Cheap-talk Evaluations in Contract Design

Ninghua Du and Quazi Shahriar*

September 5, 2018

Abstract:

This study proposes *ex-post* cheap-talk evaluations as a potential behavioral force in contract designs. Specifically, in a one shot gift-exchange game between a firm and a worker, we experimentally investigate whether a worker's preference for avoiding (seeking) written expression of disapproval (approval) from the firm can induce higher effort from the worker and thereby improve efficiency. We find that, compared to the *no-evaluation* condition, *free-form* evaluation increases both effort and efficiency significantly while *structured* evaluation does not. To identify the channel through which free-form evaluation succeeds, we run additional treatments that allow the firm to communicate its desired effort beforehand. We find that free-form evaluation still achieves the highest level of actual effort while desired efforts are similar across evaluation protocols. If the firm's desired effort serves as a proxy for the firm's expected effort, then this suggests that free-form evaluation's success is due to the saliency of these messages in the worker's preference rather than the worker's belief of the firm's expectation.

Keywords: Evaluation, Contract Design, Gift-exchange, Experiment

JEL Codes: C91, D02, D03

Acknowledgements:
We thank Nick Feltovich and two anonymous referees for valuable comments. Du thanks the National Social Science Fund of China (Grant No. 17BJY208) for financial support.

* Du: Key Laboratory of Mathematical Economics and the School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, P.R. China, email: ninghua.du@mail.shufe.edu.cn, fax: 86-21-65902167; Shahriar: San Diego State University, Department of Economics, 5550 Campanile Drive, San Diego, CA 92182, USA, email: qshahria@mail.sdsu.edu.

1. Introduction

Inefficiency problems are fairly common to arise under incomplete contracts (Grossman and Hart, 1986; Hart and Moore, 1990), as in many markets traders' obligations are specified imprecisely and one party in the transaction has the residual right to improve their own benefits at the expense of the other party. Given pecuniary and selfish preferences, the simple logic of backward induction yields a poor service and the minimum amount of payment as the prevailing outcomes of the incomplete contracts.

Since imperfect monitoring, information asymmetries, moral hazard, and transaction costs often make the use of complete contracts infeasible and formal (court-induced) solutions to incomplete contracts can be very costly, developing informal institutions to boost efficiency of the incomplete contracts becomes a potential alternative. Verbal evaluations, including social approval and disapproval, are commonly observed in real business practices. In many cases, they are costless to express and have no pecuniary payoff implications for the recipients. Theories show that people feel badly if they expect their actions to be disapproved by others, whereas they feel well if they expect their actions to be approved (Hollaender, 1990; Kandel and Lazear, 1992).[1] The existing experimental literature shows that ex-post cheap-talk evaluations can discipline economic agents to comply with social norms and to take pro-social actions through approval seeking or disapproval avoidance.[2]

[1] Besides Hollaender (1990) and Kandel and Lazear (1992), some previous theoretical studies have integrated expressions of approval and disapproval in several contexts such as peer pressure (Barron and Paulson-Gjerde, 1997), voter participation (Uhlaner, 1989; Knack, 1992) and compliance with the law (Tyler, 1990). See also Loch et al. (2000), Akerlof (1980) and Lindbeck et al. (1999).

[2] In dictator games, the amount of money being offered by the divider increases significantly when the recipient can react to the offer with ex-post unrestricted written messages (Ellingsen and Johannesson, 2008; Xiao and Houser, 2009). In public good

If pro-social behaviors are positively correlated with approvals (or negatively correlated with disapprovals) and approvals (or disapprovals) are salient enough in one's utility, providing *ex-post* evaluations may serve as a potential behavioral force to discipline one's behavior in fulfilling contractual obligations. In this study, we consider a one-shot gift-exchange game between a firm and a worker in which the firm commits to paying a wage to the worker and then the worker chooses an effort level which determines the payoffs of both the players. We investigate in a laboratory experiment whether the worker's preference for approval (disapproval) the firm may express to the worker can be salient enough in the sense that it will lead the worker to choose a higher effort level and thus improve efficiency.[3] In our laboratory setting, we implement three cheap-talk evaluation protocols: *null, structured, and free-form evaluations*. There is no evaluation provision in the *null evaluation* treatment. In the *structured evaluation* treatment, the firm passes on a structured rating (Very Good, Good, OK, Bad or Very Bad) to the worker after the worker chooses an effort level. In the *free-form evaluation* treatment, instead of a structured rating, the firm communicates an unrestricted written message.

Our experimental treatments are designed to represent the evaluation protocols commonly seen in practice.[4] The existing literature, however, does not provide any evidence establishing the superiority of any particular format of evaluation. Our experiment is intended to

games, the level of contribution increases significantly if every contributor has a chance to assign payoff neutral and costless approval (or disapproval) points to other contributors after observing the contributions (Gaechter and Fehr, 1999; Masclet et al., 2003; Rege and Telle, 2004; Dugar, 2013).

[3] The literature of gift-exchange games was started by Fehr, Kirchsteiger, and Riedl (1993). See Charness and Kuhn (2011) for a survey.

[4] For instance, both point-based rating systems (e.g. five-star rating) and open-ended feedback systems are used in e-commerce to evaluate product quality. Ghose and Ipeirotis (2011) show that, compared to ratings, text evaluations contain more information of product characteristics. These alternative evaluation protocols are also frequently used to obtain customer feedback for services such as those of health-care providers and restaurants. In workplace, besides verbal expression of approval by the employer, feedbacks include detailed information regarding the level of performance outcomes and the reason of why the employer is satisfied or dissatisfied (Stajkovic and Luthans 2003).

not only show whether an evaluation protocol is effective in enhancing efficiency but also to compare the two alternative protocols we implement. We find that, compared to the no-evaluation condition, the free-form evaluation condition increases both effort (by 61%) and efficiency (by 52%) significantly while the structured evaluation condition increases both the measures slightly (by 17% and 14%, respectively) but insignificantly.

Since *ex-post* evaluations impact decisions by changing the decision-maker's belief (Xiao and Houser, 2009), there are two categories of potential explanations behind the success of the free-form evaluations over the structured evaluations: (1) given the worker's preference for approvals (disapprovals), the worker believes that the nature of the free-form evaluations (i.e. richer, endogenous and personalized by the evaluator) makes them more salient and as a result, the worker exerts higher effort to win (avoid) positive (negative) evaluations, and (2) the worker believes that the employer anticipates higher salience of free-form evaluations and therefore expects the worker to exert higher effort and as a result, the worker behaves accordingly to live up to the employer's expectation to avoid *guilt*.[5]

The second explanation is based on worker's second-order belief and less obvious as it requires the employer to understand the reasoning behind the first explanation which is based on the worker's first-order belief. To check if the second one can possibly be an explanation, we ran three additional treatments in which we allow the firm to tell the worker the effort it expects before the worker makes the actual choice of the effort level. If the firm's proposed effort serves

[5] According to psychological game theory, a decision maker experiences guilt if he believes he let others down (Charness and Dufwenberg, 2006).

as a proxy for the firm's true expected effort, then this allows us to see if the worker takes the firm's effort expectation into account when choosing effort – the necessary condition for the second explanation. We find in the experiment that (1) the firm's proposed effort has no effect on the worker's effort choice, and (2) the average actual effort under the free-form evaluation protocol is still significantly higher than that under the other two protocols even though the average proposed effort of the firm is not much different across the three evaluation protocols. These results suggest that the success of the free-form evaluation protocol is likely due to the saliency of the firm's free-form messages the worker anticipates, but definitely not due to the worker's reading of the firm's proposed effort. The findings can be useful for further theoretical works on evaluations.

Our laboratory experiment controls for potential confounds in naturally occurring situations. First, we control for endogeneity problems. In real business practices, if the verbal evaluation does not work for an organization, that organization would not self-select itself into performing worker evaluation. Second, we control for any possible correlation between pecuniary payoffs and verbal evaluations present in the field. Third, our one shot games control for reputation from the evaluations and repeated interaction effects.[6]

This study is in the strand of the literature exploring the channels to enhance efficiency of incomplete contracts that are built on social preferences and psychological institutions. Previous studies in the literature have explored other channels such as reputation and repeated interaction

[6] There is a literature on evaluation (rating) based reputation mechanisms (Dellarocas, 2003; Cabral and Hortaçsu, 2010). However, this is not the focus of our study.

(Brown *et al*., 2004), bonus and fine (Fehr *et al*., 2007), delegation (Charness *et al*., 2012), and promise (Charness *et al*., 2013) in gift exchange games, and communication and renegotiation (Fehr *et al*., 2015; Brandts *et al*., 2016) in buyer-seller contracts with cost shocks.

2. Behavioral Hypotheses

The common structure of the gift-exchange games employed in our experiment is as follows. Each firm is matched with one worker and then they participate in the two stages described below.

- *Stage* 1: The firm proposes a wage offer w, where w is an integer between 1 and 100 (including 1 and 100).
- *Stage* 2: The worker observes w and then chooses the effort level e, where e is an integer from 1 to 10 (including 1 and 10).

The firm's earning is $90 + 10e - w$ while the worker earns $90 + w - c(e)$. The function $c(e)$ is the cost of effort given in Table 1.[7] Since the firm's surplus is $10e - w$ and the worker's surplus is $w - c(e)$, the aggregate surplus is $10e - c(e)$ which is a nonlinear transformation of the worker's effort e.

Table 1. The cost of effort.

Effort e	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Cost $c(e)$	0	1	3	5	8	12	16	20	25	30
Surplus $10e - c(e)$	10	19	27	35	42	48	54	60	65	70

[7] As far as we know, Charness *et al*. (2013) is the only study that investigates behavioral contract design in a one shot gift-exchange environment. Since we also implement our experimental design in a one shot gift-exchange environment, we borrow the cost of effort function from Charness *et al*. (2013).

Notice that the marginal benefit of increasing one unit of effort is fixed at 10. From Table 1, the marginal cost of each unit of effort is always less than 10. Therefore, the socially optimal effort level is 10. When $w > 30$ and $e = 10$, the firm and the worker share the maximum total surplus from the market. On the other hand, with purely self-regarding pecuniary preferences, the Subgame Perfect Nash Equilibrium (SPNE) predicts $w = 1$ and $e = 1$ from backward induction. If, however, the worker knows the firm can send the worker an evaluation message after observing the worker's effort and the worker's preference for approval (or disapproval) dominates the pecuniary preference, then the worker's effort choice in the presence of an evaluation protocol might be higher than the pecuniary SPNE prediction, and the effort choice can vary across evaluation protocols depending on the saliency of the preference for approval (or disapproval) across the protocols. The discussion below gives us some intuition into this argument.

Let $S(w, e)$ denote the worker's psychological payoff from the firm's strongest approval message, given w and e.[8] Without loss of generality, we only consider the evaluations with no cost to the firm and we normalize the worker's psychological payoff from a null message to be zero. In one of many possibilities, suppose when the worker chooses $e = 10$ and receives the strongest approval message from the firm, then $S(w, e) > 30$ (i.e., $S(w, e) > 30$ when $1 \leq w \leq 100$ and $e = 10$). In this case, one of the SPNE's is that the firm chooses $w = 1$ and sends out the strongest approval message if $e = 10$ and a null message if $e < 10$, and the worker chooses $e = 10$

[8] Hereafter we only discuss the approval seeking preference, since the discussion regarding to disapproval avoiding preference is symmetric.

for any wage level. The verification of the SPNE is as follows. Given the worker chooses $e = 10$ for any wage level, choosing $w = 1$ and sending out the strongest approval maximizes the firm's payoff since the evaluation costs the firm zero. On the other hand, given the firm's strategy, the worker earns $91 + S(w, e) - 30$ when $e = 10$, and $91 - c(e)$ for $e < 10$. Since $S(w, e) > 30$ when $e = 10$ and $c(e) \geq 0$ for any effort level, the worker's payoff is maximized by choosing $e = 10$. Therefore, the above strategy profile is a Nash Equilibrium (NE) of the game. Moreover, there is a proper subgame following any choice of wage w by the firm. In this subgame, the firm's choice of sending out the strongest approval when $e = 10$ and a null message when $e < 10$ is a best response to $e = 10$ since the evaluations cost the firm nothing. Meanwhile, given the firm's strategy corresponds to this subgame, the worker earns $90 + w + S(w, e) - 30$ when $e = 10$, and $90 + w - c(e)$ for $e < 10$. Since $S(w, e) > 30$ when $e = 10$ and $c(e) \geq 0$ for any effort level, the worker's payoff is maximized by choosing $e = 10$. Therefore, the strategy profile that we discuss forms an NE in the subgame following any choice of w. Finally, there is a subgame following w chosen by the firm and e chosen by the worker. In this subgame, any choice of the evaluation message is a best response by the firm since the evaluations cost zero. Now we have verified that "the firm chooses $w = 1$ and sends out the strongest approval message if $e = 10$ and a null message if $e < 10$, and the worker chooses $e = 10$ for any wage level" forms NE in all subgames of the original game. Hence the strategy profile mentioned is an SPNE.

Another possible case is that $S(w, e) > 30$ when the worker earns a non-negative share of the created aggregate surplus (i.e., $S(w, e) > 30$ when $w > c(e)$). There is then a SPNE in which the firm chooses $w > 30$ and sends out the strongest approval message if $e = 10$ and a null

message otherwise, and the worker chooses $e = 10$ if $w > 30$ and $e = 1$ otherwise. The verification procedure is similar to the discussions above. Notice that higher $S(w, e)$ induces higher maximal effort level in the set of SPNE. For example, if $S(w, e)$ is always less than 20, then the worker would never exert an effort level higher than 8. If one expects $S(w, e)$ to be higher under one evaluation protocol than another, then the effort choices could possibly be higher under the protocol with higher $S(w, e)$.

Now, consider three alternative evaluation protocols: Null Evaluation (*NE*), Structured Evaluation (*SE*) and Free-form Evaluation (*FE*). *NE* simply implements the basic gift-exchange game with no evaluation which ends right after the worker chooses *e* in stage 2. In *SE*, after the firm learns the effort level *e*, the firm sends a rating to the worker from five possible choices: Very Good, Good, OK, Bad, Very Bad.[9] The rating has no monetary consequence for either the firm or the worker. The worker knows the existence of the rating stage before making the effort choice. In *FE*, the firm sends an unrestricted message to the worker after the firm observes the worker's effort *e*. The unrestricted message also has no effect on the firm's or the worker's earnings.

In light of the findings in the literature that suggest that evaluations can promote prosocial behavior by the party being evaluated, we expect workers to choose higher effort levels when the firm is able to utilize either the structured or the free-form evaluation protocol than when the firm does not provide any evaluation to the worker. Moreover, since the message space

[9] We implement a structured effort rating rather than a gratitude rating. The effect of structured gratitude ratings could be studied in future research.

in *FE* is richer and the messages are endogenous in nature, we expect the evaluations in this protocol to be more salient in the worker's preference than those in *SE*. This, for instance, may imply that $S_{FE}(w,e) > S_{SE}(w,e)$, and based on our earlier argument, we then would expect higher effort choices in *FE* than in *SE*. We formalize our conjectures in the following hypotheses.

Hypotheses

(a) *The average effort and efficiency in SE are higher than those in NE.*

(b) *The average effort and efficiency in FE are higher than those in NE.*

(c) *The average effort and efficiency in FE are higher than those in SE.*

3. Experimental Design

To test our hypotheses, we ran an experiment that consisted of three treatments based on the three evaluation protocols discussed in the previous section – Null Evaluation (*NE*), Structured Evaluation (*SE*) and Free-form Evaluation (*FE*).

We conducted our experiment at the San Diego State University. The participants were recruited from a campus-wide list of undergraduate students who had previously responded to advertisements. None of the participants had any experience with gift-exchange game experiments. A session of our experiment lasted for only one period.[10] There were 7 sessions (112 participants): 3 sessions (40 participants) in *NE* and 2 sessions (36 participants) in each of the other two treatments.[11] No participant was permitted to participate in more than one session.

In each session, half of the participants were firms and the rest were workers. The firms

[10] See Charness et al. (2013) for a detailed discussion regarding to the advantages of one shot gift-exchange games.
[11] 55% of the total participants of 112 were female, and the average age of the participants was 19.1 years.

and workers were randomly and anonymously matched into pairs. Then the instructions were read aloud to them. (The instructions are in Appendix A.) The rest of the experiment was then carried out according to the protocol of the treatment described in the previous section. We implemented a double blind social distance protocol in all the sessions. Specifically, at the beginning of the experiment, each participant was randomly assigned an ID number which was only known by the participant; the experimenter observes the choices associated with each ID number without connecting the IDs to the participants. At the end, participants' earnings were put in separate envelopes with their IDs written on the face of the envelopes which were then on a table; on the way out each participant picked up the envelope with his/her ID number. The actual identities of the participants were never used or disclosed in any way either during or after the experiment.

The average payment was 10.9 USD (the exchange rate was 1 USD = 10 experimental dollars), and there was no show-up fee in our experiment. The payments were rounded up to the nearest quarter dollar.[12] Each session lasted for 45 minutes to an hour.

All laboratory sessions were computerized using z-Tree (Zurich Toolbox for Readymade Economic Experiments, Fischbacher, 2007) except for the evaluation task in the *FE* treatment. We wanted to make it as close to the real world evaluation process as possible in this treatment. For this reason, we asked the firms to write their messages in blank pieces of paper that they received from us. The messages were then passed on to the respective workers in an anonymous

[12] The rounding could have potentially distorted incentives for the workers, but we do not expect any significant impact because the subjects were unaware of the rounding during the time they participated in the experiment, and each subject participated in just one round and did not participate in more than one session. Even if we consider the possibility of contamination of information among subjects across sessions, based on the choices made in the experiment, we confirmed that the rounding did not affect our major findings (see footnote 15).

way, by matching the corresponding ID numbers.

4. Results

Since we implemented one-shot games in our experiment, we consider each decision in each treatment as one independent observation. Table 2 reports the summary statistics for all the treatments. Table 3 and 4 report the *p*-values from rank-sum tests and Kolmogorov-Smirnov tests, respectively.[13]

Table 2. Summary statistics (standard deviations in parentheses).

Treatments	Obs.	Wage (*w*)	Effort (*e*)	Firm Profit	Firm's Share of Surplus	Efficiency
NE	20	30.15 (21.71)	3.70 (2.77)	96.85 (18.35)	0.37 (0.32)	0.44 (0.28)
SE	18	29.06 (23.86)	4.33 (2.85)	104.28 (16.03)	0.50 (0.34)	0.50 (0.29)
FE	18	35.28 (23.18)	5.94 (2.46)	114.17 (25.83)	0.56 (0.42)	0.67 (0.21)

Table 3. Rank-sum test *p*-values for treatment differences.

	Wage (*w*)†		Effort (*e*) or Efficiency		Firm Profit†		Firm's Share of Surplus	
	SE	*FE*	*SE*	*FE*	*SE*	*FE*	*SE*	*FE*
NE	0.849	0.404	0.301	0.006***	0.224	0.021**	0.161	0.093*
SE	-	0.366	-	0.067*	-	0.102	-	0.232

Note: † denotes two-tailed tests. ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively.

We provide a detailed analysis below for different contract outcomes and features. In each of these discussions, our main focus is on the effect of evaluations and we accomplish that

[13] In Tables 3 and 4, only the tests related to wage and profit are two-tailed (non-directional) because we do not have any hypothesis about wages. Since our hypotheses about effort and efficiency are directional, all the tests related to effort and efficiency are one-tailed (directional).

by comparing the *NE* treatment, the one without evaluations, with each of the other two treatments (*SE* and *FE*) in which the firms were allowed to provide evaluations to their workers.

Wage and Effort:

The first thing we notice is that the pecuniary Subgame Perfect Nash Equilibrium does not hold in the standard gift-exchange game. In *NE*, both the average wage (30.15) and the average effort (3.70) are significantly higher than the equilibrium predictions of 1 (the two-tailed Wilcoxon signed rank test produces $p = 0.0001$ and $p = 0.0003$ for wage and effort, respectively). On the other hand, the effort choices are not as high as the optimal level ($e = 10$) either, and the difference is highly significant ($p = 0.0001$ in a two-tailed Wilcoxon signed rank test). This underscores the challenge of enforcing efficiency in a contract in which worker effort is not contractible.[14]

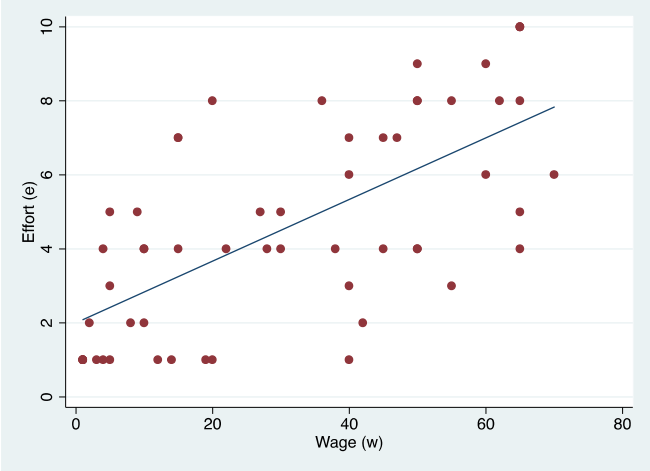
To understand how the provision of evaluation by the firms affect the wage the firms provide and the effort workers choose, we compare *SE* and *FE* to *NE*. If we focus on the average wages across the three treatments, as shown in Table 2, we find some variations across them. However, as the results of the rank-sum tests in Table 3 describe, the differences in wages were all insignificant.

If the firms' ability to evaluate their workers' effort equips the firms with additional bargaining power, then we expect to see higher effort choices by the workers in the evaluation treatments. We do actually find considerable effect of evaluation on effort choices, especially in

[14] This observation is consistent with the gift-exchange game literature we discussed in section 1.

FE. As Table 2 shows, the average effort increases from 3.70 in *NE* to 4.33 (a 17% increase over *NE*) in *SE* and to 5.94 (a 61% increase over *NE*) in *FE*. Based on the test results presented in Table 3, we find that, compared to *NE*, the increase in *SE* is not significant but that in *FE* is highly significant. Moreover, the average effort in *FE* is 37% higher than that in *SE*, which is significant at the 10% level. So, workers choose higher effort when they know that firms can provide evaluations, especially when the evaluations are free-form messages.[15]

Figure 2. Scatter-plot of wage (*w*) and effort (*e*), pooling data from *NE*, *SE* and *FE*.



Recall that in our experiment, workers choose their effort levels after observing the wages they have been offered. To gain an understanding of the nature of the relationship between wage and effort, we plot each pair of (w, e) from the three treatments in Figure 2. The positively sloped fitted line clearly shows that higher wages are associated with higher effort choices, as one would expect. Panel-(a) of Table 4 reports the results of two regressions of e on w which tell

[15] Based on the wages offered and the efforts chosen in the experiment, the rounding of payments could have affected incentives for the worker only in five of the twenty cases in *NE*. If the distortion actually happened, then the average effort level in *NE* would have been slightly lower. We have checked to make sure that this does not change our findings.

us how wages may have determined the effort choices.[16] In specification (1), we pooled the data from the three treatments. We find that the estimated coefficient of *w* is positive and highly significant.

Table 4. Regression of effort (tobit) and evaluation (ordered probit).

	Panel (a): Effort		**Panel (b): Evaluation**	
	(1)	**(2)**	**(1):** *SE*	**(2):** *FE*
Wage (*w*)	0.087 (0.013***)		–0.087 (0.031***)	–0.046 (0.018***)
Effort (*e*)			0.995 (0.284***)	0.832 (0.239***)
SE		0.235 (1.027)		
FE		3.503 (1.097***)		
*NE * w*		0.091 (0.020***)		
SE w*		0.111 (0.020***)		
*FE * w*		0.046 (0.020**)		
Female	-0.920 (0.575)	-1.197 (0.529**)	0.141 (0.589)	-0.063 (0.668)
Age	-0.001 (0.132)	-0.028 (0.119)	0.505 (0.249)	0.320 (0.339)
Constant	2.444 (2.648)	2.110 (2.456)		
Obs.	56	56	18	15
R^2	0.131	0.183	0.374	0.365
Log likelihood	–118.409	–111.419	–17.379	–16.532

Note: Standard errors are in parentheses. ***, ** and * denote statistical significance at 1%, 5% and 10% level, respectively.

In specification (2), we treat the treatments separately by including treatment dummies (NE is the reference treatment). The regression also included interaction terms between *w* and the treatment dummies to understand the effect of wage on effort across treatments. The estimated coefficients show that, consistent with our findings from the non-parametric tests for treatment differences (Table 3), effort in SE is insignificantly higher than that in NE while effort in FE is significantly higher than in NE even after controlling for wage.[17] Using a Wald test, we fail to

[16] The regressions also included gender and age of the subjects. As we mentioned earlier, we collected these data using a short questionnaire we had asked the subjects to respond to.

[17] Based on the estimated coefficients of the dummies and those of the interaction terms, effort in *FE* would be higher than in *SE* and *NE* when wage is lower than or equal to 50 and 77, respectively. 13 (out of the 18) observations in *FE* and 14 (out of the 18 observations) in *SE* had $w \leqslant 50$, while all the observations in *FE* and *NE* had $w \leqslant 77$. So, the comparison of effort between *FE* and *NE* is completely unaffected by wage.

reject the null hypothesis that the coefficients of the interactions terms are all equal to each other (p = 0.087).[18] So, not only the wages offered in the three treatments were similar and their differences were insignificant (Tables 2 and 3), the marginal effects of wage on effort across treatments were also very similar and their differences were insignificant.

Profit, Surplus Share & Efficiency:

The average firm profit was higher in the evaluation treatments than in the no-evaluation treatment (Table 2). Only the increase in *FE*, however, is significant (Table 3). The higher bargaining power of the firms that resulted from free-form evaluations is also evidenced by the share of total surplus firms enjoyed in the *FE* treatment. Recall that a worker's choice of e unilaterally determines the total realized surplus of $10e-c(e)$, and that given w, the firm's surplus $10e-w$ is also determined by e. As Table 2 shows, the average of the firm's share of surplus increases from 0.37 in *NE* to 0.50 in *SE* (a 35% increase) and to 0.56 in *FE* (a 51% increase). Once again, only the increase in *FE* is significant but at the 10% level (Table 3).

To measure overall efficiency, we calculated the proportion of the maximum possible surplus of 70 that was realized between a pair of firm-worker. The efficiency is, therefore, a nonlinear transformation of e because the realized surplus of the pair, $10e-c(e)$, is a function of e. We find from Table 3 that *NE* achieved an average efficiency of 0.44. Since e increases in *SE* and *FE*, the efficiency also increases in these two treatments. We find that the average efficiency increases to 0.50 (a 14% increase) in *SE* and to 0.67 (a 52% increase) in *FE*. Since efficiency is

[18] All the Wald tests reported in the paper are two-tailed.

solely determined by *e*, the statistical significances of the efficiency increases in *SE* and *FE* are exactly the same as those we found when analyzing the increases in *e* in these two treatments. That is, efficiency is higher in *FE* than in *NE* and *SE*, and the difference is significant at the 1% and 10% levels, respectively, but efficiency in *SE* is not significantly higher than that in *NE*. Overall, we find support for Hypotheses (b) and (c), but not for Hypothesis (a).

Evaluations:

Recall that there were five possible evaluation ratings in the *SE* treatment. For the purpose of our analysis we code them as follows: Very Good = 2, Good = 1, OK = 0, Bad = –1, Very Bad = –2. For the purpose of analysis, we also coded the free-form messages in the *FE* treatment.

To code the messages in *FE*, we recruited two coders who were naïve to our research hypotheses and did not help us with or participate in any of our sessions. The two coders worked independently, and we did not ask them to search for any specific words for coding the messages. The coders were not given any information regarding the actual choice of a specific subject who wrote the message. They, however, were given copies of the instructions because some messages were not necessarily comprehensible without the context. After reading the instructions, the coders were given a randomly ordered listing of all the messages.

We asked the coders to first categorize each message as "off-topic" or not. We instructed them to categorize a message as off-topic if the message is unrelated or irrelevant to evaluation of effort choice. If a message is not coded as off-topic, then the coders coded them using the

same five-point scale as in *SE*. If either of the coders categorized a message as off-topic, then we treated that message as off-topic. In practice, none was coded as off-topic. For all other messages, the average of the codes from the two coders were used in our analysis. Appendix B lists all the messages and their codes along with wages and efforts.

The mean evaluation score in *SE* is 0.11 while it is 0.53 in *FE* (excluding the off-topic messages). To understand the nature and the extent of association between effort and evaluation, we calculated the Spearman correlation coefficient between them. The value of the coefficient is 0.533 ($p = 0.023$) in *SE* and 0.757 ($p = 0.001$) in *FE*.[19] So, higher efforts were associated with better evaluations.

Besides efforts, the wages the firms offered could have affected the evaluations. For instance, a firm that offers a high wage is likely to expect a higher effort choice by the worker. As a result, for the same effort level, a firm that offered a higher wage may give a lower evaluation relative to a firm that offered a lower wage. Wages and evaluation ratings can thus have an inverse relationship, holding efforts constant. To explore how wages together with effort choices may have affected evaluations, we estimated ordered-probit regressions of evaluation ratings on *w* and *e*. The estimated coefficients are reported under panel-(b) in Table 4.[20] We find that, as expected, *w* affects negatively and *e* affects evaluations positively; the coefficients are significant in both the specifications under panel-(b).

[19] The *p*-values reported are for two-tailed *t*-tests.
[20] Since the evaluation codes in *SE* and *FE* are not necessarily comparable, we did not pool the data from the two treatments in our regressions.

5. Discussion

Based on the overall findings in our analysis of the data in the previous section, we can conclude that compared to the no-evaluation condition, free-form evaluation of the worker's effort by the firm increases effort and efficiency while structured evaluation does not. Is it the case that higher efforts do not obtain higher evaluations in the *SE* treatment and the effort choices in *SE* are not higher as a result? The estimation results in panel-(b) of Table 4 show that effort, in fact, has a significantly positive effect on evaluation in both the treatments. This makes us wonder why free-form evaluations were able to raise effort while structured evaluations were not, and we try to uncover the channel through which free-form evaluations succeeded in our experiment.

Several previous experimental studies show that when the format of the cheap-talk messages are structured by the experimenter, cheap talk has no impact on participants' behavior in trust and coordination games (Bracht and Feltovich, 2009; Charness and Dufwenberg, 2010; Dugar and Shahriar, 2015) while free-form cheap talk has a substantial impact (Charness and Dufwenberg, 2006; Cooper and Kuhn, 2014; Dugar and Shahriar, 2015). In all these studies, however, the cheap-talk communication precedes decision-makings whereas the evaluations in our experiment are provided *ex post*. As a result, the explanations for the success of the free-form cheap talk in the earlier studies do not fit in our case.

Taking psychological payoffs of the evaluations into account, there are two possibilities of why free-form evaluations work while structured evaluations do not. The first possibility is

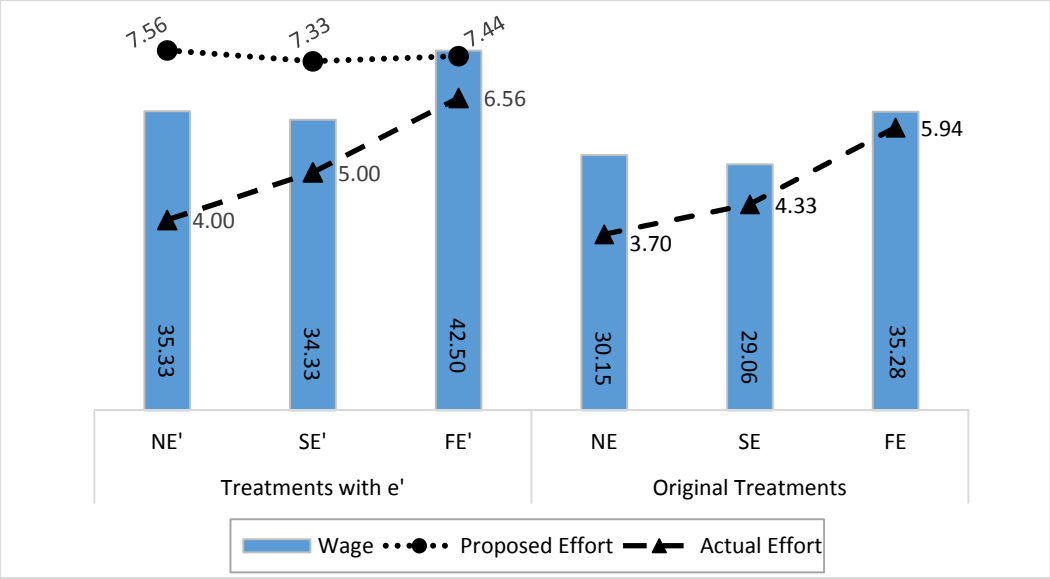
that the worker believes that the nature of the free-form evaluations (i.e. richer, endogenous and personalized by the evaluator) makes them more salient in the worker's preference and as a result, the worker exerts higher effort to win (avoid) positive (negative) evaluations. For example, a very high level of effort, say, e^* is supported as part of the SPNE in *FE* but not in *SE* because the strongest message "***Very Good*** (***Bad***)" in *SE* is not salient enough comparing with the strongest possible one in *FE*. Secondly, it also could be the case that the worker believes that the employer anticipates higher salience of free-form evaluations in the worker's preference and therefore expects the worker to exert higher effort. Acting on this belief, the worker chooses a higher effort level just to live up to the employer's expectation to avoid guilt (Charness and Dufwenberg, 2006). For instance, given a wage offer *w*, when a worker chooses an effort level $\tilde{e}$, she believes that she will be praised in *SE* but not in *FE* because the worker believes the firm's expected effort level in *SE* is $\tilde{e}$ (or less) while the worker believes the firm expects an effort level higher than $\tilde{e}$ in *FE*.

To disentangle the two explanations discussed above, we ran three additional treatments, one for each of the evaluation protocols: null, structured and free-form evaluations. The only difference between these treatments and the original ones is that the firm is now given a chance to express the effort level that the firm expects the worker to choose: the firm proposes an effort level (e') along with the wage *w* in stage 1; in stage 2, the workers are informed of not only *w* but also e' before the workers choose the real effort level *e*. e' is nonbinding in the sense that the workers are free to choose an effort level *e* different from e'. We label these three treatments as: *NE'*, *SE'* and *FE'*. Each of these three treatments was conducted the same way as its counterpart

in the previous three treatments, except for the additional decision on e' that each firm had to make. There were 6 sessions (108 participants) in total: 2 sessions (36 participants) in each of the three treatments.[21]

These treatments allow us to see (1) whether the firm's proposed effort is different across the evaluation protocols, (2) whether the free-form evaluation protocol still achieves higher effort levels than the structured evaluations protocol, and (3) whether the worker takes the firm's proposed effort into account when choosing effort. If the firm's proposed effort serves as a proxy for the firm's effort expectation and the data reject (3), then it would nullify the worker's belief about the firm's expectation as a possible explanation for the better performance of the free-form evaluation protocol.[22]

Figure 3. Treatment averages for wage, proposed and actual effort.



[21] Among these 108 participants, 61% were female and the average age was 19.7 years.
[22] There could be a distinction between the firm's stated effort and the firm's expected effort. One way to test this distinction is to ask firms to make an incentivized prediction of the effort level which we did not carry out in our experiment.

Figure 3 describes the treatment averages in the three new treatments along with those in the original three treatments. The average wages across the three new treatments with *e′* are not that different from each other and the differences are all insignificant ($p \geq 0.18$ in the corresponding rank-sum tests). We also find from Figure 3 that there were hardly any differences in the average proposed efforts across the three treatments; these differences are all also insignificant ($p \geq 0.62$ in the corresponding rank-sum tests). So, it appears that the firms do not expect differently from their workers across the three evaluation protocols.

Even though the firms' proposed efforts are similar across the treatments and the workers are informed of these proposed efforts in all three treatments, the effort choices of the workers were quite different. As Figure 3 shows, the average effort increases from 4.00 in *NE′* to 5.00 in *SE′* (a 25% increase over *NE′*) and to 6.56 in *FE′* (a 64% increase over *NE′* and a 31% over *SE′*). The increase in *FE′* is significant when compared to both of *NE′* (the rank-sum test produces $p = 0.008$) and *SE′* (the rank-sum test produces $p = 0.048$), while the increase in *SE′* is not significant when compared to *NE′* (the rank-sum test produces $p = 0.160$). Once again, we see that the firms do not pay much differently whether or not they are allowed to evaluate the workers, but the workers choose higher effort when they are given free-form evaluations of their effort choices. We also see that, for each evaluation protocol, the level of effort did not change significantly compared to the original treatment without *e′* (the rank-sum test produces $p = 0.743$ when *NE* and *NE′* are compared, $p = 0.460$ when *SE* and *SE′* are compared, and $p = 0.415$ when *FE* and *FE′* are compared).

Overall, we see that communicating the effort levels firms expect from the workers does

not change our original finding. That is, the free-form evaluation improves effort and efficiency while the structured evaluation protocol does not. Since the firms' proposed efforts are not significantly different between the two protocols, the workers' reading on the firm's proposed effort cannot possibly be an explanation for the success of the free-form evaluation protocol. This leads us to deduce that the workers' anticipation of larger psychological payoffs from free-form evaluations is the likely reason behind the success of this evaluation protocol.

6. Conclusions

We conduct one-shot gift-exchange experiment to explore the implications of approvals and disapprovals in contract design. We are unaware of any previous case of *ex-post* cheap-talk evaluations to workers in gift-exchange games.

Our three initial laboratory settings include a gift-exchange game without any evaluation, a gift-exchange game with structured evaluations, and a gift-exchange game with free-form evaluations. We find that, compared to the no-evaluation case, the efficiency (i.e., the chosen effort) increases significantly when the free-form evaluations are allowed, while such effect is not significant when the format of the evaluations are structured.

We then allow the firms to propose a nonbinding effort level (e'), where e' serves as an anchor to the worker's belief of the firm's expectation of efforts, given that proposed effort works as a proxy for expected effort. We find that free-form evaluations promote efficiency substantially regardless of whether or not the firms were allowed to propose e'. Structured evaluations, on the other hand, are always ineffective. Meanwhile, the proposed effort level e'

does not differ across different evaluation protocols, and the workers do not appear to take effort expectations of the firms into account. This suggests that the success in the free-form evaluation protocol is due to the saliency of the messages, rather than the channel of the worker's reading of the firm's proposed effort.

In sum, our experimental evidence shows that the *ex-post* expression of emotions may discipline the contractor to create surplus. For the labor market, our results suggest that free-form employer evaluations can potentially be effective in generating higher surplus.

References

Akerlof, G., 1980. "A Theory of Social Custom, of Which Unemployment May Be One Consequence." *Quarterly Journal of Economics*, 94: 749 - 775.

Barron, J., and K. Paulson-Gjerde, 1997. "Peer Pressure in an Agency Relationship", *Journal of Labor Economics*, 15: 234 - 254.

Bracht, J. and N. Feltovich, 2009. "Whatever you say, your reputation precedes you: observation and cheap talk in the trust game," *Journal of Public Economics*, 93: 1036-1044.

Brandts, J., G. Charness and M. Ellman, 2016. "Let's Talk: How Communication Affects Contract Design", *Journal of the European Economic Association*, 14: 943 - 974.

Brown, M., A. Falk and E. Fehr, 2004. "Relational Contracts and the Nature of Market Interactions", *Econometrica*, 72(3): 747 - 780.

Cabral, L. and A. Hortaçsu, 2010. "The Dynamics of Seller Reputation: Evidence from Ebay," *Journal of Industrial Economics*, 58 (1): 54 – 78.

Charness, G., R. Cobo-Reyes, N. Jiménez, J. A. Lacomba and F. Lagos, 2012. "The Hidden Advantage of Delegation: Pareto-improvements in a Gift-exchange Game", *American Economic Review*, 102, 2358 - 2379.

Charness, G., N. Du, C. Yang and L. Yao, 2013. "Promises in Contract Design", *European Economic Review*, 64: 194 – 208.

Charness, G. and M. Dufwenberg, 2010. "Bare Promises: An Experiment," *Economics Letters,* 107: 281-183.

Charness, G. and P. Kuhn, 2011. "Lab Labor: What Can Labor Economists Learn from the Lab?", Handbook of Labor Economics, Vol. 4, Part A, Ashenfelter, Orley and Card, David (eds.)

Cooper, D. and K. Kuhn, 2014. "Communication, Renegotiation, and the Scope for Collusion," *American Economic Journal: Microeconomics*, 6(2): 247-278.

Dellarocas, C., 2003. "The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science* 49 (10): 1407 – 1424.

Dugar, S., 2013. "Non-Monetary Incentives and Opportunistic Behavior: Evidence from a Laboratory Public Good Game", *Economic Inquiry*, 51(2): 1374 - 1388.

Dugar, S. and Q. Shahriar, 2015. "Restricted and Free-form Cheap-Talk and the Scope for Efficient Coordination," mimeo, San Diego State University.

Ellingsen, T. and M. Johannesson, 2008. "Anticipated Verbal Feedback Induces Altruistic Behavior", *Evolution of Human Behavior*, 29: 100 - 105.

Fehr, E., O. Hart and C. Zehnder, 2015. "How Do Informal Agreements and Renegotiation Shape Contractual Reference Points?" *Journal of the European Economic Association*, 13(1): 1 - 28.

Fehr, E., G. Kirchsteiger and A. Riedl, 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation", *Quarterly Journal of Economics*, 108(2): 437 - 459.

Fehr, E., A. Klein and K. M. Schmidt, 2007. "Fairness and Contract Design", *Econometrica,* 75(1): 121-154.

Fischbacher, U., 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10: 171- 178.

Gaechter, S., and E. Fehr, 1999. "Collective Action as a Social Exchange." *Journal of Economic Behavior and Organization*, 39: 341- 369.

Ghose, A., and P. Ipeirotis, 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics, *IEEE Transactions on Knowledge and Data Engineering*, 23: 1498 - 1512.

Grossman, S., and O. Hart, 1986. ''The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration'', *Journal of Political Economy,* 94: 691-719.

Hart, O., and J. Moore, 1990. ''Property Rights and the Nature of the Firm'', *Journal of Political Economy,* 98: 1119-1158.

Hollaender, H., 1990. "A Social Exchange Approach to Voluntary Cooperation", *American Economic Review*, 80: 1157 –1167.

Kandel, E., and E. Lazear, 1992. "Peer Pressure and Partnership", *Journal of Political Economy*, 100: 801 – 817.

Knack, S., 1992. "Civic Norms, Social Sanctions, and Voter Turnout," *Rationality and Society*, 4: 133 - 156.

Lindbeck, A., S. Nyberg, and J.W. Weibull, 1999. "Social Norms and Economic Incentives in the Welfare State", *Quarterly Journal of Economics*, 114: 1 - 35

Loch, C. H., B. A. Huberman, and S. K. Stout, 2000. "Status Competition and Performance in Work Groups", *Journal of Economic Behavior and Organization*, 43: 35 - 55.

Masclet, D., C. Noussair, S. Tucker, and M.-C. Villeval, 2003. "Monetary and Non-Monetary Punishment in the VCM." *American Economic Review*, 93: 366 - 380.

Rege, M., and K. Telle, 2004. "The Impact of Social Approval and Framing on Cooperation in Public Goods Situations", *Journal of Public Economics*, 88: 1625 - 1644.

Stajkovic, A., and F. Luthans, 2003. "Behavioral Management and Task Performance in Organizations: Conceptual Background, Meta-Analysis, and Test of Alternative Models", *Personnel Psychology*, 56: 155 - 194.

Tyler, T. R. 1990. *Why people obey the law*. New Haven: Yale University Press.

Uhlaner, C. J., 1989. "Rational Turnout: The Neglected Role of Groups", *American Journal of Political Science*, 33: 390 - 422.

Xiao, E. and D. Houser, 2009. "Avoiding the Sharp Tongue: Anticipated Written Messages Promote Fair Economic Exchange", *Journal of Economic Psychology*, 30: 393 - 404.

Appendix A: Instructions

Welcome to an economic decision experiment. If you have any question, please raise your hand and one of us will come to your desk to answer it. From now on till the end of the experiment any communication with other participants is not permitted.

You will be paid in cash at the end of the experiment. Your earnings in the experiment are calculated in experimental dollars (ED) during the experiment; then at the end of the experiment they are converted to USD at a rate of 1 USD=10 ED. **Therefore, the more ED you earn in the experiment, the more money you receive.**

You have been randomly assigned a registration number shown at the top right of this page. Each participant in this room has been assigned a different registration number and will be identified by this number. This number is to remain private during and after the experiment. No real names or identities will be used in the experiment.

Before the experiment starts, you will be given a starting balance of 90 ED. There is an even number of participants in today's session. Half of the participants will be assigned the role of an "employer" and the other half are "employee". If your registration number is odd then you are an employer and if it is even then you are an employee. Each employer will be randomly matched with one employee. The matching is anonymous, in other words, the employer (employee) will not know the identity of the employee (employer) whom he/she is matched with. Your decision can only be observed by the participant that you are matched with, and cannot be observed by others. In the experiment, each pair of employer-employee will make some decisions (using the computers in front of them) which will determine their earnings, as discussed below.

Procedures

/ Instructions in *NE*, *SE* and *FE* are in slash. /
/

- In stage 1, **the employer chooses** a wage w (w is any integer between 1 and 100, including 1 and 100) and pays the wage w to the employee immediately.
- In stage 2, the employee learns about the wage the employer has paid and then **the employee chooses an effort level** e (e is any integer between 1 and 10, including 1 and 10).

/

// Instructions in *NE'*, *SE'* and *FE'* are in double slash. //
//

- In stage 1, **the employer chooses** a wage w (w is any integer between 1 and 100, including 1 and 100) and pays the wage w to the employee immediately, and **proposes** to his/her employee an effort level e' (e' is any integer between 1 and 10, including 1 and 10) which the employer expects his/her employee to choose in stage 2.

- In stage 2, the employee learns about the wage his/her employer has paid and the effort level *e'* his/her employer has proposed, and then **the employee chooses** the real effort level *e* (*e* is any integer between 1 and 10, including 1 and 10). The real effort level *e* chosen by the employee can either be the same as the proposed *e'* or be higher or lower than the proposed *e'*.

//

(Instructions in *NE* and *NE'* are in parenthesis.)

(At this point, the experiment ends and the earnings are calculated based on *w* the employer has chosen and *e* the employee has chosen. (The next section discusses how the earnings are calculated.) The participants are paid their earnings in cash and in an anonymous manner. To do this, we will put participants' earnings in separate envelopes, write the registration numbers on the top of the envelopes and lay the envelopes on a table; on your way out you will pick up the envelope with your registration number.)

[Instructions in *SE* and *SE'* are in brackets.]

[

- In stage 3, after the employer learns about the effort level the employee has chosen, the **employer rates the employee's effort** to express satisfaction or dissatisfaction. The employer can choose from five possible ratings: Very Good, Good, OK, Bad, Very Bad. Then the employee learns about the rating the employer has chosen to give.

At this point, the experiment ends and the earnings are calculated based on *w* the employer has chosen and *e* the employee has chosen. (The next section discusses how the earnings are calculated.) The participants are paid their earnings in cash and in an anonymous manner. To do this, we will put participants' earnings in separate envelopes, write the registration numbers on the top of the envelopes and lay the envelopes on a table; on your way out you will pick up the envelope with your registration number.]

{Instructions in *FE* and *FE'* are in brace.}

{

At this point, the earnings are calculated based on *w* the employer has chosen and *e* the employee has chosen. The employer and the employee are then informed of their earnings. (The next section discusses how the earnings are calculated.)

- In stage 3, the employer will have an option to send a written message to his/her employee using the "Message Sheet". In the message, the employer is not allowed to

identify him or herself by name or number or gender or appearance. The experimenter will monitor the messages. Violations (experimenter discretion) will result in the employer receiving no money, and the employee receiving the average amount received by other employees. Other than these restrictions, the employer may say anything that he/she wishes in this message.

The experimenter will collect all the Message Sheets in a box and then pass on each employer's message to his/her employee in an anonymous way. To do this, the experimenter will put the Message Sheets upside down on a table with the employees' registration numbers written on them; each employee will come to the table and pick up the Message Sheet with his/her registration number.

At this point, the experiment ends and the participants are paid their earnings in cash and in an anonymous manner. To do this, we will put participants' earnings in separate envelopes, write the registration numbers on the top of the envelopes and lay the envelopes on a table; on your way out you will pick up the envelope with your registration number. }

How earnings are calculated

The earnings (in ED) are calculated using the following equations:

For employer:

Earnings = 90 (starting balance) + 10* effort level (e) – wage (w)

For employee:

Earnings = 90 (starting balance) + wage (w) – cost of effort level (e)

As the above equation shows, employees bear a cost for their effort. The higher the effort chosen by the employee, the higher cost the employee bears. The relationship between the effort and the cost of effort is as follows:

Effort *e*	1	2	3	4	5	6	7	8	9	10
Cost	0	1	3	5	8	12	16	20	25	30

To calculate your earnings in USD that you will be paid at the end of the experiment, your ED earnings will be divided by 10.

Calculator

To help you calculate earnings under different possibilities, we provide you an earnings calculator on the screen. You can switch between the decision screen and the calculator screen at any time during the experiment by pressing "Alt+Tab" on the keyboard. On the calculator, choose a certain wage (w) that the employer may pay and you will see the earnings of both the

employer and the employee for all possible effort levels (e) the employee may choose, given the wage w.

To check whether you understand the rules, there will be a short quiz on the computer screen before the experiment starts.

Do you have questions on the instructions and the procedure? If you have any question, please raise your hand. One of us will then come to you.

Appendix B:

Messages in the *FE* Treatment

Firm's Wage	Worker's Effort	Firm's Message	Coded Message
15	7	Thank you for all your hard work. You've done an amazing job working for me.	2
65	10	Good choice!	2
40	3	If I pay you 40 you should at least give me 6 effort level so I make $11 and you make $12. Right now I get $8 and you get $13 which is not fair. This is not how it works in actual world. Hopefully you can give me more effort so we can both leave here happy :-). I am even fine with 5 effort, I am just trying to get $10 at least haha. We should make even amount so we're fair :-).	-2
10	4	Thank you for working for my organization. Have a great day.	1
4	4	I chose 4 because I wanted to play it safe where if you chose 1. We would both still get ~$9.50 each. I don't think I will be changing because what if I increase your wage, but you decrease the effort. Possibly we can do w=65 and effort=10. We would both get 12.50. Not sure the...I want this to be fair for both.	0
70	6	Not sure what to say, but I thought you were going to put more effort.	-2
15	7	Hi. Thanks for being a great employee!	2
40	6	Cool we both got 11 dollars. Good work. If you choose effort 7 or 8, you still would have gotten 11 dollars but I would have gotten 12 or 13. I guessed that the employee would chose an amount that was even for both of us and I was right. Happy Halloween [a pumpkin picture]	2
65	10	Thanks for the effort out there! I predict we will have a highly beneficial professional relationship for the both of us.	2
45	4	Great effort. Satisfying work for the wage earned. More effort could be put in, but still enough.	2
55	3	THANKS! AIMING FOR THE BOTH OF US TO GET MORE THAN $10 BUT I GUESS NOT. HAVE FUN! ENJOY THE REST OF YOUR DAY!	-1
30	5	Thanks for the effort! I am just happy for free lunch. Cheers.	2
36	8	I have 134 ED.	0
65	10	I searched around a bit to ensure we both would be receiving hopefully the same. Thank you :).	2
5	3	YOOO. Happy Halloween.	0
50	4	Happy Halloween Fam. [a ghost picture] Rock on…Thanks for the $$$$	0
5	5	I chose 5 so that we would both have an equal amount of money at the end, surprised you chose an effort level of 5, but thanks!	2
20	8	I WILL INCREASE THE WAGE FOR THE MAXIMUM OUTPUT FOR US BOTH.	1

Messages in the *FE'* Treatment

Firm's Wage	Firm's Effort Proposal	Worker's Effort	Firm's Message	Coded Message
45	9	9	We both made good money with 45 and 9. So stick to it. :)	2
65	10	9	I see you do not want even earnings.	1
30	6	4	Employee, please work harder or you may be terminated from the workplace. - Your boss.	0
35	5	4	I tried to be fair. Have a nice day & good luck with finals!	1
30	7	7	Excellent work! You are a valuable asset to this company! Please keep up the good work!	2
45	7	7	I hope that the wage and effort proved to be beneficial as it was towards myself. Have a wonderful day!	2
20	4	1	--	0
50	8	8	($120,$120) gives the highest even pay between employer & employee. Fair for both of us! :)	2
75	5	10	Have a great day!	1
60	6	10	I offered you 60 ED and the effort # was 10. So you received $12. Good job!	2
5	7	1	--	0
50	8	8	I chose a wage of 50 and an effort of 8. So, we can both earn the same amount. Sharing is caring.	2
50	8	8	Hi. Good deal. :)	2
65	10	10	--	0
15	9	1	Compromise effort level: 4	-2
15	7	3	I agree with your effort level because it would have put us almost even. I put 15 cause I thought it would help us both.	1
60	10	10	I thought the wage was the most mutually beneficial. Thanks for not screwing me over. Thanks, your employer.	2
50	8	8	--	0