On the Impact of Incentives in Organizations - Experimental Studies on Compensation, Team Production, and Performance Feedback

Inauguraldissertation zur Erlangung des Doktorgrades der

Wirtschafts- und Sozialwissenschaftlichen Fakultät der

Universität zu Köln

2006

vorgelegt von

Diplom-Volkswirtin Kathrin Pokorny

aus Mülheim an der Ruhr

Referent: Prof. Dr. Dirk Sliwka
Korreferent: Prof. Dr. Axel Ockenfels
Tag der Promotion: 15.01.2007

Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft DFG (IR 43/1-1 and KR 2077/2-3) is gratefully acknowledged.

I am deeply grateful to my first supervisor Dirk Sliwka, for his continuous support, encouragement, advice, and humor which made life at the seminar a great experience.

I thank Axel Ockenfels for agreeing to be my second supervisor and the opportunity to use the resources at the Cologne Laboratory for Economic Research. I also thank his staff for their support.

I thank Alwine Mohnen not only for being my coauthor and colleague but also for being a very good friend.

I thank my dear colleagues, the 'boys' René Fahr, Patrick Kampkötter, Tom McKenzie, and Matthias Wibral and the 'girls' Petra Hagemann, Christine Harbring, and Petra Nieken for their open eyes and ears, substantial help in various ways and for being such a good company. I also thank Beate Ommer for looking after all of us.

I thank Anastasia Danilov, Evgenji Pechimenko, Verena Scherer, and the whole PWL student staff for their nerves and patience when dealing with me and my experiments.

I am deeply indebted to my parents for providing me with all necessary financial and emotional assets and for bearing my constant moaning.

I thank my brother Andi and my sister Tine that I can always rely on their critics to be honest.

I thank all members of the Manthei-Müske-Vogel-Clan and Achim and Melanie Pokorny for sharing the thrills and troubles with me and for keeping fingers crossed.

I thank 'Gruppe Mülheim' for keeping me from working too much on the weekends.

Finally, I thank my husband Daniel who at least suffered as much as I did for this thesis...and probably even more.

Contents

[1] This chapter is based on Pokorny (forthcoming).
[2] This chapter is based on Mohnen and Pokorny (2006a).

[3] This chapter is based on Mohnen et al. (2006).
[4] This chapter is based on Mohnen and Pokorny (2006b).

Chapter 1

Introduction

> Most of economics can be summarized in four words:
> ‘People respond to incentives’. The rest is commentary.
> [*Landsburg (1995) p. 3*]

This citation of Landsburg shows that people’s reactions to incentives constitute one of the most important characteristics of economic theory which goes along with a very specific idea of man. This typical homo economicus acts egoistically such that he always seeks to maximize his own utility. Under the assumption that people behave purely selfishly the implementation of optimal (monetary) incentives seems to be a relevant question not only in politics or social life but also in organizations.

This thesis is concerned with the impact of monetary incentives in firms. This issue has attracted a lot of attention in recent decades. Not only theoretical (for seminal papers see e.g. Holmström (1979), Grossman and Hart (1983), Hart and Holmström (1987)) but also empirical researchers (e.g. Paarsch and Shearer (1999), Lazear (2000), Shearer (2004)) have made a huge effort in exploring the optimal contract design and its actual effects on effort exertion. For this purpose the relationship between employer and employee and the resulting divergence of interests has been analyzed. However, more recently several empirical papers have provided evidence that the image of the homo economicus is all too simple and does not display all human traits which are important in practice. In particular the positive effect of incentives could not always be confirmed (e.g. Gneezy and Rustichini (2000a), Heyman and Ariely (2004), Fehr and Gächter (2002)). These deviating findings cannot be explained by standard economic theory alone as they represent psychological rather than economic aspects of decision making. Among the first to add psychological elements to standard neoclassic preferences were Geanakoplos et al. (1989) or Rabin (1993), who tried to increase the explanatory power of the economic theory.

The main objective of this thesis is to investigate whether a particular action taken by the employer indeed triggers the desired behavior of the employee and what motivational mechanisms underlie the employee's behavior when it does not. Each of the subsequent chapters deals with a specific aspect of the incentive problems that arise in organizations. These research questions are analyzed theoretically and experimentally.

Chapters 2 and 3 investigate the impact of different compensation schemes on the effort decision. One of the most intuitive and therefore widespread types of incentives is the introduction of performance-dependent compensation in the form of a piece rate. In Chapter 2 we examine the impact of varying piece rates on work performance in an economic experiment. If people react according to the standard homo economicus higher piece rates should ceteris paribus lead to higher effort levels. We use two simple real effort tasks to check the validity of this hypothesis. However, we do not find a monotonically increasing relationship between performance and strength of incentives but an inversely U-shaped pattern. This contrasts not only with standard economic theory but also with observations in a previous real effort experiment presented in Gneezy and Rustichini (2000a). To give a possible explanation for our findings we develop a simple theoretical model including reference dependency and loss aversion into the employee's preferences. This approach captures the idea of individuals who do not only consider their absolute monetary payoffs but evaluate them compared to reference wage and was first established by Kahneman and Tversky (1979). In case of reference dependent and loss averse agents there is a non-monotonic relationship between incentives and performance and therefore high powered incentives might even lead to decreasing effort levels.

However, if reference dependency and loss aversion are general traits of human preferences and hence influence individual decision making, the fixed part of the compensation also plays a crucial role in setting optimal incentives. This is especially important considering that the fixed base salary often constitutes a large part of the total compensation (see e.g. Murphy (1999)). Although simple principal agent models would lead us to predict that the fixed wage does not affect the employee's optimal effort choice, the model including reference dependency and loss aversion demonstrates a weakly decreasing relationship between the height of the fixed wage and employee performance. In contrast to this recent economic experiments in particular those exploring the gift exchange game have been able to show that individuals exert effort above the theoretically optimal level for increasing fixed wages despite the absence of any incentives (see e.g. Berg et al. (1995), Fehr et al. (1993), Gächter and Falk (2002)).

In Chapter 3 we develop alternative hypotheses on this question and investigate actual human behavior where we again use a real effort task. For this purpose we rewarded all participants with the same performance dependent piece rate but exposed them to three differently high fixed wages. Surprisingly, we observe the highest effort levels with the low and the high fixed wage. The worst performance, however, is provided under an intermediate fixed wage. This outcome can neither be explained by standard preferences nor by reference dependency and loss aversion alone. As social preferences and social norms have been shown to have an important effect on individual decision making (e.g. Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Engelmann and Strobel (2004)), we argue that it is likely that subjects in the experiment were also led by a social norm demanding a certain effort level with a very high fixed wage.

Social norms seem to play an important role in particular when people interact with each other. Chapters 4 and 5 therefore deal with incentive problems with social interaction. In team production fairness considerations often refer to the colleagues or team mates.

In Chapter 4 we use this idea to provide an explanation for the peer pressure phenomenon which has been shown to increase individual contributions in team production. For this purpose we develop a two-period model in which the two team members are inequity averse regarding their noisy contributions. We find that the effect of inequity aversion strongly depends on the informational setting. For hidden contributions the agents act as if they were purely selfish. However, when contributions are made transparent at an interim stage the agents adjust their second period effort levels according to the intermediate information they received. That is the agent with higher past contribution will decrease his future contribution and vice versa in order to counterbalance the difference between each other. Anticipating this adaption the agents exert higher effort in the first period to increase the counterpart's effort adjustment in the second. This form of peer pressure leads to a reduction in free-riding and thus to more efficient outcomes.

Similar to the previous chapters we check the theoretical results in a real effort experiment observing strong evidence for the conclusion that transparency indeed leads to a considerable effort adjustment such that inequity in contributions is reduced. Yet, the team mates' absolute effort adjustments do not seem to be equal irrespective of whether a subject contributed more or less than his counterpart in the previous period. Instead we find asymmetric adjustments which are higher for subjects who have taken the lead in the preceding period and lower for those having made lower contributions. Still, the results show that obviously the occurrence of a social norm together with transparent conditions leads to higher aggregate outputs.

Setting incentives in employer-employee relationships is often very sensitive to the agent's productivity or ability. A standard assumption in economic theory is that the agent has better information on his own ability than the principal. One of the most famous solutions to this asymmetry in information is job market signaling (see Spence (1973)). But, if an agent joins a firm directly after leaving university or changes to a new job it also seems possible that the principal or supervisor has a better perception of the agent's actual ability for the specific tasks corresponding to the job. If this is the case the agent might be interested in the information the principal has when considering which effort level to exert.

In Chapter 5 we theoretically and experimentally examine a situation in which the principal has better information about the agent's ability than the agent himself does. The principal can inform the agent about his ability by giving him performance feedback but there might be incentives for her to lie about it. Analyzing a simple signaling model yields the following main results: The principal tells the truth if there are sufficiently high costs of lying. When the principal's marginal returns on the agent's effort increases, deception is more likely to occur.

We observed ambiguity, however, when testing these theoretical results in a lab experiment with abstract effort choices. On one hand, the data provide evidence implying that subjects indeed suffer from costs of lying in this context. On the other hand, we do not find more deception with higher marginal returns on the agent's effort for the principal. Instead we observe exactly the opposite pattern: higher marginal returns on the agents' effort lead to more honest principals. This might again be induced by reference dependent preferences and loss aversion or the impact of a social norm.

The results of the four experiments presented in this thesis demonstrate that standard assumptions on the individual's preferences often do not suffice to explain human behavior. However, when setting incentives the right anticipation of the agent's reaction is critical for their success. Hence, the development of alternative theories integrating more human motives than just the maximization of one's own absolute utility seems to be important.

Chapter 2

Piece Rates and Performance[1]

2.1 Introduction

The question of optimal incentive schemes in principal agent relationships with hidden action has been the object of research for many years. A well-established result of most standard hidden action models is that higher incentives ceteris paribus lead to higher performance. Field studies and experiments present evidence for this conclusion (e.g. Lazear (2000); Paarsch and Shearer (1999); Dickinson (1999)).

Still standard results are controversial. Several experimental studies in economics were recently able to show what psychologists have already been claiming for some time namely, that the introduction of incentives does not inevitably stimulate higher effort choices. Moreover, cases exist where introducing an incentive contract even reduces effort. This phenomenon emerges in real effort experiments (Gneezy and Rustichini (2000a); Gneezy and Rustichini (2000b); Heyman and Ariely (2004)) as well as in experiments including abstract effort choices (e.g. Fehr and Gächter (2002); Irlenbusch and Sliwka (2003)).[2,3]

A frequently cited key concept explaining certain anomalies is motivation

[1] This chapter is based on Pokorny (forthcoming).

[2] For an overview see Frey and Jegen (2001).

[3] Examining the willingness of citizens to accept a nuclear waste repository in their neighborhood, Frey and Oberholzer-Gee (1997) find a similar effect in the field of politics.

crowding out (e.g. Frey (1997)). In this framework two types of motivation are specified, namely intrinsic and extrinsic motivation.[4] The core of the theory in this context is that implementing a performance dependent compensation scheme might strengthen extrinsic incentives while diminishing intrinsic motivation. As a consequence the effect of introducing a variable compensation is ambiguous. Recently economists tried to include these considerations into economic models (e.g. Bénabou and Tirole (2003); Frey (1997)). Bénabou and Tirole (2003) interpret incentives offered by the principal as a signal of task attractiveness or difficulty. In their model the principal has perfect information on the actual attractiveness of the task but the agent only observes an imperfect signal on it. If the actual attractiveness of the task is low the principal fears that the agent receives a low signal of task attractiveness and therefore ensures performance by offering higher extrinsic rewards. That way Bénabou and Tirole (2003) show that higher rewards may diminish the agent's performance as they represent the principal's pessimistic information on task attractiveness. Frey (1997) introduces a simple principal agent model including intrinsic motivation where the agent's utility increases in wage and decreases in effort. If the principal implements stronger incentives the impact of the intervention on the agent's effort choice is not clear. Frey (1997) distinguishes three effects caused by the principal's intervention namely a price effect, an enhancing effect and a crowding out effect. The price effect simply denotes the effect of higher opportunity costs of lower effort levels. Representing a positive perception of the principal's intervention the enhancing effect amplifies the impact of the price effect. In turn the crowding out effect refers to a negative assessment of the increase in incentives producing lower effort choices. Still, the net effect is hardly predictable since price and enhancing effect point in the same direction, while the crowding out effect affects the opposite. Particularly it seems feasible that for lower interventions the crowding out effect dominates the price effect resulting in reduced effort whereas higher interventions cause increasing effort choices due to the

[4] Intrinsic motivation denotes an inner drive to do things (e.g. pleasure) while extrinsic motivation describes a behavior driven by rewards outside the individual. The idea of intrinsic motivation was first introduced by Deci (1971) and DeCharms (1968).

prevailing influence of the price effect.

However, the concept of intrinsic motivation is disputed among psychologists as psychological meta studies present ambiguous results (e.g. Eisenberger and Cameron (1996); Deci et al. (1999a); Deci et al. (1999b)).[5] Furthermore, in experiments with abstract effort choices motivation crowding theory fails to provide sensible explanations, as there is no task stimulating intrinsic motivation.

Another part of the related literature examining the relationship between incentives and performance indicates the relevance of reference dependent preferences (e.g. Fehr and Götte (2005); Camerer et al. (1997)). Fehr and Götte (2005) find decreasing effort choices with stronger incentives if the reference income has been exceeded. In that case higher piece rates have a diminishing impact on work effort.

This chapter investigates the influence of varying piece rates on work performance. For this purpose we conducted a real effort experiment at the Universities of Bonn and Cologne. Real effort conditions were chosen to prevent subjects from restraining on income distribution and more importantly to generate real disutility from higher effort levels. The design of the experiment has been inspired by a real effort experiment conducted by Gneezy and Rustichini (2000a) who investigated the influence of varying piece rates on effort for an IQ test task. In one condition subjects were paid a participation fee only while subjects in the other three conditions were paid different piece rates for every correct answer in addition to the participation fee. Gneezy and Rustichini (2000a) find a V-shaped relationship between effort and intensity of incentives. Their main result is that subjects in the condition without any incentives outperformed those who were paid a very low piece rate. Moreover, the results indicate that higher piece rates lead to higher effort when only conditions with positive piece rates are considered. According to Frey (1997)'s motivation crowding theory one might argue that for the low incentive group the crowding out effect dominated the price effect while for the higher incentive condition the price effect prevailed. As another ap-

[5] For a good overview on the literature on intrinsic motivation see Kunz and Pfaff (2002).

proach Gneezy and Rustichini (2000a) offer an explanation stating that the fixed wage group thought the work on the IQ test to be part of the contract concluded with the experimenters. Introducing the variable compensation completed the contract with regard to the meaning of the fixed wage as a show-up fee.

To test the effect of different strengths of incentives on the subjects' performance we introduced four different compensation schemes and two tasks. The first task we used was an IQ test to check whether the Gneezy and Rustichini (2000a) results were robust for our design. Hence, these conditions are closely related to theirs. As motivation crowding out might be a reason for the failure of incentives the second objective of this chapter is to explore how far characteristics of the tasks influence the effect of incentives on the subjects' effort. Since the type of task might be important with regard to the ability of stimulating intrinsic motivation we introduced a counting task in addition to an IQ test. In the counting task conditions the subjects were asked to count the frequency of a particular number in blocks of random numbers which was different with respect to some features. First it did not consist of various exercises but only one task and second it required a lot of concentration. Thus, we are able to compare performance for two tasks with putative different ability of stimulating intrinsic motivation.

For the IQ test task we find two main effects: (1) Very low incentives induce higher performance than offering no incentives. (2) High piece rates reduce performance compared to very low incentives. So the data indicate an inversely U-shaped relationship between effort levels and incentive intensity contrasting Gneezy and Rustichini (2000a). The data of the counting task merely show the latter effect of lower effort levels with high powered incentives. In order to explain the experimental results we develop a simple principal agent model accounting for the agent's loss aversion.

The remainder of this chapter is structured as follows: Section 2.2 describes the experimental design and procedures and develops hypotheses on the results. Furthermore the results are presented and discussed. In Section 2.3 the model including reference dependency is introduced and developed. The last section concludes.

2.2 The Real Effort Experiment

2.2.1 Experimental Set-up and Procedures

We conducted eight different conditions with two different tasks and four different compensation schemes. The first task was comparable with the Gneezy and Rustichini (2000a) IQ test (IQ) consisting of 48 exercises extracted from a book for logical training. No special skills were necessary to answer these questions. With the other task (CN) participants were required to count the number of 'ones' and 'sevens' out of 24 blocks of random numbers put together by a computer program.[6] For both tasks the maximum number of points that could be achieved was 48.

All participants were told that they would always get 5€ as a participation fee. The participants who were assigned to the first compensation scheme received the participation fee only (no incentive = NI). The other groups earned 0.01€ (very low incentive = VLI), 0.05€ (low incentive = LI), and 0.5€ (high incentive = HI) respectively for each point scored in addition to the show-up fee.[7]

The subjects who worked on the IQ task received a point for every correct answer but 0.5 points were subtracted for any incorrect answer.[8] In the CN task an answer was valued correct if it corresponded to the correct number of 'ones' and 'sevens' respectively with a deviation of one. If the exact number of 'ones' in a block was for example 30, subjects who counted 29, 30, or 31 'ones' received one point.

After all individuals had entered the room they were requested to sit down on seats marked with a pen and paper. After everybody had sat down the supervisors handed out the exercises. Each of them was covered with a sheet with the instructions printed on.[9] The subjects had 30 minutes to work on the task. After exactly 30 minutes a bell rang and the supervisors handed out forms and asked the subjects to fill in their answers. Additionally

[6] For examples of both tasks see appendix.

[7] 1€ was about 1$ at the time of the first sessions of the experiment.

[8] That was to prevent subjects from guessing which seemed to be necessary as there were several multiple-choice questions.

[9] For a translated version of the instructions see appendix.

a questionnaire was fixed to the form. The participants were granted another 5 minutes to copy the answers into the form and complete the questionnaire. The questionnaire included several questions such as gender, age, years of study etc. The phase in which the subject answered the questionnaire was critical for the success of the experiment. Subjects might have continued working on the tasks rather than answering the questions and therefore attain higher scores. We discuss this problem again in the results chapter. The supervisors collected the forms and informed the subjects when and where they were requested to come to receive their payment. The whole procedure took about 45 minutes.

We conducted three experimental sessions. The first one took place in November 2002 at the University of Bonn. The two other sessions were run at the University of Cologne in July 2003 and October 2005 respectively.[10] In total 267 of the universities' undergraduate students of various disciplines participated in the experiment. The remuneration was paid out about one week after each session.[11]

2.2.2 Hypotheses

According to standard theoretical predictions, participants of conditions with higher incentives exhibit higher performance. Those participants who receive the fixed wage are expected to exert the lowest effort level. The monotonically increasing relation between strength of incentives and effort can be expressed by

Hypothesis 1:

$$e^{NI} < e^{VLI} < e^{LI} < e^{HI}$$

where e denotes the effort level under the respective compensation scheme. Considering motivation crowding theory things appear different. Due to the

[10] In the sessions in November 2002 and July 2003 we conducted the NI, LI and HI conditions. The VLI session took place in October 2005. For this session we used the online recruitment software ORSEE developed in Greiner (2003). We thank an anonymous referee for the suggestion to conduct a condition with a very low piece rate.

[11] In each session there were a few participants who did not come to collect their payment.

	No incentive	Very low incentive	Low incentive	High incentive
Mean	16.6	21.8	18.7	16.4
Median	17.5	20.5	19.5	15.5
Stand. dev.	8.2	9.6	9.2	9.8

Table 2.1: Descriptive statistics IQ task

multitude of different effects a precise prediction is difficult to make. Assuming that for small interventions the crowding out effect dominates the price effect but for high interventions the price effect dominates the crowding out effect we might observe a non-monotonic relation. That is, participants in the VLI condition should perform worse than those in the NI condition. Participants in the HI condition in turn should yield higher scores than those in the VLI and LI conditions. As the IQ conditions resemble the Gneezy and Rustichini (2000a) experiment a similar effect might be expected i.e. subjects in the fixed wage condition and in the high incentive condition exert significantly more effort than the very low (and low) incentive group.

Hyptothesis 2:

$$e^{NI} > e^{VLI} < e^{LI} < e^{HI}.$$

Due to the monotony of the CN task different observations are likely because compared to the IQ its potential to stimulate intrinsic motivation or work pleasure may be lower. Thus, a monotonically increasing relation between wage and effort according to Hypothesis 1 could be expected with these conditions.

2.2.3 Results

For the analysis of the results we use higher point scorings to approximate higher effort choices. First we look at the performance data of subjects who participated in the IQ test. The descriptive statistics are presented in Table 2.1. The means are illustrated in Figure 2.1 where ninety percent of the scores

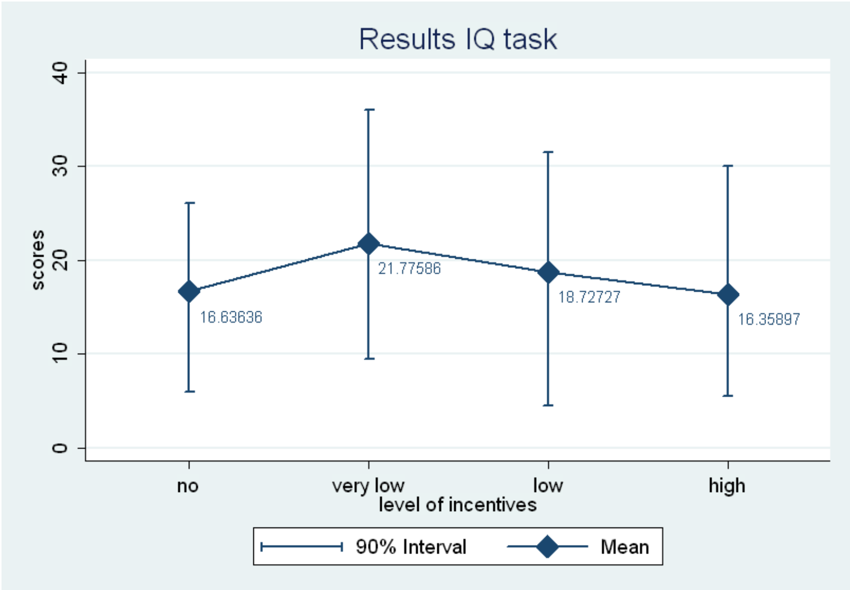


Figure 2.1: Mean scores IQ task

lie within the interval presented along the vertical lines.[12] The descriptive statistics show that subjects in the VLI condition yield the best results. For both no incentives and high incentives the efforts levels are almost equally low. All in all, Figure 2.1 depicts an inversely U-shaped relation between strengths of incentives and IQ test scores. In order to check whether the performance differences are significant we ran a pairwise comparison using a two-sided t-test. The results are displayed in Table 2.2.[13] Indeed, subjects who are offered very low incentives perform significantly better than those in the NI and the HI conditions. This provides evidence for a positive effect of very low piece rates on work effort in this context. The difference between efforts under very low incentives and low incentives is not significant at any conventional level.

Obviously, the performance increase between the no incentives and very low incentives groups contrasts the Gneezy and Rustichini (2000a) findings. Hence, we cannot confirm that a very low piece rate causes lower effort levels

[12] The entire distribution of the data is reported in Table 2.6 in the appendix.

[13] We additionally ran all comparisons using a non-parametric Mann-Whitney-U-test. The results are shown in Table 2.7 in the appendix.

	No incentive	Very low incentive	Low incentive	High incentive
No incentive		**(**)	-	-
Very low incentive			-	**(**)
Low incentive				-

**significant at 5%, * significant at 10%; - not significant
Results of one-sided test in parentheses

Table 2.2: Results t-test IQ task

	No incentive	Very low incentive	Low incentive	High incentive
Mean	25.5	26.4	25.3	22.4
Median	25	25	26	22
Stand. dev.	10.3	9	8.3	7.3

Table 2.3: Descriptive statistics CN task

compared to a fixed wage. In addition, Gneezy and Rustichini (2000a) find increasing scores with higher incentives if they have once been introduced. However, subjects in the HI condition provided significantly less effort. So our results do not only contrast standard theoretical predictions but also the Gneezy and Rustichini (2000a) results as we observe the opposite behavioral pattern

$$e^{NI} < e^{VLI} > e^{LI} > e^{HI}.$$

Therefore, the IQ task data do neither provide evidence for standard theory expressed in Hypothesis 1 nor for an undermining effect of incentives of the form described in Hypothesis 2. Let us now consider the data of the counting task. The descriptive statistics are presented in Table 2.3 and seem to be qualitatively similar to the IQ task but less pronounced. Again the subjects in the VLI condition show the highest mean scores. Figure 2.2 illustrates the inversely U-shaped relation between level of piece rate and effort exertion for the CN task. However, running a pairwise comparison between conditions we only find significant differences between the high incentive condition and the other three conditions (see Table 2.4). Again, we neither find evidence

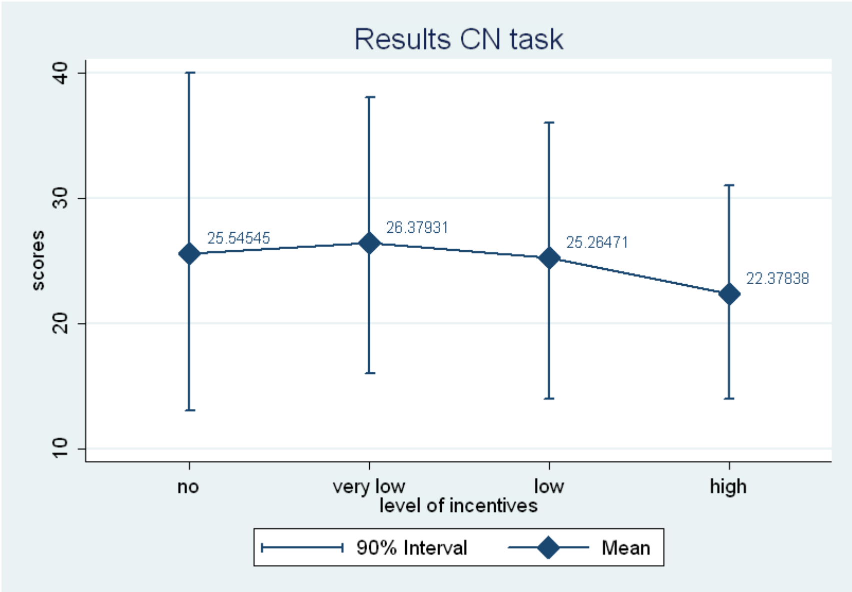


Figure 2.2: Mean scores CN task

for standard theory nor confirmation for the Gneezy and Rustichini (2000a) findings. In turn, the important discovery within the CN data seems to be the decrease in effort with a high rate of variable compensation. Table 2.4 shows a significantly negative influence of the high incentive scheme on point scorings which is line with the effect of high piece rates we observe with the IQ task.

Still, there are two issues which limit the interpretation of the data. First, during the phase in which the subjects were granted time to complete the questionnaire they might have continued working on the task instead of answering all of our questions. If they did so this might distort the results.

The second issue is that individual performance data do not only depend on effort but might be influenced by several unobserved factors such as ability to solve the tasks, the subjects' expectation about the own ability to solve the task and more importantly risk aversion for which we cannot control here. However, we have some demographic data we can use to run a robustness check on the results. Therefore we take the subjects' age ('age') and the number of semesters spent at university ('semesters') so far into a

	No incentive	Very low incentive	Low incentive	High incentive
No incentive		-	-	(*)
Very low incentive			-	**(**)
Low incentive				(*)

**significant at 5%, * significant at 10%; - not significant
Results of one-sided test in parentheses

Table 2.4: Results t-test CN task

simple regression. We also include the squares of these variables ('age^2' and 'semesters^2') to check for the impact of great deviations. In order to control for the task and the compensation scheme we include dummy variables 'no incentive' for subjects in the NI conditions, 'low incentive' for those in the LI conditions, 'high incentive' for the HI conditions and 'IQ test' for the IQ test task. Participants in the very low incentive condition serve as a reference group here. Note that subjects who did not complete the questionnaire are not included in this estimation (since we do not have their demographic data). The results of the analysis are presented in Table 2.5. The analysis demonstrates that none of the control variables has a measurable influence on the point scorings. However, even when controlling for age and semesters spent at university we observe significant poorer performance of participants in the no incentives condition. Moreover subjects in the high incentive condition yield significantly 6.2 points less than those in the very low incentive condition. Hence, we can consider the outcomes as being robust with regard to the control variables. Apart from that we find that even when we extract the performance data of the subjects from the sample who did not finish completing the questionnaire this does not alter the outcomes qualitatively.

2.2.4 Discussion

Considering the results described above we obtain two main results. First, for the IQ task subjects receiving the fixed wage only (NI) perform worse than those being paid a very low (VLI) piece rate in addition. This contradicts the Gneezy and Rustichini (2000a) results but seems to be in line

	Number of point scorings
No incentive	-3.469**
	(1.72)
High incentive	-6.196***
	(1.71)
Low incentive	-2.687
	(1.73)
IQ test	-5.731***
	(1.18)
Age	-0.452
	(1.09)
Age^2	0.00333
	(0.020)
Semesters	0.748
	(0.58)
Semesters^2	-0.0678
	(0.041)
Constant	35.35**
	(14.3)
Observations	237
R^2	0.17

Standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

Table 2.5: Results OLS regression

with standard theory predicting higher effort with stronger incentives. The second major finding is the significant decrease in point scorings with high powered incentives which occurs for both tasks. A possible interpretation of this result might be an increasing crowding out effect with increasing strength of incentives. That is, the stronger the extrinsic incentive the more intrinsic motivation is crowded out causing reduced effort choices. However, this theory cannot explain the poor results of the NI group since this group's performance should be better than the VLI group's. Another explanation might be a growing error rate with increasing incentives. Intuitively one might argue that participants feel excited imagining a very high possible income. Therefore the implementation of a high piece rate might stimulate very high effort in quantity resulting in worse quality of performance and consequently more mistakes. In that case a rather high effort in quantity would cause the worse results of the HI group compared to the VLI group. Surprisingly, the data do not confirm this conjecture. For both tasks there are no significant differences between numbers of errors between these conditions if we use a two-sided t-test.[14] Therefore this explanation cannot be robustly confirmed.

However, for us the most convincing interpretation seems to be a theory of reference dependent preferences meaning that the subjects had a certain reference income in mind when participating in the experiment. At the Universities of Bonn and Cologne experiments are run regularly. It is a common habit for students to participate and earn extra money. It is well known that participation in experiments is remunerated with an average of approximately 10€ although we did not explicitly mention this.[15] Still, we might roughly guess what the subjects had expected to earn. Note that this was not the case the session in Cologne in July 2003. By that time no permanent laboratory had been established yet and hence students were not familiar with taking part in economic experiments. Therefore subjects had to be recruited from undergraduate courses by reporting the average wage in the experiment of about 10€. This value refers to the average of earnings over

[14] If we apply a one-sided t-test we find a weakly significant increase in errors for the IQ task in HI condition compared to the VLI condition.

[15] This is an average value. Of course this value varies within the experiments depending on the role the participant plays. In ordinary jobs students earn about 10 to 15€.

all conditions. So in that case it was possible to influence the participants' expectations and generate a reference level artificially.

Due to the design of the experiment it was not possible to earn 10€ for the NI group since their payment was 5€ independent of their performance. Participants in the VLI (LI) group could earn a maximum wage of 5.48€ (7.40€) being less than the average of 10€. Still reaching this maximum income was very unlikely since the average number of points actually reached was much lower than the maximum.[16] So it can be assumed that members of the VLI and the LI group knew from the start that they would have to work hard and concentrated to get close to their reference wage. In contrast to that the HI group could reach the level of 10€ by attaining few points only.[17] Consequently the VLI (LI) group had to work a lot harder than the HI group and maybe did so to get closer to the reference wage.

2.3 An Alternative Theory

2.3.1 Incentives and Reference Dependent Preferences

None of the theories discussed above can explain the surprising outcomes of the experiment. For that reason a different type of theory has to be established. This section tries to offer one possible explanation for the experimental results with a simple model including reference dependent preferences.

The assumption of a reference level of compensation seems to be straightforward, since people might not only use their actual wage level to evaluate their utility but take a reference level (e.g. an expected wage or a rival's/workmate's wage) into account.

The relevance of reference points has been the object of research in many fields. Among the first, introducing the concept of reference points are Kahneman and Tversky (1979) analyzing decisions under uncertainty. They develop a model describing loss aversion by designing a utility function including a reference point. This utility function has a convex slope below the

[16]None of the subjects scored the maximum number of points in any condition.
[17]In the HI group subjects could earn up to a maximum of 29€.

reference point. It changes at the reference point and develops a concave slope. Hence subjects behave risk seeking below the reference point and risk averse beyond it. Evidence for the relevance of the theory was found early in many experiments (e.g. MacCrimmon and Larsson (1979)). Investigating the effect of loss aversion under risk-less choice Tversky and Kahneman (1991) extent the application of reference dependent utility. Easterlin (2001) examines the relation between happiness and income. Although he finds a positive relationship between income and happiness, income growth does not affect lifetime happiness since according to his theory aspirations grow with increasing income. Thus, the reference point, from which happiness is evaluated, changes. There are some empirical studies explaining their findings with loss aversion. Fehr and Götte (2005) find reduced effort with higher wages in a field experiment on the work habits of bicycle messengers. The messengers worked more days a month but decreased their shifts per working day that is reduced effort. On the other hand Oettinger (1999) does not find comparable effects analyzing data of stadium vendors' work participation decisions. He points out that participation is significantly higher if the expected wage can be assumed to be higher on the respective date of the game. Still, Oettinger (1999) does not observe any explicit effort levels. In a study on New York Cab Drivers Camerer et al. (1997) report decreasing numbers of working hours among Cab Drivers on high wage days. A recent experimental paper finding lower effort levels with higher variable wages for several real effort tasks is Ariely et al. (2005).

2.3.2 The Model

In this approach the assumption is made that people's utility does not only refer to the absolute height but rather to the relative height of monetary compensations. Beyond the reference point any further income growths relatively lose value. It follows that compared to the standard case less or no additional effort is rational after reaching the reference level, as costs would exceed utility gains from wage. The reference wage is defined as a point from where wages are evaluated. This might be a wage the agent expects or perceives to be appropriate for a certain task. The utility function increases linear in wage but flattens as a reference wage is reached. Thus, utility increases slower if wages exceed the reference point.

Since the model described in this section is supposed to be a theoretical approach to explain the experimental results we do not calculate the optimal incentive scheme. We rather take a wage contract comparable to those in the experiment.

Assume a utility function, which is additive-separable of the form:

$$U(w, e) = v(w) - c(e),$$

where w represents wage and e denotes the agent's exerted effort. As shown

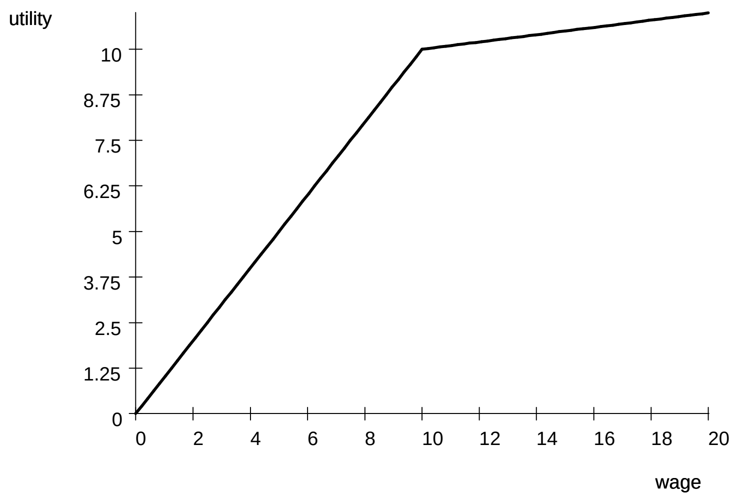


Figure 2.3: Returns on wage

in Figure 2.3 the slope of the utility function is discontinuous at value R

$$v(w) = \begin{cases} w & \text{if } w < R \\ R + (w - R)s & \text{if } w \geq R \end{cases} \tag{2.1}$$

with $0 \leq s < 1$.

The value R represents the reference wage from which the agent evaluates the actual wage.[18] Since s is strictly less than 1, marginal utility from w is smaller above the reference point than below. Assessing the situation from the reference point R the agent is in a loss situation if the first inequality is met because she stays below it. If the second inequality is met she is in a win situation. As the agent is work-averse, effort exertion is costly. Costs $c(e)$ are convex in e with $c'(e) > 0$, $c''(e) > 0$, $c(0) = 0$ and $\lim_{e \to \infty} c'(e) = \infty$. Furthermore let the wage contract be linear in e with a fixed wage α :

$$w = \alpha + \beta e.$$

The crucial question is for which value of e the agent changes from the left to the right side of the utility function. That is exactly the effort level where the agent earns a wage equal to the reference income. Let us call this decisive value the critical e

$$e^R = \frac{R - \alpha}{\beta}. \tag{2.2}$$

The optimal choice of effort is determined by the first order condition of the agent's objective function.[19] The first derivative of the corresponding objective function yields

$$U'(e) = \begin{cases} \beta - c'(e) & \text{if } e < e^R \\ s\beta - c'(e) & \text{if } e \geq e^R \end{cases}.$$

This consideration leads to three possible cases.

[18]Note, that the advantage of this utility function is that it is a very simple way to include loss aversion and reference dependency. In addition it enables us to precisely determine the reference point.

[19]Due to the concavity of the objective function the first order condition is necessary and sufficient for a maximum.

Proposition 1 *For given values of α and β the agent's optimal effort level e^* is:*

$$e^* = \begin{cases} c'^{-1}(\beta) & \text{if} \quad \beta < c'(\frac{R-\alpha}{\beta}) \\ \dfrac{R-\alpha}{\beta} & \text{if} \quad \frac{1}{s}c'(\frac{R-\alpha}{\beta}) \geq \beta \geq c'(\frac{R-\alpha}{\beta}). \\ c'^{-1}(s\beta) & \text{if} \quad s\beta > c'(\frac{R-\alpha}{\beta}) \end{cases} \tag{2.3}$$

Proof: See Appendix.

For low values of β the employee's optimal effort choice is located to the left of the critical value. The effort level is located exactly on the critical value for an intermediate β. The third row presents the employee's effort decision for sufficiently high values of β where the effort choice is located to the right of the kink. The change of e^* in the incentive parameter is presented in Figure 2.4. Figure 2.4 shows that there is an area where the optimal effort level chosen by the agent decreases in the incentive parameter. In this area the agent always chooses $e = e^R$. Since e^R decreases in β we find a downwards sloping curve in this interval. Consequently, from the principal's point of view increasing piece rates in this area are counterproductive not only because they are costly but also because they even reduce the agent's performance. However, our model is subject to some limitations which deserve mention. The behavioral predictions strongly depend on the reference wage. Obviously a reference wage is an individual's private information and hence not directly measurable. Furthermore, the reference income might vary between individuals or groups of individuals or over time. Assuming different reference wages for different agents seems to be more realistic but exacerbates predictions. The model also lacks any insights into the origin and dynamics of reference points although this might be a critical issue to make more general predictions.

Note that the model would lead us to predict a performance increase between the NI and the VLI groups. As we merely find a significant increase for the IQ task, our data do not completely confirm the predictions of the model.[20]

[20] Another issue is that in our experiment the subjects exerted positive effort even if paid

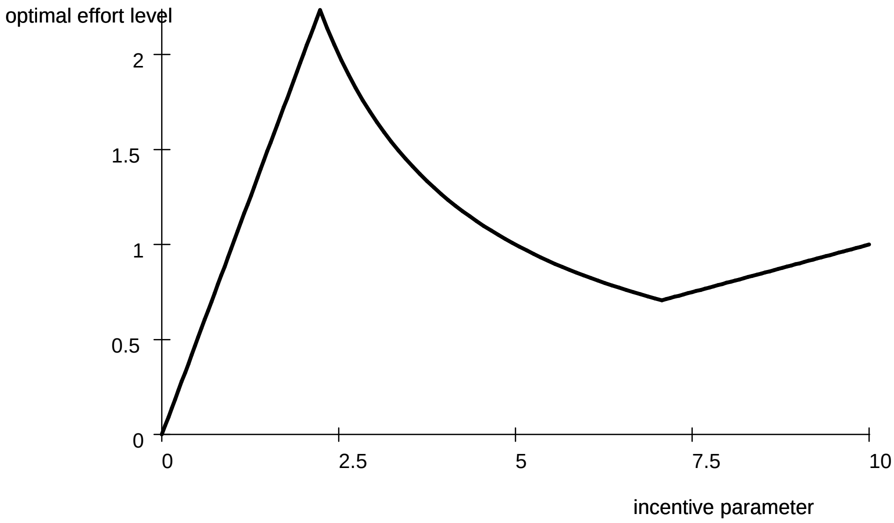


Figure 2.4: Optimal effort level

Naturally the choice of parameters drives the model. Therefore parameters β, R, s and the shape of the cost function determine whether the model generates the experimental results. Moreover, the conditions, under which this model can explain the phenomenon that occurred in the experiment, are very specific, assuming rather precise wage expectations. Still it might be a first attempt to explain the special situation in which some experiments take place.

2.4 Concluding Remarks

In this chapter the relation between incentives and effort choices has been investigated. Our results indicate two main effects: With the IQ test task subjects who were exposed to very low incentives exerted more effort than those who only earned a fixed wage. This is in contrast with Gneezy and

independent of their performance (NI group). In our model this can only be rational if we generalize the cost function, so that costs $c'(e)$ are convex in e with $c''(e) > 0$, $c(0) = 0$ and $\lim_{e \to \infty} c'(e) = \infty$ with minimum costs for a strictly positive effort level.

Rustichini (2000a) observing the opposite behavioral pattern. Secondly, with both tasks the experimental results demonstrate that effort does not necessarily increase with increasing incentives. Instead the opposite was the case in the experiment, as higher incentives led to lower performance. Since standard theory and motivation crowding theory fail to give a convincing explanation for these results a model including reference dependent utility is applied.

It remains an open question why the outcome was not similar to Gneezy and Rustichini (2000a), not even with the IQ task. Crowding out seems to be a very sensitive finding which is inherently obvious following the psychological literature on crowding out effects (see e.g. Eisenberger and Cameron (1996); Deci et al. (1999a); Deci et al. (1999b)). It would be interesting to learn under which conditions these effects are likely to appear since they might dependent on several factors such as task, composition of wages and most importantly the subjects' individual perception of the situation. Those factors might have an essential influence on the impact of incentives and should be subject to further research.

Another interesting observation is that the influence of specific task features seems to be (at least for our tasks) negligible since the outcome is qualitatively similar. A reason for that could possibly be that the differences between the tasks were too small to produce a measurable effect. The hypothesis of the IQ task generating more intrinsic motivation than the CN task cannot be confirmed. Neither the IQ task nor the CN task created a result, which can be interpreted as crowding out of intrinsic motivation. Nevertheless it is not understood that different tasks trigger off the same behavior. Still, the influence of task features cannot be resolved with these experiments.

The vital conclusion of the experiment described in this chapter is not that incentives do not work. Moreover, the implication should be that incentives might work if the agent's income is below her reference point. Beyond this point however incentives may fail to provide higher performance. Still, further evidence is necessary to confirm the hypothesis that participants, who were exposed to a higher variable pay, were quickly pleased with their wage. For this purpose it would be interesting to check whether subjects reduce their effort once they reached a certain wage level i.e. the reference wage. Unfortunately we did not observe any temporal variation in the data since we explicitly allowed subjects to switch the order of the exercises. That makes it impossible to draw any corresponding conclusions. Another option of gathering evidence for the relevance of reference points might be questioning the subjects directly (in advance to and/or after the session) what their expectations regarding the wage are. These possibilities should be considered for future research.

It remains the problematical question what practical implications can actually be derived from the experiment, since it only admits suppositions on the participants' motives. Yet, the individual wage expectations seem to be important information required to provide optimal incentives. For this purpose, further examinations of the emergence, development and measurement of reference points are necessary.

2.5 Appendix to Chapter 2

This is the first page of the CN task translated from German

Please count the number of ones and sevens

Block 1

0 1 8 5 8 7 0 4 2 4 5 7 0 8 2 4 8 5 7 4 5 8 2 1 6 2 7 6 4 4
3 8 5 2 3 9 7 9 9 3 5 5 6 0 7 1 7 5 3 0 9 6 0 6 3 3 8 5 7 9
6 4 7 6 3 2 2 4 5 8 7 7 5 1 4 2 8 7 6 6 5 3 4 9 8 3 3 0 9 8
2 8 2 6 0 4 7 6 2 0 0 2 5 3 4 5 5 4 7 4 1 1 6 0 2 5 4 3 0 0
4 4 3 1 1 9 2 5 2 4 3 5 4 3 9 8 4 0 2 7 0 8 7 0 5 3 4 3 1 1
3 4 9 8 2 6 3 3 8 2 0 9 3 4 0 0 4 5 5 5 8 9 5 5 0 1 6 4 8 6
2 6 9 0 5 8 8 5 2 1 9 5 0 1 5 3 1 7 0 3 2 4 9 6 2 4 9 7 8 3
8 6 0 5 5 2 9 9 2 3 6 6 8 1 6 3 2 3 5 6 1 4 9 3 4 4 2 1 5 1
9 4 9 3 4 3 3 2 3 0 8 5 7 5 3 7 7 4 0 3 8 3 8 3 7 2 7 7 4 3
8 0 7 2 9 2 2 2 5 4 3 8 5 7 6 6 5 8 7 7 0 0 1 8 3 7 2 0 0 4
9 8 8 6 3 9 7 5 1 5 2 2 6 8 4 8 7 6 3 9 4 9 6 0 3 7 2 6 1 5
0 8 5 2 9 1 1 9 6 7 6 6 3 8 4 3 9 7 2 3 9 8 5 3 9 8 3 2 1 9

Number of ones:	
Number of sevens:	

Block 2

8 5 2 6 4 0 5 4 7 9 2 8 6 3 7 2 0 6 5 3 7 2 9 1 1 1 7 8 4 7
6 3 9 8 0 7 8 0 5 7 9 8 1 6 2 1 3 5 9 2 6 5 6 2 7 3 1 9 3 4
0 6 9 4 3 5 6 6 3 8 4 1 2 1 0 3 4 6 9 7 3 9 1 6 7 6 2 1 9 9
1 5 3 3 3 4 0 7 3 3 6 6 0 8 6 9 0 6 6 9 2 0 1 4 6 5 2 4 7 1
5 5 2 7 4 7 9 3 6 5 0 6 8 5 3 9 4 9 8 2 7 1 6 5 9 1 1 4 8 0
0 4 6 7 9 8 5 8 6 5 4 0 4 9 9 7 8 8 9 5 1 9 1 4 8 6 7 8 0 8
0 1 1 5 2 1 0 6 4 9 5 3 0 2 1 1 6 3 2 7 3 6 7 1 8 6 8 2 7 8
1 7 5 6 9 3 6 3 3 5 0 5 4 4 1 8 2 9 0 7 2 2 6 8 4 4 3 9 6 5
0 9 5 2 6 7 0 2 0 3 7 6 9 6 4 4 5 9 0 3 1 5 3 1 9 4 9 7 5 2
7 4 3 5 3 7 5 8 4 1 6 9 5 0 0 0 8 6 6 0 3 9 2 0 3 6 2 2 2 3
8 7 2 3 1 1 0 7 6 2 9 3 4 3 5 4 8 4 0 4 3 9 8 7 0 5 8 6 1 8
9 4 1 5 2 1 8 6 7 5 9 0 0 3 9 4 1 7 6 1 3 5 8 9 5 6 1 9 8 2

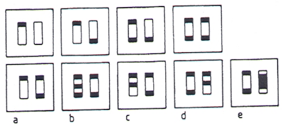
Number of ones:	
Number of sevens:	

This is the first page of the IQ task

1. *Figurenreihen fortsetzen*

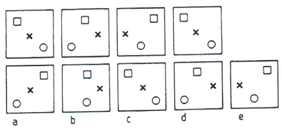
Die Lücke füllen. Mit welcher Auswahlfigur unten (a, b, c, d oder e) kann man die Figurenreihe oben richtig fortsetzen?

Beispiel:



Lösung: b

Erklärung: Hier werden in die Rechtecke schwarze Balken eingefügt: erst in das linke oben, dann in das rechte unten, dann in das linke unten, im vierten Bild rechts oben. Die Fortsetzung kann nur wie bei Lösungsvorschlag b erfolgen.

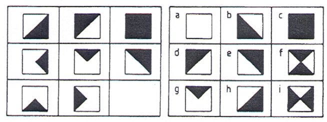


Lösung:

2. *Sinnvoll ergänzen*

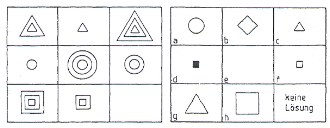
Sie sehen ein Rechteck mit 8 Figuren. Welcher der vorgegebenen 9 Lösungsvorschläge (rechts, a–i) passt in das freie 9. Feld?

Beispiel:



Lösung: b

Erklärung: Die schwarze Fläche der ersten Figur, addiert mit der schwarzen Fläche der zweiten Figur, ergibt, sozusagen als Summe, die dritte Figur. Dieses Prinzip gilt sowohl in vertikaler wie in horizontaler Richtung – ein wichtiger Hinweis für die generelle Bearbeitung dieses Aufgabentyps.



Lösung:

3. *Buchstabengruppen*

Die Unregelmäßigkeit entdecken. Welche Buchstabengruppe (a, b, c, d oder e) ist nicht wie die vier anderen Gruppen nach einer bestimmten Regel zusammengesetzt?

Beispiel:

a	b	c	d	e
CBAZ	PONM	UTSR	IHGF	ECBA

Lösung: Gruppe e

Erklärung: Das Alphabet wird hier rückwärts präsentiert, und in der Gruppe e ist der Anfangsbuchstabe E falsch, denn es müsste eigentlich das D sein.

	a	b	c	d	e
1	LNNP	DNNT	PNNP	DNNT	QNNX
2	AABA	AAAB	AAAC	AAAT	AAAU
3	CDDC	KLLK	QRRQ	UTTU	WXXW
4	YXVW	TSQR	NOLM	HGEF	EDBC
5	ORUX	ADGJ	MPSV	ILOR	LORT

4. *Zahlenreihen*

Wie heißt die nächste Zahl? Sehr häufig eingesetzt wird der Aufgabentyp »Zahlenreihen«, bei dem eine nach bestimmten Regeln aufgebaute Folge von Zahlen zu ergänzen ist.

Beispiel: 2 4 8 16 32 ?

Lösung: 64

Erklärung: Jede Zahl wird mit 2 multipliziert.

A	3	9	6	9	27	?	
B	0	−1	1	3	−1	4	?
C	2	5	11	23	47	?	
D	2	12	6	30	25	100	?
E	80	40	42	40	20	?	
F	3	8	23	68	203	?	
G	1	$^1/_2$	$^5/_2$	5	$^5/_2$	$^9/_2$	?
H	7	15	0	8	−7	?	
I	81	9	18	2	11	?	

Instructions CN task (high incentive condition)

Welcome to this experiment! Please do not talk to other participants at any time during the experiment! Your task is to count the correct number of ones and sevens in a block of random numbers. You can solve the blocks in any order you like. For every correct answer you are scored one point. An answer is valued correct if it states the exact number of ones or sevens with a deviation of one.

Example:

If the exact number of ones in a block is 30 the answers 29, 30 or 31 are valued correct.

You have 30 minutes to work on the task. After that we will hand out forms. Please copy your answers into the form. A questionnaire is fixed to the forms. Please also answer the questionnaire. You will always get a fixed participation fee of 5€ and additional 0.5€ for every point scored. The compensation is paid cash on November 26th to 28th from 10.00 to 12.00 o'clock at the Chair's library (Room 69). You will be informed about your scorings as well. Please bring along your anonymous ID-Code. The ID-Code is printed on your form. Additionally it is fixed to your form on a separate piece of paper. Please tear it off the form and bring it along when you collect your payment. If you have further questions concerning the procedures raise your hand and wait until a supervisor comes over to you. Do not ask any question aloud! Thank You for Your participation!

Instructions IQ task (high incentive condition)

Welcome to this experiment! Please do not talk to other participants at any time during the experiment! Your task is to solve the following exercises correctly. Each type of exercise is illustrated by an example. You can solve the exercises in any order you like. For every correct answer you are scored one point. For an incorrect answer you get a penalty of 0.5 points. Hence, 0.5 points will be subtracted from your score for an incorrect answer. You have 30 minutes to work on the task. After that we will hand out forms. Please copy your answers into the form. A questionnaire is fixed to the form. Please also answer the questionnaire. You will always get a fixed participation fee of 5€ and additional 0.5€ for every point scored. The compensation is paid cash on November 26th to 28th from 10.00 to 12.00 o'clock at the Chair's library (Room 69). You will be informed about your scorings as well. Please bring along your anonymous ID-Code. The ID-Code is printed on your form. Additionally it is fixed to your form on a separate piece of paper. Please tear it off the form and bring it along when you collect your payment. If you have further questions concerning the procedures raise your hand and wait until a supervisor comes over to you. Do not ask any question aloud! Thank You for Your participation!

Proof. Proposition 1: Due to the strict concavity of the objective function and the assumption on the cost function there must be a unique internal optimum. Suppose that $e^* < e^R$, then

$$e^* = c'^{-1}(\beta)$$

must hold. That is, the agent chooses an effort level to the left of the kink.

This occurs if

$$\beta < c'\left(\frac{R-\alpha}{\beta}\right).$$

The equation defines a unique cut-off value for β such that $e^* = c'^{-1}(\beta)$ iff β is smaller than the cut-off value. In case of $e^* > e^R$, the optimum is defined by

$$e^* = c'^{-1}(s\beta).$$

Consequently e^* must be located right of R which happens if

$$\beta > \frac{1}{s}c'\left(\frac{R-\alpha}{\beta}\right).$$

This cut-off value is strictly larger than the cut-off value defined by $e^* = c'^{-1}(\beta)$. It follows that in all other cases the agent chooses $e^* = e^R$. ■

CN task					IQ task			
NI	VLI	LI	HI		NI	VLI	LI	HI
6	0	10	4		-1	5	3	-2
11	14	12	12		0.5	9	3	0
12	16	13	14		3.5	9.5	3.5	1
13	17	14	14		6	11.5	4.5	5.5
15	19	16	14		6.5	12	5	5.5
15	20	16	14		7.5	13	8	6
17	21	16	15		9	14	11.5	6.5
18	21	19	16		10	15	11.5	8
18	23	19	17		12.5	15	12.5	8.5
18	23	20	18		12.5	15	12.5	10
20	24	20	18		13	15.5	14.5	10
20	24	22	19		13.5	17	15.5	10.5
21	24	23	20		14.5	19.5	16.5	11.5
22	25	23	20		16	19.5	17.5	12
23	25	24	20		16	20.5	18	12.5
24	25	25	21		16.5	21.5	19	12.5
25	28	26	21		17.5	23	19.5	12.5
25	28	26	22		18	23	21	14.5
25	29	26	22		18	23.5	21	15
26	31	27	23		19.5	27.5	22	15.5
27	31	27	23		20	29	22	17
27	32	28	23		20.5	29.5	22.5	17.5
30	36	28	25		22	30	24	17.5
31	36	28	25		22.5	31	24	18
34	37	29	25		23.5	32.5	24.5	18.5
35	38	31	26		23.5	35	24.5	21
37	38	31	27		24	36	26.5	21.5
38	40	33	27		25.5	38	28.5	21.5
39	40	35	28		26	41	31	22
40		35	28		26		31.5	22
42		36	28		26.5		32	22.5
43		38	29		26.5		32.5	23
46		40	29		33		35	25.5
		43	31					27
			34					28.5
			35					30
			41					33
								34.5
								42
33	29	34	37	# Obs.	33	29	33	39

Table 2.6: Point scorings by wage and task

	No incentive	Very low incentive	Low incentive	High incentive
No incentive		*(**)	-	-
Very low incentive	-		-	**(**)
Low incentive	-	-		-
High incentive	-	**(**)	(*)	

Results above (below) diagonal refer to IQ task (CN task)
** significant at 5%; * significant at 10%; - not significant
Results of one-sided test in parentheses

Table 2.7: Results Mann-Whitney-U-test both tasks

Chapter 3

The Role of the Fixed Wage in Setting Incentives[1]

3.1 Introduction

As we have discussed comprehensively in the previous chapter, most simple principal agent models predict that stronger incentives ceteris paribus lead to higher performance of the agent. On the other hand, it is noticeable that only a relatively small part of the total compensation is performance dependent in many firms. For example, Murphy (1999) using Towers Perrin's 1997 Worldwide Total Remuneration report shows that although the composition of executive pay varies considerably between industries, firm sizes, and countries, predominantly the greater fraction of the total compensation is covered by the base salary which constitutes a fixed compensation. This seems to be especially true for CEO payments outside the US. Due to this observation some questions arise: Does the fixed wage play an additional role in incentive schemes apart from attracting employees and ensuring the acceptance of the contract? Is there a linkage between the height of fixed payments and the strength of incentives within an incentive contract?

In contrast to Murphy (1999) indicating that variable payment plays a minor role in CEO compensation, economic theory teaches us the relevance

[1]This chapter is based on Mohnen and Pokorny (2006a).

of variable, performance-based payments. Particularly according to most simple standard agency models only the variable wage component induces incentives. A higher piece rate is followed by an increase in the effort level whereas the fixed wage only ensures the agent's participation (for seminal papers on the standard approach see, e.g. Holmström (1979), Grossman and Hart (1983) or Hart and Holmström (1987)).

The impact of exclusively paying a fixed wage on the effort decision has been examined empirically. An important approach of designing an employer-employee relationship in the lab is the gift-exchange game (e.g. Berg et al. (1995), Fehr et al. (1993)). Typically an employer offers a fixed wage to an employee and asks for a certain effort level in exchange. The desired effort level is not binding but effort exertion is costly for the employee. According to standard game theoretical predictions the employee should exert the lowest effort level possible and therefore the employer should offer the lowest wage possible at the beginning of the game. However, this is not what is observed in most experimental studies (Gächter and Falk (2002) for an overview). Generally, higher wages are paid and higher effort levels are exerted than predicted by theory. Furthermore, effort exertion significantly increases in the wage offered. These results are commonly explained by concepts of social preferences like reciprocity (Rabin (1993), Dufwenberg and Kirchsteiger (2000)) and inequity aversion or fairness (Akerlof (1982), Falk and Fischbacher (1999), Fehr and Schmidt (1999), Bolton and Ockenfels (2000)). According to reciprocity, individuals' reactions to friendly behavior are friendly as well, and vice versa. Consequently, reciprocity theories predict increasing effort levels with higher wages if a high wage is perceived as friendliness.

In the previous chapter and the related literature (e.g. Camerer et al. (1997) and Fehr and Götte (2005)) we have seen that reference dependent preferences have an impact on the effort decision. If this is the case, the height of the fixed wage has an influence on the incentive power of the piece rate. So the aim of this chapter is the investigation of the role of the fixed payment within an incentive contract. To the best of our knowledge, the effect of varying fixed wages on performance has only been examined for situations in which the fixed wage is paid exclusively. Therefore we test whether there is an impact of the fixed payment on the chosen effort level.

In our approach we first extend the analysis of reference dependent preferences and loss aversion concerning linear wage contracts. In particular we try to shed light on the question in how far the fixed wage component drives the effort decision if individuals are loss averse. For this purpose, we derive theoretical predictions on the effect of varying fixed wages on the effort decision from the model developed in Chapter 2. Secondly, we test the resulting predictions in a real effort experiment at the Universities of Bonn and Cologne. The subjects were all paid the same piece rate but the amount of the fixed payment in the three conditions is varied: low, intermediate and high. All of them were offered to work on the same calculation task and each participant in the experiment knew ex ante the precise duration of the working time and his wage contract. Note that there are no explicit principals in this experiment who benefit from the agents' effort. We do this to measure the impact of wages in a 'clean' way. Including interaction with principals might produce different motives and hence decisions.

We chose real effort conditions to create an environment which is closer to real work conditions. Laboratory experiments are almost always very abstract and this is likely to affect the results (Gneezy (2004): 4-5, 7-8). Since we have real effort costs instead of a monetary cost function in our setting we believe our results produce better implications for real work environments than in case of an abstract effort choice.

Our experimental results indicate that workers indeed care about the fixed payments. Subjects receiving the low and high fixed payments worked significantly more than those participants who were paid an intermediate fixed wage. As the principal agent model including loss aversion can explain only part of the observed behavioral pattern, we provide a different explanation for the results which focuses on a kind of social norm representing the appropriateness of the fixed wage payment.[2]

[2] Since there are no principals in the experimental setting, reciprocity does not seem to be a good explanation. However, we find a kind of reciprocity towards the height of the fixed wage which we call appropriateness.

The reminder of this chapter is structured as follows. In the next section the principal agent model with loss averse agents is analyzed with regard to the impact of the fixed wage. Section 3.3 describes the details of the experimental set-up and in Section 3.4 behavioral predictions are developed. The results are reported in Section 3.5. Section 3.6 discusses interpretations and explanations of our data. The last section concludes the chapter.

3.2 The Model

As already discussed in Chapter 2, a key notion of reference dependency and loss aversion in the context of work relationships is the evaluation of wages in relation to a reference point. That is, not only the absolute amount of wage determines the resulting utility but rather the relative amount compared to the reference wage. This reference wage might for instance be a wage which the agent considers to be appropriate for her work, an expected pay, a rival's or work mate's wage or the past periods' wage. Since wage levels below the reference wage are perceived as losses, marginal returns below it are higher than above it. It follows that compared to standard theory less or no additional effort is rational after reaching the reference point. Consequently, piece rates have a weaker impact on the effort level if the fixed wage is above the reference wage than if it is below it.

As we focus on the fixed payment, we now extend the simple model of reference dependency and loss aversion from the previous chapter to examine the influence of the fixed wage component on the agent's effort decision. Again reference dependency is introduced by a utility function with a kink at the reference point.

We use similar assumptions to those in Chapter 2: the utility function U is additive separable in utility from wage $v(w)$ and the disutility from effort exertion $c(e)$:

$$U(w,e) = v(w) - c(e),$$

where w denotes wage and e represents the effort the agent exerts. The agent evaluates her actual wage by comparing it to the reference wage R.

$$v(w) = r\begin{cases} rw & if\ w < R \\ R + (w - R)s & if\ w \geq R \end{cases}$$

with $0 \leq s < r$. The utility function $v(w)$ which is discontinuous at reference value R. Since s is strictly smaller than r, marginal returns on effort are less above the reference point than below.

As the agent is work-averse, effort exertion is costly. The cost function $c(e)$ is again assumed to be convex in e with $c'(e) > 0$, $c''(e) > 0$, $c'(0) = 0$ and $\lim\limits_{e \to \infty} c'(e) = \infty$.

Furthermore, let the wage contract be linear in e with a fixed wage α and a piece rate β:

$$w = \alpha + \beta e$$

We have seen in the previous chapter that the critical e is accurately the point where the change from higher marginal returns to lower marginal returns takes place. This is exactly the case if the wage $w = \alpha + \beta e$ equals R, so we get equation (2.2). Obviously, the critical e^R decreases in the fixed wage.

Due to the strict concavity of the objective function the first order condition is necessary and sufficient for a maximum. The first derivative of the corresponding objective function yields

$$U'(e) = \begin{cases} r\beta - c'(e) & if\ e < e^R \\ s\beta - c'(e) & if\ e \geq e^R \end{cases} .$$

Analyzing the model with regard to the fixed wage component we obtain the following proposition:

Proposition 2 *For given values of α and β the agent's optimal effort level e^* is:*

$$e^* = \begin{cases} c'^{-1}(r\beta) & if\ \alpha < R - c'^{-1}(r\beta)\beta \\ \frac{R-\alpha}{\beta} & if\ R - c'^{-1}(s\beta)\beta \geq \alpha \geq R - c'^{-1}(r\beta)\beta \\ c'^{-1}(s\beta) & if\ \alpha > R - c'^{-1}(s\beta)\beta \end{cases} \tag{3.1}$$

Proof: See Appendix.

Proposition 2 shows that if the fixed wage α is sufficiently small the agent's optimal effort level yields a wage in the left part of the utility function. The third row of the proposition determines the employee's optimal effort choice if it is located to the right of the critical value e^R. These cases occur if α is sufficiently extreme. In between these areas the optimal effort decision always equals the critical value, thus $e^* = e^R$. Then the optimal effort level results in a wage which is located exactly on the kink. The change of e^* in the fixed wage is illustrated in Figure 3.1. Figure 3.1 illustrates that for extreme

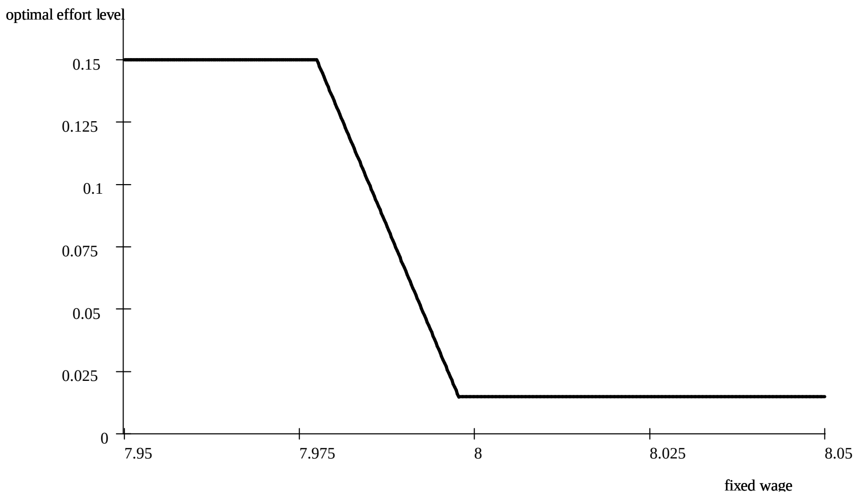


Figure 3.1: Optimal effort choice for a given fixed wage

values of the fixed wage, the agent chooses a constant effort level: for low fixed wages he works on a high level and for a sufficiently high fixed payment on a low level. This result exactly corresponds to the different marginal returns for the agent below and above the reference point. Furthermore, we see that there is an area where the optimal effort level chosen by the agent decreases in the fixed wage. In this area the agent always chooses e^R. Intuitively, if the fixed wage within the incentive contract is higher, the reference wage is already reached for lower effort levels. This makes the agent reduce his effort since he does not want to exceed the reference wage due to the lower marginal returns above it.[4] Hence, in line with Proposition 2 Figure 3.1 shows that the low effort level is even chosen for values of $\alpha < R$ as the total wage is the sum of the fixed component and the variable component.

3.3 Experimental Design and Procedures

We invited 181 undergraduate students from various faculties at the Universities of Bonn and Cologne, Germany to participate in the experiment. We conducted two sessions in July 2004. Altogether we introduced three different conditions. Participants of all conditions were seated in one single room to prevent room effects. After all individuals had entered the room subjects were randomly assigned to the conditions and requested to sit down on seats marked with pens and pieces of paper. After everybody had sat down the supervisors handed out the exercises. Each of them was covered with a sheet with the instructions printed on. Neither the instructions nor the compensation scheme were ever read out aloud. In all conditions the subjects were asked to perform the same task, which was solving relatively simple but tedious calculation exercises.[5] For all conditions we compensated the subjects with the same piece rate of 0.15€. That is, for each correct calculation the subjects were paid 0.15€. Additionally, the subjects received a fixed wage which was completely independent of the number of correct calculations. The height of the fixed wage component was 2€ (low wage), 6€ (intermediary wage) or 12€ (high wage) respectively.[6] To ensure that the participants did not solve the calculation exercise just to avoid boredom, they were additionally offered crossword puzzles and similar entertaining quizzes. We emphasized in the instructions that they would receive no variable wage for solving the quizzes. Participants in all three conditions were granted forty minutes to work on the task. After forty minutes a bell rang and the supervisors collected the exercise sheets. Moreover, questionnaires were handed

[4] Note that the agent will always exert positive effort for any positive incentive parameter β.

[5] For an example of the instructions and the task see appendix.

[6] 1€ was roughly 1.2$ at the time of the experiment.

out to each subject to collect some demographic data. Then participants were informed when and where they were asked to come and receive their payment.[7]

3.4 Hypotheses

In addition to our model of reference dependency and loss aversion, we consult standard agency models and reciprocity to derive hypotheses.

According to simple principal agent models the height of the fixed component does not influence the effort decision of the agent because only the variable component is relevant:[8]

Hypothesis 1:

$$e_L = e_I = e_H,$$

where e denotes the average effort in the respective condition.

If reciprocity is the motive for the action to take, we should expect increasing effort levels with higher wages if higher wages are perceived as a friendly action.

Hypothesis 2:

$$e_L \leq e_I \leq e_H$$

if participants act reciprocally in our experiment.

Of course reciprocity is typically closely associated to a counterpart. As there are no principals in our experiment the conventional idea of reciprocity is harder to apply here. But still the subjects might perceive the experimenter as a kind of counterpart they interact with. Although there are no obvious advantages of higher scorings for the experimenter, the participants might suffer from a bad conscience if they receive a high fixed wage but do not exert corresponding efforts. As argued in Bewley (1999) a certain amount of money

[7] The remuneration was paid off about one week after the session had taken place.

[8] It is straightforward that higher effort levels can be approximated by an increasing number of correct answers.

Fixed wage	Mean	Median	Stand. dev.	Max	Min
Low	21.068	21	9.958	47	1
Intermediary	18.094	17	7.904	47	0
High	20.207	20	9.354	42	0

Table 3.1: Descriptives statistics

goes along with a certain effort level, i.e. employees have an appropriate effort level in mind fitting to the height of their remuneration.

With regard to loss aversion and reference dependent utility the effect of varying the fixed wage depends on the individual's reference point. As shown in Figure 3.1, in our model reference dependent preferences and loss aversion lead to a weakly decreasing slope of effort in the fixed wage which is the exact opposite to the predictions derived from reciprocity theory.

Hypothesis 3:

$$e_L \geq e_I \geq e_H.$$

If the fixed wage is sufficiently small so that the agent chooses e^* according to $c'^{-1}(r\beta)$ we should observe relatively high scores. In all other cases the optimal effort level is lower than $c'^{-1}(r\beta)$. Hence, there might be a decreasing relation between scores and the fixed wage component in our experimental results if in one or two of the higher paid conditions the absolute wage exceeds the reference point.

3.5 Results

To approximate the effort the subjects exerted on the task we analyze the number of correct answers in the different conditions. Some descriptive results are presented in Table 3.1. Figure 3.2 illustrates the median number of correct answers for low wage, intermediary wage and high wage. As can be seen in Figure 3.2 the relation between exerted efforts and fixed wage components is non-monotonic in our data. For a low level of fixed wage the number of correct answers is higher than for an intermediary level. With a relatively high fixed wage the number of correct answers increases again. We

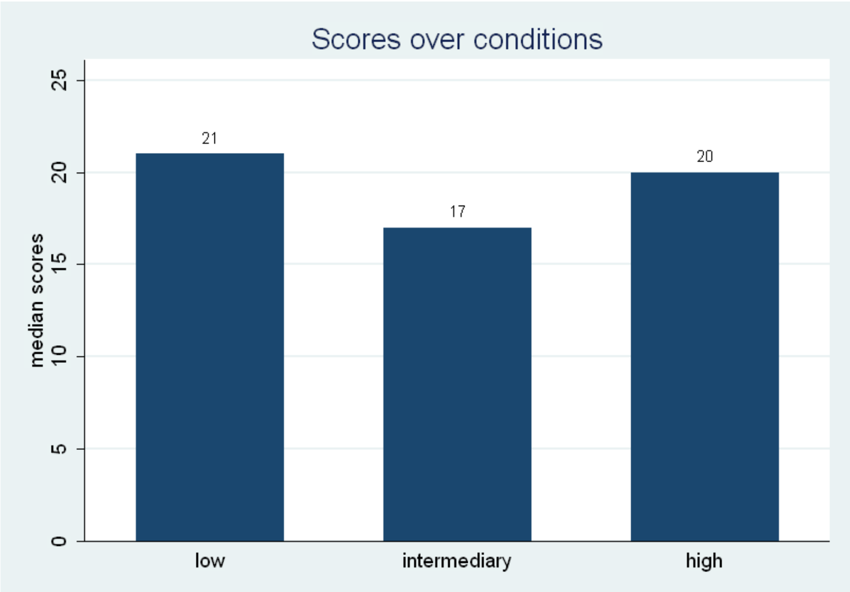


Figure 3.2: Median scores over conditions

ran a pairwise comparison with a one-sided t-test. The difference between the low and the intermediary wage conditions are significant on a 5% level ($p = 0.0340$). Comparing the intermediary and the high wage groups the difference is weakly significant ($p = 0.0895$).[2]

As the individual results of the real effort task probably do not only depend on the strength of exerted effort but also on several unobserved factors we also used parametric estimation methods to analyze our data. It is likely that other individual characteristics like e.g. calculation ability or session specific aspects like room conditions (noise, temperature) might have had an influence on the individual performance which makes the data relatively noisy, especially compared to standard lab experiments with abstract effort choice. However, the allocation of abilities in the three conditions should be normally distributed due to the large condition sizes. But for session effects

[2]Using a Mann-Whitney-U-test the difference in correct answers between the low wage and the intermediary wage groups is weakly significant ($p = 0.09$ two-tailed). The difference between the intermediary wage condition and the high wage condition is not significant ($p = 0.14$, two-tailed).

we control in the following regression analysis.[9]

We examine the results with median regression since there is a considerably large number of outliers in the data.[10] Median regression minimizes the sum of absolute deviations rather than squared deviations. Due to its lower sensitivity towards outliers we consider the median regression to be a better measure for the central tendency of the data here.[11] The results of the regressions are presented in Table 3.2. We include two wage dummies: one for the highly paid group (high fixed wage) and another for the low paid group (low fixed wage). In addition we control for the subjects' income, age and the place where the session took place ('cologne' is a dummy variable with value one if the subjects participated in the session at the University of Cologne and zero otherwise). The subjects receiving an intermediary wage served as a reference group in the regression to which the two other conditions are compared.

The coefficients of both wage dummies have a positive sign and are significantly different from zero. Model (1) shows that subjects in the high fixed wage condition yield significantly 3 correct answers more than individuals paid an intermediary fixed wage. This result is confirmed by models (2) and (3). Moreover, subjects in the low wage condition even achieve 4 correct calculations more than those in the intermediary condition in model (1). Models (2) and (3) again demonstrate a similar pattern. Hence, it follows that, deviations from the intermediary wage in both directions lead to higher performance.[12]

As the results might be driven by the individual income level we asked the subjects about the monetary amount they monthly had at their disposal (we offered three categories). The applied parametric analysis allows us to

[9] An important difference between the sessions at the University of Bonn and the University of Cologne is that this experiment was one of the first at the University of Cologne. At the University of Bonn economic experiments are conduced regularly and are well-known among students. Note that since there was only one session at each of the two universities this variable covers the time and the subject pool of the session as well.

[10] For the application of median regression, see Greene (2003).

[11] We also ran an OLS regression on the data which can be found in the appendix in Table 3.3.

[12] Note that for an OLS regression the coefficient of the high fixed wage dummy variable is not significantly different from zero.

	(1)	(2)	(3)
	Number of point scorings		
High fixed wage	3.00**	3.00*	3.39**
	(1.4)	(1.6)	(1.7)
Low fixed wage	4.00***	5.00***	4.21**
	(1.4)	(1.5)	(1.7)
Cologne	0	0	-1.21
	(1.1)	(1.3)	(1.4)
Income 2		1.00	0.91
		(1.4)	(1.5)
Income 3		1.00	0.82
		(2.1)	(2.4)
Income 4		4.00	0.70
		(4.4)	(5.1)
Age			0.30
			(0.2)
Constant	17***	16***	9.42*
	(1.1)	(1.5)	(5.3)
Observations	181	181	181
Pseudo R^2	.0245	.0325.	.0435.

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.2: Results median regression

control for these external factors. But the results show that income does not have a significant effect on the effort decision.

A second interesting question is whether different levels of the fixed wage component influence the quality of task performance or approach of solving the task. Realizing that a wage is very high, participants might feel motivated to work harder and faster. A consequence might be higher error rates. To analyze this we examine the number of given answers and the corresponding error rates. Still, we do not find significant differences in error rates for the different conditions in models (1) and (2). The average rates of mistakes are presented in Figure 3.3. Figure 3.3 shows that the error rate hardly changes

Figure 3.3: Median error rates over conditions

over conditions. Thus, we conclude that the better results of the low wage group were not triggered by a higher work quality but simply by higher work quantity.

3.6 Interpretation

As mentioned above, we find significant differences between conditions and hence we cannot confirm standard theoretical results (Hypothesis 1).

Furthermore reciprocity alone cannot explain the observations since we should find a monotonically increasing slope of effort in the fixed wage (Hypothesis 2). Still, the increase from the intermediary to the high fixed wage might be triggered by some sort of reciprocal considerations. As our experimental design does not include principals, reciprocity in the common sense is not very plausible. However, there might be some related motive behind the observed behavior which is not directly linked to a counterpart but to the wage received. Some kind of 'bad conscience' of exerting low effort although being paid a very generous fixed wage (12€) might have triggered the subjects' effort. A low effort level could be recognized as inappropriate if a very generous fixed amount is paid, although there is a piece rate. The fixed wage of 12€ is apparently very high because the average payment in such an experiment is well known to be 10€ per hour.

Loss aversion and reference dependent preferences (Hypothesis 3) can also explain only part of the experimental results since the theory might in fact predict lower effort choices for higher wages (compare low and intermediary wage) but would never predict a re-increase for high wages. Moreover, after a decrease, the effort choices should remain on the same level or diminish even more.[14]

Each of the three approaches falls short of providing clarification on its own. Combining loss aversion and the above mentioned explanation of appropriateness of wage and effort we might find another possible explanation. Bewley (1999) provides evidence that employees do not behave strategically

[14] Another straightforward interpretation is of course the occurrence of an income effect. According to the Slutsky equation a higher fixed payment leads to less effort if leisure is a normal good, which can be assumed here. Still, we do not believe in an income effect because the subjects' life income is not likely to be severely influenced by the participation in an experiment. This argument is also supported by the examination of the income variables which are insignificant in any specification of the estimation models. That is, the level of income does not significantly influence the number of correct answers in the experiment.

but feel impelled by morale. He stresses the idea that their efforts are mainly driven by the adequacy of their actual wages.[15] As there is no clearly demanded effort level in our experiment, the agent might derive a desired effort level from the height of the fixed wage he is offered. Consequently, we argue that the appropriateness of the height of the fixed payment and the effort level makes the agent increase performance with rising fixed wages.[16] Then, appropriateness of wage and effort is a kind of a social norm. This would mean that the employee transfers the fixed wage via the social norm into an appropriate level of effort.[17]

If the employee is loss averse in addition, the following interaction is possible: Given a very low wage, the social norm of appropriateness would merely require a low effort level. But the employee suffers from loss aversion if the reference point is not yet reached. So for very low wages the effort would be relatively high in order to get closer to the desired reference wage. Yet, in the region of the reference point effort decreases because the marginal returns are lower after exceeding the reference point. Consequently, for intermediate fixed wages loss aversion might overpower the norm of appropriateness which leads to lower efforts. In turn for sufficiently high fixed wages the norm of appropriateness may dominate loss aversion and as a consequence the effort choices increase. This might be a possible explanation for our experimental results but we are aware that the relevance of this approach should be tested in future research.

---

[15] Bewley (1999) conducts interviews with employees, employers, trade unionists and civil servants.

[16] As the fixed wage is the unconditional part of the compensation the agent might interpret the level of the fixed wage as a measure of trust towards him. In turn the variable payment could be considered as an instrument of control by the agent. Therefore only the height of the fixed wage should be relevant for the agent's considerations of appropriateness.

[17] We could capture this consideration in a model using a utility function similar to Fehr and Schmidt (1999). Though here it would not be inequity aversion towards a counterpart but it would be a kind of fairness/appropriateness regarding the level of the fixed wage.

3.7 Conclusion

Our experimental results indicate that workers indeed care about the fixed wages. The group with the low fixed payment as well as the group with the high fixed wage worked significantly more than those participants who received an intermediate fixed wage. Consequently, we find a U-shaped slope of effort in the fixed wage. However, at first sight it is quite surprising that an intermediate fixed wage level leads to less effort than a low or high one.

Our empirical findings cannot be explained by any of the three theories we considered beforehand, namely standard agency models, reciprocity, and reference dependent preferences. Although these three theories would lead us to predict different relationships between the level of the fixed wage and the chosen effort level, none of them alone is able to explain the experimental results.

The experiment shows that the fixed wage does have a significant influence on the individual's effort decision. Linear incentive schemes seem to be effective if the fixed wage is relatively low or high. In case of a low fixed payment it appears as if the efforts are driven by reference dependent preferences. If a sufficiently high fixed wage is paid the effort level increases in comparison to an intermediate fixed wage. Possibly a social norm of appropriateness is important in this context demanding a certain level of effort to coincide with a certain fixed wage, where not fulfilling the norm leads to disutility. The combination of two different explanatory approaches seems to be arbitrary at first glance. But typically gift exchange experiments with fixed wages provide clean evidence for reciprocity and accordingly appropriateness of wage and effort. Furthermore, the relevance of loss aversion and reference dependent preferences could be shown in previous real effort experiments on piece rates. Since the current experiment is a kind of combination of those why should we not observe both motivations together? Certainly an exact test of the validity of this approach is necessary and should be subject to further research.

Finally, the most important conclusion we draw is that the fixed wage is indeed relevant for the effort decision, and this is true within an incentive contract. Apparently there is an incentive impact of the fixed wage, so pay attention to the fixed pay!

3.8 Appendix to Chapter 3

Proof. Proposition 2: The proof of Propsition 2 is quite similar to the one of Proposition 1. We know that there must be a unique internal optimum. Suppose that, $e^* < e^R$, then $e^* = c'^{-1}(r\beta)$ must hold. Thus, the agent chooses an effort level to the left of the kink. This happens iff

$$\begin{aligned} c'^{-1}(r\beta) &< \frac{R-\alpha}{\beta} \\ &\iff \alpha < R - c'^{-1}(r\beta)\cdot\beta. \end{aligned}$$

The equation defines a unique cut-off value for α such that $e^* = c'^{-1}(r\beta)$ iff α is smaller than the cut-off value. For $e^* > e^R$, the optimum is defined by

$$e^* = c'^{-1}(s\beta).$$

Consequently e^* must be located right of e^R which occurs if

$$\alpha > R - c'^{-1}(s\beta)\cdot\beta.$$

This cut-off value is strictly larger than the cut-off value defined by $e^* = c'^{-1}(r\beta)$ since $r > s$. It follows that in all other cases the agent chooses $e^* = e^R$. ■

Instructions (intermediate wage condition)

Welcome to this experiment! Please do not talk to other participants of this experiment from now on! If you have any questions, please ask them after you read these instructions carefully!

On the sheet you see a block of figures. Every block consists of two rows. Your task at stage one is to sum up the digits per row. Thus, you calculate two sums. (in the given example row 1: Sum of 67, row 2: Sum of 63).

Example: Block 1

																Result
4	1	7	5	6	4	6	6	1	4	1	7	5	6	4	$Sum = 67$	4
8	3	1	2	2	8	8	6	1	8	3	1	2	2	8	$Sum = 63$	

At stage two you should compare if the sum in the upper row is greater or less than the sum in the lower row. If the value of the upper row is greater, you should subtract the value of the lower row from the upper row. If the value of the upper row is less or equal, than the value of the lower row, you should add both values up. This result should be filled into the column 'Result' (In the upper example: 4). Concerning the payment, only this final result is relevant!

Compensation scheme: For your participation you will receive 6€ in any case. In addition you will receive a piece rate of 0.15€ for every correct result. You will have 40 minutes to work on the task. You can split up your time individually for solving the arithmetic problem (look above) or the quizzes attached to this sheet. For solving the quizzes you will get no piece rate.

After 40 minutes, you will receive a questionnaire. Your compensation will only be paid out if you fill out the questionnaire completely. The money will be paid out from 22.07 to 05.08.2004 daily between 10 a.m. and 12 p.m. in room 119b (office: Herbert-Lewin-Str. 2). Please bring along your personal anonymous ID-code. You find the ID-Code in the right corner of this cover sheet. Additionally, you find it on a separate piece of paper attached to your cover sheet. Please tear it off and bring it with you when you come to collect your payment. If you have any further questions now, please raise your hand and wait for an instructor to come to your seat. Please don't ask any questions aloud! Thank you for your participation!

	(1)	(2)	(3)
	Number of point scorings		
High fixed wage	2.00	1.86	1.86
	(1.6)	(1.7)	(1.7)
Low fixed wage	2.88*	2.88*	2.55
	(1.6)	(1.6)	(1.7)
Cologne	-2.28*	-2.26*	-1.90
	(1.3)	(1.4)	(1.4)
Income 1		0.77	0.81
		(1.5)	(1.5)
Income 2		0.75	0.25
		(2.3)	(2.3)
Income 3		4.69	-1.18
		(5.4)	(6.8)
Age			0.33
			(0.2)
Constant	19.3***	18.8***	11.0*
	(1.3)	(1.5)	(5.9)
Observations	181	181	181
R^2	0.04	0.04	0.05

Standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

Table 3.3: Results OLS regression

Chapter 4

The Impact of Transparency in Team Production[1]

4.1 Introduction

In chapters 2 and 3 we have analyzed the impact of individual incentives on the agent's effort decision. In many real world contexts however work results cannot be assigned to a particular person but merely to a group of people or a team. In that case any kind of performance dependent compensation is accompanied by the danger of free-riding and consequently inefficient effort levels. This problem has been discussed comprehensively in the theoretical and empirical literature in recent decades (e.g. Holmström (1982), Alchian and Demsetz (1972) and Newhouse (1973), or Prendergast (1999) for a survey).

However, it is sometimes claimed that peer pressure effects induce team members to work harder and reduce free-riding. But not much attention has been given so far to the precise mechanism that triggers peer pressure. In this chapter, we provide a theoretical explanation for peer pressure effects based on agents' inequity aversion. In a second step, we test the implications of our theory in another real effort experiment. One of the key objectives of our investigation is to study the impact of the informational structure on the

[1]This chapter is based on Mohnen et al. (2006).

development of peer pressure in teams.

We first consider a simple model in which two agents can contribute to a certain task and both are paid based on the total output of the team. Individual contributions depend on the effort exerted and a random component. We compare two settings: In the first setting, both agents can observe their respective contributions only after the end of the game. In the second setting each team member obtains intermediate information on the contribution of his team partner at an interim stage. We assume that both agents are inequity averse over their respective contributions, i.e. each dislikes contributing more or less to the final outcome than his team partner. We show that inequity aversion indeed affects the outcome, provided that interim information on the respective contributions is given. The key mechanism is the following: When it has turned out that one of the agents has contributed more to a task than his team partner an inequitable situation arises. The team partner will have an incentive to reduce this inequity by exerting more effort in the following period. On the other hand, the agent who contributed more has an incentive to reduce his effort level. Hence, an interdependency arises between past outcomes and future contributions. But this has an important effect on the incentives to contribute early in the game as a higher contribution in the present raises the coworker's willingness to raise his effort in the future. Hence, this effect increases the marginal return of effort and leads to higher effort levels early in the game if the respective contributions are made transparent.

However, if individual contributions remain intransparent and become known only at the end of the game, we show that inequity aversion has no impact on equilibrium effort levels as inequity averse agents act exactly like purely selfish agents would.

We test several hypotheses derived from the model in a real effort experiment. Each of the participants in the experiment was requested to perform a tedious counting task on a computer screen which was similar to CN task described in Chapter 2. Teams of two participants were remunerated with a linear piece rate based on the sum of correct answers of both participants. In the first condition participants learned the correct answers which their team partner attained at intermediate stages of the game. In the other condition they received that information only at the end of the experiment. The main hypotheses derived from the model are confirmed: Participants yield more correct answers when knowing that their team mates would be informed about the intermediate outcomes. In addition, we find evidence that the difference between outcomes within the team in the previous period indeed has a strong impact on individual outcomes in subsequent periods. Interestingly we observe an asymmetry in the reaction not predicted by our simple model: Whereas an agent who has contributed more than his colleague in the previous period reduces his effort strongly, having contributed less only leads to a weak increase in future effort.

Among the first introducing the idea of peer pressure in economic team models are Kandel and Lazear (1992). In their approach, team members can choose certain actions that raise the cost of a reduction in individual productive effort for the other team members. However, the mechanism by which the costs of productive effort are increased is not studied endogenously in the model. Barron and Gjerde (1997) find that the existence of peer pressure can result in the optimality of lower powered incentives. Che and Yoo (2001) show that team incentives are more effective in an infinitely repeated game as team members can sanction past behavior of their colleagues. Knez and Simester (2001) find that the introduction of a company wide team incentive scheme at Continental Airlines raised productivity significantly. They explain their result by claiming that mutual monitoring and peer pressure effects counterbalanced free-riding. Backes-Gellner et al. (2006) investigate a Kandel-Lazear type peer pressure model and find a concave relationship between team size and peer pressure which they confirm by empirically analyzing the effort exerted in groups of founders.

There is some previous experimental evidence on peer pressure in teams. Falk and Ichino (2006) show that subjects working at the same time in the same room work harder than subjects in a control treatment with one person working in a room all alone. In contrast to our study they remunerate their subjects with a fixed wage. Hence, increasing the other's incentives to contribute more which will be beneficial in the future can be no motive in their experiment. Moreover, the paper does not give a theoretical explanation of their results within an economic model. Sausgruber (2005) examines peer effects between teams rather than within teams. He finds that the average contribution of the other team in the previous period is positively correlated to the contribution of a person to the own team output in the current period.

Our investigation is of course also strongly related to the literature on social preferences. Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) show that many experimental results can be explained when allowing for the possibility that some agents' utility decreases in the inequality of payoffs between agents. Recently a literature has emerged analyzing the incentive effects of several contractual forms when agents are inequity averse. Examples are for instance Biel (2004), Englmaier and Wambach (2002), Demougin and Fluet (2003), Grund and Sliwka (2005), or Itoh (2004).[2] Most closely related to the theoretical part of our investigation is Huck and Biel (2006) who give a rationale for leadership behavior in a game in which two agents are remunerated with a team incentive scheme. They show that output is higher when one of the agents (i.e. the leader) can act as a first mover as his effort level influences his follower when the latter is inequity averse. Masclet (2002), however, allows for punishment of shirking colleagues which is used in order to achieve equity.

The remainder of the chapter is organized as follows: In Section 4.2 we present the model where the theoretical results regarding the intransparent and transparent setting are developed in Sections 4.2.1 and 4.2.2 respectively. Furthermore the settings are compared in Section 4.2.3. The experimental design and hypotheses are described in 4.3.1 and 4.3.2. Section 4.3.3 deals with the empirical results. The last section concludes.

4.2 The Model

To examine the effects of transparency and inequity aversion on contributions to a team task, we consider a model set-up with two agents as team members

[2] For an overview see for instance Englmaier (2004).

and analyze their effort decisions in a two-period model, $t = 1, 2$. Two risk neutral agents $i = 1, 2$ are working in team.[3] An agent i's contribution y_{it} to the team output in period t consist of his effort level but is additionally affected by some random noise component. Hence, in period t i's contribution to the team output is given by

$$y_{it} = e_{it} + \varepsilon_{it} \tag{4.1}$$

where e_{it} denotes the effort exerted by agent i in period t. The random components ε_{it} are identically and independently distributed and follow a normal distribution $\varepsilon_{it} \sim N(0, \sigma_{\varepsilon}^{2})$.

The agents' costs of effort in each period are represented by a quadratic cost function, $\frac{c}{2} e_{it}^{2}$ with $c > 0$. Both agents receive a fixed wage as well as a team bonus $\beta < 1$ paid at the end of period 2 and hence the wage of an agent i is given by

$$w_i = \alpha + \beta \sum_{t=1}^{2} (y_{it} + y_{jt}) \text{ for } i = 1, 2. \tag{4.2}$$

Given this compensation scheme, the variable payment depends on the team output over both periods, i.e. the sum of the contributions of both team members in both periods.

Furthermore, we assume that both agents are inequity averse over their respective contributions. It is important to note that we assume that agents care for inequity in total contributions and not in effort levels alone. We consider this as important for the following reason. Actual effort levels are not directly observable in many real world work contexts but only output measures which are also affected by external factors beyond the direct control of individual employees. Thus, the utility loss from inequity probably is to a large part driven by differences in observed individual output contributions rather than chosen effort levels alone. Hence, the utility function of an agent

[3] We assume risk neutrality to find clear evidence for the inequity aversion effect. Otherwise, both forces, the risk aversion and the inequity aversion would interact.

over both periods is given by

$$u_i = w_i - \eta \left(\sum_{t=1}^{2} (y_{it} - y_{jt}) \right)^2 - \frac{c}{2} e_{i1}^2 - \frac{c}{2} e_{i2}^2, \tag{4.3}$$

where $\eta \geq 0$ denotes the degree of inequity aversion, y_{it} is the agent's own contribution and y_{jt} that of his respective colleague. Thus, the utility of agent i equals his remuneration less his effort costs in both periods and an expression reflecting the utility loss from inequity in contributions within the team. It is obvious that the agents do not suffer from inequity aversion when both contribute equally over both periods. But if the sum of contributions over both periods is not the same for both agents, the inequity aversion leads to a loss in utility. Since it seems reasonable, that bigger differences lead to more than a proportional effect on this disutility component than small differences in contributions, we model the influence of the differences in contributions as a quadratic function. For simplicity, we assume that both team members have the same degree of inequity aversion η.

In our analysis we compare two cases. In the first case both agents get the information about their respective contribution to the team output at an interim stage after the first period. One possible interpretation would be that they closely work together in a transparent environment and therefore can observe their respective contributions after each period. Another explanation would be that feedback is given by a central manager after the first period who informs both agents about each others' contributions. We call this the *transparency* case. In the second case this intermediate information is not given but agents learn their respective contributions only after the second period. This case is labeled *intransparency* case. Payoff functions are exactly the same in both cases.[4]

[4] As we are interested in the impact of inequity aversion on effort provision, we focus on the agents' decisions rather than on the princpal's. Therefore we do not analyze the principal's optimal contracting. However, if agents are protected by limited liability, it can easily be shown that the principal will choose $\alpha = 0$ and $\beta = 1/2$ in both cases we analyze below.

4.2.1 Intransparency

We first solve the game for the intransparency case where contributions are unobservable at the interim stage after period 1. Here, both agents cannot observe the y_{i1} between periods. We solve the game by backward induction. Note that in period 2 each agent knows his own effort level from period 1 and in equilibrium can infer the equilibrium effort level of the other player. We denote the equilibrium strategy of an agent i in period t by $\hat{e}_{it}$. Thus, in period 2 each agent i maximizes

$$E_i\left[\alpha+\beta\sum_{t=1}^{2}\left(y_{it}+y_{jt}\right)-\eta\left(\sum_{t=1}^{2}\left(y_{it}-y_{jt}\right)\right)^2\right]-\frac{c}{2}e_{i2}^2. \tag{4.4}$$

The agent maximizes his expected utility consisting of his expected material payoff, his expected 'inequity costs' and the costs of his effort. Using that $E\left[x^2\right]=V\left[x\right]+\left(E\left[x\right]\right)^2$ the expression for the expected inequity costs can be simplified and we obtain

$$E\left[\eta\left(\sum_{t=1}^{2}\left(e_{it}+\varepsilon_{it}-\hat{e}_{jt}-\varepsilon_{jt}\right)\right)^2\right]=4\eta\sigma_{\varepsilon}^2+\eta\left(\sum_{t=1}^{2}\left(e_{it}-\hat{e}_{jt}\right)\right)^2. \tag{4.5}$$

It is interesting to note that inequity costs are not only influenced by possible difference in the agents' effort levels but also by the noise affecting their respective contributions. The noisier the environment, the larger is the probability that contributions differ even when effort levels coincide. Hence, (4.4) can be rewritten as

$$\alpha+\beta\sum_{t=1}^{2}\left(e_{it}+\hat{e}_{jt}\right)-4\eta\sigma_{\varepsilon}^2-\eta\left(\sum_{t=1}^{2}\left(e_{it}-\hat{e}_{jt}\right)\right)^2-\frac{c}{2}e_{i2}^2. \tag{4.6}$$

To maximize the expected utility we differentiate (4.6) with respect to e_{i2} and obtain agent i's reaction function

$$e_{i2}^*\left(e_{i1},\hat{e}_{j1},\hat{e}_{j2}\right)=\frac{\beta-2\eta\left(e_{i1}-\sum_{t=1}^{2}\hat{e}_{jt}\right)}{c+2\eta}.$$

This is the best response of agent i given his own effort level in the first period e_{i1} and the strategies of his colleague j in both periods. An agent's effort level in period 2 will be higher, the higher the equilibrium strategies of his counterpart j, and the lower, the higher his own actual effort in period 1 has been. Both effects are stronger the higher the inequity aversion parameter η. The effects are due to an agent's potential utility loss from a 'guilty conscience' when he contributes less than his colleague. But for a given total effort exerted by his colleague, the more an agent has contributed in the past, the less he has to work in the present to achieve equitable contributions.

In the intransparency case an agent knows in the first period, that his contributions will not be observed by the other agent before the second period ends. Hence, the agent cannot react to his counterpart's contribution. But an agent i will anticipate that his first period effort choice will affect his own second period reaction $e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2})$.

Hence, agent i maximizes in period 1

$$\begin{aligned} \max_{e_{i1}} \alpha + \beta \left(e_{i1} + e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2}) + \sum_{t=1}^{2} \hat{e}_{jt} \right) - 4\eta\sigma_\varepsilon^2 \qquad (4.7) \\ -\eta \left(e_{i1} + e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2}) - \sum_{t=1}^{2} \hat{e}_{jt} \right)^2 - \frac{c}{2}\left(e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2})\right)^2 - \frac{c}{2} e_{i1}^2. \end{aligned}$$

Solving this optimization problem, we obtain the following first period reaction function

$$e_{i1} = \frac{\beta + 2\eta \left(\sum_{t=1}^{2} \hat{e}_{jt} \right)}{c + 4\eta}.$$

By solving for the subgame perfect equilibrium, we obtain the following result:

Proposition 3 *When agents cannot observe their respective contributions at the interim stage after $t = 1$, equilibrium effort levels are independent from the agent's degree of inequity aversion and are given by*

$$e_{it} = \frac{\beta}{c} \text{ for } i = 1, 2 \text{ and } t = 1, 2.$$

Proof: See Appendix.

It is important to note that without transparency, the introduction of inequity aversion over contributions does not alter the equilibrium predictions compared to a case where agents are purely selfish. As the agents do not receive any information about their colleague's contribution to the team result after the first period, agents cannot adapt their effort choices contingent on their observation of their colleague's contribution.

4.2.2 Transparency

In the transparency case, the intermediate contributions to the team output of both agents are observed. We examine how the observability of y_{i1} and y_{j1} influences the agents' decisions. Note that the agents are symmetrically informed at the beginning of period 2 as both have observed the contributions in the first period y_{1t} and y_{2t}. Hence, in period 2 agent i maximizes his expected utility

$$\begin{aligned} &\max_{e_{i2}} \alpha + \beta \left(y_{i1} + y_{j1} + e_{i2} + \hat{e}_{j2} \right) \\ &-\eta E \left[\left(y_{i1} - y_{j1} + e_{i2} + \varepsilon_{i2} - \hat{e}_{j2} - \varepsilon_{j2} \right)^2 \middle| \, y_{i1}, y_{j1} \right] - \frac{c}{2} e_{i2}^2. \end{aligned} \tag{4.8}$$

This can be simplified and, from the first order condition we get the following reaction function

$$e_{i2}(\Delta y_{i1}, \hat{e}_{j2}) = \frac{\beta - 2\eta \left(\Delta y_{i1} - \hat{e}_{j2} \right)}{2\eta + c}.$$

where $\Delta y_{i1} := y_{i1} - y_{j1}$.

Note that e_{i2} is increasing in $\hat{e}_{j2}$ and y_{j1} and decreasing in y_{i1}. Besides the monetary motive, the agent's effort choice is guided by his objective to minimize the expected disutility from inequity. An agent works the harder, the more effort is exerted by his colleague in equilibrium in the current period and the higher the contribution of his colleague in the previous period has been. He works the less, the more he himself has contributed in the previous period.

Substituting the optimal strategies of his colleague e_{j2} and solving for e_{i1} leads to the following result:

Proposition 4 *In the transparency case, the equilibrium effort levels in period 2 are*

$$e_{i2}(\Delta y_{i1}) = \frac{\beta}{c} - \frac{2\eta}{4\eta + c}\Delta y_{i1} \text{ for } i = 1, 2 \tag{4.9}$$

The second period effort of an agent i is strictly decreasing in the difference of first period contributions $\Delta y_{i1} = y_{i1} - y_{j1}$.

Proof: See Appendix.

Now, the efforts exerted in the second period depend upon the actual output in the first period. When the first period contributions of both agents differ, the agent with the higher contribution exerts a lower and the one with the lower contribution chooses a higher second period effort level than in the case without transparency (or the case where all agents are purely selfish). An agent whose first contribution has turned out high, feels less obliged to exert higher effort levels in the second period. On the other hand, an agent whose first period contribution has been low, strives to make up for the differences in contributions by exerting a higher effort level in the second period. Only for $\Delta y_1 = 0$ both agents' effort levels coincide.

In period 1 agent i maximizes

$$\max_{e_{i1}} E\left[\alpha + \beta \sum_{t=1}^{2}(y_{it} + y_{jt}) - \eta\left(\sum_{t=1}^{2}(y_{it} - y_{jt})\right)^2\right] - \frac{c}{2}e_{i1}^2 - E\left[\frac{c}{2}e_{i2}^2\right] \tag{4.10}$$

taking the effect of e_{i1} on the equilibrium strategies in period 2 into account.

Again, as the situation is symmetric, agent j's reaction function is symmetric, too. After substituting the reactions function of j in i, we get the following result:

Proposition 5 *In the transparency case, the agents choose identical effort levels in period* 1 *given by*

$$e_{i1} = \frac{\beta}{c} \cdot \frac{6\eta + c}{4\eta + c} \text{ for } i = 1, 2. \tag{4.11}$$

Efforts are increasing in the degree of the agents' inequity aversion.

Proof: See Appendix.

The first period equilibrium effort is increasing in the degree of the agents' inequity aversion η. The more inequity averse the agents are, the higher will be the first period effort level.

4.2.3 A Comparison of both Settings

It is straightforward to compare the effort decisions in both settings. When interim contributions are unobservable, inequity averse agents act exactly like purely selfish agents and equilibrium effort levels coincide and are equal to β/c for both periods. However, in the transparent case the agents in general choose different second period effort levels to counterbalance inequitable outcomes from the first period. But note that from equation (4.9) we know that

$$e_{i2}(\Delta y_{i1}) + e_{j2}(-\Delta y_{i1}) = \frac{2\beta}{c}.$$

Hence, aggregate effort levels do not differ between both settings in the second period. The effort adaptions chosen to counterbalance either higher or lower contributions from the first period exactly offset each other. The agent who contributed more works less and the one with the smaller contribution increases his effort level by exactly the same amount.

However, comparing the first period efforts for both informational settings yields the following result:

Proposition 6 *In the transparent setting both agents exert higher first period effort than in the intransparent setting as*

$$\frac{\beta}{c} \cdot \frac{6\eta + c}{4\eta + c} > \frac{\beta}{c}$$

for $\eta > 0$. This leads to higher expected utilities for the agents when contributions are made transparent at the interim stage.

Proof: See Appendix.

In the first period, both team members always choose the same effort level within each situation but in the transparency case, they both work harder. Hence, inequity aversion helps to reduce the free-rider problem even in a finitely repeated game, but it does so only when intermediate contributions are transparent. As higher effort levels by an agent in the present increase his team partner's costs to free-ride in the second period due to a utility loss caused by the inequity aversion it is rational to increase one's own effort in the first period. Hence, this mechanism gives an explanation for observed peer pressure effects based on social preferences.

Furthermore note that by exerting higher efforts in the transparency case the agents attain more efficient outcomes. As it can easily be shown they never exceed the effort level chosen under full cooperation. Therefore the agents are always better off in a situation with transparent contributions. Note also that the principal will always prefer a situation in which the agents observe each others' contributions at the interim stage providing him with higher efforts in the first period and thus higher total effort. In the following section we test whether these theoretical results are valid by examining the behavior of participants in a real effort experiment.

4.3 Experimental Evidence

4.3.1 Experimental Design and Procedures

Subjects were randomly assigned to a team consisting of two persons each. The team partners were never exchanged, thus each participant kept his counterpart for the entire duration of the experiment. Altogether 7 periods were played each of which lasted 8 minutes. In all periods the participants were offered to work on a simple counting task which was similar to the CN task in Chapter 2 but the subjects solved it on the computer. Periods 1 and 7 (i.e. the first and the last period) were different from the other periods. In these periods subjects were paid according to their individual performance in the counting task. For each correct answer they received a piece rate of 0.14€ which was directly transferred to their individual account. We will call these periods the individual periods in the following. The individual periods at the very beginning and at the end of the experiment were introduced to provide individual ability measures for this specific task for each participant. In addition we are able to observe whether learning or fatigue influenced the results.[5]

However, in all other periods (i.e. periods 2, 3, 4, 5 and 6) the participants had the choice between two options at each point in time: First, they could again work on counting the correct number of sevens in blocks of random numbers. In that case 0.14€ were transferred to the team account for each correct answer. After each period the amount of the team account was equally divided between the two team members and we therefore call these periods the team periods. But, alternatively, the subjects could push a time-out button at any time during the team periods which caused the screen to be blocked for 25 seconds. During this time they were not able to continue with the counting task but were paid 0.10€ to their individual account. We introduced the time-out button to make sure that subjects had significant opportunity costs of working on the counting task.

Whenever an answer in the calculation task was incorrect 0.01€ were

[5] For instance this might be worthwhile if we want to distinguish between fatigue and an endgame effect when performance happens to decrease in the last team period.

subtracted from the individual account irrespective of the type of period (individual or team period). After each team period every subject was informed about

- the number of blocks he had worked on in the previous period
- the number of correct answers he had given in the previous period
- the number of times he pushed the time-out button in the previous period.

In order to check for the impact of the team mate's performance on the individual effort, we introduced two conditions. Subjects in the intransparent condition merely received the information described above. In the transparent condition however participants were additionally informed about the number of correct answers provided by their respective counterpart.

Altogether 208 students of various faculties participated in the experiment. We recruited all subjects using an online recruitment system by Greiner (2003). We conducted 7 sessions in May 2006 at the Cologne Laboratory for Economic Research, University of Cologne and used the experimental software z-tree by Fischbacher (1999) for programming the experiment. Each subject was seated in an own cabin with a computer terminal. The participants were not given any oral instructions, instead all informations were described on a sheet of paper laid out in each cabin.[6] Additionally, the structure of the payoff function and other important features of the experimental design were repeated on the computer screen before each period started. At the end of the experiment all subjects were informed about the sum of their earnings in the individual periods and the team periods. Additionally, they were all paid a fixed show up fee of 2.50€. On average the subjects earned approximately 16€. The whole procedure took about 90 minutes.

4.3.2 Hypotheses

From the theoretical model developed in Section 4.2 we derive three main hypotheses on the subjects' behavior. Proposition 4 indicates that subjects

[6] For the printed instructions see appendix.

will adapt their counting effort levels according to the interim information.

Hypothesis 1: *In the transparency condition the subjects having made the higher team contribution in the previous period compared to the team mate will decrease their effort in the next period and vice versa.*

Hypothesis 1 captures the effect that the agents dislike differences in contributions between team members. Hence, the one who has taken an early lead will reduce and the one who has fallen behind will feel obliged to raise his effort in order to counterbalance the previous period outcome difference. Moreover, we have seen in Proposition 5 that transparency leads to a reduction in free-riding causing higher first period effort choices. Therefore we hypothesize:

Hypothesis 2a: *In periods 2 to 5 outcomes are higher in the transparent than in the intransparent setting.*

This hypothesis illustrates the agents' incentives to exert higher efforts to increase the team mate's costs of free-riding. This effect is likely to persist as long as there is at least one subsequent period. However, in the last period the outcome cannot influence the team mate's future behavior.

Hypothesis 2b: *In the last team period (period 6), efforts decline in the transparent condition and the difference between the outcomes in the two informational settings diminishes.*

The aggregate results should therefore be equal for both settings in the last period.

4.3.3 Results

Let us first consider the subjects' effort adjustment depending on the relation of their own contribution to the team output compared to the counterpart's contribution. We approximate the participants' effort by the number of correct answers given (score)[7] to investigate this relation. Still, we have two

[7] The expressions in parentheses are the variable labels used in the following analysis.

further effort measures to check the robustness of the score results which are on the one hand the number of blocks the subject has worked on irrespective of the correctness of the result (blocks), and on the other hand the number of time outs the subject has taken (time outs).[8] As effort adjustment according to the counterpart's contribution is not possible for the intransparent condition we restrict the analysis of the effort adjustment to the data from the transparent condition. Let us first examine how a subject's change in scored points from one period to the next is influenced by the score difference within his team in the previous period. According to Hypothesis 1 we expect that those subjects having taken a lead over their team mate in the previous period will reduce their effort in the next, while those team members who have fallen behind will increase their effort. To analyze this, we measure the difference in scores within in the team by taking the first subject's score less the second subject's score (team difference). So for the subject who produced the higher score, the team difference variable has a positive sign and for the subject with the lower performance the team difference variable has a negative sign. Second, we approximate the participants' effort adjustment by the score change variable. The score change is defined by the difference between a subject's score in the current period less the same subject's score in the previous period. Thus, if a subject increased the score between two consecutive periods, the score change is positive while it is negative if the subject decreased the score. We predict a negative relationship between the team difference in the previous periods and the score change from the previous to the subsequent periods. Figure 4.1 shows this relationship for periods 3, 4, 5, and 6. We skip period 2 because there has not been a precedent team period as period 1 is an individual period. According to Figure 4.1, indeed, this negative relationship seems to exist. We test the results illustrated in the previous figure by applying a fixed effects estimation which we restrict to periods 2 to 6.[9] The models in Table 4.1 examine the influence of the team difference on the change between periods in the three performance measures

[8]Note that the relation between the measures score and blocks and effort can be assumed to be positive and that it should be negative for time outs.

[9]Period 2 is used only as a baseline for the changes in period 3.

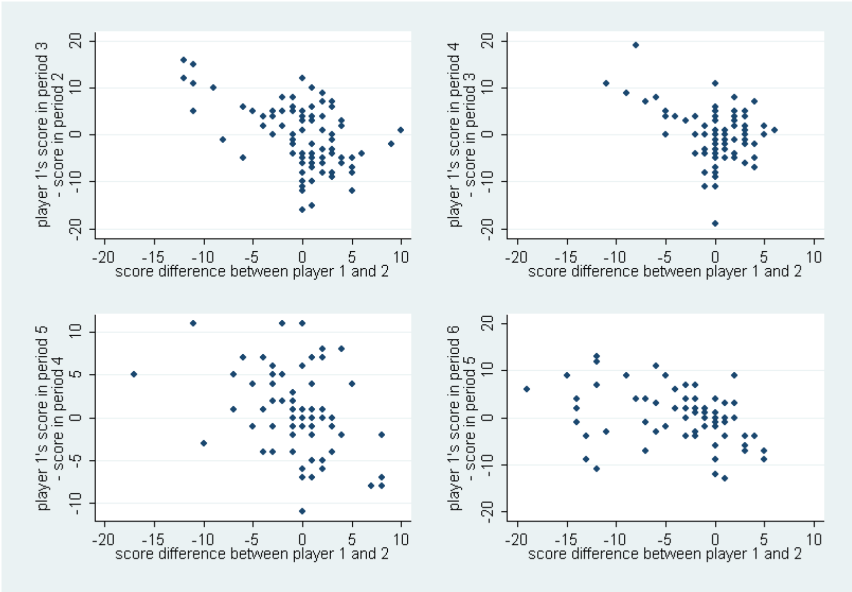


Figure 4.1: Relation between team difference in $t-1$ and score change

(score, blocks and time out).[10] Considering model (1) we observe a highly significantly negative influence of the team difference on the score change. From that we may conclude that the more a subject's score exceeded the counterpart's score in the past period, the more the subject decreased his effort in the subsequent period. In turn the participants raised the score if confronted with a negative team difference. The models for block change (2) and time out change (3) confirm this conclusion. Model (2) demonstrates that higher performance differences within the team do not only lead to a decrease in the number of correct answers between the current and the previous period but also to a decrease in the number of blocks worked on. In turn a higher performance difference in the team causes the usage of the time-out button to rise (see model (3)). It seems that the subject with the higher outcome feels free to take time outs more often to counterbalance the team difference. This is in line with our theoretical predictions.

The estimations in Table 4.2 refer to the influence of the sign of the team

[10]The variables block change and time out change are defined (analogously to score change) as the differences between the values of the respective measure in the current period less the value of the measure in the previous period.

	(1) Score change	(2) Blocks change	(3) Time outs change
Team difference	-0.563***	-0.649***	0.539***
in t-1	(0.054)	(0.073)	(0.061)
Constant	-2.739***	-2.852***	3.068***
	(0.39)	(0.53)	(0.44)
Observations	352	352	352
Number of subjects	88	88	88
R^2	0.36	0.27	0.27

Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Fixed effects estimation (period dummies included)

Table 4.1: Fixed effects estimation I

difference illustrating the difference between being the one who has had the lead in the previous period (higher performer) and the one with the lower team contribution (lower performer). The higher performer variable is defined as the maximum of the subject's score less the counterpart's score and zero. Analogously the lower performer variable is defined as the maximum of the counterpart's score less the respective subjects score and zero. Thus, in one period either the variable lower performer or the variable higher performer is positive, both are zero if both agents have contributed equally to the team output in the previous period. Regarding models (1) and (2) we observe a significantly negative impact of being the higher performer. This is very intuitive because the more a person exceeded his coworker's score in the past period, the more he reduces the current effort to enforce equality in contributions. This effect occurs for the score change and the blocks change. The coefficient for time outs change is of course positive as taking more time outs indicates lower effort on the task. Hence, the analyses of all three performance measures provide evidence for our theoretical predictions. For the lower performer variable all coefficients show the opposite sign. Hence, the

	(1) Score change	(2) Blocks change	(3) Time outs change
Higher performer in t-1	-0.987***	-1.092***	0.932***
	(0.081)	(0.11)	(0.093)
Lower performer in t-1	0.139*	0.207*	-0.145
	(0.081)	(0.11)	(0.093)
Constant	2.338***	0.398	0.518
	(0.51)	(0.72)	(0.59)
Observations	352	352	352
Number of subjects	88	88	88
R^2	0.45	0.33	0.34

Standard errors in parentheses, *** p < 0.01, ** p < 0.05, * p < 0.1
Fixed effects estimation (period dummies included)

Table 4.2: Fixed effects estimation II

impact of performing less than the counterpart seems to be consistent with our theory as well but is less pronounced in effect size and level of significance. This is particularly interesting with regard to our model in which we assume that differences in contributions are equally detrimental for advantageous and disadvantageous inequality. In contrast to that the subjects' behavior implies that they suffer less from being the lower performer than from being the higher performer and therefore make stronger effort adjustments in the latter case. Yet, taking the results together we may conclude that Hypothesis 1 is at least partly confirmed. Although we have seen that subjects adapted their effort according to their team mates' scores, it is not clear whether this behavior leads to a net increase in scores. Therefore we compare the scores between the transparent and the intransparent condition. In Figure 4.2 the point scorings over periods are displayed. Again we only consider the team periods 2 to 6. The figure indicates that subjects in the intransparent condition yield lower scores than those in the transparent con-

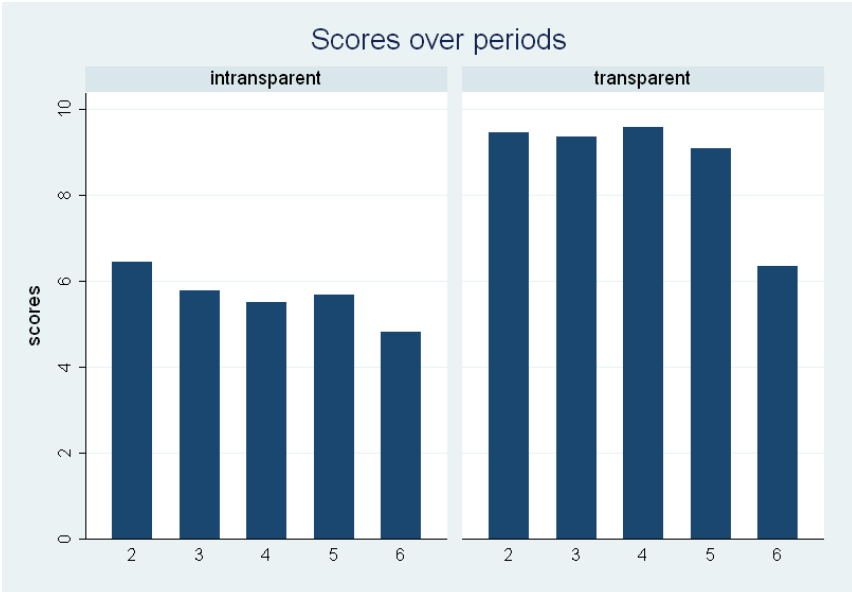


Figure 4.2: Scores over periods by condition

dition. In addition there seems to be a considerable endgame effect as scores drop in the transparent condition in the last period.[11] In the intransparent condition the effect seems to occur as well but appears to be less pronounced. We analyze each period separately with a non-parametric Mann-Whitney-U-test to check whether the differences between conditions are significant. If we take the mean score per team for each period we obtain the following results which are illustrated in Table 4.3. For periods 2 to 5 we find significantly higher scores for those subjects who were informed about their team mate's performance between periods. This confirms our Hypothesis 2a. Hence, we may conclude that indeed transparent conditions in team production enforce higher effort levels and reduce free-riding. However, we cannot conclude from that result alone whether the underlying motive reflected by our theory of inequity aversion is valid. It could as well be that participants in the transparent condition developed a feeling of commitment or team spirit as they got more information about their team mates. Yet, considering period 6

[11] The difference between period 5 and 6 is highly significant for the transparent condition if we use the Wilcoxon matched pairs signed rank test for dependent samples (absolute z-value 4.720).

Periods	Transparent vs. intransparent condition
2	(3.467)***
3	(3.360)***
4	(3.075)***
5	(2.259)**
6	(0.965) not significant

*** $p < 0.01$, *** $p < 0.05$, * $p < 0.1$
Absolute value of z-statistic in parentheses

Table 4.3: Results Mann-Whitney-U-test

we do not find significant differences between the conditions anymore. This seems to be an endgame effect implying that the subjects in our experiment did not exert more effort due to a stronger feeling of commitment or team spirit. Rather than that they rationally anticipated the positive impact of a currently higher effort on the team mate's future effort. Of course, in the last period this motive became irrelevant as there was no subsequent period. Therefore the experimental results provide evidence for the validity of the theoretical approach in Section 4.2. One might also argue that fatigue caused the scores to decline in period 6. However, using the Wilcoxon matched pairs signed rank test for dependent samples, we find a highly significant increase in scores in period 7 which is an individual period (absolute z-value 5.774). This indicates that subjects might not have suffered from fatigue but indeed did not have incentives to stick to the high effort level in period 6 providing evidence for Hypothesis 2b.

4.4 Discussion

In this chapter we have theoretically and experimentally analyzed the impact of inequity aversion on effort exertion on a team task in order to provide a possible mechanism for peer pressure. We have shown that not only in theory but also in the experiment the effect strongly depends on the informational setting in which the task is solved.

The theoretical results suggest that when there is no intermediate information about the team mate's contribution to the team output, the agents' inequity aversion does not influence in the choice of effort. Hence, the players always act identically to purely selfish ones. However, if the team mate's contribution can be observed at an interim stage, the agents deviate from the purely selfish choice to minimize their costs of inequity in contributions. Therefore they adjust their effort levels in the second period such that the agent having worked less increases and the other one decreases his effort.

The subjects in our experiment indeed adapted their effort according to the interim information they received about their counterparts' contribution such that they counterbalanced contributions to increase equity. However, in contrast to our model they did not seem to adjust in a symmetric way. According to the observed behavioral pattern it seems that subjects who exerted higher performance and therefore were in a disadvantageous position strongly decreased their effort in the subsequent period. However, those participants having had an advantage due to having been the lower performing team member still felt the need to reduce inequity by increasing effort but to a lower extend. This led to a less pronounced absolute adjustment for the lower performers. The notion of an asymmetric effect of advantageous and disadvantageous inequity has also been pointed out by Fehr and Schmidt (1999) which seems to be reflected in this experiment.

Another conclusion from the theoretical model is that the later period effort adjustment, being anticipated in the earlier period, leads to higher first period effort choices. That is, agents work harder in the early in the game to increase the team mate's costs of free-riding and therefore his future effort adjustment. Hence, the total sum of efforts is greater in the transparent than in the intransparent case.

Indeed our experimental data provide evidence for a positive impact of transparent conditions in team production on effort provision. Thus, we can confirm that transparency of individual contributions might help to reduce free-riding. However, when the time horizon is fixed in which a team works together in the same formation, the cooperation within the team might break down near the end.

All in all the data seem to confirm the theoretical results but yet there are some limitations to the interpretation. For simplicity we neglected the question of individual abilities in the model and in the experiment. In the model we assumed that agents are equally productive. This assumption is of course violated in the experiment in which the subjects certainly had different abilities to solve the counting task. Still, including abilities into the model might provide interesting insights especially if each agent has more accurate information about his own ability than about his team mate's. Then, the team members might try to use their first period effort as a signal of their ability.

Another important issue is that the model is limited to two periods while the experiment lasted for 5 team periods so that the experimental design does not reflect the theoretical set-up precisely. Yet, the greater number of periods was necessary in the experiment to give the participants the chance to learn and to understand the experimental design while analyzing the model over 5 periods would add much complexity but probably little to the results.

The model as well as the experiment show that the agents are better off in a setting in which the counterpart's contributions are revealed. Thus, we may conclude that transparency leads to more efficient outcomes and should therefore be preferred from a social welfare point of view. Hence, the main practical implication seems to be that transparent work environments might be desirable in firms when employees work in groups. If workers adhere to a social norm evoking feelings of guilt or anger respectively for differences in team contributions, transparency might enhance the chance to develop beneficial mutual monitoring and in turn yield an effective reduction in free-riding.

4.5 Appendix to Chapter 4

Proof. Proposition 3: In period 1 an agent maximizes (4.7) taking into account his own second period reaction function $e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2})$ and the equilibrium strategies of his team partner in both periods. As e_{i2}^* will be chosen to maximize exactly the same expression given e_{i1} we can use the Envelope Theorem. Partially differentiating (4.7) with respect to e_{i1} we obtain

$$\beta - 2\eta \left(e_{i1} + e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2}) - \sum_{t=1}^{2} \hat{e}_{jt} \right) - ce_{i1} = 0.$$

Substituting $e_{i2}^*(e_{i1}, \hat{e}_{j1}, \hat{e}_{j2}) = \frac{\beta - 2\eta\left(e_{i1} - \sum_{t=1}^{2} \hat{e}_{jt}\right)}{c+2\eta}$ yields

$$\begin{aligned}
\beta - 2\eta \left(e_{i1} + \frac{\beta - 2\eta\left(e_{i1} - \sum_{t=1}^{2} \hat{e}_{jt}\right)}{c+2\eta} - \sum_{t=1}^{2} \hat{e}_{jt} \right) - ce_{i1} = 0 &\Leftrightarrow \\
\beta - 2\eta \left(\frac{\beta + 2\eta\left(\sum_{t=1}^{2} \hat{e}_{jt}\right)}{c+2\eta} - \sum_{t=1}^{2} \hat{e}_{jt} \right) = \left(c + 2\eta - \frac{4\eta^2}{c+2\eta} e_{i1} \right) e_{i1} &\Leftrightarrow \\
\beta - 2\eta \left(\frac{\beta - c\left(\sum_{t=1}^{2} \hat{e}_{jt}\right)}{c+2\eta} \right) = \left(\frac{c^2 + 4\eta}{c+2\eta} \right) e_{i1} &\Leftrightarrow \\
e_{i1} = \frac{\beta + 2\eta \sum_{t=1}^{2} \hat{e}_{jt}}{c+4\eta}&.
\end{aligned}$$

The second period effort therefore becomes:

$$e_{i2} = \frac{1}{c+2\eta} \left(\beta - 2\eta \frac{\beta + (2\eta - c - 4\eta) \sum_{t=1}^{2} \hat{e}_{jt}}{c+4\eta} \right) = \frac{\beta + 2\eta \sum_{t=1}^{2} \hat{e}_{jt}}{c+4\eta}.$$

Hence, it follows that $e_{i1} = e_{i2} = \left(\beta + 2\eta \sum_{t=1}^{2} \hat{e}_{jt}\right) / (c + 4\eta) =: e_i$. We can substitute the reaction functions of agent j and obtain $e_{i1} = e_{i2} = \frac{\beta}{c}$. ■

Proof. Proposition 4: Using that $E[x^2] = V[x] + (E[x])^2$, the expression for the expected inequity costs can be simplified and becomes

$$\eta E\left[(y_{i1} - y_{j1} + e_{i2} + \varepsilon_{i2} - \hat{e}_{j2} - \varepsilon_{j2})^2\right] = \eta\left[2\sigma_\varepsilon^2 + (y_{i1} - y_{j1} + e_{i2} - \hat{e}_{j2})^2\right].$$

Hence, (4.4) can be simplified and agent i maximizes

$$\max_{e_{i2}} \alpha+\beta\left(y_{i1}+y_{j1}+e_{i2}+\hat{e}_{j2}\right)-2\eta\sigma_{\varepsilon}^{2}-\eta\left(y_{i1}-y_{j1}+e_{i2}-\hat{e}_{j2}\right)^{2}-\frac{c}{2}e_{i2}^{2}.$$

From the first order condition we obtain the reaction function

$$e_{i2}(\Delta y_1,\hat{e}_{j2})=\frac{\beta-2\eta\left(\Delta y_{i1}-\hat{e}_{j2}\right)}{2\eta+c}. \tag{4.12}$$

From where we get the equilibrium strategies. ■

Proof. Proposition 5: In period 1 agent i maximizes

$$E\left[\alpha+\beta\sum_{t=1}^{2}\left(y_{it}+y_{jt}\right)-\eta\left(\sum_{t=1}^{2}\left(y_{it}-y_{jt}\right)\right)^{2}\right]-\frac{c}{2}e_{i1}^{2}-E\left[\frac{c}{2}\left(e_{i2}(\Delta y_{i1})\right)^{2}\right] \tag{4.13}$$

where

$$\begin{aligned}
& E\left[\alpha+\beta\sum_{t=1}^{2}\left(y_{it}+y_{jt}\right)-\eta\left(\sum_{t=1}^{2}\left(y_{it}-y_{jt}\right)\right)^{2}\right] \qquad (4.14)\\
= \ & \alpha+\beta\cdot\left(e_{i1}+\hat{e}_{j1}+E\left[e_{i2}(\Delta y_{i1})+e_{j2}(-\Delta y_{i1})\right]\right)\\
& -\eta E\left[\left(e_{i1}+\varepsilon_{i1}-\hat{e}_{j1}-\varepsilon_{j1}+e_{i2}(\Delta y_{i1})+\varepsilon_{i2}-e_{j2}(-\Delta y_{i1})-\varepsilon_{j2}\right)^{2}\right]
\end{aligned}$$

As

$$\begin{aligned}
E\left[e_{i2}(\Delta y_{i1})+e_{j2}(-\Delta y_{i1})\right] &= \frac{2\beta}{c} \text{ and}\\
E\left[e_{i2}(\Delta y_{i1})-e_{j2}(-\Delta y_{i1})\right] &= -\frac{4\eta}{4\eta+c}\Delta y_{i1}.
\end{aligned}$$

(4.14) is equivalent to

$$\begin{aligned}
&\alpha+\beta\left(e_{i1}+\hat{e}_{j1}+\frac{2\beta}{c}\right)\\
&-\eta E\left[\left(e_{i1}+\varepsilon_{i1}-\hat{e}_{j1}-\varepsilon_{j1}-\frac{4\eta}{4\eta+c}\Delta y_{i1}+\varepsilon_{i2}-\varepsilon_{j2}\right)^{2}\right]
\end{aligned}$$

where the expected value of the 'inequity costs' are

$$
\begin{aligned}
& V\left[e_{i1}+\varepsilon_{i1}-\hat{e}_{j1}-\varepsilon_{j1}-\tfrac{4\eta}{(4\eta+c)}\Delta y_1+\varepsilon_{i2}-\varepsilon_{j2}\right] \\
& +\left(e_{i1}-\hat{e}_{j1}+E\left[-\tfrac{4\eta}{(4\eta+c)}\Delta y_1\right]\right)^2 \\
= \; & V\left[\left(1-\tfrac{4\eta}{(4\eta+c)}\right)(\varepsilon_{i1}-\varepsilon_{j1})+\varepsilon_{i2}-\varepsilon_{j2}\right] \\
& +\left(e_{i1}-\hat{e}_{j1}-\frac{4\eta}{(4\eta+c)}(e_{i1}-\hat{e}_{j1})\right)^2 \\
= \; & \left(\left(1-\tfrac{4\eta}{(4\eta+c)}\right)^2+1\right)2\sigma_\varepsilon^2+\left(1-\tfrac{4\eta}{(4\eta+c)}\right)^2(e_{i1}-\hat{e}_{j1})^2 . \qquad (4.15)
\end{aligned}
$$

Using the reaction function given in Proposition 4, the expected second period effort costs are

$$
\begin{aligned}
E\left[\frac{c}{2}\left(e_{i2}(\Delta y_{i1})\right)^2\right] &= E\left[\frac{c}{2}\left(\frac{\beta}{c}-\frac{2\eta}{4\eta+c}\Delta y_1\right)^2\right] \qquad (4.16) \\
&= \frac{c}{2}\left[2\sigma_\varepsilon^2\left(\frac{2\eta}{4\eta+c}\right)^2+\left(\frac{\beta}{c}-\frac{2\eta}{4\eta+c}(e_{i1}-\hat{e}_{j1})\right)^2\right].
\end{aligned}
$$

Hence, putting it all together, by substituting expected inequity costs (4.15) as well as expected second period effort costs (4.16) the agent's optimization problem (4.13) becomes

$$
\begin{aligned}
& \max_{e_{i1}} \alpha+\beta\left(e_{i1}+\hat{e}_{j1}+\frac{2\beta}{c}\right)-\eta\left[\left(1-\frac{4\eta}{4\eta+c}\right)^2+1\right]2\sigma_\varepsilon^2 \\
& -\eta\left(1-\frac{4\eta}{4\eta+c}\right)^2\left[e_{i1}-\hat{e}_{j1}\right]^2-\frac{c}{2}e_{i1}^2 \\
& -\frac{c}{2}\left[\left[\frac{2\eta}{4\eta+c}\right]^2 2\sigma_\varepsilon^2+\left(\frac{\beta}{c}-\frac{2\eta}{4\eta+c}(e_{i1}-\hat{e}_{j1})\right)^2\right]
\end{aligned}
$$

The first order condition is

$$
\beta-2\eta\left(1-\frac{4\eta}{4\eta+c}\right)^2[e_{i1}-\hat{e}_{j1}]-c\left(\frac{\beta}{c}-\frac{2\eta}{4\eta+c}(e_{i1}-\hat{e}_{j1})\right)\left[-\frac{2\eta}{4\eta+c}\right]-ce_{i1}=0
$$

leading to the reaction function

$$e_{i1} = \frac{\frac{6\eta+c}{4\eta+c}\frac{\beta}{c} + 2\eta\frac{2\eta+c}{(4\eta+c)^2}\hat{e}_{j1}}{1 + 2\eta\frac{2\eta+c}{(4\eta+c)^2}}.$$

Solving for the first period equilibrium strategies yields the given result. ■

Proof. Proposition 6: Agent i's utility is

$$\alpha + 3\frac{\beta^2}{c} - 4\eta\sigma_\varepsilon^2. \tag{4.17}$$

in the intransparent setting and

$$\begin{aligned} &\alpha + \frac{2\beta^2}{c}\left(\frac{6\eta+c}{4\eta+c} + 1\right) - 2\eta\sigma_\varepsilon^2\left[\left(1 - \frac{4\eta}{4\eta+c}\right)^2 + 1\right] \\ &-c\left(\frac{2\eta}{4\eta+c}\right)^2\sigma_\varepsilon^2 - \frac{\beta^2}{2c}\left[\left(\frac{6\eta+c}{4\eta+c}\right)^2 + 1\right] \end{aligned} \tag{4.18}$$

in the transparent setting. Checking whether (4.18) is greater than (4.17), we obtain

$$\left(8\eta^3 + 3\eta^2 c\right)2\sigma_\varepsilon^2 + \left(3\eta^2 + \eta c\right)\frac{\beta^2}{c} > 0.$$

As all $\eta \geq 0$ and $c > 0$ the inequation is always valid. ■

Instructions (intransparent condition)

Welcome to this experiment:

Please read these instructions carefully:

- All decisions you make in this experiment are anonymous
- At the end of the experiment you will also be paid anonymously
- During the experiment no communication is allowed

Procedure:

This experiment consists of 7 periods, each of which lasts for 8 minutes. Periods 1 and 7 differ from rounds 2 to 6.

Periods 1 and 7 are played as follows:

You are requested to count the number of the 'Sevens' (the digit 7) in a block of random numbers. You enter this number in the corresponding box and confirm your choice with OK.

Payoff for period 1 and period 7:

For each correct answer you receive 14 Cent.

ATTENTION: For each wrong answer 1 Cent will automatically be subtracted from your individual account.

Periods 2 to 6 of this experiment are played as follows:

You will be randomly assigned to another participant where the two of you form a team. This assignment remains the same for the whole experiment. That is that you play with the same counterpart from period 2 to period 6. However, you and your counterpart stay anonymous.

In periods 2 to 6 you can choose between two options:

1. You can work on the block of random numbers and count the numbers of the 'Sevens' in this block (analogously to round 1 and 7).

OR

2. You can take a time out by pressing the 'time out' button. If you press the 'time out' button the counting task is blocked and you cannot continue working on any block of random numbers for 25 seconds.

Payoff for period 2 to period 6:

1. For each correct answer in the counting task, 14 Cent will be paid into your team account. For each correct answer your counterpart gives, 14 Cent will also be paid into your the team account.

The team account will be equally divided between you and your counterpart at the end of each period.

2. Each time you press the 'time out' button 10 Cent will be paid into your individual account.

ATTENTION: For each wrong answer 1 Cent will automatically be subtracted from your individual account.

After each period you will see a table on the screen displaying the following information:

1. The number of blocks you finished in the previous period
2. The number of correct answers you gave in the previous period
3. The number of time outs you took in the previous period

The table with the details for the particular period looks as follows:

Screenshot

Your counterpart receives the same information about his own performance. Neither you nor your counterpart learn about the number of correct answers the other team member gave in the respective round. After the last period (period 7) you will be informed about your total payoff. In addition to that you receive a show up fee of 2.50€. Please stay at your seat until we call on your cabin number. Please bring along these instructions and your cabin number. Thank you for your participation!

Chapter 5

Performance Feedback and the Role of Honesty[1]

5.1 Introduction

In light of various religious commandments and folk wisdom sayings from around the world, most societies seem to adhere to ethical standards that proscribe lying. However, standard economists claim that the completely selfish homo economicus uses dishonesty whenever it serves to maximize his own utility. In many situations in everyday and economic life we depend on information provided by other people which we cannot prove to be correct. In turn, losing trust in other people's statements due to negative experiences with deception, this might result in considerable costs for the respective organization or society. While society may bear the cost of individuals' lies, for the liar dishonesty seems to be particularly attractive in situations characterized by a conflict of interest. Evidence for this has been found in studies on cheap talk games (for theoretical models see Park (2005) and his literature overview; for experimental studies see e.g. Charness (2000); Crawford (1998) provides a survey of experiments with cheap talk). In a recent experimental study Gneezy (2005) examined deception in situations with a conflict of interest. He finds that the fraction of subjects utilizing lies strongly depends

[1]This chapter is based on Mohnen and Pokorny (2006b).

on the consequences that arise from lying. If it is more beneficial for them, subjects lie more often. Yet Gneezy (2005) found that subjects also care about the consequences the lies have on their counterpart. Thus, if deception strongly reduces their counterpart's payoff, they will refrain from lying more frequently than if there is little impact on their counterpart's payoff.

With regard to organizations we often discuss the conflict of interest between principal and agent. When information asymmetries arise, these conflicts can have negative effects on social welfare. One possible information asymmetry might be the principal having information on the agent's productivity or ability to accomplish a certain task. Bénabou and Tirole (2003) develop a similar approach by assuming that the principal has better information on the agent's effort costs. In this context they examine the role of bonus payments as a signal of task attractiveness. Related to Bénabou and Tirole (2003) we make the assumption, that if an agent enters a new job his employer has a better perception of how capable the employee is to fulfill his job than the agent himself does. As he is new, he typically does not know how well-matched his abilities are with the tasks corresponding to the job. Due to more experience an employer gets a good impression about the employee's actual ability for a certain task when observing him working on the job. Hence, the employee depends on the employer's feedback concerning his actual ability. Many firms try to overcome the described asymmetry by conducting feedback talks so that the employee receives the information gathered through observation. However, as we will show, the employer might not give feedback on the employee's ability honestly. The employer, anticipating how the staff will react to the feedback information, might have an incentive to give strategic rather than truthful feedback.[2] Therefore we question whether an employee can always trust the information he is given in feedback discussions.

[2] In the literature we find a huge amount of psychological studies on feedback, for an overview see Latham and Locke (1990), Latham and Locke (1991) or Kluger and DeNisi (1996). Whereas in the economic literature, feedback in an employer-employee relationship is - to the best of our knowledge - so far not well analyzed. Mohnen and Pokorny (2005) provide a simple economic model on the agent's reaction to feedback and test their implications using an experimental data set.

Suppose there is a set of heterogenous employees in which some demonstrate high and some demonstrate low ability for the work. Then the question arises: should the employer be honest with the full range of employees and truthfully reveal their individual abilities to them. To answer this question, it is important to consider the employee's reaction to this feedback, taking into account that his boss can decide whether to give feedback or not and whether to lie or not. Whenever employees are interested in their production output they produce, their own ability to do the task becomes an important piece of information for them. For example, a candid assessment of their ability might be important to an employee if he is paid an output-contingent wage as in chapters 2 and 3 or if the output influences his future career opportunities.

Obviously, in these cases, people will adapt their effort levels if they learn that they are more or less productive because their marginal returns are higher or lower than they expected a priori. Therefore giving truthful feedback does not lead to higher profits for the employer in every case. Since the employer's interest is to encourage her staff to work hard, she would have an incentive to tell each of them that they have high ability if they believed her feedback information naively. The rationale behind this is the following: Given an performance-dependent wage, an employee will work more if he believes that he has a high ability because high achievers have higher marginal returns than those with low ability. Thus, the employer might not be well advised to give every employee honest feedback on their ability, if in some cases, honest feedback would yield lower returns for her.

In this chapter, we try to shed light on this conflict of interest in two steps. First, we analyze a two-period signaling game in which the principal knows her agent's ability perfectly after the first period while the agent does not. We assume that there are two types of agents, those with high ability and those with low ability, where the agent's ability is equivalent to his work productivity. In both periods the agent makes an effort decision and his wage is contingent on the output he produces. The agent's output is affected by his actual ability for the task and the effort he makes. After the first period, the principal can give positive (ability is high), negative (ability is low) or no feedback to the agent. If deception (i.e. giving positive feedback although the true ability is low and vice versa) does not generate any costs, no separating equilibria exist in the game. As merely pooling equilibria can be established, the ability information will never be revealed and consequently the agent will choose an optimal effort level according to his a priori expectation of his ability. Introducing positive costs of lying makes separating equilibria feasible provided that these (signaling) costs are sufficiently high. In these equilibria, the agent's actual ability is revealed directly or indirectly by the feedback information of the principal. Thus, our (1.) main result is: If the costs of lying are sufficiently high, separating equilibria are feasible and therefore the feedback information is informative. Additionally, we examine the impact of the height of the principal's marginal return on the agent's effort and obtain the (2.) main result: If the principal benefits less from the agent's effort, i.e. her marginal returns are lower for each effort unit exerted by the agent, she is less likely to lie. As the returns from deception decrease with the marginal returns from the agent's effort, she has a lower incentive to lie for any given costs of lying.

In the second part of the chapter we test these theoretical results in a laboratory experiment in which the principals might have incentives to give positive feedback to their agents' independent of the true ability. Similar to the theoretical model we matched one principal to one agent for two periods. Deviating from the previous chapters, we do not use a real effort task. Instead the agents chose an abstract effort level and had to bear the respective monetary costs of effort in each of the periods.[3] After the first period the principal was informed about both, the agent's actual ability and his output, and decided whether she wanted to give feedback to the agent or not. If she decided to give feedback she could choose between positive (ability is high) or negative (ability is low) feedback. We did not introduce monetary costs of lying. Thus, any deception did not directly affect the principal's monetary payoff in the experiment. Hence, the costs of lying are an individual's psychological costs.

[3]However, a replication of this experiment with a real effort task might be very interesting.

According to our theoretical framework, we ran two main conditions: a baseline and a low return condition. In the latter, we decreased the principal's returns on the output produced by the agent compared to the baseline condition. This allowed us to examine whether the principals lie less often if lying becomes less attractive monetarily.

Our experimental data give evidence for the following conclusions: Principals give positive feedback less frequently if matched to a low ability agent than if matched to a high ability agent. This indicates that they do consider the true ability when giving feedback and therefore the existence of costs of lying. In contrast to our theoretical results and to Gneezy (2005) we do not find more deception with higher returns on the agents' efforts. Instead we observe exactly the opposite pattern with decreasing frequency of deception with higher returns on lying. So apparently lower returns do not enhance honesty but decrease the fraction of honest feedback.

The remainder of the chapter is structured as follows: In Section 5.2 we analyze the signaling model. Section 5.3 describes the experimental design and procedures. We derive hypotheses in Section 5.4 and present the experimental results in Section 5.5. The last section concludes the chapter.

5.2 Theory

5.2.1 Model Set-up

We analyze a two-period model with riskneutral parties. A principal employs an agent whose ability a is unknown to both of them ex ante. In each period, the agent produces an output y_t which is linear in his chosen effort level e_t with $y_t(a, e_t) = a \cdot e_t$, $t = 1, 2$. Effort is assumed to be observable but not verifiable. The agent's ability a to produce the output y_t can be either low or high, i.e. $a \in \{a_L; a_H\}$ with $E[a] = \bar{a}$ and $a_L < \bar{a} < a_H$. The prior distribution is such that high ability a_H occurs with probability τ and low ability a_L with probability $1-\tau$. Further, we assume quadratic effort costs of the form $c(e_t) = \frac{1}{2}e_t^2$ and a linear wage $w(y_t) = \beta \cdot y_t$ is paid by the principal in each period where $\beta > 0$. Hence, the agent's utility in period t is given by

$$U_t^A = \beta \cdot y_t(e_t, a) - \frac{1}{2}e_t^2 \tag{5.1}$$

and the principal's net payoff on the agent's effort equals

$$W_t^P = (\theta - \beta) \cdot y_t(e_t, a) \tag{5.2}$$

where $\theta > \beta$. The parameter θ represents the returns on the agent's output for the principal. Hence, $\theta - \beta$ is the net return on output after the variable compensation has been paid. Naturally, the players interests are not completely in line because the principal wants her agent to exert as much effort as possible but the agent has to bear the respective costs of effort.

When only the prior information about his actual ability is available to him, the agent would choose an effort level e^* according to

$$e_t^*(\tau) = \beta\left[\tau \cdot a_H + (1 - \tau) \cdot a_L\right] = \beta \cdot \bar{a} \equiv \overline{e} \tag{5.3}$$

in order to maximize the expected value of his periodic utility function (5.1). Note that the agent's effort decision depends on his ability which is unknown to him ex ante.

However, in this setting the principal has the opportunity to give feedback to her agent after period $t = 1$. The principal learns whether the agent's actual ability is low with a_L or high with a_H. The idea behind this is the following: Due to more experience the principal is able to determine if her agent has high or low ability a. Given this information the principal can decide if and how she wants to give feedback on ability to the agent. For this purpose we assume that the principal can select one option from the following set $f \in \{f_0; f_L; f_H\}$, where f_0 denotes the case in which the principal gives no feedback at all. Furthermore, f_L states that the actual ability is low which is negative feedback and f_H specifies that the actual ability is high which is positive feedback. Recall that the agent cannot verify the principal's feedback until the game ends.

This set-up constitutes a signaling game with three possible actions from which the principal can choose. After observing the principal's feedback signal, the agent might update his beliefs about his own ability, namely calculate a posterior belief that his actual ability is high. We will refer to this posterior probability as $\pi(f)$ in the following (off equilibrium beliefs are denoted by $\widetilde{\pi}(f)$). As a consequence, the agent will base his second period effort decision on this updated information. The key question is, can the agent trust the information revealed in the feedback signal or not?

When giving feedback, the principal can deceive the agent in two ways: First, she can tell the agent that the actual ability is low when it is in fact high. Second, she can tell him that the actual ability is high even though it is low. The latter option is much more attractive for the principal if the agent naively trusts the principal's feedback. Nevertheless let us assume that in both cases the principal has to bear constant costs $l \geq 0$ if she lies to her agent.[4] These costs are equal for both cases of deception. If she gives no feedback no costs of lying arise. The crucial issue is whether it is in the principal's best interest to tell the truth.

Finally, we summarize the time structure of the game:

- In $t = 1$ the agent chooses an effort level e_1.
- Afterwards the principal observes the agent's effort and his ability and decides whether or not to give feedback to her agent. The available set of feedback signals is $f \in \{f_0; f_L; f_H\}$, i.e. giving no feedback, or claiming the actual ability is high or low. By assumption deception generates costs $l \geq 0$ for the principal. The agent receives the signal f.
- Given this feedback information f, the agent makes a second effort decision e_2 in period $t = 2$.[5]
- Both parties receive their respective payoffs.

[4] Note that a standard homo economicus would surely feature $l = 0$.

[5] Besides, we assume that the agent cannot quit the contract after the first period. He must stay in the firm in period $t = 2$. The same is true for the principal, she cannot lay off her agent and replace him with another employee. With this assumption we want to capture the feature that the labor market does not provide arbitrary job alternatives. Otherwise, one could think of an agent leaving the firm after receiving a negative feedback signal f_L.

This situation constitutes a dynamic game with asymmetric information. Thus, the equilibrium concept is a perfect Bayesian equilibrium characterized by

- the principal's strategy described by a combination of $f \in \{f_0; f_L; f_H\}$ for both outcomes of the actual ability,
- a posterior probability $\pi(f)$ that the actual realization of the agent's ability is high characterizing the agent's beliefs after having observed the principal's feedback decision, and
- the agent's strategy denoted by an effort level in each period, i.e. e_1 for period 1 and $e_2(\pi)$ for period 2.

The results are developed in the subsequent section.

5.2.2 Theoretical Results

It turns out that the agent's equilibrium strategy in $t = 1$ is easy to analyze: In equilibria in which the principal plays a strategy which is not contingent on the agent's first period effort, (5.3) maximizes the agent's utility in period 1. Then, the optimal effort choice is based on the agent's prior belief $e_1^* = e_1^*(\tau)$. A strategy contingent on e_1 can only be part of an equilibrium if the principal is indifferent between two feedback decisions. As will become clear, this does not arise in the generic cases. Hence, we focus on those equilibria in which the principal does not condition her feedback on the agent's first period effort.

The agent's effort decision in $t = 2$ depends on his updated information, i.e. $e_2^* = e_2^*(\pi)$. Consider a situation in which the agent was naive such that he believed that the feedback information would always be correct. Then he would choose

$$\beta \cdot a_H \equiv e_H$$

when receiving positive feedback and

$$\beta \cdot a_L \equiv e_L$$

when hearing negative feedback. However, the principal can act strategically when giving feedback. Therefore, we examine which perfect Bayesian equilibria exist in the described set-up. In the following we focus on the second period as the effort level in the first period is (5.3). In the appendix, a detailed derivation of the existing equilibria can be found.[6]

Proposition 7 *A separating equilibrium exists in which all types give honest feedback when the costs of lying are sufficiently high according to*

$$l \geq \beta \cdot a_L \left[a_H - a_L \right] (\theta - \beta) \tag{5.4}$$

i.e. the principal gives positive feedback f_H if the actual ability is high and negative feedback f_L otherwise. If the costs of lying are zero, there are no separating equilibria in the game in which the principal gives honest feedback for one or both ability realizations.

Proof: See Appendix.

When the described equilibrium exists, the principal's returns on deception are too low compared to the corresponding costs and therefore honesty pays off for her. As the costs of lying represent the signaling costs in this game, separating equilibria never exist for zero costs of lying. The simple intuition is the following: when there are zero costs of lying a principal employing a low ability agent will always imitate the signal of a principal employing a high ability agent because with no costs of lying every signal is free. Consequently, only pooling equilibria are feasible. However, if costs of lying are sufficiently high a separating equilibrium exists in which the principal always tells the truth. Only in the latter case does the feedback indeed contain valuable information.

We obtain a similar separating equilibrium under the same conditions in which the principal gives positive feedback to a high ability agent but no feedback to a low ability agent. That is, the principal does not lie but avoids truthful feedback for low ability agents.

If the costs of lying decrease, deception becomes more attractive to the principal in the case of a low ability agent. Thus, for sufficiently low costs, we

[6] For simplicity we do not consider any hybrid equilibria.

find that all principals will give positive feedback, i.e. claim that the agent's actual ability is high independent of the true realization.

Proposition 8 *With sufficiently low costs of lying, namely*

$$l \leq \beta \cdot a_L \left[a_H - a_L\right] (\theta - \beta) \left(\tau - \widetilde{\pi}(f)\right) \tag{5.5}$$

there is a pooling equilibrium in which the principal always reports that the true ability is high, i.e. gives positive feedback.

Proof: See Appendix.

As decreasing costs of lying represent lower signaling costs, reporting high ability when the true ability is low becomes more attractive. Then the principal is more likely to pretend that the actual ability is high even if it is low. Note that the separating equilibrium described in Proposition 7 and the pooling equilibrium in Proposition 8 can never coexist for given costs of lying l as long as $\tau < 1$. That is, we have separating equilibria for higher and pooling equilibria for lower costs of lying.

An important question with regard to the described setting is whether the revelation of information pays off in terms of efficiency. For an efficiency analysis we compare the ex ante expected social welfare of the no feedback pooling equilibrium with a separating equilibrium. In the no feedback pooling equilibrium the principal does not have to bear any costs of lying, so it is the most beneficial pooling equilibrium for him. As all pooling equilibria lead ex ante to the same expected utility for the agent, the no feedback pooling equilibrium is the most efficient one. Furthermore, neither the principal's nor the agent's ex ante expected utility differs across the two separating equilibria. Hence, both separating equilibria are equally efficient and we can select any of them for comparison to the no feedback pooling situation. Comparing the outcomes of the two types of equilibria, yields the following result:

Proposition 9 *In any separating equilibrium the ex ante expected utility of both agent and principal is higher than in any pooling equilibrium. Hence, the social welfare is higher as well.*

Proof: See Appendix.

In a separating situation the true ability is revealed, enabling the agent to adjust the optimal effort level to the new information, i.e. he chooses a high effort level when the true ability turns out to be high and a low one when it turns out to be low. In contrast to that, he does not learn the ability information in the pooling equilibrium. So the agent exerts an effort corresponding to the ex ante expected value of ability. This effort level is from an ex post point of view either too high (when the true ability is low) or too low (when the true ability is high). It follows that the agent is better off in a separating equilibrium. The principal also benefits from the separating equilibrium which is driven by the agent's productivity. Ex ante the expected effort exerted by the agent is equal in the pooling and the separating equilibria. However, in the separating equilibria the agent exerts high second period effort when he is more productive and low second period effort when he is less productive. So altogether the principal profits more from the higher effort level expended by an agent with high ability than he suffers from the lower effort level demonstrated by an agent with low ability making the separating equilibria more beneficial than the pooling equilibria. From that we can draw the conclusion that even the principal is better off in a separating equilibrium. Thus, both parties benefit from separating equilibria and consequently social welfare is higher compared to any pooling equilibrium.[7]

Moreover, we consider the case in which the principal's return on the agent's effort can vary. Then, the parameter θ in equation (5.2) is different for various kinds of jobs. Hence, we obtain the following effect on the separating equilibrium described in the following Corollary:

[7] Here the princpal faces a commitment problem. However, a credible commitment of the principal to always honestly report the agent's true ability would have a positive impact on social welfare.

Corollary 10 *If the principal's return θ increases, the separating equilibrium in which both types of principals give truthful feedback becomes less probable as condition (5.4) gets more restrictive.*

Since the returns on lying increase relatively to the constant costs of lying, it is more likely that the principal deceives the agent if the true ability is low.

To analyze how people actually behave in this situation, we ran a lab experiment in order to test our theoretical results. Below we describe our experimental set-up.

5.3 Experimental Design and Procedures

In our experiment we used a two-player feedback game.[8] In this game we had two types of players, namely principals and agents. They played in matched groups, each consisting of one principal and one agent.[9] All subjects played 15 rounds and a round consisted of two periods each. The agent's ability was determined by a random draw before each round and remained constant over the two periods of a round. It was common knowledge, that the ability could take value 2 or 6 with equal probability hence $\tau = 0.5$. In the first period the agent was asked to choose an effort level out of the set $\{1, 2, 3, 4, 5, 6, 7, 8\}$. For each effort level the agent had to bear the respective costs presented by the quadratic cost function $c(e) = \frac{1}{2}e^2$. After the first period the principal observed the agent's first period decision and learned his actual ability i.e. the true realization of the random draw. Then the principal could choose one out of three messages to provide information to the agent:

- The actual ability is 2.
- The actual ability is 6.
- No message.

[8] Mohnen and Pokorny (2005) present a similar game in a lab experiment but as they merely analyze the agent's effort decision as a reaction to performance feedback they do not introduce principals.

[9] Each player kept his role, either principal or agent, over all rounds.

After the agent had received the selected message the second period started. The agent was again asked to select an effort level out of the set described above with the corresponding costs. After that both players were informed about their individual payoffs in the first and the second period and the total payoff of the particular round. The agent's payoff for each period of a round was calculated according to

$$\pi_A = a \cdot e - \frac{1}{2}e^2 + 16$$

with a being the actual ability and e the selected effort level. Thus, the agent was compensated by a linear incentive contract with a piece rate $\beta = 1$ and a fixed payment of 16. For the principal the payoff for each period was

$$\pi_P = (\theta - \beta) \cdot a \cdot e.$$

Hence, the output produced by $a \cdot e$ is multiplied by a return factor θ.

Note that payoffs of all players only depend on the effort choice as there are no direct monetary costs of feedback for the principal, neither for giving true feedback information nor for lying. Thus, if there are costs of lying they would be only indirect personal or psychological costs. Obviously, it is in the principal's best interest that the agent exerts higher effort levels. However, for each possible realization of ability there is an inner solution maximizing the agent's objective function. If the agent knows that the true ability level is 6 (2), his best answer is to choose an effort level of 6 (2). A priori the agent's best effort choice is 4 as both values of the ability are equally probable.

The parameterization varied between conditions to test our theoretical results. Hence, in the baseline condition the principals' net return $\theta - \beta$ was equal to 2.[10] In the low return condition we set the return factor θ to a lower level, i.e. $\theta - \beta = 1.5$. In both conditions principals and agents were matched

[10] In one session of the baseline condition we had 30 subjects so no subject was to play with the same counterpart twice since we played 15 rounds. In the other session of the baseline condition we only had 28 subjects. In that case it happened that subjects were matched who had been counterparts before. However, as no one ever knew to whom he was matched and whether and when he was matched to that person again, we do not believe that this had an influence on our results.

to a new anonymous counterpart after each round. We also tested a third condition, which we call partners condition, without new matching. We will refer to this in Section 5.5.4.

Altogether 172 students of various faculties participated in the experiments. For the recruitment of the participants we again used the online recruitment system by Greiner (2003). All sessions were run at the Cologne Laboratory for Economic Research, University of Cologne in November 2005 and January 2006. We used the experimental software z-tree by Fischbacher (1999) for programming the experiment. The procedure was such that the participants were given oral instructions by the experimenter which were accompanied by several presentation slides including screen shots. After that each subject went into an own cabin in which the computer terminal was placed. In addition to the oral instructions the subjects received short printed instructions which were laid out in each cabin.[11] After the last round of the experiment subjects were requested to complete a questionnaire including an instrument measuring the individual's Machiavellianism developed by Henning and Six (1977).[12] In the experiment we used tokens where each token was worth 0.167€. For paying the subjects, one round was drawn by lot. The sum of both periods' profits of that round was paid out at the end of the sessions. Additionally, all subjects received a show up fee of 2.50€ independent of the number of tokens gained. Subjects earned approximately 12€ on average. The whole procedure took about 90 minutes.

5.4 Hypotheses

In the first periods only the agents make decisions. We predict that they choose the optimal effort level of 4 which corresponds to the a priori expected value of the ability. This should be the case in all conditions (Hypothesis 1).

Featuring pooling and separating equilibria, the signaling game does not provide clear predictions for how the subjects might react in the second periods unless we know the actual costs of lying l. In the case where individual

[11] For the printed instructions see appendix.

[12] The details concerning this instrument are described in Section 5.5.2.

costs of lying are relatively low, we are more likely to observe a pooling equilibrium in which principals lie to their agents when their true ability is low. That is, regardless of the true ability, the principals report that the true ability is 6 and there are no differences in feedback behavior between rounds with different abilities. In this case, in line with Proposition 8, the agents should stick to an effort level of 4 in the second period according to the a priori expected value of their ability (Hypothesis 2).

Conversely, if the costs of lying are relatively high, the principals might act honestly even if the true ability is low (Proposition 7). In that case we should observe differences in the principals' feedback behavior between those rounds when they encounter agents with an ability of 2 and those with a true ability of 6. Hence, the agents should adjust their effort decision in period $t = 2$. After receiving the feedback that the actual ability is 2, the optimal effort choice should be 2, and when being told that the actual ability is 6, the second period effort should equal 6 (Hypothesis 3).

Note that apart from the costs of lying, principals should ex ante prefer the separating equilibrium with an expected second period payoff of $20\,(\theta - \beta)$ for the principal whereas in the pooling equilibrium the expected second period payoff is $16\,(\theta - \beta)$.

According to Corollary 10 and Gneezy (2005) we should observe more honesty when the principals' return to the agents' effort is lower. Consequently, there should be differences in the feedback behavior of subjects in the baseline and the low return condition. And in turn, the agents should be more likely to trust their principals resulting in more frequent adjustments of their second period effort levels in the low return condition compared to the baseline condition results (Hypothesis 4).

5.5 Empirical Findings

5.5.1 Descriptive Results

Let us first consider some descriptive results to get a first impression of the empirical outcomes. With regard to the principals the interesting question

	Baseline	Low return
No feedback	.047	.067
Negative feedback (lie)	.022	.008
Positive feedback	.931	.925

Table 5.1: Feedback decisions across conditions when ability was high

	Baseline	Low return
No feedback	.310	.152
Negative feedback	.227	.210
Positive feedback (lie)	.463	.649

Table 5.2: Feedback decisions across conditions when ability was low

is obviously: what kind of feedback did the respective subject send to her agent? As each principal has three options regarding this decision, namely no feedback, negative feedback and positive feedback, we differentiate between these three options in the analysis. Still, it is important to consider under which circumstances the principal chose the respective option, i.e. if the true ability was low or high. Hence, we give statistics on the feedback behavior conditional on the true ability. Table 5.1 presents the fraction of no feedback, negative feedback and positive feedback, for the baseline and the low return conditions over all subjects and all rounds given that the true ability was high. According to Table 5.1 giving positive feedback was chosen much more often than the other two options if the true ability was high. That is, in both conditions almost all subjects gave feedback truthfully. Hence, subjects obviously consider positive feedback to be the most attractive option to choose independent of the condition they are in. In the low ability case, the results look different as shown in Table 5.2. In this case most principals gave positive feedback although the true ability is low which is deceptive. Moreover, the no feedback message is sent much more often than it is with high ability indicating that the principals use this option to avoid telling lies.

Let us now consider the agents' effort decisions. In the first period of each round subjects do not have any information on the actual ability realization. But agents can condition their second period effort decision on the feedback

	Baseline	Low return
1st period effort	3.96 (1.64)	3.87 (1.52)
2nd period effort/no feedback	2.89 (1.30)	3.58 (1.72)
2nd period effort/negative feedback	2.43 (1.08)	2.52 (1.46)
2nd period effort/positive feedback	5.04 (1.79)	4.37 (2.07)

Table 5.3: Agents' effort decisions (standard deviation in parentheses)

information received between periods. According to our theory, this feedback might make a difference and therefore we present the second period efforts separately for positive and negative feedback in Table 5.3. In the first periods the average effort level chosen by the agents was almost 4 which is close to the theoretical prediction as described in Hypothesis 1.[13] In the second periods the selected efforts indeed differ depending on the feedback observed. The second period effort seems to be higher when positive feedback was given compared to the cases when no or negative feedback was given. In the following sections we will analyze the data in detail and interpret the results.

5.5.2 The Existence of Costs of Lying

Do Principals Suffer from Costs of Lying?

Our first question is: do costs of lying even exist? As discussed above, the principals might have an incentive to report that the agent's ability is high even though his true ability is low. If the agents trust their principals naively, the latter benefit from lying in low ability cases. In contrast there is no obvious advantage to deception in high ability cases.

To examine whether principals indeed suffer from costs of lying, we considered the feedback decision separately depending on whether the actual ability is low or high in the baseline condition. The first column of Table 5.1 supports the argument we offered in the last paragraph. Around 93% of the principals gave positive and thus honest feedback if their agent showed high ability in the baseline condition. For the low ability case (Table 5.2,

[13] Testing whether the means are different from 4 with a two-sided t-test, we find significant differences for neither condition.

first column), the fraction of honest feedback is much lower. Here only about 23% of the agents received honest negative feedback. Principals gave positive feedback substantially more often to high ability than to low ability agents, despite the incentive to always give positive feedback independent of the actual ability. Hence, we suppose that principals do consider the true ability value when deciding what feedback to give and there are at least some honest principals among our subjects. Indeed, there seem to be costs of lying, although here they are not monetary but instead subjective psychological costs.

However, not all of the principals always give positive feedback when the true ability is high. There are two possible causes, one technical and one theoretical, for this scenario: First, it suggests that some of the subjects did not understand the design of the game and therefore made mistakes accordingly. Second, they might have consciously chosen to lie or to give no feedback for high ability because they thought that the agents would not believe the positive feedback anyway.

To further analyze our data, we use variables for each principal in the sample reflecting the individual feedback behavior. As we want to grasp information from all rounds, we computed the fraction of cases in which a particular principal chose to give positive feedback when the true ability was low (recall that the mean of this variable over all subjects is .46, see Table 5.2, first column). However, if the actual ability is low, positive feedback is, of course, a deception. We also determined the fraction of positive feedback messages for each principal when the true ability was high which gives a rate of truthful feedback to the agents (recall that the mean of this variable over all subjects is .93, see Table 5.1, first column).

Significant differences between these fraction variables for each individual would imply that there is some obstacle deterring principals from deceiving their agent's if the true ability is low. As these two variables are dependent, we analyze the differences between these kinds of feedback behavior by applying the Wilcoxon matched pairs signed rank test for dependent samples. The results indicate that principals make use of the positive feedback option highly significantly more often if the true ability is high (absolute z-value

4.648).[14] From that we may conclude that subjects indeed consider the actual ability and have a preference for being honest to their agent. This result may imply that at least some principals choose strategies corresponding to a separating equilibrium which is described in Hypothesis 2.[15]

Are There Certain Types of Principals?

According to the theory the costs of lying play a crucial role in the principals' feedback decisions. Consequently it is important to learn more about these costs. It would be interesting to know whether principals select a certain strategy at the beginning and stick to it all over the game such that we can determine certain types of principals. If so, we could distinguish players who lie whenever it seems profitable to them from those who never lie. Figure 5.1 shows the percentage of lies across all low ability rounds per principal. Obviously, there are some principals in both conditions whose fraction of dishonest feedback is zero. Hence, they never lied when the agent's actual ability was low throughout the game. A zero rate of lying does not mean, however, that the respective subject always revealed the true ability as cases in which the subject gave no feedback are included. In the baseline condition there are more principals who never lied than in the low return condition. In turn, in the low return condition we find far more principals with a lie rate of one. These are subjects who lied each time when the actual ability was low. We might claim that these principals have lower costs of lying and therefore played the pooling equilibrium (Hypothesis 2). So we have some subjects who are steadfast regarding the strategy they chose such that they either never lied or always lied with low ability. However, there are other types of players that apparently changed their strategy over rounds. They have fractions of lies strictly greater than zero but strictly smaller than one. The individual behavior of those mixing subjects is presented in Figure 5.3 and

[14] For the low return and partners condition we find a similar relationship with absolute z-values of 3.572 and 4.510 respectively.

[15] With Proposition 7 and the parametrization of the experiment, we can calculate that false feedback (i.e. deception) in case that the true ability is low increases the principals' payoff by 16. Consequently, we can compute that the disultility from lying must be higher than 16 for subjects giving feedback honestly.

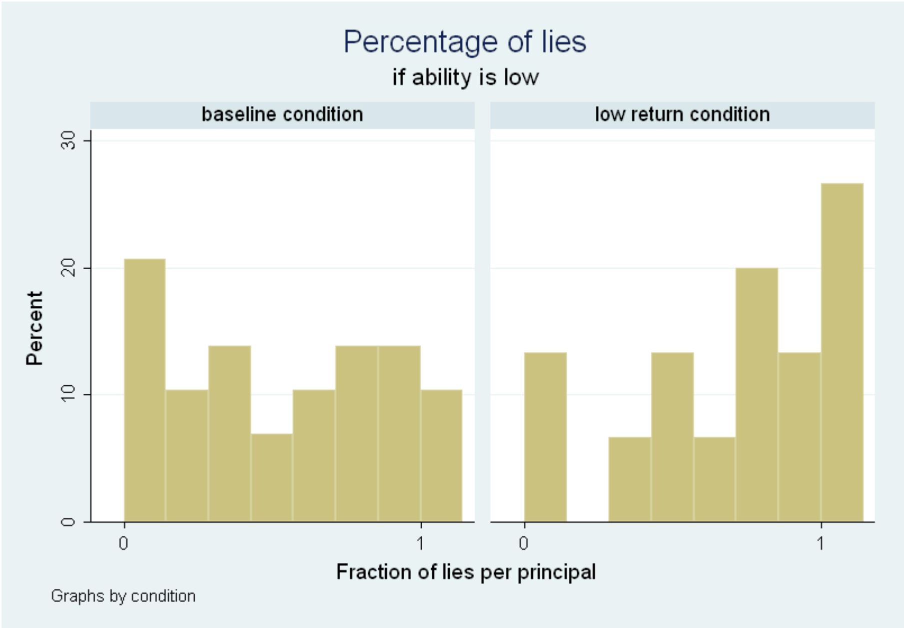


Figure 5.1: Fraction of lies per principal (if ability was low)

5.4 in the appendix. It seems that there is no systematic behavioral pattern in the data. Thus, it remains unclear why they changed their strategy.

To further analyze the types of principals we observed in the experiment, we look at the subjects' Machiavellianism scores to find out whether the degree of Machiavellianism can explain the individual's feedback behavior. In social psychology the notion of Machiavellianism denotes personality traits reflecting manipulativeness, egoism and the attitude that the end justifies the means and was first established by Christie and Geis (1970). Psychological tests to measure Machiavellianism have been used before to predict behavior in trust games (see Gunnthorsdottir et al. (2002)). In order to measure the subjects' Machiavellianism we used an instrument by Henning and Six (1977) consisting of 18 items which are answered on a 6-point response scale (where 1 denotes strong disagreement and 6 strong agreement). The human traits that are covered by Machiavellianism suggest that individuals with higher scores have lower costs of lying and hence we use it as a proxy for the individual costs of lying. We test the influence of the Machiavellianism score on the

	Lie
Machiavellianism score	0.338**
	(0.14)
Partners	-0.894***
	(0.26)
Low return	0.465*
	(0.24)
High ability	-1.664***
	(0.37)
Constant	-1.693***
	(0.45)
Observations	1290
Number of subject	86

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Round dummies included

Table 5.4: Results random effects probit estimation

decision to lie with random effects probit estimation. The results are shown in Table 5.4. 'Partners' is a dummy variable with value one if the subject was in the partners condition and zero otherwise. Analogously 'low return' is a dummy variable with value one for the low return condition and zero otherwise. The dummy variable 'high ability' takes the value one if the actual ability in the respective round was high and zero if it was low. The model in Table 5.4 shows that indeed a subject's probability to lie increases with a higher Machiavellianism score. Thus, we may conclude that the test we used measures to a certain extend the psychological costs of lying a subjects suffers from.

Do Agents Consider the Feedback Information?

In the theory section we showed that feedback can be an informative signal on the agents ability if there are sufficiently high costs of lying. The experimental data provide evidence that these psychological costs might exist in laboratory settings as well. However, whether and how signaling works out strongly depends on the agents' beliefs. Those beliefs in turn have an impact on the agents' reaction, i.e. they influence what the agents consider to be their best response in terms of their second period effort. Hence, we tested the impact of the type of feedback observed by the agents on their second period effort decision. To do so, we looked at the baseline condition effort data and compare the means of second period effort conditional on the previous feedback. According to Table 5.3, there was a mean effort of 2.98 for no feedback, 2.43 for negative feedback, and 5.04 for positive feedback for the baseline condition. As each agent chose his second period effort 15 times, the observations are dependent and therefore, we again use the Wilcoxon matched pairs sign rank test. The difference between second period efforts after observing no feedback versus negative feedback is weakly significant (absolute z-value 1.826). Thus, we might draw the conclusion that agents on average believe that low feedback is a stronger indicator for low ability than no feedback. The effort levels selected by the agents after receiving no feedback and positive feedback messages respectively, demonstrates a highly significant difference (absolute z-value 4.348). Thus, positive feedback increases second period efforts considerably compared to no feedback. We find a stronger result between second period efforts while comparing negative and positive feedback. Hence, the Wilcoxon matched pairs sign rank test again shows a highly significant result (absolute z-value 4.212).[16]

Altogether the results support the conclusion that agents use the feedback information when adjusting their second period efforts and the change corresponds to the message they received from their principals. In particular, it indicates that some agents and some principals (as shown in the previous chapter) play a separating equilibrium. Otherwise the agents' effort adjustment would not differ between positive and negative feedback. Thus, agents seem to believe that they receive different feedback information for different actual abilities and in turn, they react differently to positive, negative and no feedback respectively.

[16]The test results for all conditions are presented in Table 5.8 in the appendix.

5.5.3 Does Deception Increase with Higher Returns?

Although we showed in the last section that costs of lying appear in both the baseline and the low return condition there may be differences in the strength of the cost effect. It seems especially intuitive that deception rises as it becomes more beneficial. Gneezy (2005) and Corollary 10 imply that the probability of observing deception is higher when the returns on lying increase. To test whether the theoretical result is valid, we compare the occurrence of lies in the baseline and low return condition. As the principals' return on the agents' effort is higher in the baseline condition principals should have stronger incentives to make their agents increase their effort. Hence, if the principals expect lying (i.e. giving positive feedback even if the actual ability is low) to have a positive impact on the agents' effort decision, they would be more likely to lie in the baseline than in the low return condition. To test this we used the fraction of dishonest feedback that each principal sent to his agents compared to all feedback decisions.

Table 5.1 shows that, when the true ability was high, the fractions of lies were very low for both conditions. This is not surprising as deception is not very attractive in that case (mean .02 and .01, see Table 5.1, second row). Running a pairwise comparison with the Mann-Whitney-U-test for those rounds with high ability the data do not highlight any significant differences. When the actual ability was low, principals in the baseline condition lied less (mean .46, see Table 5.2, third row, first column) than those in the low return condition (mean .65, see Table 5.2, third row, second column), a result that contrasts the theoretical prediction in Corollary 10. Here, we find a weakly significant difference (absolute z-value 1.900) between the baseline and the low return condition applying the Mann-Whitney-U-test. Consequently, we conclude that there is at least some weak evidence in the data that increasing returns on lying diminish false feedback, and this contradicts our Hypothesis 4. Hence, the separating equilibrium is less likely to occur in the low return condition than in the baseline condition as described in Corollary 10.

Our model and the results in Gneezy (2005) would lead us to predict higher lying rates for the baseline condition than for the low return condition because higher returns feature the pooling equilibrium in which all types report high ability. As we observe the opposite pattern we conclude that cutting returns from lying does not necessarily enhance honesty.[17]

The question of why this might be the case is important. There are several possible explanations of which we will discuss two. The first refers to social preferences, in particular inequity aversion (Bolton and Ockenfels (2000), Fehr and Schmidt (1999)). According to the inequity aversion theory, inequity in payoffs leads to disutility. As discussed in Chapter 4 this approach has been shown to explain several deviations from purely rational behavioral patterns (e.g. Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Engelmann and Strobel (2004)). In our design principals and agents were treated and paid off inequitably. On average, the principals earned more than the agents and both parties could compute this as they knew their counterparts' complete payoff function. As the marginal returns on the agents' effort were higher for the principals in the baseline condition there was more inequality in payoffs in the baseline than in the low return condition. Deception, however, decreased the agents' payoff in favor of the principals'. Therefore principals might have refrained from lying to their agents because they realized that they were in a better position anyway and thus they tried to reduce inequality in the baseline condition.

Another possible explanation might be that the principals have reference dependent preferences. Recall that reference dependency suggests that individuals do not only appreciate the absolute height of the payoffs but also its height compared to a reference point. Payoffs smaller (greater) than the reference point are perceived as losses (gains). Loss aversion then captures the idea that marginal utility gains below the reference point are higher than above it. That is, utility gains above the reference point lose relative value. Due to the higher returns on effort, the principals earn more in the baseline condition than in the low return condition for a given effort level of the agent. If our subjects were loss averse and the reference point was on average ex-

[17] Besides, giving no feedback seems to be used to avoid lying if the true ability is low. It appears that principals prefer giving no feedback to deceiving and saying the truth. Yet, most of the agents adequately react to this by decreasing their effort.

	Low ability	High ability
No feedback	.175	.065
Negative feedback	.646	.014 (lie)
Positive feedback	.180 (lie)	.921

Table 5.5: Feedback decisions in the partners condition

ceeded in the baseline condition, they might have refrained from lying as the costs of lying were too high compared to the lower utility gains they could have realized. In turn, as they were on average below the reference point, principals in the low return condition could realize higher marginal returns and therefore it was worth to bear the costs of deception.

5.5.4 Long Term Relationships

Unlike in our experiment, however, employer-employee relationships are typically long term relations. Therefore the opportunity to build up reputation might play an important role in real life situations (e.g. Sobel (1985), Wilson and Sell (1997) or Hermalin (1998) study reputation in cheap talk games). So we introduced a partners condition where we did not newly match subjects for all 15 rounds and as a consequence allowed principal agent pairs to build up a reputation over time. As in the baseline condition, we chose $\theta - \beta = 2$ in the partners condition. Hence, we compare these two conditions to figure out the effect of long term relationships.

For the partners condition, it seems plausible that the principals act more honestly compared to those in the baseline condition as they can build up reputation to be reliable partners over rounds (see Sobel (1985) for a theory of credibility). The results are displayed in Table 5.5. For the baseline condition, we find that when the true ability was low most principals gave positive (i.e. deceptive) feedback. In contrast, in the partners condition, 65% of all feedback is negative for low ability which means that most subjects gave honest feedback.

Due to reputational concerns we would expect principals to suffer more from lying in the partners condition. However, in both conditions, baseline

	Baseline	Partners
1st period effort	3.96 (1.64)	3.69 (1.54)
2nd period effort/no feedback	2.89 (1.30)	2.53 (1.16)
2nd period effort/negative feedback	2.43 (1.08)	2.54 (.97)
2nd period effort/positive feedback	5.04 (1.79)	5.74 (1.83)

Table 5.6: Agents' effort decisions (standard deviation in parentheses)

and partners, principals refrained from giving positive feedback highly significantly more often if the true ability was low (partners condition: absolute z-value 4.510).

5.5.5 More Honesty and Trust in Long Term Relationships?

Intuitively one might argue that lies destroy trust and this is likely to affect the results in future rounds. In this section, we investigate how deception influences the agents' trust in the principals' feedback messages. From the experimental literature (e.g. Camerer and Weigelt (1988), Bohnet and Huck (2004), Gächter and Falk (2002) or Bolton and Ockenfels (2004), for an overview see Andreoni and Croson (2004)) we know that reputation is an important and powerful mechanism in enhancing cooperation and trust. Hence, we examine whether the fraction of dishonest feedback is lower when reputation can be built up. For this purpose we compare the principals' feedback behavior in the baseline and the partners condition. For the aggregate fraction of lies we find that subjects in the baseline condition lie more (.46) than those in the partners condition (.18) if the actual ability is low. Running a Mann-Whitney-U-test we find that the difference is highly significant (absolute z-value 2.904). This implies that from an aggregate point of view long term relationships as applied in the partners condition indeed encourage subjects to act more honestly. When compared, the difference between the dishonesty of feedback in the baseline and the partners condition for high ability is very small and insignificant (baseline .022 vs. partners condition .014). Whether this increase in honesty facilitated by long term partnership is useful in that the principals substantially benefit, depends on the agents' trust in the feedback information. Intuitively, one might argue that more honesty should induce agents to trust more. The descriptive statistics are presented in Table 5.6. Looking at the agents' effort reactions we find that agents react rather reluctant if given no feedback. In the baseline condition they choose an average effort level of 2.89 and 2.53 in the partners condition. This difference between conditions is small and insignificant but supports the conclusion that the agents interpreted no feedback as a signal of low ability. For negative feedback, the agents selected an average effort level of 2.43 in the baseline and 2.54 in the partners condition. Again, there are no significant differences between the conditions. However, agents in the baseline condition responded with an average effort level of 5.04 to positive feedback while subjects in the partners condition even exerted an average of 5.74. This difference indicates that agents in the partners condition trust more than those in the baseline and this is significant (absolute z-value 2.568) applying the Mann-Whitney-U-test. Hence, it seems that in the baseline condition the increased number of deceptions leads to more distrust in positive feedback resulting in lower average second period effort levels. This therefore seems to be the obverse of deception. Lies enable the principals to realize higher profits in rounds with low ability but the agents learn about that and positive feedback becomes less credible even if it is true.

In order to disentangle these effects, we use the data of the baseline and partners condition to explain the agents' second period effort by several independent variables in a random effects regression. The estimation results are presented in Table 5.7. 'Baseline' is a dummy variable with value one for the baseline condition and zero for the partners condition. 'Positive feedback' is a dummy variable with value one if the respective agent was given positive feedback in the respective round and zero otherwise. In our opinion positive feedback seems to be the most attractive and therefore the least credible action the principal can take. Analogously 'no feedback' is a dummy variable with value one if the agent was given no feedback in the respective round and zero otherwise. When the agent was given negative feedback in a round both feedback dummies (positive and no feedback) would be zero. To check for past experiences, we included a variable 'fraction of past lies' which is the proportion of lies in previous rounds that the agent experienced up to the current round. Hence, it is the running sum of lies divided by the number of previous rounds, the agent experienced in the past.

Model (1) shows that the condition does not influence the effort decision in the second periods per se. We may conclude that the agents indeed react to the positive feedback information observing that the variable 'positive feedback' leads to a highly significant increase in second period effort which is not surprising. The effect is similar but less pronounced for 'no feedback'. However, agents also seem to be sensitive toward past experiences. The coefficient for 'fraction of past lies' is negative and highly significant demonstrating that agents choose lower levels of second period effort if they have been lied to more often in the past. Obviously, agents refrain from higher effort levels fearing that the feedback information might be untrue in the current round again. Still, past experiences should only have an impact on the second period effort if positive feedback is given. Giving no feedback does not contain any information at all where lying by giving negative feedback (in case the actual ability is high) is not very attractive and therefore is rarely used. So in the latter cases of feedback one cannot draw any conclusions from previous lies. To check whether this is the case, we include interaction terms in model (2) controlling for the effect of 'fraction of past lies' for positive feedback and no feedback. The coefficient for the interaction of positive feedback and the fraction of lies in previous rounds is negative and highly significant. This indicates that indeed past lies destroy the credibility of positive feedback. This is not the case for no feedback where the coefficient for the interaction term is small and insignificant. Note also that the direct effect of 'fraction of past lies' disappears in this regression model. However, some variables might have fairly different implications for the particular agent depending on the condition he is in. To control for the effect that (some of) the independent variables might have different impacts in the two conditions, we introduce several interaction terms.

In models 3 and 4 we show specifications including these interaction terms in addition. The most important difference between the two conditions seems

	(1)	(2)	(3)	(4)
		Effort period 2		
Baseline	-0.0953	-0.0217	-0.217	-0.156
	(0.25)	(0.25)	(0.33)	(0.35)
Positive feedback	2.957***	3.403***	3.487***	3.513***
	(0.15)	(0.16)	(0.17)	(0.18)
No feedback	0.499***	0.189	-0.0115	0.0269
	(0.16)	(0.19)	(0.23)	(0.25)
Fraction of past lies	-2.454***	-0.663	-1.919***	-1.620*
	(0.34)	(0.52)	(0.71)	(0.89)
Positive feedback*fraction of past lies		-3.513***	-3.262***	-3.621***
		(0.57)	(0.62)	(0.94)
No feedback*fraction of past lies		0.895	0.691	0.210
		(0.76)	(0.79)	(1.24)
Positive feedback*baseline			-0.165	-0.246
			(0.25)	(0.30)
No feedback*baseline			0.416	0.308
			(0.34)	(0.40)
Fraction of past lies*baseline			1.800***	1.301
			(0.65)	(1.10)
Positive feedback*baseline*				0.626
fraction of past lies				(1.25)
No feedback*baseline*				0.811
fraction of past lies				(1.61)
Constant	2.906***	2.882***	2.983***	2.962***
	(0.25)	(0.25)	(0.26)	(0.26)
Observations	840	840	840	840
Number of subjects	56	56	56	56

Standard errors in parentheses, round dummies included
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5.7: Results of random effects estimation

to be the interaction of 'fraction of past lies' and 'baseline' where the coefficient is positive and highly significant. This outcome is intuitively clear since 'fraction of past lies' is more meaningful if based on one single principal rather than on many ever changing ones. In the partners condition, the agent can conclude from a high number of previously experienced lies that his principal is not a very honest partner since he remains the same person all over the game. Still, the conclusion should be less strong in the baseline condition as experiences with past principals do not necessarily lead to good predictions of the behavior of future principals. Although it is useful to update the expectation of the behavior of the whole 'principal population', it is less meaningful with regard to the behavior of a particular individual whose behavior has not been observed beforehand.

To conclude, we have seen that a lie has two major effects. First, in the short run a lie leads to higher second period efforts in the round in which it was told. Second, it diminishes future impacts to positive feedback in the long run. In case of the partners matching this does harm the liar's own future utility. In the baseline condition future principals suffer from a negative long term effect if matched to an agent who has been lied to in the past similar to an external effect of deception. These results generate another important issue: in this context, is lying beneficial for the principal after all if we consider both effects together, and what is the effect of deception on social welfare?

5.5.6 Is More Honesty Profitable and if so, for Whom?

Let us first consider the agents' payoffs displayed in Figure 5.2. The left part of the figure shows that the agents' mean payoff in the first periods is quite similar across the baseline and the partners conditions. Using the Mann-Whitney-U-test we do not find any significant differences. However, the left part of the figure also illustrates the agents' mean payoff in the second periods after the principals have made their feedback decisions. Obviously agents in the partners condition (which has been shown to enhance honesty) yield on average higher payoffs than those in the baseline and this is highly significant with the Mann-Whitney-U-test (absolute z-value 2.715). If we run the same analysis for total payoffs (sum of payoff in first and second periods) we find a similar but less pronounced result (absolute z-value 1.656). These outcomes support the hypothesis that the agents in total benefit more from honesty in the relationship as true feedback allows them to make more optimal effort decisions in the second periods. With regard to the principals' payoffs, the right

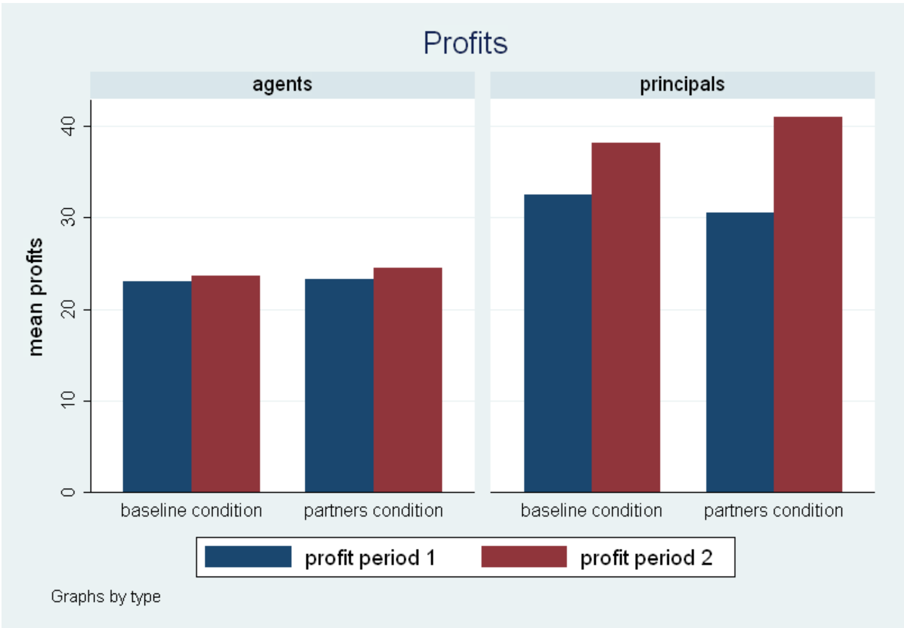


Figure 5.2: Mean profits for principals and agents

part of Figure 5.2 shows the mean payoffs in the first and second periods aggregated over all rounds. Again, the first period payoffs hardly differ between the two conditions and this is insignificant at any conventional level. Yet, when considering the second periods, we observe weakly significantly higher payoffs in the partners condition (absolute z-value 1.657). This implies that even the principals benefit from the lower level of deception occurring in the partners condition. However, in terms of total payoffs the difference is no longer significant. Nevertheless, we find weak evidence for a positive effect of reputation and resulting honesty on the principals' aggregate payoffs in the second periods.

The outcomes described above are interesting as deception in the short run seems to be more attractive. But, as we have shown in the previous chapter, deception causes the credibility of feedback to decrease. This in turn induces agents to reject positive feedback even if true. Therefore the positive effect of true positive feedback (when the actual ability is high) diminishes. Hence, more honesty indirectly leads to an increase in payoffs for the principals. Finally, we may conclude that honesty seems to raise social welfare as it not only features a strong positive impact on the agents' second period payoff but also a weak positive impact on the principals' second period payoff. Both parties seem to benefit from reputational systems enhancing honesty.

5.6 Summary and Discussion

In this chapter we have theoretically and experimentally examined the role of the principal's feedback honesty in a framework in which the agent's actual ability is uncertain. As the agent's monetary payoff is output-based his optimal effort level increases with his ability. The model posits that the principal (who is the only one to know the true ability) might prefer to lie to the agent about his ability. In particular she might tell the agent that he has high ability even when, in reality, the agent has low ability because there are sufficiently low costs associated with lying. In that case we obtain a pooling equilibrium and the true ability is not revealed. In turn, for higher costs of lying separating equilibria become feasible.

The experimental outcomes show that some principals seem to play separating strategies which indicates that costs of lying indeed exist. Further, we are able to distinguish the steadfastly honest from the steadfastly deceiving as well as mark the principals who vacillate. Hence, we conclude that for some people deception is associated with costs, even if the costs are not monetary but instead are driven by a reluctance to break religious or ethical rules or the threat of a bad conscience. Although deception is widespread, the agents adjust their effort according to the received feedback information and thus increase the principals' payoffs in the respective round. Still, past experiences of deception undermine the agents' trust in the feedback they receive. That is, the impact of the feedback information on the agents' effort decreases with the number of lies the latter has previously experienced. Hence, we might say that deception has serious and measurable external effects.

Surprisingly, decreased incentives to lie do not induce principals to act more honestly. Instead, we observe the opposite pattern as deception rates even rise. This contradicts not only our theoretical prediction but also observations in Gneezy (2005). As lower incentives to lie come along with lower payoffs of the principals, the more aggressive lying behavior might be explained by alternative approaches such as inequity aversion or reference dependency and loss aversion.

Additionally, we find strong evidence that long term relationships enhance honesty due to reputational concerns. In long term partnerships the principals make use of deception less often. Furthermore the agents react much more sensitively toward past experiences of deception if they interact with the same principal continuously. In this case the principals harm themselves by lying to their agent. This shows that indeed lies have a negative long term effect in continuous partnerships leading to more honesty.

Finally, both, our theoretical results and our experimental data show that being honest is worthwhile in the long run. Then, both parties benefit if the true ability is revealed. So, if the principal was able to commit himself to being honest to the agent in the future this would lead to an increase in welfare.

Even if we neglect any social or ethical aspects of honesty and deception, employers and employees are ex-ante better off with truthful performance feedback because it yields higher economic payoffs. Thus, we can indeed conclude that honesty is the best policy.

5.7 Appendix to Chapter 5

Proof. Proposition 7: First we will refer to the separating equilibrium in which the principal is always honest, so we can describe her strategy as (f_L, f_H). The expression to the left (right) of the comma refers to the principal's choice of action if the true ability is low (high) in the following. Hence, if the agent receives negative feedback he will know that his true ability is low. Therefore he will choose $e^* = \beta a_L \equiv e_L$. Analogously, if the agent observes positive feedback f_H he will be sure that his actual ability is high and exert an effort level according to $e^* = \beta a_H \equiv e_H$. When the agent observes no feedback f_0 off the equilibrium path his reaction function is defined by $e^*(\widetilde{\pi}) = \beta \left[\widetilde{\pi} a_H + (1 - \widetilde{\pi})\, a_L\right]$ and therefore a function of his belief $\widetilde{\pi}$, that the true ability is high. To ensure that the principal indeed plays (f_L, f_H) without incentive to deviate in equilibrium, his strategy must be optimal. Thus, in case of $a = a_L$,

$$\begin{aligned} (\theta - \beta) a_L e_L &\geq (\theta - \beta) a_L e_H - l \\ l &\geq a_L \left[e_H - e_L\right] (\theta - \beta) \end{aligned} \tag{5.6}$$

must be valid. This is the case, if costs of lying l are sufficiently high. In the case of $a = a_H$ the corresponding condition is:

$$\begin{aligned} (\theta - \beta) a_H e_H &\geq (\theta - \beta) a_H e_L - l \\ l &\geq -(\theta - \beta) a_H \left(e_H - e_L\right) \end{aligned} \tag{5.7}$$

Condition (5.7) is always met as the right hand side is strictly negative, but the costs of lying are at least zero. Yet, the principal can also deviate off the equilibrium path by choosing f_0. If the true ability is low $a = a_L$, she faces the following condition:

$$\begin{aligned} (\theta - \beta) a_L e_L &\geq (\theta - \beta) a_L \left[\widetilde{\pi} e_H + (1 - \widetilde{\pi})\, e_L\right] \\ \widetilde{\pi} &\leq 0 \end{aligned} \tag{5.8}$$

Condition (5.8) can only be met when the agent's belief that the true ability is high $\widetilde{\pi}$ equals zero. In the case where the true ability is high we obtain the condition:

$$\begin{aligned}(\theta-\beta)a_H e_H &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ 1 &\geq \widetilde{\pi}\end{aligned} \tag{5.9}$$

It is obvious that this condition is always met. Therefore the respective separating equilibrium does exist if conditions (5.6) and (5.8) hold. However, when there are no costs of lying (that is $l=0$) the inequality (5.6) is never fulfilled and this equilibrium could not exist. ■

Proof. Proposition 8: In the pooling equilibrium described in Proposition 8 the principal always chooses to give positive feedback f_H regardless of the actual ability. Hence, the agent chooses $e^*=\overline{e}$ in equilibrium. For $a=a_L$ the principal's utility is $(\theta-\beta)a_L\overline{e}-l$. When she deviates and plays f_L or f_0 she receives $(\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]$. Thus she faces the condition:

$$\begin{aligned}(\theta-\beta)a_L\overline{e}-l &\geq (\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ \left[e_H-e_L\right](\theta-\beta)\,(\tau-\widetilde{\pi})\,a_L &\geq l\end{aligned} \tag{5.10}$$

For $a=a_H$ the principal yields $(\theta-\beta)a_H\overline{e}$ in equilibrium and $(\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]$ if she deviates to f_0:

$$\begin{aligned}(\theta-\beta)a_H\overline{e} &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ \tau &\geq \widetilde{\pi}\end{aligned} \tag{5.11}$$

But, she can as well choose f_L off the equilibrium path. Then, the corresponding condition is:

$$\begin{aligned}(\theta-\beta)a_H\overline{e} &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]-l\\ \left[e_H-e_L\right]\left[\tau-\widetilde{\pi}\right](\theta-\beta)a_H &\geq -l\end{aligned} \tag{5.12}$$

So in equilibrium conditions (5.10) to (5.12) must hold. ■

Proof. Proposition 9: To show that the separating equilibria are efficient, we compare the sums of the ex ante expected utilities of agent and principal in a separating equilibrium with the most efficient pooling equilibrium. The latter is the pooling equilibrium in which no feedback is given as no signaling costs arise. It does not matter which separating equilibrium is selected for comparison as they all result in the same utilities for agent and principal, respectively. Moreover, we only consider the second period because the first periods outcomes are the same for all types of equilibria. The sum of both parties' ex ante expected utility is

$$\begin{aligned}
& E\left[U_2^A\right] + E\left[W_2^P\right] \\
= \; & \tau\left[\beta^2 a_H \overline{a}\right] + (1-\tau)\;\left[\beta^2 a_L \overline{a}\right] - \frac{1}{2}\beta^2\overline{a}^2 \\
& +\tau\left[\left(\theta-\beta\right) a_H \beta\overline{a}\right] + (1-\tau)\left[\left(\theta-\beta\right) a_L \beta\overline{a}\right] \\
= \; & \beta^2\overline{a}\underbrace{\left(\tau a_H + (1-\tau)\;a_L\right)}_{\overline{a}} - \frac{1}{2}\beta^2\overline{a}^2 + \left(\theta-\beta\right)\overline{a}\beta\left(\tau a_H + (1-\tau)\,a_L\right) \\
= \; & \frac{1}{2}\beta^2\overline{a}^2 + \left(\theta-\beta\right)\beta\overline{a}^2
\end{aligned} \tag{5.13}$$

in the most efficient pooling equilibrium. In a separating equilibrium it is:

$$\begin{aligned}
& E\left[U_2^A\right] + E\left[W_2^P\right] \\
= \; & \tau\left[\left(\beta a_H\right)^2 - \frac{1}{2}\beta^2 a_H^2\right] + (1-\tau)\;\left[\left(\beta a_L\right)^2 - \frac{1}{2}\beta^2 a_L^2\right] \\
& + \left(\theta-\beta\right)\beta\left(\tau a_H^2 + (1-\tau)\,a_L^2\right) \\
= \; & \frac{\beta^2}{2}\left(\tau a_H^2 + (1-\tau)\,a_L^2\right) + \left(\theta-\beta\right)\beta\left(\tau a_H^2 + (1-\tau)\,a_L^2\right)
\end{aligned} \tag{5.14}$$

Social welfare is ex ante higher in the separating equilibrium if the following inequation indicating that (5.14) is greater than (5.13) is valid:

$$\begin{aligned} & \frac{\beta^2}{2}\left(\tau a_H^2+(1-\tau)\, a_L^2\right)+(\theta-\beta)\,\beta\left(\tau a_H^2+(1-\tau)\, a_L^2\right) \\ > & \frac{1}{2}\beta^2\overline{a}^2+(\theta-\beta)\,\beta\overline{a}^2 \end{aligned} \tag{5.15}$$

$\Leftrightarrow$

$$\frac{\beta^2}{2}\left(\tau a_H^2+(1-\tau)\, a_L^2-\overline{a}^2\right)+(\theta-\beta)\,\beta\left(\tau a_H^2+(1-\tau)\, a_L^2-\overline{a}^2\right)>0$$

For (5.15) to be greater than zero, the terms in brackets have to be greater than zero because by assumption $(\theta-\beta)>0$ and $\beta>0$:

$$\begin{aligned} \tau a_H^2+(1-\tau)\, a_L^2-\overline{a}^2 &> 0 \\ \tau a_H^2+(1-\tau)\, a_L^2 &> \overline{a}^2 \\ \tau a_H^2+(1-\tau)\, a_L^2 &> [\tau a_H+(1-\tau)a_L]^2 \end{aligned}$$

According to Jensen's inequality, this is true. ■

In addition we will analyze the remaining three equilibria existing in the game in the following:

1. We discuss the candidate (f_0, f_H). Hence, the principal will not deviate in the equilibrium for $a=a_L$ if condition (5.6) is met. In the case of $a=a_H$ the corresponding condition is:

$$\begin{aligned} (\theta-\beta)a_He_H &\geq (\theta-\beta)a_He_L \qquad (5.16) \\ e_H &\geq e_L \end{aligned}$$

   Condition (5.16) is always met. Yet, the principal can deviate from the equilibrium path by choosing f_L. Then for $a=a_L$:

$$\begin{aligned} (\theta-\beta)a_Le_L &\geq (\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\, e_L\right] \\ \widetilde{\pi} &\leq 0 \end{aligned} \tag{5.17}$$

must hold. However, condition (5.17) can only be met when the agent's belief $\widetilde{\pi}$ that the true ability is high equals zero. Considering the case where the true ability is high we obtain the condition:

$$\begin{aligned}(\theta-\beta)a_H e_H &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]-l\\ (\theta-\beta)a_H\,(1-\widetilde{\pi})\left[e_H-e_L\right] &\geq -l \qquad (5.18)\end{aligned}$$

It is obvious that this condition is always met if $l>0$. Therefore the respective separating equilibrium exists if conditions (5.6) and (5.17) are met.

2. Suppose a pooling equilibrium in which all types choose to give no feedback f_0. For $a=a_L$ the principal's utility is $(\theta-\beta)a_L\overline{e}$. When she deviates to f_L, she receives $(\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]$. Thus, she faces the condition:

$$\begin{aligned}(\theta-\beta)a_L\overline{e} &\geq (\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ \tau &\geq \widetilde{\pi} \qquad (5.19)\end{aligned}$$

But she can as well play f_H off the equilibrium path. Then the corresponding condition is:

$$\begin{aligned}(\theta-\beta)a_L\overline{e} &\geq (\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]-l\\ l &\geq -(\tau-\widetilde{\pi})\left[e_H-e_L\right](\theta-\beta)a_L \qquad (5.20)\end{aligned}$$

For $a=a_H$ the principal yields $(\theta-\beta)a_H\overline{e}$ in equilibrium and $(\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]-l$ if she deviates to f_L:

$$\begin{aligned}(\theta-\beta)a_H\overline{e} &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]-l\\ l &\geq -(\tau-\widetilde{\pi})\left[e_H-e_L\right](\theta-\beta)a_H \qquad (5.21)\end{aligned}$$

If the principals deviates to f_H

$$\begin{aligned}(\theta-\beta)a_H\overline{e} &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ \tau &\geq \widetilde{\pi}\end{aligned} \tag{5.22}$$

must be met. Hence, in equilibrium conditions (5.19) to (5.22) must hold.

3. Consider a pooling equilibrium in which f_L is always chosen. If $a=a_L$ this is optimal compared to f_0 when

$$\begin{aligned}(\theta-\beta)a_L\overline{e} &\geq (\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ \tau &\geq \widetilde{\pi}\end{aligned} \tag{5.23}$$

holds. But, the principal can as well give positive feedback off the equilibrium path. Then, the corresponding condition is:

$$\begin{aligned}(\theta-\beta)a_L\overline{e} &\geq (\theta-\beta)a_L\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]-l\\ l &\geq -(\tau-\widetilde{\pi})\left[e_H-e_L\right](\theta-\beta)a_L\end{aligned} \tag{5.24}$$

For $a=a_H$ the principal yields $(\theta-\beta)a_H\overline{e}-l$ in equilibrium and $(\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]$ if she deviates to f_0 or f_H:

$$\begin{aligned}(\theta-\beta)a_H\overline{e}-l &\geq (\theta-\beta)a_H\left[\widetilde{\pi}e_H+(1-\widetilde{\pi})\,e_L\right]\\ \left[e_H-e_L\right](\theta-\beta)(\tau-\widetilde{\pi})\,a_H &\geq l\end{aligned} \tag{5.25}$$

must be met. Hence, in equilibrium (5.23) to (5.25) must hold.

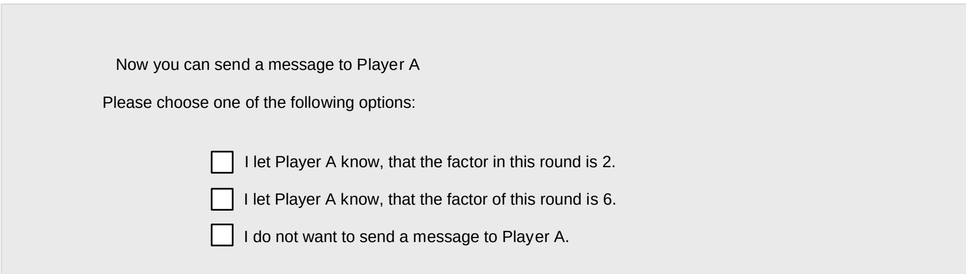
Short Instructions (baseline condition)

Before each round starts, one Player A is matched to one Player B. Each round consists of two periods. A factor is randomly drawn before every round starts. The factor is unknown to Player A as well as to Player B, but can only take the value 2 or the value 6. In 50 percent of the cases the factor is 2 and in 50 percent of the cases it is 6. Then the first round starts:

Player A is asked to select a number between 1 and 8. For each number he has to bear the specific cost of the number which are directly subtracted from his account.

Effort	1	2	3	4	5	6	7	8
Costs of Effort	0.5	2.0	4.5	8.0	12.5	18.0	24.5	32.0

After that, Period 1 is finished and Player B gets to know which factor was drawn beforehand. Player B is then asked to send a message to Player A. In this message Player B can tell that the factor is 2 or 6. But she can also decide not to send any message. Look at the screenshot:



The selected message will be send to Player A and Period 2 begins.

In Period 2, Player A is again asked to choose a number between 1 and 8. It is important to note that the random factor from Period 1 remains the same in Period 2.

Random draw of the factor

1. Period: Player A chooses a number between 1 and 8
Player B gets to know the actual value of the factor (which is either 2 or 6) Player B selects a message and sends it to Player A
2. Period: Player A receives the message and chooses between 1 and 8

The players will be told the number of tokens from the two periods

For each period, both types of players receive their tokens depending on the number that was chosen by Player A.

For Player A the tokens per period are calculated as follows:
Tokens for Player A: 16 + Factor*Number – Cost of the number

For Player B the tokens per period are calculated as follows:
Tokens for Player B: 2*Factor*Number

In total, there are 15 rounds consisting of the 2 periods described above. Before every new round starts, the factor 2 or 6 is again randomly drawn and you will be matched to a new player to which you have not been assigned to before. One of the rounds will be drawn by lot for your compensation of the experiment. Both periods of this round will be paid off. One token is equals 0.167 Euro.

	No. vs. pos.	No. vs. neg.	Pos. vs. neg.
Baseline condition	4.348***	1.826***	4.212***
Low return condition	1.458	2.351**	4.288***
Partners condition	3.922***	0.848	4.542***

Absolute z-values, *** p < 0.01, ** p < 0.05, * p < 0.1

Table 5.8: Results Wilcoxon sign rank test

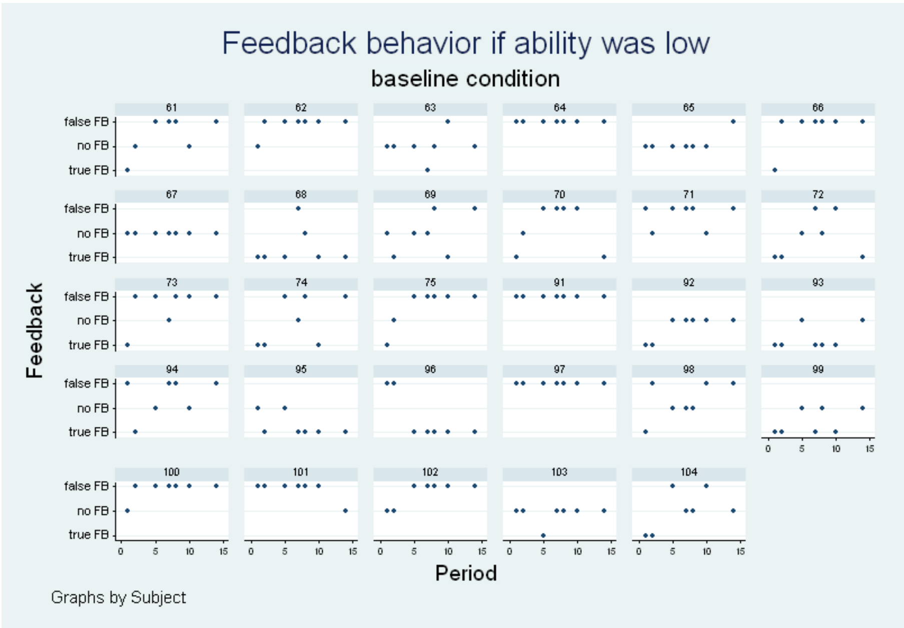


Figure 5.3: Feedback decisions baseline condition (for low ability)

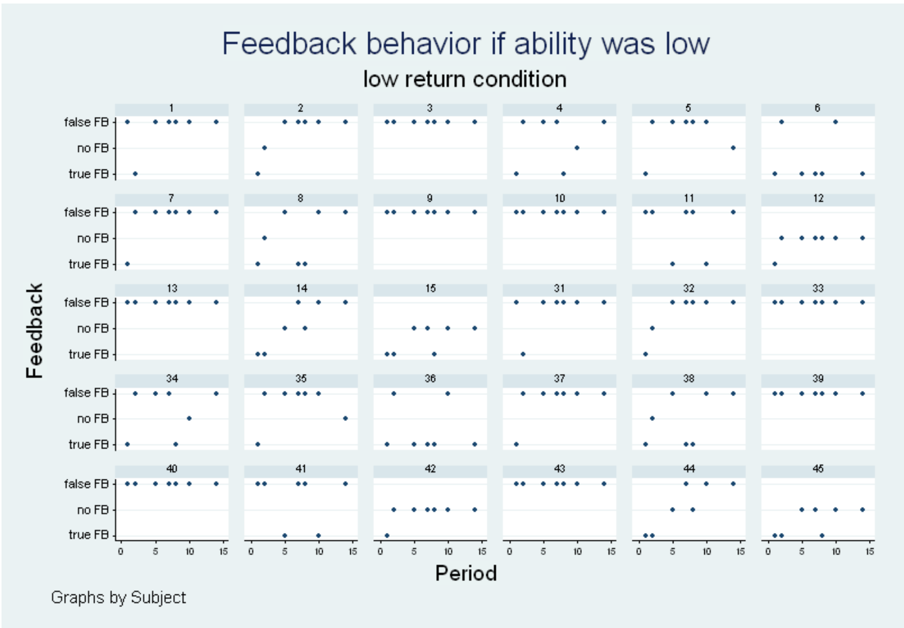


Figure 5.4: Feedback decisions low return condition (for low ability)

Bibliography

Akerlof, G. (1982): Labor Contracts as Partial Gift Exchange. Quarterly Journal of Economics, 97, pp. 543–69.

Alchian, A. and Demsetz, H. (1972): Production, Information Costs, and Economic Organization. American Economic Review, 62, pp. 777–795.

Andreoni, J. and Croson, R. (2004): Partners versus Strangers - Random Rematching in Public Goods Experiments. In: Plott, C. and Smith, V. (Ed.) Handbook of Experimental Economic Results.

Ariely, D., Gneezy, U., Loewenstein, G. and Mazar, N. (2005): Large Stakes and Big Mistakes. FRB of Boston Working Paper No. 05-11.

Backes-Gellner, U., Mohnen, A. and Werner, A. (2006): Team Size and Effort in Start-Up-Teams - Another Consequence of Free-Riding and Peer Pressure in Partnerships. Working Paper, University of Cologne and University of Zurich.

Barron, J. M. and Gjerde, K. P. (1997): Peer Pressure in an Agency Relationship. Journal of Labour Economics, 15, pp. 234–254.

Berg, J., Dickhaut, J. and McCabe, K. (1995): Trust, Reciprocity and Social History. Games and Economic Behavior, 10, pp. 122–142.

Bewley, T. (1999): Why Wages Don't Fall During a Recession. Harvard University Press.

Biel, P. R. (2004): Inequity Aversion and Team Incentives. Working Paper University College London.

Bénabou, R. and Tirole, J. (2003): Intrinsic and Extrinsic Motivation. Review of Economic Studies, 70, pp. 489–520.

Bohnet, I. and Huck, S. (2004): Repetition and Reputation - Implications for Trust and Trustworthiness When Institutions Change. AER Papers and Proceedings, 94, pp. 362–366.

Bolton, G. and Ockenfels, A. (2000): ERC - A Theory of Equity, Reciprocity and Competition. American Economic Review, 90, pp. 166–193.

Bolton, Katok E., G. and Ockenfels, A. (2004): How Effective Are Electronic Repuation Mechanisms? An Experimental Investigation. Management Science, 50, pp. 1587–1602.

Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R. (1997): Labor Supply of New York City Cabdrivers: One Day at a Time. Quarterly Journal of Economics, 112, pp. 407–441.

Camerer, C. and Weigelt, K. (1988): Experimental Tests of a Sequential Equilibrium Reputation Model. Econometrica, 56, pp. 1–36.

Charness, G. (2000): Self-Serving Cheap Talk - A Test of Aumann's conjecture. Games and Economic Behavior, 33, pp. 177–194.

Charness, G. and Rabin, M. (2002): Understanding Social Preferences with Simple Tests. Quarterly Journal of Economics, 117, pp. 817–869.

Che, Y.-K. and Yoo, S.-W. (2001): Optimal Incentives for Teams. American Economic Review, 91, pp. 525–54.

Christie, R. and Geis, F. (1970): Studies in Machiavellianism. Academic Press.

Crawford, V. (1998): A Survey of Experiments on Communication via Cheap Talk. Journal of Economic Theory, 78, pp. 286–298.

DeCharms, R. (1968): Personal Causation: The Internal Affective Determinants of Behavior. New York Academic Press.

Deci, E. L. (1971): Effects of Externally Mediated Rewards on Intrinsic Motivation. Journal of Personality and Social Psychology, 18, pp. 105–115.

Deci, E. L., Koestner, R. and Ryan, R. M. (1999a): A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. Psychological Bulletin, 125, pp. 627–668.

Deci, E. L., Koestner, R. and Ryan, R. M. (1999b): The Undermining Effect is Reality After All-Extrinsic Rewards, Task Interest, and Self-Determination: Reply to Eisenberger, Pierce, and Cameron (1999) Lepper, Henderlong, and Gingras (1999). Psychological Bulletin, 125, pp. 692–700.

Demougin, D. and Fluet, C. (2003): Inequality Aversion in Tournaments. Discussion Paper Humbold University Berlin.

Dickinson, D. L. (1999): An Experimental Examination of Labor Supply and Work Intensities. Journal of Labor Economics, 17, pp. 638–670.

Dufwenberg, M and Kirchsteiger, G. (2000): Reciprocity and Wage Undercutting. European Economic Review, 44, pp. 1069–1078.

Easterlin, R. A. (2001): Income and Happiness: Towards a Unified Theory. The Economic Journal, 111 (July), pp. 465–484.

Eisenberger, R. and Cameron, J. (1996): Detrimental Effects of Reward: Reality of Myth? American Psychologist, 51, pp. 1153–1166.

Engelmann, D. and Strobel, M. (2004): Inequality Aversion, Efficiency and Maximin Preferences in Simple Distribution Experiments. American Economic Review, 94, pp. 857–869.

Englmaier, F. (2004): A Survey on Moral Hazard, Contracts, and Social Preferences. In: Agarwal, B. and Vercelli, A. (Ed.) Psychology, Rationality and Economic Behavior: Challanging standard assumptions (forthcoming).

Englmaier, F. and Wambach, A. (2002): Contracts and Inequity Aversion. CESifo Working Paper 809, Universität München.

Falk, A. and Fischbacher, U. (1999): A Theory of Reciprocity. Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 6.

Falk, A. and Ichino, A. (2006): Clean Evidence on Peer Pressure. Journal of Labor Economics, 24, p. 39Ű57.

Fehr, E. and Gächter, S. (2002): Do Incentive Contracts Crowd Out Voluntary Cooperation?, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 34.

Fehr, E. and Götte, L. (2005): Do Workers Work More When Wages are High? IZA Discussion Paper No. 125.

Fehr, E., Kirchsteiger, G. and Riedl, A. (1993): Does Fairness Prevent Market Clearing? An Experimental Investigation. Quarterly Journal of Economics, 108, pp. 437–460.

Fehr, E. and Schmidt, K. M. (1999): A Theory of Fairness, Competition, and Cooperation. Quarterly Journal of Economics, 114, pp. 817–868.

Fischbacher, U. (1999): z-Tree - Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual. Working Paper No. 21, Institute for Empirical Research in Economics, University of Zurich., working Paper Nr. 21, Institute for Empirical Research in Economics, University of Zurich.

Frey, B. S. (1997): Markt und Motivation. Wie ökonomische Anreize Die (Arbeits-) Moral Verdrängen. Vahlen, München.

Frey, B. S. and Jegen, R. (2001): Motivation Crowding Theory: A Survey of Empirical Evidence. Journal of Economic Surveys, 15 (5), pp. 589–611.

Frey, B. S. and Oberholzer-Gee, F. (1997): The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. American Economic Review, 87, pp. 746–755.

Gächter, S. and Falk, A. (2002): Reputation and Reciprocity: Consequences for the Labour Relation. Scandinavian Journal of Economics, 104, pp. 1–26.

Geanakoplos, J., Pearce, D. and Stacchetti, E. (1989): Psychological Games and Sequential Rationality. Games and Economic Behavior, 1, pp. 60–79.

Gneezy, U. (2004): Do High Wages Lead to High Profits? An Experimental Study of Reciprocity Using Real Effort. Working Paper, University of Chicago, Graduate Business School.

Gneezy, U. (2005): Deception: The Role of Consequences. American Economic Review, 95, pp. 384–394.

Gneezy, U. and Rustichini, A. (2000a): Pay Enough or Don't Pay at All. Quarterly Journal of Economics, 115, pp. 791–810.

Gneezy, U. and Rustichini, A. (2000b): A Fine is a Price. Journal of Legal Studies, 29, pp. 1–18.

Greene, W. H. (2003): Econometric Analysis. Upper Saddle River, Prentice-Hall, 5th Edition.

Greiner, B. (2003): An Online Recruitment System for Economic Experiments. In: Kremer, K. and Macho, V. (Ed.) Forschung und wissenschaftliches Rechnen, GWDG Bericht 63, Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung, pp. 79–93.

Grossman, S.J. and Hart, O.D. (1983): An Analysis of the Principal-Agent-Problem. Econometrica, 51, pp. 7–45.

Grund, C. and Sliwka, D. (2005): Envy and Compassion in Tournaments. Journal of Economics and Management Strategy, 14, pp. 187–207.

Gunnthorsdottir, A., McCabe, K. and Smith, V. (2002): Using the Machiavellianism instrument to predict trustworthiness in a bargaining game. Journal of Economic Psychology, 23, pp. 49–66.

Hart, O. D. and Holmström, B. (1987): The Theory of Contracts. Bewley, T., Cambridge, Cambridge University Press.

Henning, H.-J. and Six, B. (1977): Konstruktion einer Machiavellismus Skala. Zeitschrift für Sozialpsychologie, 8, pp. 185–198.

Hermalin, B. E. (1998): The Boy Who Cried Wolf: Honesty in Repeated Games of Asymmetric Information. Working Paper, Cornell University.

Heyman, J. and Ariely, D. (2004): Effort for Payment: A Tale of Two Markets. Psychological Science, 15 (11), pp. 787–793.

Holmström, B. (1979): Moral Hazard and Observability. The Bell Journal of Economics, 10, pp. 74–91.

Holmström, B. (1982): Moral Hazard in Teams. Bell Journal of Economics, 13, pp. 324–340.

Huck, S. and Biel, P. R. (2006): Endogenous leadership in teams. Journal of Institutional and Theoretical Economics (forthcoming).

Irlenbusch, B. and Sliwka, D. (2003): Incentives, Decision Frames and Motivation Crowding Out - An Experimental Investigation. Mimeo, Universität Bonn.

Itoh, H. (2004): Moral Hazard and Other-Regarding Preferences. Japanese Economic Review (forthcoming).

Kahneman, D. and Tversky, A. (1979): Prospect theory: An analysis of decisions under risk. Econometrica, 47, pp. 313–327.

Kandel, E. and Lazear, E. P. (1992): Peer Pressure and Partnerships. Journal of Political Economy, 100, pp. 801–17.

Kluger, A. N. and DeNisi, A. (1996): The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. Psychological Bulletin, 119, pp. 254–284.

Knez, M. and Simester, D. (2001): Firm-Wide Incentives and Mutual Monitoring at Continental Airlines. Journal of Labor Economics, 19, pp. 743–772.

Kunz, A. and Pfaff, D. (2002): Agency Theory, Performance Evaluation, and the Hypothetical Construct of Intrinsic Motivation. Accounting, Economics and Society, 27, pp. 275–295.

Landsburg, S. E. (1995): The Armchair Economist: Economics And Everyday Experience. Free Press.

Latham, G. P. and Locke, E. A. (1990): A Theory of Goal Setting & Task Performance. Prentice-Hall: Engelwood Cliffs, New Jersey.

Latham, G. P. and Locke, E. A. (1991): Self-Regulation through Goal-Setting. Organizational Behavior and Human Decision Processes, 50, pp. 212–247.

Lazear, E. P. (2000): Performance Pay and Productivity. American Economic Review, 90 (5), pp. 1346–62.

MacCrimmon, K. R. and Larsson, S. (1979): Utility Theory: Axioms versus Paradoxes. Allais, M. and Hagen, O., Dordrecht/Boston, Reidel, pp. 333–409.

Masclet, D. (2002): Peer Pressure in Work Teams: The Effects of Inequity Aversion. Working Paper, University of Lyon.

Mohnen, A. and Pokorny, K. (2005): Eine ökonomische Analyse von Feedback - theoretische Implikationen und experimentelle Evidenz. Working Paper, University of Cologne.

Mohnen, A. and Pokorny, K. (2006a): Reference Dependency and Loss Aversion in Employer-Employee Relationships - A Real Effort Experiment on the Incentive Impact of the Fixed Wage. Working Paper, University of Cologne.

Mohnen, A. and Pokorny, K. (2006b): Is Honesty the Best Policy? An Experimental Study on the Honesty of Feedback in Employer-Employee Relationships. Working Paper, University of Cologne.

Mohnen, A., Pokorny, K. and Sliwka, D. (2006): Transparency, Inequity Aversion and the Dynamics of Peer Pressure in Teams - Theory and Evidence. Working Paper, University of Cologne.

Murphy, K. J. (1999): Executive Compensation. In: Ashenfelter, O. and Card, D. (Ed.) Handbook of Labor Economics, Volume 3, North-Holland, Amsterdam.

Newhouse, J.P. (1973): The Economics of Group Practice. Journal of Human Resources, 8, pp. 37–56.

Oettinger, G. S. (1999): An Empirical Analysis of the Daily Labor Supply of Stadium Vendors. Journal of Political Economy, 107 (2), pp. 360–392.

Paarsch, H. J. and Shearer, B. S. (1999): The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting Industry. Journal of Human Resources, 34, pp. 643–667.

Park, I.-U. (2005): Cheap Talk Referrals of Differentiated Experts in Repeated Relationships. Rand Journal of Economics, 36, pp. 391–411.

Pokorny, K. (forthcoming): Pay - But Don't Pay Too Much: An Experimental Study on the Impact of Incentives. Journal of Economic Behavior and Organization.

Prendergast, C. J. (1999): The Provision of Incentives in Firms. Journal of Economic Literature, 37, pp. 7–63.

Rabin, M. (1993): Incorporating Fairness Into Game Theory and Economics. American Economic Review, 83, pp. 1281–1302.

Sausgruber, R. (2005): Testing for Team Spirit - An Experimental Study. Working Paper, University of Innsbruck.

Shearer, B. (2004): Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment. Review of Economic Studies, 71, pp. 513–534.

Sobel, J. (1985): A Theory of Credibility. Review of Economic Studies, 52, pp. 557–573.

Spence, M. (1973): Job Market Signaling. Quarterly Journal of Economics, 88, pp. 355–74.

Tversky, A. and Kahneman, D. (1991): Loss Aversion in Riskless Choice: A Reference Dependent Model. Quarterly Journal of Economics, 106 (4), pp. 1039–1061.

Wilson, R. K. and Sell, J. (1997): "Liar, Liar ..." - Cheap Talk and Reputation in Repeated Public Goods Settings. Journal of Conflict Resolution, 41, pp. 695–717.