(0.468)	(0.485)	(0.503)
Mannheim	0.556	0.427	0.779***	0.744**
	(0.500)	(0.498)	(0.417)	(0.439)
Morning	0.407	0.317	0.389	0.268*
	(0.494)	(0.468)	(0.490)	(0.446)
Afternoon	0.235	0.402**	0.274	0.451***
	(0.426)	(0.493)	(0.448)	(0.501)
Evening	0.358	0.280	0.337	0.280
	(0.482)	(0.452)	(0.475)	(0.452)
Group size	6.852	7.024	7.421***	7.512***
	(0.760)	(0.994)	(0.833)	(0.633)
Duration of working period 1	96.585	102.693***	105.426***	102.248***
	(6.237)	(7.516)	(4.127)	(5.100)
Mannheim university	0.259	0.305	0.411**	0.451**
	(0.441)	(0.463)	(0.495)	(0.501)
Ludwigshafen university	0.210	0.134	0.232	0.146
	(0.410)	(0.343)	(0.424)	(0.356)
Heidelberg university	0.506	0.549	0.305***	0.390
	(0.503)	(0.501)	(0.463)	(0.491)
Baseline performance[1]	15.873	16.628	16.191	16.528
	(3.066)	(3.754)	(4.193)	(3.745)

Note: The table reports means for each group. Standard deviations 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 performance is measured as correct entries per minute in working period 1.

Table 3: Treatment Effects on Productivity in Working Period 2

	I	II	III	IV
TC All	0.839*	0.818*	0.845*	0.828*
	(0.439)	(0.433)	(0.440)	(0.434)
TC Best3	1.272***	1.163***		
	(0.386)	(0.408)		
TC Best1	0.941**	0.882**		
	(0.385)	(0.386)		
TC Best3 - Rec.			0.647	0.515
			(0.518)	(0.535)
TC Best3 - Non-Rec.			1.698***	1.588***
			(0.407)	(0.430)
TC Best1 - Rec.			0.043	-0.046
			(0.761)	(0.757)
TC Best1 - Non-rec.			1.100***	1.046***
			(0.377)	(0.382)
Controls				
Baseline performance	Yes	Yes	Yes	Yes
Demographics		Yes		Yes
Other controls		Yes		Yes
Observations	627	627	627	627
Sessions	91	91	91	91
R^2	0.781	0.784	0.783	0.787

Note: This table reports OLS coefficient estimates (standard errors clustered by session are reported in parentheses). The dependent variable is a worker's performance in the second working period measured by the number of correct database entries per minute. *TC All* represents the treatment where all subjects in a work group received a thank-you card after the first working period. *TC Best3* and *TC Best1* refer to sessions where thank-you cards were assigned to either the best three or the best performing subjects, respectively. Treatment *Control* is omitted and serves as the reference category. All columns include controls for *baseline performance*, its *squared* and *cubic term*, as well as the performance *rank* of participants in working period 1. Column II includes additional controls for *demographics* such as gender, university and field of study as well as *other controls* such as location, time of day, duration of working period 1 and group size. Column III allows for differential treatment effects for recipients (*Rec.*) and non-recipients (*Non-Rec.*) of thank-you cards. Column IV additionally includes all control variables. Note that all regressions also include dummies for and observations from four other treatments, which are reported in Bradler and Neckermann (2013) and Bradler et al. (2013). The data are included to obtain more accurate estimates of the coefficients for the control variables. Significance levels are denoted as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Supplementary Tables – not for publication

Table 4: Simple Diff-in-Diff Estimation of Treatment Effects on the Difference in Productivity between Working Periods 2 and 1

	I	II	III	IV
TC All	0.820*	0.804*	0.820*	0.802*
	(0.474)	(0.462)	(0.475)	(0.463)
TC Best3	1.315***	1.277***		
	(0.385)	(0.411)		
TC Best1	0.925**	0.915**		
	(0.368)	(0.372)		
TC Best3 - Rec.			0.626	0.557
			(0.512)	(0.524)
TC Best3 - Non-Rec.			1.755***	1.730***
			(0.410)	(0.446)
TC Best1 - Rec.			0.431	0.399
			(0.710)	(0.700)
TC Best1 - Non-rec.			1.001***	1.001***
			(0.359)	(0.368)
Controls				
Demographics		Yes		Yes
Other controls		Yes		Yes
Observations	627	627	627	627
Sessions	91	91	91	91
R^2	0.082	0.092	0.096	0.108

Note: This table reports OLS coefficient estimates (standard errors clustered by session are reported in parentheses). The dependent variable is the difference in correct entries per minute between working period 2 and working period 1. *TC All* represents the treatment where all subjects in a work group received a thank-you card after the first working period. *TC Best3* and *TC Best1* refer to sessions where thank-you cards were assigned to either the best three or the best performing subjects, respectively. Treatment *Control* is omitted and serves as the reference category. Column I and II shows overall treatments effects. Column III and IV allows for differential treatment effects for recipients (*Rec.*) and non-recipients (*Non-Rec.*) of thank-you cards. Columns II and IV include addtional controls for *demographics* such as gender, university and field of study as well as *other controls* such as location, time of day, duration of working period 1 and group size. Note that all regressions also include dummies for and observations from four other treatments, which are reported in Bradler and Neckermann (2013) and Bradler et al. (2013). The data are included to obtain more accurate estimates for the coefficients for the control variables. Significance levels are denoted as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Treatment Effects on Total Productivity and Error Rates in Working Period 2

	Total Productivity		Error Rate	
	I	II	III	IV
TC All	0.688	0.697	-0.589**	-0.592**
	(0.432)	(0.432)	(0.262)	(0.264)
TC Best3	1.167***		0.140	
	(0.427)		(0.258)	
TC Best1	0.846**		-0.188	
	(0.402)		(0.204)	
TC Best3 - Rec.		0.425		-0.136
		(0.550)		(0.275)
TC Best3 - Non-Rec.		1.661***		0.305
		(0.446)		(0.318)
TC Best1 - Rec.		-0.109		-0.304
		(0.780)		(0.235)
TC Best1 - Non-rec.		1.021**		-0.169
		(0.396)		(0.219)
Controls				
Baseline performance	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes
Observations	627	627	627	627
Sessions	91	91	91	91
R^2	0.780	0.783	0.623	0.624

Note: This table reports OLS coefficient estimates (standard errors clustered by session are reported in parentheses). The dependent variable is a worker's performance in the second working period measured by the number of correct database entries per minute. *TC All* represents the treatment where all subjects in a work group received a thank-you card after the first working period. *TC Best3* and *TC Best1* refer to sessions where thank-you cards were assigned to either the best three or the best performing subjects, respectively. Treatment *Control* is omitted and serves as the reference category. In columns I and II, the dependent variable is a worker's total number of entries made in the second working period. In columns III and IV, the dependent variable is a worker's error rate in the second working period. The error rate is calculated by dividing total errors by total entries entered. All columns include controls for *baseline performance*, its *squared* and *cubic term*, as well as the performance *rank* of participants in working period 1. *Baseline performance* is measured as the number of entries in working period 1 in columns I and II, and the error rate in working period 1 in columns III and IV, respectively. Further, all regressions include controls for *demographics* such as gender, university and field of study as well as *other controls* such as location, time of day, duration of working period 1 and group size. Column II and IV allow for differential treatment effects for recipients (*Rec.*) and non-recipients (*Non-Rec.*) of thank-you cards. Note that all regressions also include dummies for and observations from four other treatments, which are reported in Bradler and Neckermann (2013) and Bradler et al. (2013). The data are included to obtain more accurate estimates for the coefficients for the control variables. Significance levels are denoted as follows: * p < 0.1, ** p < 0.05, *** p < 0.01.