

IZA DP No. 8311

Employee Recognition and Performance: A Field Experiment

Christiane Bradler
Robert Dur
Susanne Neckermann
Arjan Non

July 2014

Employee Recognition and Performance: A Field Experiment

Christiane Bradler

ZEW Mannheim

Robert Dur

Erasmus University Rotterdam, Tinbergen Institute, CESifo and IZA

Susanne Neckermann

Erasmus University Rotterdam, ZEW and Tinbergen Institute

Arjan Non

Maastricht University

Discussion Paper No. 8311

July 2014

IZA

P.O. Box 7240

53072 Bonn

Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Employee Recognition and Performance: A Field Experiment^{*}

This paper reports the results from a controlled field experiment designed to investigate the causal effect of unannounced, 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. Our results are consistent with workers having a preference for conformity and being reciprocal at the same time.

JEL Classification: C93, M52

Keywords: employee motivation, recognition, reciprocity, conformity, field experiment

Corresponding author:

Robert Dur
Erasmus University Rotterdam
Department of Economics H9-15
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
E-mail: dur@ese.eur.nl

^{*} We gratefully acknowledge comments and suggestions by a Department Editor, an Associate Editor, three anonymous referees, Iwan Barankay, Gary Charness, Tore Ellingsen, Dirk Engelmann, Guido Friebel, David Gill, Michael Kosfeld, Steve Levitt, John List, Michel Maréchal, Dina Pommeranz, Ingrid Rohde, Marie Claire Villeval, and numerous seminar and conference participants. We thank Ann-Kathrin Koessler for excellent research assistance. The experiment has been conducted within the ethical guidelines of our home institutions.

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. Stajkovic and Luthans (2003) provide a meta-analysis of studies conducted in organizations, and report strong positive effects of recognition on employees’ performance. 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, what is the effect of exclusive recognition for the best-performing workers on performance of workers who do not receive recognition?

In this paper we take up these questions by conducting a large-scale field experiment in a controlled work 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 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

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, 2013).

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>.

workers' effort. From the way recognition was provided, workers could deduce that no further thank-you cards 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 stark contrast to Kosfeld and Neckermann (2011), the possible provision of recognition was not announced at the start of the three-hour working period.⁶

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 so far or to the best three performers so far 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 time constraints. This resembles that, in practice, managers are often constrained in the number of people they can praise in a thoughtful, personal way. Lastly, we did our best to keep perceived monitoring constant across all treatments.

Our results are as follows. First, we find that providing public recognition has a positive effect on subsequent performance of workers. When all employees in a work group receive a thank-you card, subsequent performance increases by about 5.2% (or 0.27 standard deviations) as compared to the untreated control. When public recognition is more exclusive, such that the best 3 out of 8 employees in a work group receive a thank-you card, the estimated effect is even bigger and amounts to a 7.3% (or 0.38 standard deviations) performance increase. Remarkably, employees who do *not* receive a thank-you card are mainly responsible for this performance increase. Their performance rises by more than 10% (or 0.52 standard deviations) relative to comparable others in the control group. Recipients of a thank-you card, on the other hand, show a performance increase of only 3.3% (or 0.17 standard deviations). When recognition is even scarcer and limited to the single best-performing employee, we find a slightly weaker average treatment effect of about 5.6% (or 0.29 standard deviations), and a similar pattern with respect to the source of the response, with non-recipients increasing performance by about 6.6% (or 0.34 standard deviations), while recipients' response is close to zero. Hence, we find that unannounced recognition can be a powerful tool to raise performance, particularly of those employees who do not get it!

As we will argue in more detail below, the underlying mechanisms that explain our results best are that workers have a preference for conformity and are reciprocal at the same time. In the treatments with scarce recognition, non-recipients learn that their performance belongs to the bottom of the distribution, implying that it is more likely that they do not conform to the apparent work norm in the group. Consequently,

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 usually remain a surprise.

⁶ Thus, while we use the same symbol (a thank-you card), our study differs fundamentally from Kosfeld and Neckermann (2011). While they use the card as an award in a tournament, which is announced before the start of the work, our study looks at the effect of unannounced recognition provided after some time of work on subsequent work performance. We further differ from Kosfeld and Neckermann (2011) by varying the scarcity of recognition.

they feel a need to increase performance. The smaller positive response of non-recipients in the treatment with only one card as compared to the treatment with three cards is in line with this argument, because the signal of relatively low performance to non-recipients is weaker when recognition is very scarce. The positive response to the treatment with recognition for all workers suggests that reciprocity also plays a role. Reciprocity might further explain why we do not observe that recipients of recognition in treatments with scarce recognition reduce their performance, which we would expect if conformity preferences would be the sole mechanism. Overall, our experimental data suggest that recognition stimulates performance in a workplace setting and that exclusivity can be an effective tool to motivate relatively bad performers to catch up. However, that is only true up to a point: if recognition becomes too scarce, its effectiveness diminishes.

Our study contributes to several literatures. First, we contribute to a substantial body of evidence in the business and psychology literature on the power of employee recognition, in particular Stajkovic and Luthans (2003) and Grant and Gino (2010). A unique feature of our study is that we exogenously vary the scarcity of recognition, ranging from providing recognition to all workers in a work group to recognizing only the best worker. The issue of how exclusive or inclusive recognition should be is of clear relevance for practitioners. Moreover, varying scarcity of recognition allows us to further our understanding of the driving forces of workers' behavior, such as reciprocity and conformity preferences. We also contribute to the growing literature in economics documenting the power of nonfinancial work interventions. Most closely related are the papers looking at nonfinancial gift-exchange, employee awards, and performance feedback. A prominent recent example is Kube et al. (2012), who find that nonfinancial gifts are more effective at increasing worker performance than financial ones. In contrast to that study, we provide recognition that has no material value and vary its scarcity. Moreover, we provide recognition for a job well done after some time of working, while Kube et al. provide gifts at the start of the job. In our treatments with scarce recognition, it is provided conditional on past performance, which is another important departure from most gift-exchange studies. Other recent papers have studied nonfinancial awards as incentives (Kosfeld and Neckermann, 2011; Gubler et al., 2013; Anderson et al., 2013; Ashraf et al., 2013, 2014). We complement this literature by studying recognition that is not announced in advance, but comes as a surprise. Lastly, we contribute to the literature on the effects of relative performance feedback on workers' performance (Blanes i Vidal and Nossol, 2011; Barankay, 2011a,b; Delfgaauw et al., 2013).⁷ A distinguishing feature of our study is that we look at *ex post* effects of coarse 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. Moreover, in our study the relative performance feedback always comes with a clear signal of the employer's appreciation of the employee's performance.

Our main and novel findings are that unannounced recognition in the workplace is most effective for

⁷ Other contexts in which relative performance feedback have been studied include student test scores (Bandiera et al., 2009; Azmat and Iriberry, 2010; Tran and Zeckhauser, 2012), contributions to an online community (Chen et al., 2010), and real-effort in the lab (Charness et al., 2014; Freeman and Gelber, 2010; Kuhn and Tymula, 2012).

intermediate levels of scarcity, and that non-recipients are mainly responsible for the performance increase. These findings have interesting management implications. In practice, managers may worry about making recognition scarce as they fear negative responses from the non-recipients. Our results suggest that this fear is unwarranted. Further, our results suggest that scarce recognition can be a cost-effective tool to motivate the workers at the bottom of the productivity distribution. This finding is particularly informative for organizations who care strongly about the performance of their least productive employees. For instance, in mass production it is often the productivity of the least motivated employee on the assembly line that determines the quality of the end product and the length of the production time. Lastly, our evidence suggests that workers care about conforming to a prevailing work norm and that scarce recognition can be a powerful way to communicate this norm to workers.⁸ Hence, the mechanism that we isolate should be particularly relevant and active whenever workers face some uncertainty about the prevailing work norm. Such situations are commonplace, for instance in organizations with high employee turnover and regularly changing team composition.

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.

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.⁹

Reciprocity of employees First, workers may consider the thank-you card a gift from the employer and feel inclined to reciprocate. A natural way to reciprocate is to increase subsequent work effort. 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.

⁸ Conformity preferences have been found to be important drivers of behavior in lab experiments on work motivation (Danilov and Sliwka, 2013; Bicchieri and Xiao, 2009) as well as in several other contexts, such as energy consumption (Schultz et al., 2007; Nolan et al., 2008), charitable donations (Frey and Meier, 2004; Shang and Croson, 2009; Alpizar et al., 2008), tax compliance (Bobek et al., 2007), and voter turnout (Gerber and Rogers, 2009).

⁹ See the working version of this paper for a formal exposition of the theoretical arguments (Bradler et al., 2013a).

(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. Naturally, 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. Insofar as this holds, recipients should respond the stronger, the smaller the number of cards.

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 reciprocity discussed above, particularly for the response of non-recipients. According to conditional altruism, provision of public recognition should affect effort of recipients and non-recipients identically, as all workers receive the same signal on employer kindness. Effort responses should vary across our treatments as the scarcity of recognition affects perceived kindness. In particular, providing just a few cards may be seen as a weaker signal of kindness than providing all workers with a thank-you-card. Hence, conditional altruism predicts that recognition raises the performance of all workers and that this increase is the stronger the more cards are provided.¹⁰

Conformity preferences of employees Third and last, in both treatments with scarce recognition, preferences for conformity may lead to subsequent performance changes. 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), then this coarse rank information should decrease effort of recipients and increase effort of non-recipients. We expect weaker effects for non-recipients in the treatment where only a single thank-you-card is provided to the best performer than in the treatment where the three best-performing workers receive a thank-you card. The reason is that non-recipients receive more information about their relative performance when they learn that they are among the bottom 5 out of 8 in the group than when they learn that they are not the best in the group. Conversely, recipients of the card

¹⁰ Other models that formalize intentions yield similar predictions, e.g. Falk and Fischbacher (2006) and Gul and Pesendorfer (2010). Alternatively, the thank-you card may affect workers’ beliefs about how important the task is to the employer (as in, for example, Benabou and Tirole, 2003, and Swank and Visser, 2007) or about whether the principal is worth impressing (as in the model of Ellingsen and Johannesson, 2008). These theories yield similar predictions.

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.

Table 1 summarizes our theoretical predictions.

3 Design and Implementation

3.1 Background

In order to estimate the causal effect of recognition on employee performance, we conducted a natural field experiment. We follow the field-experimental methodology as laid out by Harrison and List (2004), Levitt and List (2009), and List and Rasul (2011).¹¹ This means that we observed workers in a controlled work environment and that workers 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 recruited students at different universities in and around Mannheim via student job centers, flyers, and notes on bulletin boards. The announcement informed people that the research center needed short-term student helpers for a three-hour data-entry job and that workers 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 workers that the job was one-time, that a second participation 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 workers 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 employees. In contrast to Kosfeld and Neckermann (2011), provision of

¹¹ We conducted two additional experiments in the same setup. The results are reported in Bradler and Neckermann (2013) and Bradler et al. (2013b). All regressions reported below include observations from and dummies for the treatments in these experiments to obtain more accurate estimates for the coefficients of the control variables. All results are robust to not including these data in the regression analyses.

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

¹³ In case of a second participation, workers 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.

thank-you cards was 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 employees in the work session depending on the treatment. After this interruption, employees worked for another hour.

3.2 Implementation

Upon arrival, workers 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 they 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 workers received a short introduction to their employer (ZEW) and to the project that the surveys stemmed from. The pen-and-paper 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 workers 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 workers, questionnaires were stacked in a high, non-transparent box placed in front of each worker, 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 employees that she was working outside and was available for queries at any time. She also said that they could take breaks whenever necessary. These measures (individual breaks, payment in advance, internet access, and absence of supervisor) were taken to give workers substantial leeway in the amount of time they spend on filling the database. They are not uncommon for this type of job at this research institute. Furthermore, a collective break was avoided to minimize possible group effects and communication between workers.¹⁵ Before the start of the instructions, the research assistant publicly made note about which workstation was used by whom. Thus, workers were aware of the fact that login codes of a workstation could be linked to their person, which should help to hold perceived monitoring constant across treatments.

Eight workers were assigned to each work session. However, since some workers did not show up, the average number of workers 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, employees 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

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 workers and differences in the number of questions that were asked during the introduction. The standard deviation of the duration of working

after employees' official start of work), the research assistant entered the room and told workers that she had just checked the server with the help of the IT-department to ensure that the data were transmitted correctly from each computer to the central data base. She informed workers that this was necessary as there had been server problems in the past.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. After this intervention, employees 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.

3.3 Treatments

As described above, the treatment interventions took place after working period 1. The different scripts can be found in Appendix A. In all treatments, including the control treatment, the research assistant informed workers that the data were transmitted correctly from all computers to the central server, as there previously were some software problems. 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 workers' 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 worker. 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*), workers were informed that the research assistant had only three cards available, and had just 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,

period 1 is 6.6 min. Differences in the lengths of working period 1 are mainly due to delays in the introductory period.

¹⁷ The data transmission could only be checked with the help of the IT department who had access to the server. Therefore, we had to request whether the data transfer was successful in every session, and workers were informed about this to make them aware of the fact that we can observe the data transfer from each computer. As technical problems occurred in the first sessions, we decided to give this information in all later sessions too.

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.

¹⁹ We measured performance in terms of the number of finished questionnaires in working period 1. Workers were told that the IT department had informed the research assistant which three logins had performed best. Hence, in all treatments, including the control, workers knew that the research assistant herself could not directly observe the performance of workers in real-time.

the worker 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 worker.²⁰ This excuse renders it unlikely that workers expect additional cards at the end of working period 2 and makes sure that workers 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 workers who received cards comparable.²¹

Since sessions took place over the course of several months, we cannot rule out that workers in later sessions heard about treatments in earlier sessions. We took two measures to address this concern. First, we conducted most sessions of the more salient treatments (*TC Best3*, *TC Best1*) subsequent to the less salient treatments (*Control*, *TC All*). However, we also had a few later sessions of the less salient treatments (*Control*, *TC All*) to be able to control for time effects. Second, we elicited how people learned about the job in the feedback form at the end of the employment. Only a handful of workers indicated that they had learned about the job from a friend. The exclusion of these workers 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*.²² 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.²³

²⁰ Even though it does not seem to be a lot of work for the president to sign a few more cards, note that in total we employed more than 300 people. Although workers were not aware of the total number of people working for us, they were able to recognize from the recruitment process that several dozens of workers had to be recruited for the job, as they had the choice between at least 3 different dates and between 3 different shifts (morning, afternoon, and evening). Note also that the president of the institute is generally known to be a busy man. For instance, he frequently appears on TV and writes for several newspapers. Moreover, at the time, he was the president of the Council of Academic Advisors to the German government. Therefore, we are confident that employees found the explanation for the scarcity of recognition reasonable.

²¹ 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.

²² 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 observations as these individuals 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).

²³ 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 workers 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).

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 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.²⁴

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

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

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, linearly as well as squared and cubic.²⁵ 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.²⁶ The results of both specifications are similar. As already suggested by Figure 2, recognition has a positive and significant effect on subsequent performance. Recognition which is given to all workers (*TC All*) increases performance by about 0.8 correct entries per minute, which amounts to an increase of 5.2% (or 0.27 standard deviations). Scarce recognition has a larger estimated effect on performance, as a Wald test on the joint significance of *TC Best3* and *TC Best1* as compared to *TC All* ($p=0.03$, one-sided) indicates. *TC Best3* increases performance by 1.2 clicks per minute, which corresponds to a performance increase of 7.3% (or 0.38 standard deviations) relative to the control group. The increase in performance in *TC Best1* is smaller than in *TC Best3*, and amounts to 0.9 correct entries per minute (5.6% or 0.29 standard deviations).

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. Our results also

²⁴ Table B.1 in Appendix B presents the raw means, standard deviations as well as the p-values of the Mann-Whitney rank sum test comparing the recognition treatments with the control treatment.

²⁵ This amounts to a standard diff-in-diff estimation when the coefficient for the linear term is one ($\delta_1 = 1$) and the coefficients of the quadratic and cubic terms are zero ($\delta_2 = \delta_3 = 0$). Our specification is more flexible and allows us to control for mean reversion, for which we find some indication in our data. However, results are very similar if a simple diff-in-diff estimation is applied, that is, regressing treatment dummies on the difference in performance from working period 1 to 2 without controlling for baseline performance. Estimation results are reported in Table B.2 in Appendix B.

²⁶ Our results are robust to the inclusion of further controls for different time effects, such as the day of the week or whether the session took place during the final exam period, during the semester, or during semester break. We also ran specifications including interaction terms of treatments with gender and find no evidence for gender-specific treatment effects.

suggest that recognition works best when it is scarce, but not too scarce, even though it should be noted that the differences in performance between *TC Best3* and *TC All* (Wald test, p-value=0.22, one-sided) as well as between *TC Best3* and *TC Best1* (Wald test, p-value=0.19, one-sided) are not statistically significant.

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 of the patterns in the data. 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 seem to be mainly driven by the non-recipients rather than the recipients.

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’. 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. In *TC Best3*, performance of non-recipients is enhanced by 1.59 clicks per minute which is a 10% increase (or 0.52 standard deviations) compared to the performance of similar others in the control group. In *TC Best1*, the improvement is comparatively lower; non-recipients’ performance increases by 1.05 clicks per minute which amounts to a 6.6% higher performance (or 0.34 standard deviations) compared to the control group. Both these differences are statistically significantly different from zero and economically meaningful. The performance differences between non-recipients in *TC Best3* and *TC Best1* are statistically significant at the 10% level (Wald test: p=0.06, one-sided). Recipients, by contrast, do not show a statistically significant difference in performance in *TC Best3*, and even a slightly lower performance in *TC Best1*, as compared to similar others in control.²⁷ The difference between recipients and non-recipients is significant at the 5% level in *TC Best3* (Wald test: p=0.02, one-sided), and at the 10% level in *TC Best1* (Wald test: p=0.06, one-sided).²⁸

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 also allows us to analyze treatment effects separately for quantity and quality, where quality is measured as the error rate (percentage of incorrect entries). We find no significant treatment effects for quality levels except for a smaller error rate in *TC*

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 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. 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 a small increase in *TC Best3* as compared to control. Hence, it seems likely that employee behavior is also partly driven by reciprocity, which explains these positive level-effects. This reciprocity may either stem from an inclination to return a gift or from signaling, or a combination of the two. The data do not allow us to distinguish between these forms of reciprocity in a meaningful way.

5 Conclusion

In a controlled workplace context, we have shown that unannounced provision of public recognition to employees causes a statistically and economically significant increase in performance. Further, our results suggest 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. However, these differences are not statistically significant. Interestingly, 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 apparently 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 decrease performance, as conformity would suggest. Moreover, we find that recognition for all workers in a group increases performance. A

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), leaving little room for improvement. Treatment effects on total productivity and quality are reported in Table B.3 in Appendix B.

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. (2013b) do not find any evidence for incentive effects of *announced* thank-you cards to the three best performing workers in an identical setup.

natural candidate to reconcile these findings is that, in addition to conformity preferences, workers are also reciprocal and hence increase performance in 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 provision 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 organization. For instance, when organizational performance is particularly determined by the best performing workers, scarce recognition may have limited or even adverse effects. In such cases, providing recognition to all employees is likely to be a better alternative, as doing so is more likely to boost performance of top performers.³⁰

The current experiment raises a number of issues 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.³¹ Interesting follow-up treatments would either only provide relative performance information or use scarce rewards, but hand them out randomly, independent of past performance. Another interesting follow-up study could look at whether the effects persist over time and when recognition is provided repeatedly. Moreover, to assess the robustness of the results, it might be fruitful to replicate the design with more experienced workers and in a setting where other incentives are present. As an example from a different context, Hoogveld and Zubanov (2014) replicate our study in an undergraduate Micro-Economics course. They provide unexpected public recognition in a random subset of tutorial groups to the 30% best students of the group. Their findings correspond to ours, in that recognition has a particularly positive effect on the performance of non-recipients.

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. Such future studies will also be helpful for managers who wonder to what extent the expression of one's gratefulness can be used strategically and repeatedly. We hope that the findings of our paper will encourage

³⁰ Unreported regressions show that the estimated effect of recognition in *TC All* is rather homogeneous across the performance distribution, i.e. the estimated effect size for the three best workers in a work session is comparable to the bottom five. Note that this is somewhat exploratory, as we do not have a clear ex-ante hypothesis on heterogeneous treatment effects.

³¹ 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 find 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 significant performance responses from the recipients in all of their treatments, which is in line with the current study. Ashraf et al. (2014) and Schultz et al. (2007) also disentangle the effects of recognition and feedback in different contexts. Schultz et al. (2007) also find a positive response of individuals with below median performance which is entirely driven by feedback, while Ashraf et al. (2014) report negative responses of underperforming individuals to feedback. Both find positive responses to recognition.

researchers and organizations to join forces and set up field experiments so as to learn about the effects of different forms of recognition in a wide variety of workplace contexts.

References

- Akerlof, G. A., 1982. Labor contracts as partial gift exchange. *Quarterly Journal of Economics* 97 (4), 543–69.
- Alpizar, F., Carlsson, F., Johansson-Stenman, O., 2008. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in costa rica. *Journal of Public Economics* 92(5-6), 1047–1060.
- Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J., 2013. Steering user behavior with badges. In: *Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee*, pp. 95–106.
- Ashraf, N., Bandiera, O., Jack, K., 2013. No margin, no mission? A field experiment on incentives for public service delivery. Unpublished Manuscript, London School of Economics.
- Ashraf, N., Bandiera, O., Lee, S. S., 2014. Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization* 100, 44–63.
- Azmat, G., Iriberry, 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. 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.
- Benabou, R., Tirole, J., 2003. Intrinsic and extrinsic motivation. *The Review of Economic Studies* 70 (3), 489–520.
- Bernheim, B. D., 1994. A theory of conformity. *Journal of Political Economy* 102 (5), 841–877.
- Bicchieri, C., Xiao, E., 2009. Do the right thing: but only if others do so. *Journal of Behavioral Decision Making* 22 (2), 191–208.
- 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.

- Bobek, D. D., Roberts, R. W., Sweeney, J. T., 2007. The social norms of tax compliance: Evidence from Australia, Singapore, and the United States. *Journal of Business Ethics* 74 (1), 49–64.
- Bradler, C., Dur, R., Neckermann, S., Non, A., 2013a. Employee recognition and performance: A field experiment. Centre for European Economic Research Discussion Paper No. 13-017.
- Bradler, C., Dur, R., Neckermann, S., Non, A., 2013b. Reward expectancy and performance: A field experiment. Mimeo.
- Bradler, C., Neckermann, S., 2013. Money meets recognition: A field experiment on worker effort. Mimeo, Centre for European Economic Research, Mannheim.
- 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., 2014. The dark side of competition for status. *Management Science* 60(1), 38–55.
- 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.
- Danilov, A., Sliwka, D., 2013. Can contracts signal social norms? experimental evidence. IZA Discussion Paper.
- 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., 2013. Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics* 31(2), 305–326.
- Dur, R., 2009. Gift exchange in the workplace: Money or attention? *Journal of the European Economic Association* 7 (2-3), 550–560.
- Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: The human side of incentive theory. *American Economic Review* 98 (3), 990—1008.
- Falk, A., Fischbacher, U., 2006. A theory of reciprocity. *Games and Economic Behavior* 54 (2), 293–314.
- 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.
- Frey, B. S., Meier, S., 2004. Social comparisons and pro-social behavior: Testing conditional cooperation in a field experiment. *American Economic Review* 94 (5), 1717–1722.
- Gerber, A. S., Rogers, T., 2009. Descriptive social norms and motivation to vote: everybody's voting and so should you. *The Journal of Politics* 71 (01), 178–191.
- 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.
- Gubler, T., Larkin, I., Pierce, L., 2013. The dirty laundry of employee award programs: Evidence from the field. Harvard Business School NOM Unit Working Paper 13-069.
- Gul, F., Pesendorfer, W., 2010. Interdependent preference models as a theory of intentions. Working Paper, Princeton University.
- 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.
- Hoogveld, N., Zubanov, N., 2014. The power of (no) recognition: Experimental evidence from the university classroom. IZA Discussion Paper.
- Kosfeld, M., Neckermann, S., 2011. Getting more work for nothing? Symbolic awards and worker performance. *American Economic Journal: Microeconomics* 3 (3), 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), 1644–62.
- Kube, S., Maréchal, M. A., Puppe, C., 2013. Do wage cuts damage work morale? Evidence from a natural field experiment. *Journal of the European Economic Association* 11(4), 853–870.
- 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.
- Levitt, S., List, J. A., 2009. Field experiments in economics: The past, the present, and the future. *European Economic Review* 53 (1), 1–18.
- List, J. A., Rasul, I., 2011. Field experiments in labor economics. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, 1st Edition. Vol. 4. Elsevier, pp. 103–228.
- 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.
- Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., Griskevicius, V., 2008. Normative social influence is underdetected. *Personality and social psychology bulletin* 34 (7), 913–923.
- Non, A., 2012. Gift-exchange, incentives, and heterogeneous workers. *Games and Economic Behavior* 75 (1), 319–336.
- Rhoades, L., Eisenberger, R., August 2002. Perceived organizational support: A review of the literature. *Journal of Applied Psychology* 87 (4), 698–714.
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms. *Psychological Science* 18, 429–434.
- Shang, J., Croson, R., 2009. A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal* 119 (540), 1422–1439.
- 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.

Swank, O. H., Visser, B., 2007. Motivating through delegating tasks or giving attention. *Journal of Law, Economics, and Organization* 23(3), 731–742.

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.

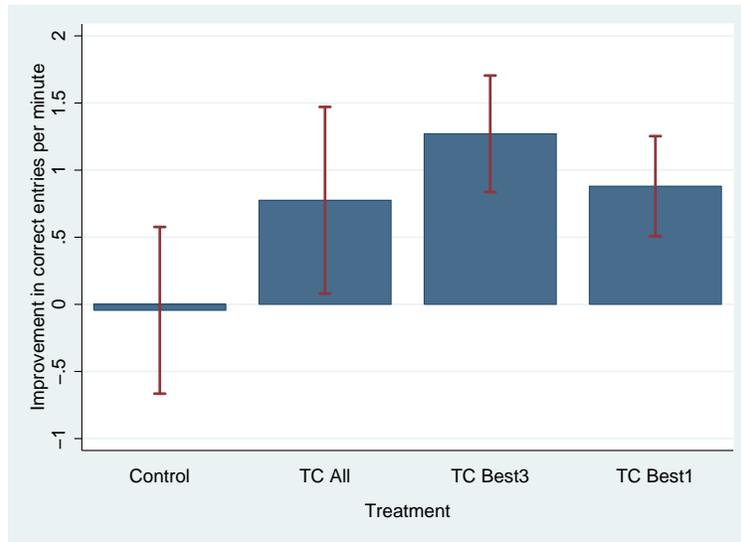
Figures

Figure 1: Thank-you Card



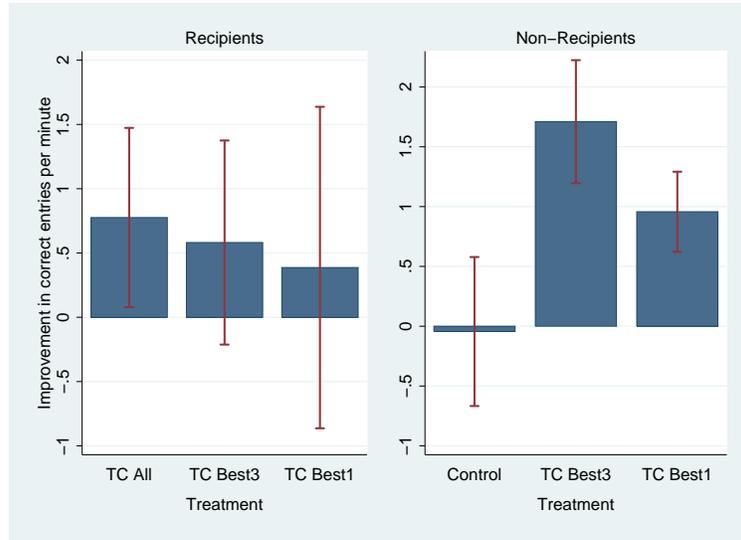
Note: The translation of the text on the card is: "Dear John Public, we would like to thank you cordially for your dedicated work. With your support for this project's comprehensive data-entry, you provide an important basis for our research."

Figure 2: Improvement in correct entries per minute by treatment



Note: The bars indicate the average change in correct entries from working period 1 to 2 by treatment. The whiskers depict the 95% confidence intervals based on cluster-adjusted standard errors (clustered on session level).

Figure 3: Improvement in correct entries per minute by treatment and performance group



Note: The bars indicate the average change in correct entries from working period 1 to 2 by treatment and performance group. The whiskers depict the 95% confidence intervals based on cluster-adjusted standard errors (clustered on session level). *TC All* is included in the left panel as all workers in this treatment receive a card. The *Control* treatment is included in the right panel as nobody receives a card.

Tables

Table 1: Theoretical Predictions

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

Note: R denotes recipients of recognition, N denotes non-recipients. + and - represent the direction of the predicted performance change. In cases where the theories allow predicting the relative magnitude of the effect in comparison to other treatments, this is indicated with the $<$ or $>$ signs. The first line, for example, reads as follows: Reciprocity theory predicts that recipients of thank you cards respond positively. The size of this effect is either equal across all treatments or increases in size the more exclusive the card is, i.e. the lower the number of recipients relative to 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 (0.494)	0.549 (0.501)	0.558 (0.499)	0.659 (0.477)
Economics major	0.383 (0.489)	0.317 (0.468)	0.368 (0.485)	0.512* (0.503)
Mannheim	0.556 (0.500)	0.427 (0.498)	0.779*** (0.417)	0.744** (0.439)
Morning	0.407 (0.494)	0.317 (0.468)	0.389 (0.490)	0.268* (0.446)
Afternoon	0.235 (0.426)	0.402** (0.493)	0.274 (0.448)	0.451*** (0.501)
Evening	0.358 (0.482)	0.280 (0.452)	0.337 (0.475)	0.280 (0.452)
Group size	6.852 (0.760)	7.024 (0.994)	7.421*** (0.833)	7.512*** (0.633)
Duration of working period 1	96.585 (6.237)	102.693*** (7.516)	105.426*** (4.127)	102.248*** (5.100)
Mannheim university	0.259 (0.441)	0.305 (0.463)	0.411** (0.495)	0.451** (0.501)
Ludwigshafen university	0.210 (0.410)	0.134 (0.343)	0.232 (0.424)	0.146 (0.356)
Heidelberg university	0.506 (0.503)	0.549 (0.501)	0.305*** (0.463)	0.390 (0.491)
Baseline performance ¹	15.873 (3.066)	16.628 (3.754)	16.191 (4.193)	16.528 (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* (0.439)	0.818* (0.433)	0.845* (0.440)	0.828* (0.434)
TC Best3	1.272*** (0.386)	1.163*** (0.408)		
TC Best1	0.941** (0.385)	0.882** (0.386)		
TC Best3 - Rec.			0.647 (0.518)	0.515 (0.535)
TC Best3 - Non-Rec.			1.698*** (0.407)	1.588*** (0.430)
TC Best1 - Rec.			0.043 (0.761)	-0.046 (0.757)
TC Best1 - Non-rec.			1.100*** (0.377)	1.046*** (0.382)
Controls				
Baseline performance	Yes	Yes	Yes	Yes
Demographics	No	Yes	No	Yes
Other controls	No	Yes	No	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 workers 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 workers, 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 workers 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. (2013b). The data are included to obtain more accurate estimates of the coefficients for the control variables. Significance levels are denoted as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix A: Treatment Scripts (Original in German)

Control:

"May I bother you for a moment? With the help of our IT department, we did a brief check of the 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? With the help of our IT department, we did a brief check of the 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? With the help of our IT department, we did a brief check of the 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? With the help of our IT department, we did a brief check of the 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."

Appendix B: Supplementary tables

Table B.1: Improvement in correct entries per minute after treatment intervention by treatment

	Mean (sd)	Mann-Whitney rank-sum test vs. Control p-value
Control	-0.044 (1.964)	
TC All	0.776 (2.374)	0.006***
TC Best3	1.271 (1.820)	0.000***
TC Best1	0.880 (1.811)	0.005***

Note: Standard deviations are reported in parentheses. Significance levels are denoted as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: 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* (0.474)	0.804* (0.462)	0.820* (0.475)	0.802* (0.463)
TC Best3	1.315*** (0.385)	1.277*** (0.411)		
TC Best1	0.925** (0.368)	0.915** (0.372)		
TC Best3 - Rec.			0.626 (0.512)	0.557 (0.524)
TC Best3 - Non-Rec.			1.755*** (0.410)	1.730*** (0.446)
TC Best1 - Rec.			0.431 (0.710)	0.399 (0.700)
TC Best1 - Non-rec.			1.001*** (0.359)	1.001*** (0.368)
Controls				
Demographics	No	Yes	No	Yes
Other controls	No	Yes	No	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 workers 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 workers, 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 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. 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. (2013b). The data are included to obtain more accurate estimates for the coefficients of the control variables. Significance levels are denoted as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Treatment Effects on Total Productivity and Error Rates in Working Period 2

	Total Productivity		Error Rate	
	I	II	III	IV
TC All	0.688 (0.432)	0.697 (0.432)	-0.589** (0.262)	-0.592** (0.264)
TC Best3	1.167*** (0.427)		0.140 (0.258)	
TC Best1	0.846** (0.402)		-0.188 (0.204)	
TC Best3 - Rec.		0.425 (0.550)		-0.136 (0.275)
TC Best3 - Non-Rec.		1.661*** (0.446)		0.305 (0.318)
TC Best1 - Rec.		-0.109 (0.780)		-0.304 (0.235)
TC Best1 - Non-rec.		1.021** (0.396)		-0.169 (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 workers 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 workers, 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 workers 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. (2013b). The data are included to obtain more accurate estimates for the coefficients for the control variables. Significance levels are denoted as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.