Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Employee Recognition and Performance: A Field Experiment

Christiane Bradler, Robert Dur, Susanne Neckermann, Arjan Non

To cite this article:
Christiane Bradler, Robert Dur, Susanne Neckermann, Arjan Non (2016) Employee Recognition and Performance: A Field Experiment. Management Science 62(11):3085-3099. http://dx.doi.org/10.1287/mnsc.2015.2291

MANAGEMENT SCIENCE
Vol. 62, No. 11, November 2016, pp. 3085–3099
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

https://doi.org/10.1287/mnsc.2015.2291

Employee Recognition and Performance: A Field Experiment

Christiane Bradler
Centre for European Economic Research (ZEW), 68161 Mannheim, Germany, christianebradler@yahoo.de

Robert Dur
Erasmus University Rotterdam, 3000 DR Rotterdam, Netherlands; Tinbergen Institute, 3000 DR Rotterdam, Netherlands; Munich Society for the Promotion of Economic Research (CESifo), 81679 München, Germany; and Institute for the Study of Labor (IZA), 53113 Bonn, Germany, dur@ese.eur.nl

Susanne Neckermann
Erasmus University Rotterdam, 3000 DR Rotterdam, Netherlands; Centre for European Economic Research (ZEW), 68161 Mannheim, German; Tinbergen Institute, 3000 DR Rotterdam, Netherlands, neckermann@ese.eur.nl

Arjan Non
Maastricht University, Research Centre for Education and the Labour Market (ROA), 6200 MD Maastricht, Netherlands; and Munich Society for the Promotion of Economic Research (CESifo), 81679 München, Germany, j.non@maastrichtuniversity.nl

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 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.

Data, as supplemental material, are available at https://doi.org/10.1287/mnsc.2015.2291.

Keywords: employee motivation; recognition; reciprocity; conformity; field experiment
History: Received May 30, 2013; accepted March 9, 2015, by John List, behavioral economics. Published online in *Articles in Advance* February 9, 2016.

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 *1,001 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" (p. 1).[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]

Although 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 address these questions by conducting a large-scale field experiment in a controlled

[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" (p. 1).

[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), which shows 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, for example, Rhoades and Eisenberger (2002), Cropanzano and Mitchell (2005), and Wagner and Harter (2006).

work environment. During the time period from 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 took 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 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 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. Furthermore, we make clear that the management recognizes the work input from all workers, but can only recognize a subset of the employees explicitly because of 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 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 three out of eight 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 nonrecipients increasing performance by about 6.6% (or 0.34 standard deviations), whereas recipients' response is close to zero. Hence, we find that unannounced recognition can be a powerful tool to raise performance, particularly for 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, nonrecipients 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, they feel a need to increase performance. The smaller positive response of nonrecipients in the treatment with only one card as compared to the treatment with three cards is in line with this

[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; Kube et al. 2012, 2013). The short-term nature of the job enables us to minimize possible career concerns, which greatly facilitates the interpretation of the results. Note that temporary workers are quite prevalent nowadays and "represent an interesting middling ground between subjects in one-shot experiments and full-time employees in firms" (Al-Ubaydli et al. 2015, p. 540).

[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, accessed December 31, 2015).

[5] This is often true in practice, even in organizations with a culture for praise. Although recognition in general might be expected in such organizations, the particular situation, timing, and kind of recognition usually remain a surprise. This is not without reason: Rogers and Frey (2015) argue that habituation to stimuli (such as recognition) is less likely if the stimulus is presented at unpredictable time intervals and is not presented in the same way every time.

[6] Thus, although we use the same symbol (a thank-you card), our study differs fundamentally from Kosfeld and Neckermann (2011). Whereas 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.

argument, because the signal of relatively low performance to nonrecipients 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 were the sole mechanism. Following this interpretation, for recipients the negative effect generated by the preference for conformity is neutralized by the positive effect generated by positive reciprocity concerns. 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, whereas Kube et al. (2012) provide gifts at the start of the job.[7] 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; Anderson et al. 2013; Ashraf et al. 2014a, b; Bareket-Bojmel et al. 2016; Gubler et al. 2016). 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 (Barankay 2011a, b; Blanes i Vidal and Nossol 2011; Delfgaauw et al. 2013).[8] 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 intermediate levels of scarcity, and that nonrecipients are mainly responsible for the performance increase. These findings have interesting management implications. In practice, managers may worry about making recognition scarce because they fear negative responses from the nonrecipients. Our results suggest that this fear is unwarranted. Furthermore, 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 that 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.[9] 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

[7] Ockenfels et al. (2015) also have a treatment in which they hand out gifts after some time. They study how the timing of financial gifts affects performance and find that, holding the total wage bill constant, workers' performance is about 11% higher when the wage is increased in two steps as opposed to a single increase at the outset.

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

[9] Conformity preferences have been found to be important drivers of behavior in lab experiments on work motivation (Bicchieri and Xiao 2009, Danilov and Sliwka 2016) as well as in several other contexts, such as energy consumption (Schultz et al. 2007, Nolan et al. 2008, Allcott and Rogers 2014), charitable donations (Frey and Meier 2004, Alpizar et al. 2008, Shang and Croson 2009), tax compliance (Bobek et al. 2007), and voter turnout (Gerber and Rogers 2009).

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 §3. Section 4 lays out the empirical strategy and presents the results of our experiment. Finally, §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.[10]

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 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. (2012), who show that nonpecuniary 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 more strongly when the number of cards is smaller.

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 nonrecipients. According to conditional altruism, provision of public recognition should affect effort of recipients and nonrecipients 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 stronger the more cards are provided.[11]

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, nonrecipients learn that they belong to the lower end of the performance distribution. If workers have a preference for conformity (e.g., 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 nonrecipients. We expect weaker effects for nonrecipients 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 nonrecipients receive more information about their relative performance when they learn that they are among the bottom five out of eight in the group than when they learn that they are not the best in the group. 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. Table 1 summarizes our theoretical predictions.

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

[11] Other models that formalize intentions yield similar predictions, for example, Falk and Fischbacher (2006) and Gul and Pesendorfer (2016). Alternatively, the thank-you card may affect workers' beliefs about how important the task is to the employer (as in, e.g., Benabou and Tirole 2003, 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.

Table 1 Theoretical Predictions

	Card for all		Card for best 3		Card for best 1
Reciprocity	R: +	≤	R: + N: 0	≤	R: + N: 0
Conditional altruism	+	>	+	>	+
Conformity	R: 0		R: − N: +	> >	R: − N: +

Notes. Recipients of recognition are denoted by R, and N denotes nonrecipients. The + and − signs represent the direction of the predicted performance change. The < and > signs indicate cases where the theories allow predicting the relative magnitude of the effect in comparison to other treatments. 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 nonrecipients.

3. Design and Implementation

3.1. Background

To estimate the causal effect of recognition on employee performance, we conducted a natural field experiment following the methodology laid out by Harrison and List (2004), Levitt and List (2009), and List and Rasul (2011).[12] This means that we observed workers in a controlled work environment and that workers were not aware 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 hundred completed 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.[13] Students could sign up for the job online and were subsequently contacted by phone to arrange the 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,[14] 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 locations and three times of day (morning, noon, and afternoon).[15] 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 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. Because 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 (e.g., peer effects as in Falk and Ichino 2006), questionnaires were stacked in a high, nontransparent 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

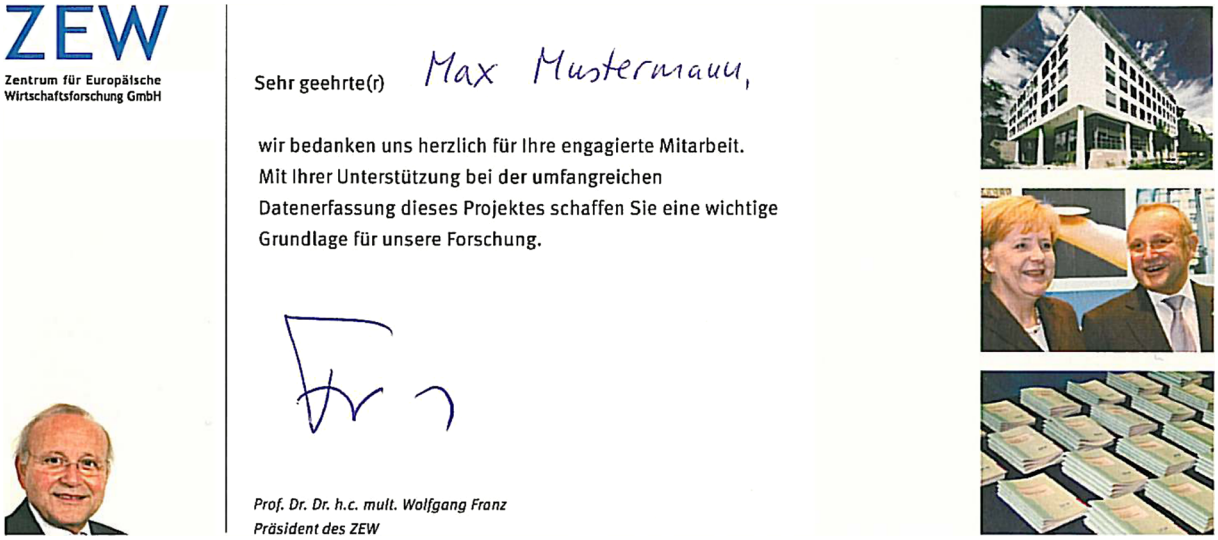
[12] We conducted two additional experiments in the same setup. The results are reported in Bradler and Neckermann (2013) and Bradler et al. (2013b). All regressions reported below include observations from and dummies for the treatments in these experiments to obtain more accurate estimates for the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

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

[14] 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.

[15] 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 location effects, but dropping these controls does not affect the results. All sessions were conducted by the same research assistant.

Figure 1 (Color online) Thank-You Card



Notes. 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." (Photos on the card by Dichiser/ZEW.)

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 spent 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.[16] Before the start of the instructions, the research assistant publicly made note about which workstation was used by whom. Thus, workers were aware that the 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 (standard deviation of 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.[17] Subsequent to working period 1 (i.e., two hours 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 database. She informed workers that this was necessary as there had been server problems in the past.[18] 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.[19] 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.

[16] 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.

[17] 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 period 1 is 6.6 minutes. Differences in the lengths of working period 1 are mainly due to delays in the introductory period.

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

[19] The mean length 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 A. In all treatments, including the control treatment, the research assistant informed workers that there had been some software problems in the past, but that in this session all the data were transmitted correctly from all computers to the central server. Nothing else happened in the *Control* treatment, so 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 Best*3), workers were informed that the research assistant had only a limited number of cards available, and that it had just been decided to hand these out to those three persons who had performed best until then.[20] The treatment "Thank-you card for the best 1" (*TC Best*1) was identical to *TC Best*3 except only the one worker who performed best in the first working period received a thank-you card.

The scarcity of cards was explained by saying that the president did not manage to sign a card for each worker.[21] This excuse renders it unlikely that workers expect additional cards at the end of working period 2 and ensures that workers knew that there was a practical reason for making recognition scarce. Because 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.[22]

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 Best*3, *TC Best*1) 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 Best*3, and 82 in *TC Best*1.[23] There are statistically significant differences between treatments in some of the observables. However, as shown in Table 3, 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.[24]

We first examine the main treatment effects. Figure 2 gives a first impression. It shows the average

[20] 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.

[21] 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 three dates and three 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.

[22] 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 Best*3. We control for group size in all regressions. The results are not sensitive to this.

[23] In total, 363 people participated in 48 sessions. We dropped four sessions (two of *Control* and two of *TC All*) from the analysis because of severe server breakdowns during working time. Moreover, we excluded two observations as these individuals were hardly able to work on the task because of 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).

[24] 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 (standard deviation of 10.8).

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)
Location 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)
Student in Mannheim	0.259	0.305	0.411**	0.451**
	(0.441)	(0.463)	(0.495)	(0.501)
Student in Ludwigshafen	0.210	0.134	0.232	0.146
	(0.410)	(0.343)	(0.424)	(0.356)
Student in Heidelberg	0.506	0.549	0.305***	0.390
	(0.503)	(0.501)	(0.463)	(0.491)
Baseline performance[a]	15.877	16.628	16.191	16.528
	(3.067)	(3.754)	(4.193)	(3.745)

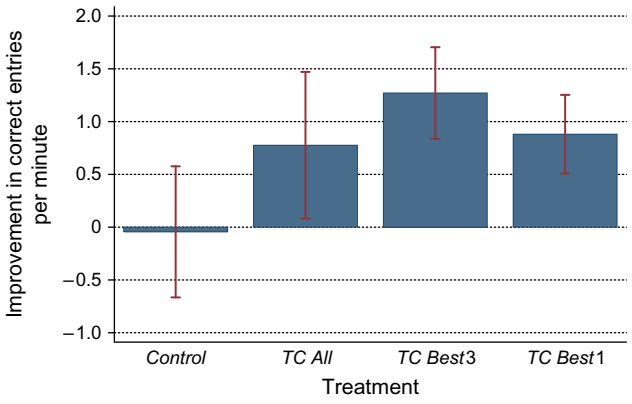
Notes. The table reports means for each group. Standard deviations are displayed in parentheses.

[a]Baseline performance is measured as correct entries per minute in working period 1.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. Significance levels indicate a difference of means (compared to the control group).

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 Best*3. Also, performance improves slightly more in *TC Best*1 than in *TC All*, suggesting that recognition motivates and that exclusive recognition works better than praising everyone.[25]

Figure 2 (Color online) Improvement in Correct Entries per Minute by Treatment



Notes. 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).

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.[26] 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 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.[27] The results of both specifications are similar. As suggested by Figure 2, recognition has a positive and significant effect on subsequent performance. Recognition that is given to all workers (*TC All*) increases performance by about 0.8 correct

[25] 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.

[26] By controlling for baseline performance in a linear, quadratic, and cubic form, our specification is more flexible than a simple diff-in-diff estimation and allows us to control for mean reversion, for which we find some indication in our data. A standard way to detect mean reversion is to run a simple linear regression for the control group where the standardized performance in period 2 is regressed on the standardized performance in period 1 (Gelman and Hill 2006). This yields a significant coefficient of the standardized performance in period 1 of 0.85, which indicates some mean reversion. However, since the magnitude of mean reversion is low, we find very similar results 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.

[27] 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.

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 Best*3	1.272***	1.163***		
	(0.386)	(0.408)		
*TC Best*1	0.941**	0.882**		
	(0.385)	(0.386)		
*TC Best*3—*Rec.*			0.647	0.515
			(0.518)	(0.535)
*TC Best*3—*Nonrec.*			1.698***	1.588***
			(0.407)	(0.430)
*TC Best*1—*Rec.*			0.043	−0.046
			(0.761)	(0.757)
*TC Best*1—*Nonrec.*			1.100***	1.046***
			(0.377)	(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	0.776	0.780	0.779	0.783

Notes. This table reports ordinary least squares (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 Best*3 and *TC Best*1 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 nonrecipients (*Nonrec.*) 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.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

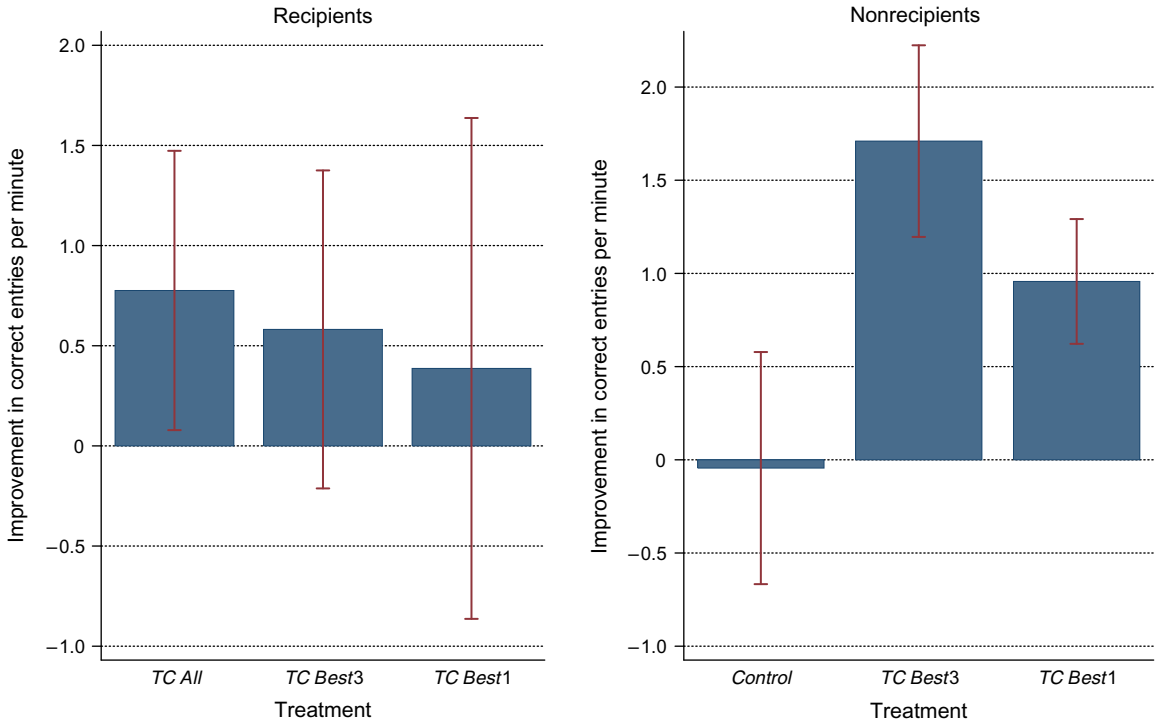
entries per minute, which amounts to an increase of 5.2% (or 0.27 standard deviations). Scarce recognition has a larger estimated effect on performance, as a Wald test on the joint significance of *TC Best*3 and *TC Best*1 as compared to *TC All* ($p = 0.06$, two-sided) indicates. *TC Best*3 increases 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 Best*1 is smaller than in *TC Best*3, 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 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 Best*3 and *TC All* (Wald test, p-value = 0.44, two-sided) as well as between *TC Best*3 and *TC Best*1 (Wald test, p-value = 0.37, two-sided) are not statistically significant.

To shed light on underlying mechanisms and to assess the relevance of the different theories discussed in §2, we next look at recipients and nonrecipients of recognition separately. Recall that the theories have distinct predictions regarding the responses of recipients and nonrecipients (see Table 1). Figure 3 gives a first impression of the patterns in the data. The figure shows performance improvements separately for recipients and nonrecipients. 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 Best*1 because of the small number of observations in this cell. Interestingly, in all treatments where the research assistant was forced to differentiate between employees, the nonrecipients subsequently improve by more than the recipients do. The difference is most pronounced for the *TC Best*3 treatment, where the improvement of nonrecipients 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 nonrecipients rather than the recipients.

An obvious point of criticism on these simple comparisons between recipients and nonrecipients 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 nonrecipients in each treatment. The results confirm the descriptive analysis in Figure 3: positive treatment effects are largely driven by nonrecipients. In *TC Best*3, performance of nonrecipients 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 Best*1, the improvement is comparatively lower; nonrecipients' 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 of these differences are statistically significantly different from zero and economically meaningful. The performance differences between nonrecipients in *TC Best*3 and *TC Best*1 are

Figure 3 **(Color online) Improvement in Correct Entries per Minute by Treatment and Performance Group**



Notes. 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 because all workers in this treatment receive a card. The *Control* treatment is included in the right panel because nobody receives a card.

close to significant at the 10% level (Wald test: $p = 0.11$, two-sided). Recipients, by contrast, do not show a statistically significant difference in performance in *TC Best*3, and even a slightly lower performance in *TC Best*1, as compared to similar others in control.[28] The difference between recipients and nonrecipients is significant at the 5% level in *TC Best*3 (Wald test: $p = 0.03$, two-sided), and marginally significant in *TC Best*1 (Wald test: $p = 0.11$, two-sided).[29]

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 §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 nonrecipients is well in line with conformity theory.[30] Moreover, the fact that the improvement in productivity of nonrecipients is larger in *TC Best*3 than in *TC Best*1 nicely fits with the predictions of conformity preferences. At the same time, however, conformity theory cannot explain the increase in performance in *TC All*. Furthermore, in isolation, conformity theory would predict that recipients of scarce recognition reduce their performance, although we find a small increase in *TC Best*3 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

[28] 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 three performers in the first working period) do increase their performance significantly and substantially subsequent to receiving a combination of recognition and money.

[29] The data also allow 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 All* (significant at the 10% 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.

[30] 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 nonrecipients that we observe here is caused by nonrecipients' expectation of and striving for another round of cards at the end of working period 2. A further reason to be skeptical 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.

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. Furthermore, 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 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 nonrecipients. 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, nonrecipients apparently feel inclined to improve performance so as to adhere to the apparent work norm. In line with this interpretation, we find weaker (but still positive) responses of nonrecipients 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 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 provided to a relatively large group of best-performing workers. 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, since doing so is more likely to boost performance of top performers.[31]

The current experiment raises a number of issues that remain open for future research. Since the performance increase was primarily driven by nonrecipients, a similar effect might be reached at lower cost with pure relative performance information.[32] Interesting follow-up treatments would either only provide relative performance information or use scarce rewards, but hand them out randomly, independent of past performance. Furthermore, in our experiment we justified the scarcity of recognition by pointing to managerial time constraints. It would be interesting to see whether behavioral patterns change when it is stated that recognition is made scarce deliberately, i.e., when the manager states that she wishes to provide recognition only to the best-performing workers. This may dramatically change the signal provided through scarce recognition about the manager's characteristics. Moreover, to assess the robustness of the results, it would be fruitful to replicate the design with more experienced workers and in a setting where other incentives are present. Stajkovic and Luthans (2003), Ashraf et al. (2014a, b), Neckermann et al. (2014), and Bareket-Bojmel et al. (2016) report positive effects of recognition in such settings. Interestingly, like us, Bareket-Bojmel et al. (2016) find evidence for a reciprocal response of employees to the provision of thank-you cards. As an example from a different context, Hoogveld and Zubanov (2014) replicate our study in an undergraduate microeconomics 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 nonrecipients.

Another interesting follow-up study could look at how long the effects persist and whether the effects

[31] 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 estimated effect size for the bottom five.

[32] 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 nonrecipients increase their effort similarly in all treatments. This suggests that the information per se rather than the rewards drives the response of nonrecipients. Interestingly, the authors do not find significant performance responses from the recipients in any of their treatments, which is in line with the current study. Schultz et al. (2007) and Ashraf et al. (2014b) also disentangle the effects of recognition and feedback in different contexts. Schultz et al. (2007) find a positive response of individuals with below-median performance that is entirely driven by feedback, whereas Ashraf et al. (2014a) report negative responses of underperforming individuals to feedback. Both find positive responses to recognition.

remain of similar size when recognition is provided repeatedly. There are two reasons why we think this is particularly relevant in our context. First, we study recognition that comes as a surprise, which raises the issue whether the surprise element can be maintained. Clearly, when recognition is provided repeatedly in the exact same manner, the effect may soon disappear. Such recognition is no longer personal and thoughtful, and workers will most likely start to doubt the manager's motives. Managers therefore need to think about different ways of recognizing workers to ensure that their recognition remains effective.[33] In line with this, Rogers and Frey (2015) argue that treatment effects of stimuli are generally more likely to show up repeatedly when a surprise element is maintained, for instance when the treatment is provided irregularly, when the time intervals between repetitions are long, and when the treatment is not presented in the same way every time (see also McSweeney 2004). Second, our findings suggest that scarce recognition is particularly effective because it provides information on the work norm. Can we expect workers to react to this information when the same information is provided repeatedly? Recent evidence in the context of energy conservation suggests that regular information provision can be remarkably powerful even after many repetitions. Allcott and Rogers (2014) show, using a field experiment, that consumers continue to respond to monthly feedback on their energy consumption after more than two years of treatment. This effect comes on top of the posttreatment effect: consumers who are discontinued from treatment after two years continue to consume less energy than the control group, but consume more than the group that still receives feedback. Apparently, feedback is either still informative after two years, or people need frequent reminders of what they know already.

More generally, it would be helpful if future studies provide further evidence about the effects of different forms of recognition in different circumstances. Given the wide variety in those two dimensions, the results of a single study on the effects of recognition, by their nature, may not generalize far beyond the specific setting in which they are found.[34] This underlines the need for future studies. We hope that the findings of our paper will encourage 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.

[33] The success of books such as *1001 Ways to Reward Employees* (Nelson 2005) suggests that managers acknowledge this need.

[34] See Al-Ubaydli and List (2015) for an insightful discussion on the generalizability of experimental results in economics.

Supplemental Material

Supplemental material to this paper is available at https://doi.org/10.1287/mnsc.2015.2291.

Acknowledgments

The authors gratefully acknowledge comments and suggestions by the 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. The authors thank Ann-Kathrin Koessler for excellent research assistance. The experiment has been conducted within the ethical guidelines of our home institutions. A. Non gratefully acknowledges funding from the Netherlands Organisation for Scientific Research (NWO) [Project 452-10-006].

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 Best*1):
"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	Mann–Whitney rank-sum test vs. control
Control	−0.044	
	(1.964)	
TC All	0.776	0.006***
	(2.374)	
*TC Best*3	1.271	0.000***
	(1.820)	
*TC Best*1	0.880	0.005***
	(1.811)	

Notes. Standard deviations are reported in parentheses.
***$p < 0.01$.

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*	0.804*	0.820*	0.802*
	(0.474)	(0.462)	(0.475)	(0.463)
*TC Best*3	1.315***	1.277***		
	(0.385)	(0.411)		
*TC Best*1	0.925**	0.915**		
	(0.368)	(0.372)		
*TC Best*3—*Rec.*			0.626	0.557
			(0.512)	(0.524)
*TC Best*3—*Nonrec.*			1.755***	1.730***
			(0.410)	(0.446)
*TC Best*1—*Rec.*			0.431	0.399
			(0.710)	(0.700)
*TC Best*1—*Nonrec.*			1.001***	1.001***
			(0.359)	(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

Notes. 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 Best*3 and *TC Best*1 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. Columns I and II show overall treatments effects. Columns III and IV allow for differential treatment effects for recipients (*Rec.*) and nonrecipients (*Nonrec.*) 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.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

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	0.697	−0.589**	−0.592**
	(0.432)	(0.432)	(0.262)	(0.264)
*TC Best*3	1.167***		0.140	
	(0.427)		(0.258)	
*TC Best*1	0.846**		−0.188	
	(0.402)		(0.204)	
*TC Best*3—*Rec.*		0.425		−0.136
		(0.550)		(0.275)
*TC Best*3—*Nonrec.*		1.661***		0.305
		(0.446)		(0.318)
*TC Best*1—*Rec.*		−0.109		−0.304
		(0.780)		(0.235)
*TC Best*1—*Nonrec.*		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

Table B.3 (Continued)

	Total productivity		Error rate	
	I	II	III	IV
Observations	627	627	627	627
Sessions	91	91	91	91
R^2	0.780	0.783	0.475	0.478

Notes. 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. *TC All* represents the treatment where all workers in a work group received a thank-you card after the first working period. *TC Best*3 and *TC Best*1 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. Furthermore, 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 nonrecipients (*NonRec.*) 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.

$^{**}p < 0.05$; $^{***}p < 0.01$.

References

Akerlof GA (1982) Labor contracts as partial gift exchange. *Quart. J. Econom.* 97(4):543–569.

Al-Ubaydli O, List JA (2015) On the generalizability of experimental results in economics: With a response to Camerer. Fréchette GR, Schotter A, eds. *Handbook of Experimental Economic Methodology*, Chap. 20 (Oxford University Press, New York).

Al-Ubaydli O, Andersen S, Gneezy U, List JA (2015) Carrots that look like sticks: Toward an understanding of multitasking incentive schemes. *Southern Econom. J.* 81(3):538–561.

Allcott H, Rogers T (2014) The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *Amer. Econom. Rev.* 104(10):3003–3037.

Alpizar F, Carlsson F, Johansson-Stenman O (2008) Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *J. Public Econom.* 92(5–6):1047–1060.

Anderson A, Huttenlocher D, Kleinberg J, Leskovec J (2013) Steering user behavior with badges. *Proc. 22nd Internat. Conf. World Wide Web* (International World Wide Web Conferences Steering Committee, Geneva), 95–106.

Ashraf N, Bandiera O, Jack K (2014a) No margin, no mission? A field experiment on incentives for public service delivery. *J. Public Econom.* 120:1–17.

Ashraf N, Bandiera O, Lee SS (2014b) Awards unbundled: Evidence from a natural field experiment. *J. Econom. Behav. Organ.* 100:44–63.

Azmat G, Iriberri N (2010) The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *J. Public Econom.* 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, London.

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

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

Bareket-Bojmel L, Hochman G, Ariely D (2016) It's (not) all about the Jacksons—Testing different types of short-term bonuses in the field. *J. Management*. Forthcoming.

Benabou R, Tirole J (2003) Intrinsic and extrinsic motivation. *Rev. Econom. Stud.* 70(3):489–520.

Bernheim BD (1994) A theory of conformity. *J. Political Econom.* 102(5):841–877.

Bicchieri C, Xiao E (2009) Do the right thing: But only if others do so. *J. Behavioral Decision Making* 22(2):191–208.

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

Blau PM (1964) *Exchange and Power in Social Life* (John Wiley & Sons, New York).

Bobek DD, Roberts RW, Sweeney JT (2007) The social norms of tax compliance: Evidence from Australia, Singapore, and the United States. *J. Bus. Ethics* 74(1):49–64.

Bradler C, Neckermann S (2013) Money meets gratitude: A field experiment on worker effort. Mimeo, Centre for European Economic Research (ZEW), Mannheim, Germany.

Bradler C, Dur R, Neckermann S, Non A (2013a) Employee recognition and performance: A field experiment. ZEW Discussion Paper 13-017, Centre for European Economic Research, Mannheim, Germany.

Bradler C, Dur R, Neckermann S, Non A (2013b) The power of surprise: The impact of spontaneous and announced recognition on employee performance. Mimeo, Centre for European Economic Research (ZEW), Mannheim, Germany.

Cameron J, Pierce W (1994) Reinforcement, reward, and intrinsic motivation—A meta-analysis. *Rev. Educational Res.* 64(3): 363–423.

Charness G, Masclet D, Villeval M-C (2014) The dark side of competition for status. *Management Sci.* 60(1):38–55.

Chen Y, Harper FM, Konstan J, Li SX (2010) Social comparisons and contributions to online communities: A field experiment on movielens. *Amer. Econom. Rev.* 100(4):1358–1398.

Cropanzano R, Mitchell MS (2005) Social exchange theory: An interdisciplinary review. *J. Management* 31(6):874–900.

Danilov A, Sliwka D (2016) Can contracts signal social norms? Experimental evidence. *Management Sci.*, ePub ahead of print February 29, http://dx.doi.org/10.1287/mnsc.2015.2336.

Deci EL, Ryan RM (1971) Effects of externally mediated rewards on intrinsic motivation. *J. Personality Soc. Psych.* 18(1):105–115.

Delfgaauw J, Dur R, Sol J, Verbeke W (2013) Tournament incentives in the field: Gender differences in the workplace. *J. Labor Econom.* 31(2):305–326.

Dur R (2009) Gift exchange in the workplace: Money or attention? *J. Eur. Econom. Assoc.* 7(2–3):550–560.

Ellingsen T, Johannesson M (2008) Pride and prejudice: The human side of incentive theory. *Amer. Econom. Rev.* 98(3):990–1008.

Falk A, Fischbacher U (2006) A theory of reciprocity. *Games Econom. Behav.* 54(2):293–314.

Falk A, Ichino A (2006) Clean evidence on peer effects. *J. Labor Econom.* 24(1):39–58.

Fehr E, Kirchsteiger G, Riedl A (1993) Does fairness prevent market clearing? An experimental investigation. *Quart. J. Econom.* 108(2):437–59.

Fischer P, Huddart S (2008) Optimal contracting with endogenous social norms. *Amer. Econom. Rev.* 98(4):1459–1475.

Fisher RJ, Ackerman D (1998) The effects of recognition and group need on volunteerism: A social norm perspective. *J. Consumer Res.* 25(3):262–275.

Freeman RB, Gelber AM (2010) Prize structure and information in tournaments: Experimental evidence. *Amer. Econom. J.: Appl. Econom.* 2(1):149–164.

Frey BS, Meier S (2004) Social comparisons and pro-social behavior: Testing conditional cooperation in a field experiment. *Amer. Econom. Rev.* 94(5):1717–1722.

Gelman A, Hill J (2006) *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge University Press, Cambridge, UK).

Gerber AS, Rogers T (2009) Descriptive social norms and motivation to vote: Everybody's voting and so should you. *J. Politics* 71(1):178–191.

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

Grant AM, Gino F (2010) A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior. *J. Personality Soc. Psych.* 98(6):946–955.

Gubler T, Larkin I, Pierce L (2016) Motivational spillovers from awards: Crowding out in a multitasking environment. *Organ. Sci.* 27(2):286–303.

Gul F, Pesendorfer W (2016) Interdependent preference models as a theory of intentions. *J. Econom. Theory* 165(September):179–208.

Harrison GW, List JA (2004) Field experiments. *J. Econom. 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. *J. Eur. Econom. Assoc.* 8(4):817–837.

Holton V, Dent F, Rabbetts J (2009) Motivation and employee engagement in the 21st century: A survey of management views. Report, Ashridge Business School, Hertfordshire, UK.

Hoogveld N, Zubanov N (2014) The power of (no) recognition: Experimental evidence from the university classroom. IZA Discussion Paper 7953, Institute for the Study of Labor, Bonn, Germany.

Kosfeld M, Neckermann S (2011) Getting more work for nothing? Symbolic awards and worker performance. *Amer. Econom. J.: Microeconomics* 3(3):1–16.

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

Kube S, Maréchal MA, Puppe C (2012) The currency of reciprocity—Gift exchange in the workplace. *Amer. Econom. Rev.* 102(4):1644–1662.

Kube S, Maréchal MA, Puppe C (2013) Do wage cuts damage work morale? Evidence from a natural field experiment. *J. Eur. Econom. Assoc.* 11(4):853–870.

Kuhnen CM, Tymula A (2012) Feedback, self-esteem and performance in organizations. *Management Sci.* 58(1):94–113.

Levine DK (1998) Modeling altruism and spitefulness in experiments. *Rev. Econom. Dynam.* 1(3):593–622.

Levitt S, List JA (2009) Field experiments in economics: The past, the present, and the future. *Eur. Econom. Rev.* 53(1): 1–18.

List JA, Rasul I (2011) Field experiments in labor economics. Ashenfelter O, Card D, eds. *Handbook of Labor Economics*, 1st ed. Vol. 4 (Elsevier, New York), 103–228.

McSweeney FK (2004) Dynamic changes in reinforcer effectiveness: Satiation and habituation have different implications for theory and practice. *Behav. Analyst* 27(2):171–188.

Neckermann S, Yang X (2012) The value of saying thank you: An experiment on the performance consequences of unexpected rewards. Mimeo, Erasmus University Rotterdam, Rotterdam, Netherlands.

Neckermann S, Cueni R, Frey BS (2014) Awards at work. *Labour Econom.* 31(December):205–217.

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

Nolan JM, Schultz PW, Cialdini RB, Goldstein NJ, Griskevicius V (2008) Normative social influence is underdetected. *Personality Soc. Psych. Bull.* 34(7):913–923.

Non A (2012) Gift-exchange, incentives, and heterogeneous workers. *Games Econom. Behav.* 75(1):319–336.

Ockenfels A, Sliwka D, Werner P (2015) Timing of kindness—Evidence from a field experiment. *J. Econom. Behav. Organ.* 111:79–87.

Rhoades L, Eisenberger R (2002) Perceived organizational support: A review of the literature. *J. Appl. Psych.* 87(4):698–714.

Rogers T, Frey E (2015) Changing behavior beyond the here and now. Keren G, Wu G, eds. *The Wiley Blackwell Handbook of Judgment and Decision Making* (John Wiley & Sons, Chichester, UK).

Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V (2007) The constructive, destructive, and reconstructive power of social norms. *Psych. Sci.* 18(5):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. *Econom. J.* 119(540):1422–1439.

Sliwka D (2007) Trust as a signal of a social norm and the hidden costs of incentive schemes. *Amer. Econom. Rev.* 97(3):999–1012.

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

Swank OH, Visser B (2007) Motivating through delegating tasks or giving attention. *J. Law, Econom., Organ.* 23(3):731–742.

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

Wagner R, Harter JK (2006) 12: *The Elements of Great Managing* (Gallup Press, New York).

Wiley C (1997) What motivates employees what motivates employees according to over 40 years of motivation surveys. *Internat. J. Manpower* 18(3):263–280.