

THE EFFECTS OF FINANCIAL AND RECOGNITION INCENTIVES ACROSS WORK CONTEXTS: THE ROLE OF MEANING

MICHAEL KOSFELD, SUSANNE NECKERMANN and XIAOLAN YANG*

We manipulate workers' perceived meaning of a job in a field experiment and interact meaning of work with both financial and recognition incentives. Results show that workers exert more effort when meaning is high. Money has a positive effect on performance that is independent of meaning. In contrast, meaning and recognition interact negatively. Our results provide new insights into the stability of incentive effects across important work contexts. They also suggest that meaning and worker recognition may operate via the same motivational channel. (JEL C93, J33, M12, M52)

I. INTRODUCTION

The question of how and to what degree individual incentives affect performance has captured the interest of researchers for the last several decades. Researchers addressing this question initially focused primarily on financial incentives, though in recent years increasing attention has been devoted to nonfinancial incentives, as well.¹ Interestingly, and particularly

relevant for our purposes, this literature does not paint an altogether consistent picture. Consider, for instance, research on the effectiveness of relative performance feedback. Azmat and Iriberry (2010), Blanes i Vidal and Nossol (2011), and Tran and Zeckhauser (2012) indicate that such feedback can have a substantial positive effect on performance, while Barankay (2012) reaches the conclusion that relative performance feedback can backfire. Likewise, findings vary widely across papers exploring the impact of recognition. Ashraf, Bandiera, and Lee (2014) and Kosfeld and Neckermann (2011), for example, find that recognition works well as an incentive, while recognition works less well in Hammermann and Mohnen (2014), and has detrimental effects in Gubler, Larkin, and Pierce (2013). These disparate findings suggest that there exist important differences across populations and situations that influence the effectiveness of rewards. Despite the importance of understanding how and when the context in which an incentive is implemented influences its effectiveness, research to date has made little progress in answering this question.²

*Financial support from the DFG-ANR project "Understanding organisations - The complex interplay of incentives and identity" (FR 2822/1-1) and the Alfons und Gertrud Kassel-Stiftung is gratefully acknowledged. This article was previously circulated under the title "Knowing that you matter, matters! The interplay of meaning, monetary incentives, and worker recognition."

Kosfeld: Professor, Faculty of Economics and Business Administration, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, 60323 Frankfurt, Germany; CEPR; CESifo; IZA; and ZEF. Phone +49 69 798 34823, Fax +49 69 798 35201, E-mail kosfeld@econ.uni-frankfurt.de

Neckermann: Assistant Professor, Department of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3000 DR Rotterdam, The Netherlands; ZEW; Tinbergen Institute; and CREMA. Phone +31 010-4081393, Fax +31 010-4089161, E-mail neckermann@ese.eur.nl

Yang: Associate Professor, College of Economics, Academy of Financial Research, Zhejiang University, Zheda Road NO.38, Hangzhou, China. E-mail yxlan@zju.edu.cn

1. The literature is much too large to be cited in its entirety. As a starting point, see, for example, Lazear (2000) and Shearer (2004) on the effects of piece rates, and Harbring and Irlenbusch (2003) and Delfgaauw et al. (2013) on tournament incentives; with respect to nonfinancial incentives, there is work on relative performance feedback and recognition (both are discussed below) as well as on other nonpecuniary motivators, including mission motivation (e.g., Fehrler and Kosfeld 2014, Carpenter and Gong 2016), social identity (e.g., Chen and Li 2009, Masella et al. 2014), as well as

intrinsic and image motivation (e.g., Gneezy et al. 2011 and Kamenica 2012 for surveys).

2. While we are not aware of any study that has looked at interaction effects among different rewards and the work context in particular, there are a few studies that look at interaction effects among different motivators or between incentives and the nature of the task. Camerer and Hogarth (1999) compile a meta-analysis of 74 experiments with varying performance-based financial incentives. They find that incentive effects interact with participants' cognitive capital and task requirements. Charness and Grieco (2014) document that incentive effects vary with the nature of the task. Stajkovic and Luthans (2003) present a meta-analysis of behavioral management studies, analyzing if and how different rewards (money, social recognition, and feedback) interact

While jobs naturally vary along a number of dimensions that could possibly interact with the effectiveness of rewards, one interesting and relevant starting point, which has received relatively little attention in the literature so far, is the perceived meaningfulness of work. While there is not one universally accepted definition of “meaningful work,” it is commonly understood as a task that is recognized by others and/or has some point or purpose (Ariely, Kamenica, and Prelec 2008, 672).³ This article assesses the effects of both a financial incentive (a piece rate) and a nonfinancial incentive (symbolic work recognition) across two meaning of work conditions (high vs. low). We opted for a financial incentive as financial incentives are the primary means of motivating and compensating employees in the workplace. Also, most economic research on incentives focuses on financial rewards (Prendergast 1999). We added a nonfinancial incentive because recent evidence shows that worker recognition, for example, via symbolic awards such as “Employee of the Month,” also plays an important role in labor relationships (Ashraf, Bandiera, and Lee 2014; Bradler et al. 2016; Gubler, Larkin, and Pierce 2013; Kosfeld and Neckermann 2011). A key difference between worker recognition (the nonfinancial incentive) and meaning (the context factor) is that meaning of work relates to the task, whereas worker recognition recognizes the worker.

The few existing studies on meaning of work document that meaning has a powerful main effect on performance. Ariely, Kamenica, and Prelec (2008) find in a lab experiment that participants’ reservation wage strongly depends on the perceived meaning of the task. In their study, meaning is manipulated by the experimenter either explicitly acknowledging or ignoring a participant’s performance. Chandler and Kapelner (2013) conduct a study on Amazon’s MTurk and compare the effect of meaning on participation and effort across different cultures (India

with one another. Similarly, Lourenço (2016) reports the findings from a field experiment with retail sales representatives that looks at interaction effects among this same set of motivators (money, recognition, and feedback). Kvaloy, Nieken, and Schöttner (2015) find that the effectiveness of piece rates interacts with motivational talk, and Bradler and Neckermann (2016) show that gifts that combine recognition and money only work well when they come with a personal touch.

3. The notion of meaning is related to the concept of *task significance* (cf. Hackman and Oldham 1976) capturing the degree “to which a job provides opportunities to improve the welfare of others” (Grant 2008, 110).

and United States). This study varies meaning by advertising the task either neutrally as an image labeling task or meaningfully as a contribution to cancer research, and finds an effect of meaning on participation, but not on effort. Finally, Grant (2008) documents a significant increase in the performance of fundraising callers when the latter receive positive information about the meaning of their job; in this case, meaning refers to the impact that the raised funds have on the recipients. The interpretation of these findings is that individuals care about “what they do” and that they respond to changes in perceived job meaning in both labor supply and effort provision.

By studying both financial and nonfinancial incentives in situations with high and low perceived meaning of work, our study not only sheds light on the robustness of these incentives across one exemplary context but also allows us to gain a better and broader understanding of the role of meaning itself. It is, for example, unclear what mechanisms are responsible for the positive effects of meaning documented in the studies above, that is, *why* individuals care about what they do and what impact their work has on others. One possible mechanism is that individuals care about the different outcomes of their job per se. Another mechanism is that individuals do not care about outcomes directly but are more concerned about their social- and self-image that is associated with their job and its outcomes (Bénabou and Tirole 2006).⁴ Understanding the precise mechanisms is important in order to integrate the role of meaning in current behavioral economic models and to derive relevant policy implications for firm owners and managers.

In order to address these issues, we conducted a field experiment in collaboration with a large research center in Hangzhou, China, and hired 413 students for a one-time data entry job. Based on a 2×3 design, we vary both the meaning of the task and the provision of incentives independently. In the *high meaning* condition, workers are told that their work is important for a research project. In the *low meaning* condition, workers are told that they will be entering data for a quality check and that these data will most likely never actually be used. Independently of this, workers are either paid a fixed wage (*baseline*), a fixed wage plus piece rate of 1 Yuan (*monetary incentive*), or a fixed wage with the hardest working individual additionally receiving a symbolic

4. While Grant (2008) suggests that the first mechanism is in place, the results in Ariely, Kamenica, and Prelec (2008) seem more in line with the second mechanism.

award (smiley button) at the end of the session (*recognition*).

Our main results show that, first, in line with previous evidence, meaning significantly increases performance. The average effect size is about 14%. In the low meaning condition, the effect of meaning is significantly larger than the effect of what was deemed a typical piece rate in this setting (8%) and somewhat smaller than the effect of recognition in the form of the smiley button (19%). The large impact of recognition is in line with the evidence cited above that shows that recognition has substantial motivating power albeit the rewards used to confer recognition are often negligible in financial terms. Second, we find no significant interaction between meaning and monetary incentives, that is, both motivators have similarly positive effects independent of whether the other incentive is present or not. One might have been concerned that the financial incentive would crowd out intrinsic motivation in the high meaning case.⁵ However, we find no evidence for crowding out in our data. This result is in line with recent evidence from Ariely, Bracha, and Meier (2009), who document that monetary incentives increase task performance both when intrinsic motivation is low and when it is high as long as effort is exerted in private, which was the case in our study. Third, we find a strong and negative interaction between meaning and recognition. Recognition substantially increases performance only if workers' perceived meaning of the task is low. Performance in this treatment is as high as in the baseline condition with high meaning, suggesting that recognition provides meaning to an otherwise meaningless task. Recognition has no further effect when workers' perceived meaning is high. Similarly, meaning does not increase performance in the presence of recognition. We can rule out that the ineffectiveness of the recognition rewards is driven by pure ceiling effects, as performance under high meaning and monetary incentives is substantially higher than under high meaning and recognition incentives.

Overall, our results corroborate the importance of meaning with regard to work effort. They also show that in connection with different levels of meaning, monetary incentives are strong

and reliable motivators, while recognition may or may not yield positive effects. This finding may help to explain the diverging results in the growing literature on nonfinancial rewards as outlined above. Even though the presence and absence of meaning cannot alone explain the differences in findings on relative performance feedback and nonfinancial awards, our result suggests that recognition in particular, and potentially other forms of nonfinancial incentives more broadly, might be sensitive to changes in the environment.

Finally, the interactions observed in our data provide suggestive evidence for the underlying mechanism through which both meaning and recognition affect performance. Our results are inconsistent with the assumption that, at least in our case, meaning and recognition influence workers' motivation independently via separate channels. Rather, our findings suggest that both motivators rely on the same channel. One possible candidate for such a channel is image seeking. The model of Bénabou and Tirole (2006), for example, predicts that tasks or rewards with a bearing on an individual's social or self-image work less well in the presence of other motivators that also have image value. The intuition is that the marginal benefit of image rewards declines as motivation is more and more ascribed to image motivation. In consequence, image rewards can have substitutive, rather than additive effects, which is exactly what we find. While we do not rule out that there may be other cases and settings where meaning matters directly, it is interesting to see that meaning may work exclusively via an image-seeking channel, that is, that there are cases in which meaning and recognition can be perfect substitutes. This has important implications for workplace design and reward policies, but also for economic models on behavior in the workplace.

The article proceeds as follows. Section II explains the design and treatments of our field experiment. Section III derives behavioral hypotheses. Section IV presents the results. Finally, Section V concludes.

II. EXPERIMENTAL DESIGN

A. Basic Set-Up

In collaboration with the Social Survey Research Center of Zhejiang University in Hangzhou, China, we conducted a field experiment to assess the impact of meaning on work performance in relation to and in combination

5. The literature on crowding out suggests that the provision of financial incentives can be counterproductive because they may reduce an employee's intrinsic motivation to perform the task. Because meaningful tasks are more likely to stimulate intrinsic motivation, they seem susceptible to such crowding out effects.

with monetary incentives and worker recognition. The experiment took place between November 2010 and March 2011. At that time, the research center had just received more than 400 surveys from a large-scale study on the parent-child relationship of migrant workers in China that needed to be filed into an electronic database. In the name of the research center, we hired students for a one-time, 2-hour data entry job where they could earn a fixed pay of Yuan 50 (\$8).⁶ The job offer was announced on the research center's website and posted on campus bulletin boards. Students registered for the job via email.

Upon arrival, students were seated in front of different workstations. The workstations were arranged with sufficient space between them to ensure that individuals felt unobserved. In each session, students received a short introduction to the research center and the particular project the survey was about. Additionally, a short lesson was given on how answers from the surveys were to be filed into the database. During the introduction, it was emphasized that the job was one-time, that a second participation was not possible, and that the research center had no job vacancies to fill at the moment. Students used a web interface to file the surveys. The web interface was set up like an online version of the survey the students had in front of them. Each survey consisted of 151 questions, 14 of which were free-text fields. The remaining questions were multiple choice questions.

Every student received 20–30 questionnaires that were stacked in a nontransparent box in front of them, with a second box next to it to deposit the completed questionnaires. This minimizes peer effects as it prevents participants from comparing their work with each other (cf. Falk and Ichino 2006; Mas and Moretti 2009). Nobody completed all the surveys provided to them. When students started working, the research assistant left the room informing them that she was working outside, was available for queries at any time, and that they could take breaks whenever necessary. Five minutes before the end of a session, the research assistant came back into the room asking the students to stop working and to close the database. Subsequently, participants were asked to fill out a short questionnaire to provide feedback to the research center regarding the job. The questionnaire included questions about satisfaction with the job and the fixed wage, the

quantity and quality of the work accomplished, prior experience with entering data, and whether full concentration had been required.

During work, students could use the internet whenever they wanted, for example, to check emails. Together with the fixed wage, these measures (visible internet access and absence of a supervisor in the room) were taken to give participants in the baseline condition a certain leeway in how much time to spend on filing the surveys. Furthermore, a collective break was avoided to minimize group effects and communication between participants. Overall, the task was quite exhausting and monotonous. Therefore, we do not believe that students had a lot of intrinsic motivation to perform the task in the absence of any meaning.

B. Treatments

We implemented a 2×3 design, in which both the meaning of the task (*low meaning* vs. *high meaning*) and the provision of incentives (*baseline* vs. *monetary incentive* vs. *recognition*) were varied independently. A detailed script of each treatment condition is given in the Appendix S1, Supporting Information.

First, with regard to incentives, we implemented as *baseline* a no-incentive condition, in which students earned only the fixed wage of Yuan 50 independent of work output. Next, in the *monetary incentive* condition, students earned the fixed wage of Yuan 50 plus a piece rate of Yuan 1 per survey. Finally, in the *recognition* condition, students earned the fixed wage of Yuan 50 and they were told that the person who entered most data in their workgroup would be awarded a smiley button being handed over by the research assistant in front of all students at the end of the work session. The motivation for this treatment manipulation comes from popular awards such as “Employee of the Month,” in which top performers’ effort is symbolically recognized in a publicly observable manner.

Second, with regard to the manipulation of meaning, we had two treatments, one high and one low in meaningfulness of work. In the *high meaning* condition, students were informed that the data that they enter would be used directly for the analyses of the research project. In contrast, in the *low meaning* condition, students were told that the data had been entered already and that they would enter data only for a quality check and that it was unlikely that their data would ever actually be used for the analysis of the project.

6. The hourly wage for a student job at Zhejiang University ranges between Yuan 20 and 30.

We opted for this manipulation, as it does, in our view, capture an important aspect of the meaningfulness of a task at work: the degree to which an employee's output, be it data entry, the repair of a machine, or a research report for a superior, will ever actually be used. Based on our own experience working with firms, many employees report having to do things that they feel no one is ever going to look at or to do work that makes little sense to them and other employees. In that sense, we consider the setting of our experiment to be both realistic and relevant.

Note that our low meaning condition gives subjects conflicting signals about the usefulness of their work, as they were, on the one hand, hired for a job and asked to do something, and, on the other hand, implicitly told that their work would most likely not be used. Again, we consider such situations not uncommon in many organizational contexts. In firms, individuals are employed to do a certain job and they receive a wage (with or without financial and nonfinancial incentives just like in this study). Therefore, in firms, just like in this field experiment, there is some baseline perception on the side of the workers that their work should, at least to some extent, be useful to the organization. Still, more often than not, supervisors are too slow to read a report before it outlives its usefulness, or a certain piece is repaired or produced and never actually utilized. Individuals frequently have a sense that their work will not be used, but they do not know for sure. In that sense, the conflicting signals between being asked to do something and the notion that the work might not actually be used is not uncommon in the workplace. Of course, we cannot rule out that individual subjects in the low meaning condition might have thought that their work would be used for the research analysis and was, hence, important. Nevertheless, perceived meaningfulness should, on average, be lower in the low meaning than in the high meaning condition. Our considerations are also reflected in the terms "low" and "high" meaning: We prefer to contrast high meaning with low meaning rather than high meaning with no meaning, which we consider unrealistic for most workplace scenarios.⁷ Our manipulation

7. Had we had more subjects at our disposal, we would have loved to run a neutral treatment, in which we simply tell people to enter the data without either highlighting the importance of their work or emphasizing its possible uselessness. Because we were less interested in studying the pure effect of meaning in this article, as this had been done elsewhere

of meaning differs from previous work and therefore complements these studies.⁸

In total, 413 students participated in the study across the six treatment conditions. Two hundred students participated in the low meaning condition (baseline 86, monetary incentives 67, recognition 47), and 213 participated in the high meaning condition (baseline 59, monetary incentives 53, recognition 101).

III. BEHAVIORAL HYPOTHESES

One research question of our study is whether the meaning of a task affects performance. Informed by previous findings, we hypothesize that in the absence of any incentives average performance is higher when the perceived meaning of the task is high compared to when the perceived meaning is low.

HYPOTHESIS 1 (Meaning): In the absence of recognition or monetary incentives, average performance is higher in the high meaning condition compared to the low meaning condition.

With regard to the effect of monetary incentives, our hypothesis is that piece rates increase average performance, because students care about monetary income and have an incentive to increase effort as long as the monetary benefit of Yuan 1 exceeds the marginal cost of completing another survey.

The next question is whether and how the effect of monetary incentives depends on meaning. In recent years, a large literature both in psychology and economics has shown that extrinsic incentives such as piece rates may generate important interaction effects with subjects' intrinsic motivation to perform a task, leading to a potential crowding-out of intrinsic motivation, and, possibly, a performance reduction (Deci, Koestner, and Ryan 1999; Gneezy, Meier, and Rey-Biel 2011). An individual's concern for a good image, vis-à-vis himself and others, has

before, we decided for more power on the different interaction treatments, in lieu of a third meaning condition.

8. Our meaning interventions relate to the importance and relevance of the work for the organization, while existing studies focus on the sustainability and visibility of output and the explicit recognition by a supervisor (in an experimental task that has no meaning in and of itself) (Ariely, Kamenica, and Prelec 2008), the intrinsic meaningfulness of a task (image labeling versus part of cancer research) (Chandler and Kapelner 2013), or the extent to which a worker's effort affects the wellbeing of others (Grant 2008).

been identified to be a key channel (Bénabou and Tirole 2006).⁹ Based on these results, it is possible that monetary incentives may crowd out intrinsic motivation in our experiment. This is particularly relevant for the high meaning condition, where participants' intrinsic motivation is expected to be high relative to the low meaning condition (cf. Hypothesis 1). However, students always work in private and their performance is never revealed to others. Any image effect can therefore only be driven by an individual's concern about her self-image. Ariely, Bracha, and Meier (2009) show that under such conditions, monetary incentives tend to *not* crowd out behavior, while they do generate negative effects if behavior is made public (and thus social image concerns come into play), suggesting that self-image concerns have less motivational power than similar concerns about one's social image.¹⁰ We therefore expect crowding-out effects in our experiment, if they exist, to be small.

HYPOTHESIS 2 (Monetary Incentives): *Monetary incentives increase average performance in both meaning conditions; because performance is private, crowding-out effects are expected to be small.*

When it comes to worker recognition, our point of comparison is Kosfeld and Neckermann (2011), who show that symbolic awards with little or no material value significantly increase worker performance. The suggested channel for the positive effect is the nonmaterial benefit for the receiver of an award in the form of recognition and status both from peers and the award-giving institution, here the employer. Based on these results, we expect a positive effect of the symbolic award on average performance.

The interaction effect with regard to meaning is a priori ambiguous. On the one hand, it is conceivable that meaning increases the award's value, in so far as receiving an award for a meaningful task may be associated with a higher status and recognition than receiving an award

for a relatively meaningless task. In this case, meaning should interact positively with recognition. On the other hand, meaning also has a recognition element itself. One source of meaning, for example, rests in the recognition of the work as valuable or important (see, e.g., Ariely, Kamenica, and Prelec 2008). Such recognition and acknowledgment may be explicit, implicit, or even internalized as is the case, for example, for types of behaviors that are obviously deemed desirable. This suggests that meaning and recognition may operate, at least partly, via the same channel. To the extent that this is true, recognition can be expected to work less well when the perceived meaning of the task is high as compared to when the perceived meaning is low. The reason is that, according to image-reward theory, the marginal benefit of image rewards declines as image rewards increase, because behavior is more and more ascribed to image motivation (Proposition 4, Bénabou and Tirole 2006).

HYPOTHESIS 3 (Recognition): *Recognition increases average performance. The interaction effect with meaning is ex-ante ambiguous: (a) if meaning increases the value of awards, the effect of awards should be higher in the high meaning condition; (b) if, however, the meaning of a task has image value itself, symbolic awards work less well in the high meaning condition.*

Note that Hypothesis 3 implies that our data will allow us to draw first conclusions about the underlying channels through which meaning and recognition operate. A positive (or zero) interaction effect indicates that meaning and recognition enter a person's utility function separately, possibly even reinforcing each other (Hypothesis 3a). In contrast, a negative interaction effect suggests that meaning and recognition operate via the same channel, for example, image rewards (Hypothesis 3b). We now turn to the results of our field experiment.

IV. RESULTS

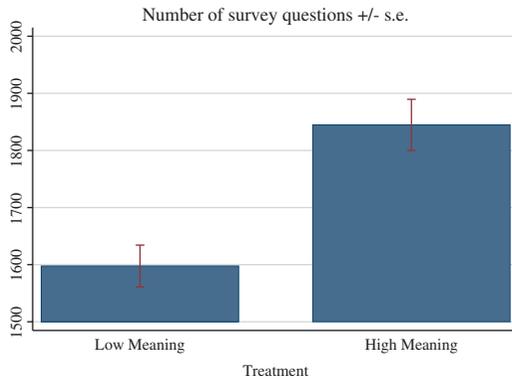
In the results section, we proceed as follows. We first consider the effect of meaning in the absence of either monetary incentives or recognition. This allows us to tie our study to the existing literature and serves as an important check for our context manipulation—a performance difference between the high and low meaning condition indicates that our meaning intervention was successful in changing the work environment. We

9. Two arguments underlie this logic. First, individuals like to think of themselves (and be seen by others) as intrinsically motivated and/or altruistic. Second, while individuals may remember their actions and the presence or absence of extrinsic rewards, they may not remember their own motivation. Hence, the motivation of subjects that engage in a certain activity in an attempt to signal to themselves and others that they are altruistic/intrinsically motivated, might be reduced when financial rewards are introduced. See also Bowles and Polanía-Reyes (2012) for a more general discussion.

10. This conclusion is also corroborated by recent evidence from van der Weele and von Siemens (2013).

FIGURE 1

Effect of Meaning in the Absence of Incentives



then compare the effectiveness of monetary and recognition incentives across the high and low meaning conditions.

Figure 1 shows average performance in terms of the total number of survey questions entered in our low and high meaning baseline treatments without incentives. The data tell a clear story. Increasing the meaning of the task when agents face no explicit incentives raises performance from 1598 questions in the low meaning condition to 1845 questions in the high meaning condition (Mann-Whitney test, $p = .001$).¹¹ Thus, we find strong support for Hypothesis 1 corroborating the importance of a worker's perceived meaning of his task with regard to intrinsically motivated effort.

The average effect size of meaning is 15.5% (247 questions) based on raw data. If we compare this to the effect of monetary incentives and recognition alone, we find that meaning has a significantly larger effect than our monetary incentive but the effect is similar to, and not significantly different from, the effect of the recognition that the institute provided. Consider Table 1, which shows the absolute and relative increase in performance in our three main treatment manipulations based on raw data, taking the low meaning/baseline condition as a benchmark. On average, monetary incentives increase performance by 138 questions or 8.7% (Mann-Whitney test, $p = .008$). This is significantly less than the effect of meaning (Mann-Whitney test, $p = .049$). In comparison, recognition increases performance even by 291 questions, or 18.2%

11. All tests in our analysis are two-sided.

TABLE 1

Effect of Meaning versus Monetary Incentives versus Recognition

	High Meaning	Monetary Incentives	Recognition
Average performance increase	247	138	291
Percentage	15.5%	8.7%	18.2%

Notes: Performance is measured as the number of survey questions entered into the database within the 115-minute work period. Effects are based on raw data with the low meaning/baseline condition as the benchmark.

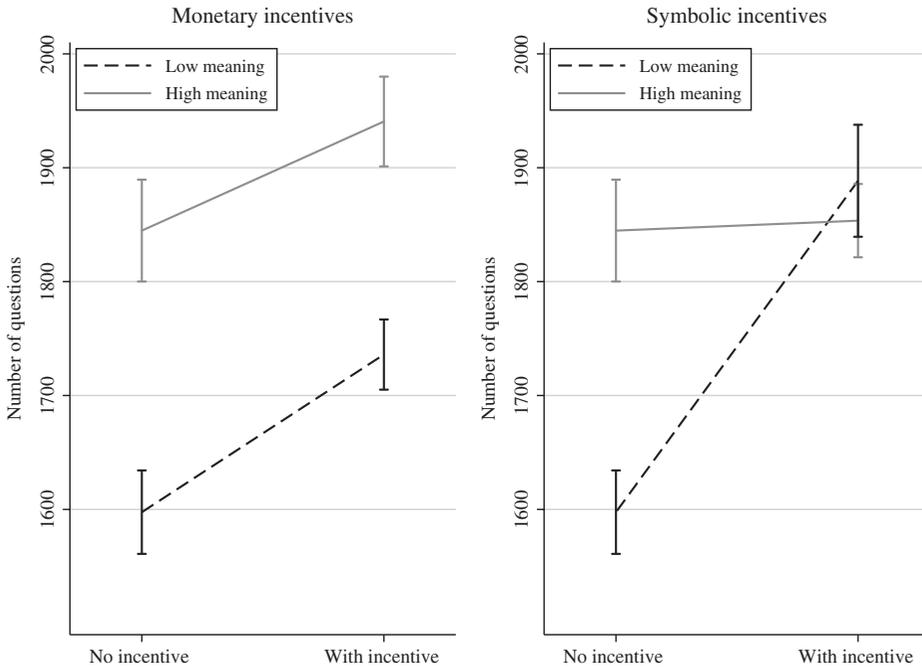
(Mann-Whitney test, $p = .001$); however, the differential effect compared to meaning is not statistically significant (Mann-Whitney test, $p = .809$). Overall, the data suggest that meaning and recognition have similarly strong effects while the effect of monetary incentives (1 Yuan per survey) is positive but relatively weaker.

Let us now turn to the interaction effects. Figure 2 shows average performance in all six treatments.¹² The left panel compares students' performance with and without monetary incentives in both meaning conditions; the right panel shows the same comparison for recognition.

Two important observations can be made from Figure 2. First, monetary incentives increase performance both when meaning is low (black dashed line) and when meaning is high (gray solid line). We have already seen that students enter significantly more questions in the low meaning condition if they are offered a piece rate compared to if they are paid only a fixed wage. Figure 2 shows that a similar positive effect can be observed when the perceived meaning of the task is high. In this case, average performance increases from 1,845 questions under the fixed wage to 1,941 questions under piece rates, which constitutes an increase of about 5% based on raw data (Mann-Whitney test, $p = .140$). Second, recognition increases performance only when the perceived meaning of the task is low. Again, we have seen above that performance in the low meaning condition increases by about 18% if students are recognized by a symbolic award. In contrast, if the meaning of the task is high, the introduction of a symbolic award has no sizable effect on performance. Students' average performance in this case is 1,854 with award and 1,845

12. See Table S1 for an overview of the means and standard deviations in all treatments.

FIGURE 2
Effect of Monetary Incentives and Recognition Conditional on Meaning



without award, which is not statistically different (Mann-Whitney test, $p = .989$). The difference in performance in the high meaning condition between monetary and recognition incentives is marginally significant (Mann-Whitney test, $p = .083$).

The above findings are confirmed by a regression analysis with individual performance as the dependent variable. Table 2 displays the results. The omitted category is the low meaning condition without incentives. We include dummies for the high meaning condition, the provision of monetary incentives, and recognition, as well as for the interaction between the two incentive schemes and the high meaning condition, respectively. The first column includes no controls, columns II and III add controls for group size and gender, as well as for answers to the feedback form students filled out at the end of the job. Comparing the three columns in Table 2 shows that adding control variables has no notable effect on our results.

Taking the full model in column III as a benchmark, we see that all treatments elicit substantial and statistically significant performance increases over the baseline condition. In

particular, increasing only the perceived meaning of the task raises performance by about 230 survey questions on average compared to performance in the situation with low meaning and no additional incentives. In comparison, paying a piece rate of 1 Yuan increases average performance by 150 questions. While the difference between these two effects is considerable, it is no longer statistically significant once controls are included ($p = 0.143$). Paying a piece rate has a similarly positive effect if the perceived meaning of the task is high. In this case, average performance increases by 127 questions, which is less than in the low meaning condition, but the differential effect is not statistically significant as the interaction between high meaning and monetary incentives shows. By the same token, increasing meaning in the presence of monetary incentives also increases performance. The average effect size is 208 questions. Overall, the results confirm our first two Hypotheses 1 and 2: Meaning and monetary incentives are both significant motivators and there are no observable crowding-out effects.

In our setting, by far the largest effect comes from symbolic awards, that is, worker

TABLE 2
Treatment Effects on Performance

	I	II	III
High meaning	247.238*** (70.367)	241.262*** (69.691)	231.043*** (68.232)
Monetary incentive	138.397** (52.518)	143.698*** (52.409)	150.555*** (52.074)
Recognition	290.995*** (63.931)	314.093*** (71.395)	308.254*** (61.394)
High meaning \times monetary incentive	-42.590 (79.549)	-50.226 (78.531)	-23.098 (82.372)
High meaning \times recognition	-282.267*** (93.021)	-293.532*** (95.093)	-275.482*** (85.259)
Constant	1597.558*** (49.090)	1308.645*** (188.703)	1246.688*** (187.663)
Controls (group size, gender)	No	Yes	Yes
Controls (questionnaire)	No	No	Yes
Observations	413	413	410
Sessions	55	55	55
Adj. R^2	0.110	0.114	0.150

Notes: The table reports ordinary least square estimates of treatment effects on performance. Performance is measured as the number of survey questions entered into the database within the 115-minute work period. Standard errors in parentheses account for clustering at the session level. The omitted category is the low meaning/baseline condition. Column I contains no controls. Column II contains controls for group size and gender. Group size is the number of students who worked together in a group. Column III adds answers to the questionnaire (whether the individual was happy with the job, with the fixed wage, the quantity of work accomplished, the quality of work accomplished, whether full concentration had been required to complete the job, and whether the individual had prior experience with entering data). These variables might be endogenous to treatment.

Asterisks next to coefficients indicate significance at the 10%, 5%, and 1% levels.

recognition. In the low meaning condition, average performance increases by more than 300 questions when a symbolic award is offered. Taking controls into account, we see that this effect is significantly stronger than the effect of monetary incentives ($p = .002$) but not significantly different from meaning alone ($p = .249$). In contrast, if meaning is high, the effect of recognition is basically zero. Average performance with symbolic award is about 33 questions higher than without, which is not statistically different, and the interaction between meaning and symbolic incentives is negative and highly significant.

In much the same way as recognition does not increase average performance when meaning is high, an increase in meaning has no effect in the presence of recognition, either. Column III in Table 2 suggests that performance is actually reduced by about 40 questions but the difference is not statistically significant ($p = .415$).

Our findings suggest that, at least in settings similar to ours, meaning and recognition entail substitutive effects as is indicated by the negative interaction between the two variables. This speaks against the idea that both motivators operate separately from each other and enter an individual's utility function via different,

possibly additive, channels. Rather, it suggests that recognition and meaning are substitutes and work via the same motivational channel. As was laid out above, this is in line with image-reward theory (Bénabou and Tirole 2006) that predicts that incentive effects of image rewards, while positive, are inherently limited due to the declining marginal benefit of these rewards. Interpreted through this lense, our negative interaction effect is consistent with both meaning and recognition having image value and therefore, operating via the same channel. In sum, our data reject Hypothesis 3a and confirm Hypothesis 3b.

Finally, we checked for gender effects and do not find any. We also find no systematic differences with respect to the number of mistakes students make across treatments, which is probably due to the fact that the number of mistakes in our set-up is in general extremely low: 2/3 of the students have an error rate of less than 1% and 93% have an error rate of less than 5%.

V. CONCLUSION

This article studies how financial and non-financial incentives interact with meaning of work as a context factor. Our results, which are

based on a field experiment in collaboration with the Social Survey Research Center of Zhejiang University, corroborate previous findings on the importance of meaning for workers' willingness to exert effort. The results confirm that providing workers with a heightened awareness of the meaning of their job is a powerful tool to increase performance. In the setting of this study, meaning outperforms monetary incentives and works almost equally well as worker recognition in the absence of meaning. This has important and far-reaching implications for the workplace, suggesting that it might be more cost-effective for managers to talk to workers and spend time on the factory floor communicating the value of their work than to institute complex compensation schemes to incentivize performance.

Nevertheless, we also find that monetary incentives do increase performance in a statistically and economically significant way. Hence, even though nonmonetary rewards may be more cost-effective at times, monetary incentives work well and are also very robust to different contexts (here in situations with high and low perceived task meaning). The fact that we do not observe a crowding-out effect in our data is in line with recent evidence from Ariely, Bracha, and Meier (2009), who also find that monetary incentives do not reduce performance in image-motivated tasks as long as effort is exerted in private.

The finding that worker recognition in the form of symbolic awards matches the impact of meaning and that the combination of meaning and recognition together does not lead to higher performance than either in isolation suggests that both meaning and recognition work via the same motivational channel, with substitutive effects on performance. One explanation is that both meaning and recognition serve as image rewards (Bénabou and Tirole 2006). The result also documents that, depending on the context, recognition can either boost performance or be relatively ineffective, thus pointing to a sensitivity of recognition in particular and, possibly, nonfinancial rewards more generally.

Our study presents a first step in understanding important context dimensions, like meaning. The interactions in our design allow us to draw first conclusions about the underlying preference structure. Future research needs to look at further incentives and contexts and their interactions as well as test the robustness of the current set of findings over time and across different settings.

REFERENCES

- Ariely, D., E. Kamenica, and D. Prelec. "Man's Search for Meaning: The Case of Legos." *Journal of Economic Behavior & Organization*, 67, 2008, 671–77.
- Ariely, D., A. Bracha, and S. Meier. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review*, 99(1), 2009, 544–55.
- Ashraf, N., O. Bandiera, and S. S. Lee. "Awards Unbundled: Evidence from a Natural Field Experiment." *Journal of Economic Behavior and Organization*, 100(April), 2014, 44–63.
- Azmat, G., and N. Iriberrí. "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students." *Journal of Public Economics*, 94(7–8), 2010, 435–52.
- Barankay, I. "Rank Incentives: Evidence from a Randomized Workplace Experiment." University of Pennsylvania Working Paper, 2012.
- Bénabou, R., and J. Tirole. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5), 2006, 1652–78.
- Blanes i Vidal, J., and M. Nossol. "Tournaments without Prizes: Evidence from Personnel Records." *Management Science*, 57(10), 2011, 1721–36.
- Bowles, S., and S. Polanía-Reyes. "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature*, 50, 2012, 368–425.
- Bradler, C., and S. Neckermann. "The Magic of the Personal Touch: Field Experimental Evidence on Money and Gratitude as Gifts." Mimeo, 2016.
- Bradler, C., R. Dur, S. Neckermann, and A. Non. Forthcoming. "Employee Recognition and Performance: A Field Experiment." *Management Science*, 2016.
- Camerer, C. F., and R. M. Hogarth. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty*, 19(1-3), 1999, 7–42.
- Carpenter, J., and E. Gong. "Motivating Agents: How Much Does the Mission Matter?" *Journal of Labor Economics*, 34(1), 2016, 211–36.
- Chandler, D., and A. Kapelner. "Breaking Monotony with Meaning: Motivation in Crowdsourcing Markets." *Journal of Economic Behavior & Organization*, 90, 2013, 123–33.
- Charness, G., and D. Grieco. "Creativity and Financial Incentives." University of California, Santa Barbara, Working Paper, 2014.
- Chen, Y., and S. X. Li. "Group Identity and Social Preferences." *American Economic Review*, 99(1), 2009, 431–57.
- Deci, E. L., R. Koestner, and R. M. Ryan. "A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological Bulletin*, 125(6), 1999, 692–700.
- Delfgaauw, J., R. Dur, J. Sol, and W. Verbeke. "Tournament Incentives in the Field: Gender Differences in the Workplace." *Journal of Labor Economics*, 31(2), 2013, 305–26.
- Falk, A., and A. Ichino. "Clean Evidence on Peer Effects." *Journal of Labor Economics*, 24(1), 2006, 39–58.
- Fehrler, S., and M. Kosfeld. "Pro-Social Missions and Worker Motivation: An Experimental Study." *Journal of Economic Behavior & Organization*, 100, 2014, 99–110.
- Gneezy, U., S. Meier, and P. Rey-Biel. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives*, 25(4), 2011, 191–210.
- Grant, A. "The Significance of Task Significance: Job Performance Effects, Relational Mechanisms, and Boundary Conditions." *Journal of Applied Psychology*, 93(1), 2008, 108–24.

- Gubler, T., I. Larkin, and L. Pierce. "The Dirty Laundry of Employee Award Programs: Evidence from the Field." Harvard Business School Working Paper 13-069, 2013.
- Hackman, J. R., and G. R. Oldham. "Motivation through the Design of Work: Test of a Theory." *Organizational Behavior and Human Performance*, 16, 1976, 250–79.
- Hammermann, A., and A. Mohnen. "The Pric(ze) of Hard Work—Different Incentive Effects of Non-Monetary and Monetary Prizes." *Journal of Economic Psychology*, 43, 2014, 1–15.
- Harbring, C., and B. Irlenbusch. "An Experimental Study on Tournament Design." *Labour Economics*, 10(4), 2003, 443–64.
- Kamenica, E. "Behavioral Economics and Psychology of Incentives." *Annual Review of Economics*, 4(1), 2012, 427–52.
- Kosfeld, M., and S. Neckermann. "Getting More Work for Nothing? Symbolic Awards and Worker Performance." *American Economic Journal: Microeconomics*, 3(3), 2011, 86–99.
- Kvaloy, O., P. Nieken, and A. Schöottner. "Hidden Benefits of Reward: A Field Experiment on Motivation and Monetary Incentives." *European Economic Review*, 76, 2015, 188–99.
- Lazear, E. P. "Performance Pay and Productivity." *American Economic Review*, 90(5), 2000, 1346–61.
- Lourenço, S. M. "Monetary Incentives, Feedback, and Recognition—Complements or Substitutes? Evidence from a Field Experiment in a Retail Services Company." *The Accounting Review*, 91(1), 2016, 279–97.
- Mas, A., and E. Moretti. "Peers at Work." *American Economic Review*, 99(1), 2009, 112–45.
- Masella, P., S. Meier, and P. Zahn. "Incentives and Group Identity." *Games and Economic Behavior*, 86, 2014, 12–25.
- Prendergast, C. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1), 1999, 7–63.
- Shearer, B. "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *Review of Economic Studies*, 71(2), 2004, 513–34.
- Stajkovic, A. D., and F. Luthans. "Behavioral Management and Task Performance in Organizations: Conceptual Background, Meta-analysis, and Test of Alternative Models." *Personnel Psychology*, 56(1), 2003, 155–94.
- Tran, A., and R. Zeckhauser. "Rank as an Inherent Incentive: Evidence from a Field Experiment." *Journal of Public Economics*, 96(9-10), 2012, 645–50.
- van der Weele, J., and F. von Siemens. "Bracelets of Pride and Guilt? An Experimental Test of Self-Signaling in Charitable Giving." CESifo Working Paper, No. 4674, 2013.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Descriptive statistics

Table S1. Descriptive statistics: means and standard deviations