No. 57

UNIVERSITY OF COLOGNE

WORKING PAPER SERIES IN ECONOMICS

PEER PRESSURE IN MULTI-DIMENSIONAL WORK TASKS

FELIX EBELING
GERLINDE FELLNER
JOHANNES WAHLIG

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

Peer Pressure in Multi-Dimensional Work Tasks

Felix Ebeling, Gerlinde Fellner and Johannes Wahlig*

November 15, 2012

Abstract

We study the influence of peer pressure in multi-dimensional work tasks theoretically and in a controlled laboratory experiment. Thereby, workers face peer pressure in only one work dimension. We find that effort provision increases in the dimension where peer pressure is introduced. However, not all of this increase translates into a productivity gain, since the effect is partly offset by a decrease of effort in the work dimension without peer pressure. Furthermore, this tradeoff is stronger for workers who run behind in the dimension of peer pressure. Finally, we analyze the optimal group composition to harness peer pressure. Effort in the dimension of peer pressure and overall productivity seem to be unaffected by group composition, but the effort reduction in the dimension that is not subject to peer pressure is stronger when workers' skills are highly diverse. Hence, it seems like optimal group composition depends on work environment. While existing literature recommends maximizing worker-groups' skill diversity in one-dimensional work tasks, our results suggest to mix similar workers in multi-dimensional tasks.

JEL Classification: D03, D2, J21

Keywords: Peer Effects, Multi Tasking, Incentives, Laboratory Experiment

*Felix Ebeling, University of Cologne, Department of Economics, ebeling@wiso.uni-koeln.de, Gerlinde Fellner, University of Ulm, Institute of Economics, gerlinde.fellner@uni-ulm.de, Johannes Wahlig, University of Cologne, Department of Economics, wahlig@wiso.uni-koeln.de. We thank Bernd Irlenbusch, Sebastian Lotz and Axel Ockenfels for helpful comments. Furthermore, we thank the participants of the ESA European Conference 2012. We also gratefully acknowledge the financial support from the German Science Foundation (through Gottfried Wilhelm Leibniz Price of the DFG, awarded to Axel Ockenfels).

1. Introduction

Lately, several research papers discuss the opportunity to use psychological incentives, such as peer pressure, to increase work productivity. In particular, peer effects might serve as a substitute for explicit monetary incentives. The most parsimonious definition of peer effects at the work place is 'the shift of productivity of individual *i* when productivity of individual *j* changes and all else remains equal' (Falk and Ichino 2006, p.40). Recent empirical research also provides recommendations how to compose workforce to optimally exploit such peer effects (see, for instance, Mas and Moretti 2009, Bandiera et al. 2010).

So far, peer pressure has mainly received attention in one-dimensional work tasks. However, many work environments are characterized by a multiplicity of demands and facets, such as quantity and quality. Peer pressure can, however, by definition be only exerted in work dimensions that are subject to social comparisons, i.e. when monitoring by peers is possible. In multi-dimensional work tasks, however, social monitoring will be limited to work dimensions that are easily observable.

In this paper, we argue that peer pressure does not necessarily unfold strictly positive effects but might also have severe drawbacks. Our research centers on the presence of peer effects in more complex work tasks consisting of multiple dimensions. Especially in multidimensional work tasks, peer pressure in a single dimension may crowd out effort in other dimensions, which is alike to the theoretical results of Holmstrom and Milgrom (1991) for monetary incentives in multidimensional work tasks.

Based on Holmstrom and Milgrom (1991), we provide a theoretical framework for this argument and put it to test in an experiment. Our results show that psychological incentives in the workplace seem to have similar drawbacks as pecuniary incentives:

Incentives in an observable work dimension crowd out effort in a non-observable dimension. Consequently, we recommend a more cautious handling of psychological incentives when considering more complex work tasks. In a further step, we investigate how to organize a possibly heterogeneous workforce into work teams to optimally harness peer effects.

The paper is organized as follows. In section 2, we will present an overview of the related literature. By addressing the question of how peer pressure affects performance in multi-dimensional work tasks, we combine two, by now unlinked, sub-branches of research on incentives, that is, research on peer pressure and research on multidimensional incentive problems. To point out the contribution of our paper, we give a brief overview of the existing literature in these areas. In section 3, a theoretical framework is presented that derives the partially detrimental effects of peer pressure. Section 4 explains the setup of an experiment that tests the theory and presents the main hypotheses. Section 5 presents the results and section 6 concludes with a discussion.

2. RelatedLiterature

2.1 Peer Effects

The analysis of peer effects has a long tradition in psychological and sociological research. The positive effect of the presence of peers and the observability of own behavior on effort and performance is denoted by social facilitation (Zajonc, 1965). In economic research, peer effects are distinguished according to their source. In particular, they might arise due to technological or psychological reasons. In former case, peer effects are purely based on rational considerations (Gould and Winter 2009). In contrast, psychological peer effects are based on phenomena as shame, social pressure or contagious enthusiasm that are represented in individuals' preferences. Due

to our research focus, in the following we only consider literature about psychological peer effects.

Economists have only recently started to look at psychological peer effects in a variety of domains, like education (Sacerdote 2001, Zimmerman 2003, Graham 2008), sports (Guryan et al. 2009), crime (Glaeser et al. 1996), public good contribution (Falk et al. 2010) and classical work environments.[1] What this literature has in common is the focus on externalities of individual behavior or abilities on others. Researchers are interested in the question whether such externalities exist, and if so, whether efforts by different individuals are complements or substitutes. An illustrative example is provided by Sacerdote (2001). He uses the random assignment of college dorm roommates to measure peer effects in educational outcomes. Sacerdote finds a positive influence of roommates on each other: an increase in the roommate's first year average grade results in a significant increase of student's first year average grade.

We are concerned with psychological peer effects at the workplace. The existing research on peer effects at the workplace presents mixed results. More specifically, the prevalence of peer effects at the workplace seems to depend on the payment regime. While there is clear evidence for peer effects under fixed wages regime (Falk and Ichino 2006, Mas and Moretti 2009), the picture is less clear for other payment regimes, such as piece-rates. Guyran et al. (2009) suggests that this might be due to a crowding out of psychological incentives by monetary incentives: the more salient the monetary incentive, the less important are psychological incentives for performance. Their own findings from professional golf tournaments as well as laboratory findings by Eriksson et al. (2009) support this view. Both articles do not find evidence for peer effects in

[1] Furthermore we are aware of research on peer effects in welfare participation (Bertrand et al. 2000), unemployment insurance take-up (Kroft 2008) and retirement planning (Duflo and Saez 2003).

tournaments or piece-rate regimes.[2] However, this crowding-out rule might only be valid for individual incentives, since in case of revenue sharing between workers, there is a clear rationale for crowding-in (Kandel and Lazear 1992): the more own payoff depends on the performance of the co-worker, the stronger the peer pressure. Empirical evidence by Chan et al. (2010) supports this view.

We focus on a work environment of fixed wages, which applies to a large variety of jobs. Fixed wages are low powered individual monetary incentives. Intuitively appealing and corroborated by previous research results, such environments leave space for psychological incentives. To the best of our knowledge, there are two existing studies that are closely related to ours in that also analyze peer effects in fixed wage work environment. Falk and Ichino (2006) conducted a controlled field experiment where participants, consisting of high school students, received a fixed payment for stuffing letters into envelopes. The authors identify significant peer effects that increase productivity. Furthermore, in their study less productive workers were more affected by peer effects. They conclude that '"bad apples" gain quality from "good apples", but do not damage the latter one' (Falk and Ichino, 2006, p. 54). Another similar study was done by Mas and Moretti (2009) who investigated peer effects for supermarket cashiers who received fixed wages. They identified positive productivity spillovers of faster co-workers on slower ones and thus derived the recommendation to mix workers with a maximum of skill diversity to increase productivity. However, both studies focus on one-dimensional work tasks, or at least measure output only in the work dimension that is subject to peer pressure. In contrast, we focus explicitly on a work task that consists of multiple dimensions where peer monitoring and thus peer pressure can take effect only

[2] In contrast, Bandiera et al. (2010) do find evidence for peer effects in a piece-rate environment. However, they are less general and only observable for co-workers who are socially tied to each other.

in one dimension. This approach links research on peer pressure to multidimensional incentive problems.

2.2 Multidimensional Incentive Problems

Holmstrom and Milgrom (1991) were first to theoretically analyze the optimal payment scheme for workers in multidimensional work tasks.[3] They provide a rationale for a lack of incentives in contracts even if principals are able to monitor certain work dimensions. As long as there are some work dimensions that are excluded from monitoring, explicit incentives on the monitored dimensions might crowd out effort in non-monitored dimensions. Similarly, Pendergast (1999) emphasizes "dysfunctional behavioral responses" that arise when monetary incentives meet missing permanent holistic measures of the workers' contribution (p.8). This theory is supported by empirical evidence. For instance, Paarsch and Shearer (2000) compare quantity and quality data of workers planting trees under different payment regimes. When workers are paid according to the number of trees they have planted, quantity clearly increases in comparison to fixed wages. However, this productivity increase is partly offset by a reduction in the quality of work. Johnson et al. (2012) came to similar results when scrutinizing the effects of a payment system change for bus drivers in Chile. When wage system changes from fixed-wage system to per-passenger system, waiting time reduced by 13%, but accidents increased by 67%. The results of both papers are very similar to what we observe in our experiment, even though we do not change monetary incentives, but introduce peer pressure in one work dimension.

Let us now turn to the theoretical arguments on how peer pressure might affect effort choice in a multidimensional work task.

[3] Similarly, work might consist of several tasks.

3. A simple model of peer pressure in a multidimensional work task

The primary goal of this section is to give some formal arguments on how workers' effort choice in a multidimensional work task reacts to peer pressure. In line with the literature, we assume that peer pressure arises when workers can compare themselves to their peers in terms of output in a specific work dimension. However, in a work task that consists of multiple dimensions, workers might be able to observe others' output or effort in some, but not in all dimensions. In the simple case of a work task that has a quantity and a quality dimension, it is plausible to assume that quantity output can be more easily observed and monitored than quality output. Thus, peer pressure is present in one work dimension, but not in others, but might still have spill-over effects to effort in work-dimensions that are not subject to peer monitoring.[4]

The baseline of our model is similar to Holmstrom&Milgrom (1991). We consider a work task that consists of *n* dimensions and is paid with a fixed wage *F*. A worker makes the choice of a vector of efforts $\vec{t} = (t_1, \dots, t_n)$ at strictly convex personal costs $C(t)$. We shall suppose that effort in the various work dimensions is perfectly substitutable in the agent's cost function, hence$C(t_1 + \cdots + t_n)$.Contrary to many other models, we shall *not* suppose that all work is unpleasant. A worker on the job may take pleasure in working up to some limit. Incentives are only required to encourage work beyond that limit. Formally, we assume that there is some $t > 0$ such that $C'(t) \leq 0$ for$t \leq \bar{t}$and $C(\bar{t}) = 0$.[5] From the F.O.C. of our utility function $U = F - C(t_1 + \cdots + t_n)$ follows an effort provision of $\vec{t}^* = t_1^* + \cdots + t_n^* = \bar{t}$.

[4]Note that in the model, we speak of 'effort' provided in the different work dimensions, where in fact, only 'output' is observable. If we assume that work output in a specific dimension is an increasing, strictly monotonic function of effort in that dimension, the argument made for effort changes directly translates to changes in observed output.

[5] The cost function is U-shaped with its minimum at $\bar{t}$.

However, contrary to Holmstrom and Milgrom (1991) we do not introduce a monetary incentive in our model, but a psychological incentive that arises from peer observability in dimension i. Formally, peer pressure is modeled as in Kandel and Lazear (1992) by introducing a negative term $P(\cdot)$in the utility function. Thereby, peer pressure decreases in provided effort in dimension i, i.e.,$^{\partial P}/_{\partial t_i} < 0$. This changes workers' utility function to $U = F - C(t_1 + \cdots + t_n) - P(t_i)$.

Deriving the F.O.C. for this new utility function for t_i and t_{-i}, where $-i \neq i$, reveals a change of effort provision. Accumulated effort rises to $\vec{t}^{**} > \overline{t}$, with $t_i^{**} = \vec{t}^{**}$ and $t_{-i}^{**} = 0$. Hence, our model providesa corner solution. Effort in non-peer-observable dimensions is entirely crowded out, while effort in the peer dimension rises even above the previous level of accumulated effort.

4. Experimental design and hypotheses

The advantages of starting the investigation of peer pressure in multidimensional work tasks in a lab setting are twofold. First, it allows creating a work task that consists of not more and not less than two equally important dimensions, one that is observable and one that is not. As we ill show further on, the observability of only one dimension represents a realistic work situation. Second, the anonymous setting in the lab allows control over other factors that might serve as incentives to exert effort, like firm (group) affiliation or close personal relations between co-workers.

The work task that we implemented in the experiment required participants to describe pictures on the computer screen during a period of 25 minutes. For that, subjects earned a fixed wage of 5 Euros.[6] The pictures were selected from a pool of more than 200 million pictures under a creative common license from flikr.com which can be used

[6] Before the 25 minutes started, there was a learning phase of three pictures. In this way participants got to know the handling of describing pictures.

costless for non-commercial purposes (for a typical picture, see Appendix A). Describing digital pictures is a task that is easily executable by humans, but difficult for machines. Nowadays such "Human Intelligence Tasks" are typically done in online micro-labor markets as Amazon's Mechanical Turk (Horton 2010) or Indian data entry firms (Kaur et al. 2010). Thus, this task represents a realistic work setting. Participants were asked to describe each picture with at least one label and at most ten labels. Within this range, participants could freely decide how many labels they wanted to assign to a pictures. We did not give any advice about the appropriate number of labels, but we instructed them that the descriptions are needed for research and teaching purposes. When subjects considered a picture as sufficiently described, they could press a button to get to the next picture on their screen. This work task is ideally suited for our purposes, since it clearly comprises two work dimensions that are clearly measurable: quantity, as the number of pictures, and quality as the (average) number of labels assigned to a picture.

To study the effect of peer pressure on work output, we let subjects experience the same work task in two subsequent phases of 25 minutes each but introduce two different between-subjects treatments, the peer treatment and the control treatment. For reasons explained below, we refer to the first work phase of each treatment as the *no-feedbackworkphase* and the second one as the *feedback work phase*.

In the first work phase, the control and the peer treatment did not differ. All participants worked individually and received no feedback about the work of other participants, hence *no-feedback work* phase. This ensures that peer pressure is absent in this phase, since subjects could neither observe (nor be observed by) others. At the end of the first work phase, subjects were asked to subjectively asses how hard they worked. More precisely, subjects had to give answers to the following three questions on a seven point

scale: How much effort did you exert? How stressed do you feel now? How exhausted did you get? A value 1 corresponded to "not at all" and a value of 7 to "very strong".[7]

In the second work phase, subjects had to complete exactly the same task as before, i.e. describing (different) pictures for a period of 25 minutes for a fixed wage of 5 Euros, but now they worked under a *feedback*-condition. The peer and control treatment differed with respect to the kind of feedback they received.

In the *peer treatment* workers received feedback information about their own current quantity output (i.e., the number of pictures they have described so far in this work phase), as well as the quantity output by another participant, their co-worker. More precisely, the information about the number of pictures described so far in this phase was exchanged in real time between two matched workers.[8] During the whole work phase, subjects could see in the upper right corner of the computer screen the number of pictures that they described so far in this phase, as well as the number of pictures that a co-worker had described so far. As soon as one of the workers finished a picture, the picture counter increased by one. Thus, subjects could observe others and compare themselves only with respect to the quantity dimension. It was not revealed how many labels the co-worker has assigned to the pictures.

Rather particular for a laboratory experiment, the co-worker team, within which information was exchanged, got to know each other before starting the work. This was done after instructing participants for this phase by letting them stand up and allocating co-workers to laboratory seats next to each other. However, participants were not allowed to talk to each other. We decided to implement this unusual design feature basically for two reasons. First, workplaces are almost never completely anonymous.

[7] These questions are similar to the questions Falk and Dohmen (2011) used to find out how hard participants worked in their experiment.
[8] Hence, picture count starts at zero in the feedback-work phase.

Hence, our setting becomes more realistic and improves external validity of our results. Second, in previous studies about peer pressure (especially Falk and Ichino 2006, Mas and Moretti 2009, and Bandiera et al. 2010) workers also know each other. Due to non anonymity, our results are more comparable with theirs.

In the *control treatment,* subjects' feedback consisted only of their own quantity. During the whole phase subjects could see in the upper right corner of the computer screen how many pictures they have described so far. Similarly to the first work phase, subjects described pictures without any information exchange with other participants, but only got feedback, on their own quantity.[9] To keep the *control treatment* as similar as possible to the *peer treatment*, individuals were randomly reseated before the second work phase started.

In both treatments, subjects had to again submit a self-assessment on how hard they worked and on how stressed they felt at the end of the feedback work phase. Altogether, 181 subjects participated in the experiment that was conducted computerized with zTree (Fischbacher, 2007) in the Cologne Laboratory for Economic Research. 122 took part in the *peer treatment* and 59 in the *control treatment*. Subjects were students of the University of Cologne, 93 males and 88 females, from a wide variety of fields of study and were recruited with ORSEE (Greiner 2004). One session lasted between 75 and 85 minutes. The average payoff was 14.1 Euro.

Clearly, peer pressure works via a process of social comparison. To learn whether social preferences moderate the effect of peer pressure, we elicited an indicator for social preferences by using the social value orientation (SVO) slider measure (Murphy et al.

[9] In one session of the control treatment participants received no feedback in the second work phase. This was done to reveal possible experimenter demand effects: subjects might perceive feedback information about a specific work dimension as a hint about the experimenter's preferences for this particular dimension. However, there was no difference across control sessions with and without feedback about own quantity. Hence, we decided to pool the data for the control treatment and stick to the term "feedback work phase".

2011) at the beginning of the experiment. Subjects were asked to complete the six primary SVO Slider items (see Appendix B), that allows a classification of preferences from altruistic to egoistic / competitive. In each of the six items, subjects had to choose an allocation that assigns money to themselves and another, randomly chosen participant in the experiment. After completing this task, a dice was rolled to select one of the items to become payment relevant at the end of the experiment. The dice roll was done by one randomly chosen participant and visible to all participants. Participants within a session were randomly matched into pairs to allocate the payoffs according to the choice in the selected item.[10]

In light of the theory on peer effects in multidimensional tasks, presented in the previous section, we derive three main hypotheses.

First, peer pressure introduced by the exchange of information on quantity between two workers will lead to an increase of quantity output. However, as pure learning effects are expected to increase output as well from the first to the second work phase, we compare the difference in the quantity (i.e., the number of pictures) between the no-feedback to the feedback work phase across the peer and control treatment.

> ***Hypothesis 1.*** *In the peer treatment, output in the quantity dimension will increase more from first to the second work phase than in the control treatment.*

The effort increase in one work dimension due to peer pressure will crowd out effort, and thus output, in the other work dimension that is not observable. We therefore state

[10] Thus, every participant received a payoff from the SVO slider task as an allocator as well as a receiver.

Hypothesis 2. *In the peer treatment, output in the quality dimension will decrease more from the first to the second work phase than in the control treatment. In particular, under the assumptions made on individuals' effort cost function, output is reduced to the minimum, i.e. one label per picture.*

Still, the effect of peer pressure on overall output that can be measured by the aggregate number of labels assigned is predicted to be positive, since the quantity increase due to peer pressure overcompensates the quality decrease.

Hypothesis 3.*In the peer treatment, productivity will increasemore from the first to the second work phase than in the control treatment.*

5. Results

We split up our main findings into three parts. In section 5.1, we investigate the general effects of peer pressure on workers' output and productivity. In section 5.2, we report how peer pressure individually affects different types of workers. In section 5.3, we analyze our data with regard to the optimal exploitation of peer effects.

5.1 Effort Provision and Peer Pressure

Our first result concerns the effect of peer pressure on effort provision in the two different work-dimensions, quantity and quality. In each work phase quantity is measured by the number of described pictures, whereas quality is measured by average number of labels per pictures. Furthermore we are interested in the impact of peer pressure on productivity, which is measured by the aggregate number of labels (or output in the quantity times output in the quality dimension).

In the no-feedback phase, subjects described on average 30.3 pictures (SD=1.77) with 5.5 labels (SD=0.26) in the control treatment and 33.1 pictures (SD=1.68) with 5.3 labels (SD=0.15) in the peer treatment. Hence, as to be expected, Mann Whitney U-test show

no difference between quantity as well as quality output in the first work phase (p=.61 for pictures and p=.66 for labels). Note also, that although the fixed wage provides no monetary incentives to work at all, subjects obviously exhibit significant work effort which lends strong supports to the theoretical assumption that not all work is unpleasant.

Figure 1 summarizes the changes of quantity, average quality and productivity between the no-feedback and feedback work-phase for both the control and the peer treatment. In the control treatment, quantity increases by an average of 10.4 pictures, while in the peer treatment the increase was 18.2 pictures. At the same time, quality decreases, in the control treatment by an average of 0.32 labels and in the peer treatment by 0.75 labels per picture. Nevertheless, overall productivity increases on average by 31.3 labels in the control treatment and 46.4 labels in the peer treatment[11]. Note that also in the control treatment all dimensions significantly change, which can be considered the result of a learning effect. They represent a base level to evaluate the impact of peer pressure. Results from Mann-Whitney U test reveal that the differences in the peer treatment always exceed the differences in the control treatment, for the quantity and quality dimension (p<.01 and p=.04) as well as for the overall productivity (p<.01).[12] These first results confirm our expectations: peer pressure leads to an increase in quantity (supporting Hypothesis 1), and a decrease in quality, although we do not observe a complete crowding out of quality, as Hypothesis 2 suggests. On average, workers in the peer treatment assign an average of 4.5 labels (SD=0.16) to a picture in the feedback work phase, which is significantly more than the minimum of 1 (95% confidence interval is [4.22;4.87]). Hence, effort dimensions seem to be substitutes and

[11] All differences are significantly different from 0 (according to 95% confidence intervals around the mean observed differences).
[12] Cumulative distribution functions of the differences in quality, quantity and productivity between the no feedback and feedback work phase can be found in Appendix B.

quantity crowds out quality. Still overall productivity increases confirming the expectation formulated in Hypothesis 3.

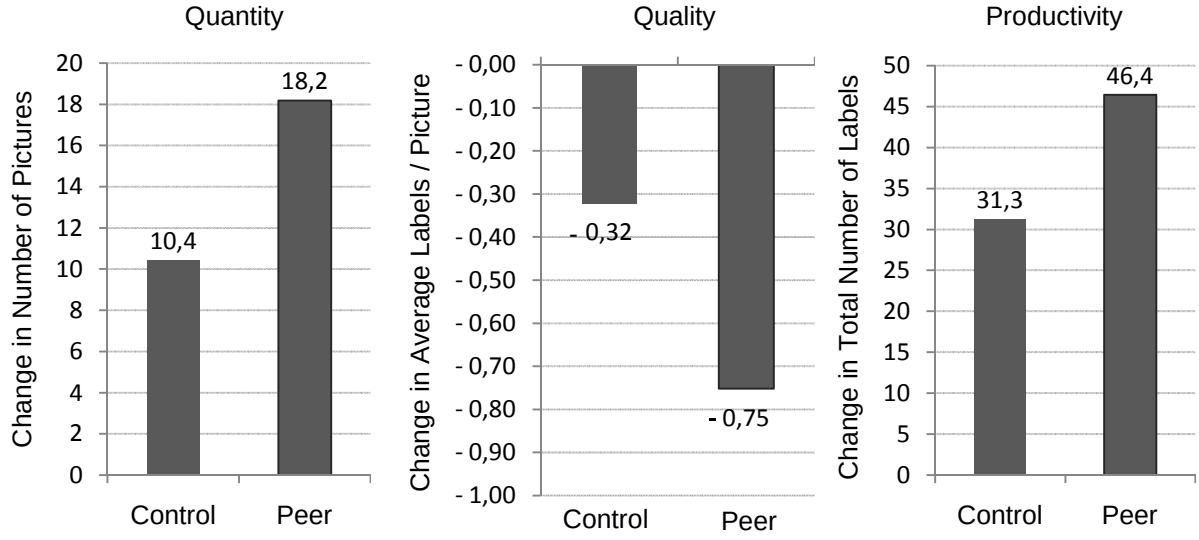


FIGURE 1. MEAN CHANGE FROM NO-FEEDBACK TO FEEDBACK PHASE

Notes: The figure shows, for both treatments, the mean change of quantity (=number of described pictures), quality (= average labels per picture) and productivity (=total number of labels) from the no-feedback to the feedback work phase. In case of quality it reports the unweighted mean of subjects' mean number of labels.

To substantiate our first results and control for (ex-ante) heterogeneity of workers we additionally conduct regression analyses, where the dependent variables are the differences in quantity, quality and productivity between the no-feedback and feedback-phase and the main independent variable is a dummy for the peer treatment. Moreover, we add the absolute amount of quality, quantity and productivity in the no-feedback phase as independent variables in specifications (2) to (4). Tables 1 to 3 present the results that confirm a stronger increase of quantity and productivity and a stronger decrease of quality in the peer treatment. Furthermore, regressions reveal the importance of ex-ante differences between workers. Independent of the treatment, ex ante more productive workers increase quantity more strongly, while workers who provide ex ante more quality stronger decrease quality, but also stronger increase productivity.

TABLE 1. QUANTITY DIFFERENCE BETWEEN TREATMENTS

	(1)	(2)	(3)	(4)
Peer treatment	7.76***	7.72***	7.47***	7.21***
	(2.63)	(2.64)	(2.57)	(2.46)
Quality-NF		-0.29		
		(0.79)		
Quantity-NF			0.11	
			(0.09)	
Productivity-NF				0.08***
				(0.02)
Observations	181	181	181	181
R-squared	0.0477	0.0487	0.0597	0.1155

Notes: OLS Regressions with "difference in number of pictures between feedback and no-feedback phase" as dependent variable. Robust standard errors in square brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01. "Peer" represents binary variable for peer treatment, "Quality-NF" represents variable for average number of labels per picture in the no-feedback work phase, "Quantity-NF" represents variable for number of picture in the no-feedback work phase, "Productivity-NF" represents variable for total number of labels in the no-feedback work phase

TABLE 2. QUALITY DIFFERENCE BETWEEN TREATMENTS

	(1)	(2)	(3)	(4)
Peer treatment	-0.43**	-0.45***	-0.45**	-0.42**
	(0.18)	(0.17)	(0.17)	(0.17)
Quality-NF		-0.14***		
		(0.05)		
Quantity-NF			0.01	
			(0.00)	
Productivity-NF				0.00
				(0.00)
Observations	181	181	181	181
R-squared	0.0327	0.0826	0.0438	0.0386

Notes: OLS Regressions with "difference in average number of labels per picture between feedback and no-feedback phase" as dependent variable. Robust standard errors in square brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01.

TABLE 3. PRODUCTIVITY DIFFERENCE BETWEEN TREATMENTS

	(1)	(2)	(3)	(4)
Peer treatment	15.17** (5.59)	16.27*** (5.64)	16.72** (5.69)	14.55*** (5.58)
Quality-NF		6.57*** (1.85)		
Quantity-NF			-0.56* (0.26)	
Productivity-NF				0.09 (0.07)
Observations	181	181	181	181
R-squared	0.0295	0.1105	0.0826	0.0437

Notes: OLS Regressions with "difference in total number of labels between feedback and no-feedback phase" as dependent variable. Robust standard errors in square brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01.

5.2 Individual Adjustment to Peer Pressure

As we have already established how peer pressure affects behavior on aggregate, we want to gain a more precise picture on how workers of different abilities adjust to peer pressure. Since we cannot observe ability directly, we proxy it by the quantity provided in the no-feedback work phase. Thereby, our experimental design allows measuring the influence of absolute and relative output provisionin the no-feedback work phase on adjustment processes in the feedback phase. An *absolutely* high output implies that a worker's output (in the no-feedback work phase) is high in comparison to all other workers' output (in the no-feedback work phase), a *relatively* high output implies his output (in the no-feedback work phase) is high in comparison with the co-worker's output (in the no-feedback work phase) he is paired with later on. Thereby, we focus on relative differences in the quantity dimension. The idea is that a worker who provides less (more) quantity in the no-feedback work phase will soon run behind (ahead) her co-

worker in the feedback phase. Peer pressure then might affect the behavior of the different workers differently. As the subsequent analysis shows, the ex ante relative differences in quantity are indeed the most important variable for adjustment processes.

Figure 2 visualizes the reaction of different workers to peer pressure. In the three graphs, the abscissa represents the absolute difference in quantity between the worker and her co-worker in the no feedback phase (before they were matched). For example, a worker with a value of 40 described 40 pictures more in the no-feedback phase than her co-worker. We label this difference as "Heterogeneity" since it describes how workers differ in their quantity provision before the exchange of peer information. The ordinate again represents the respective absolute differences in quantity, quality and productivity between the no-feedback and the feedback phase.

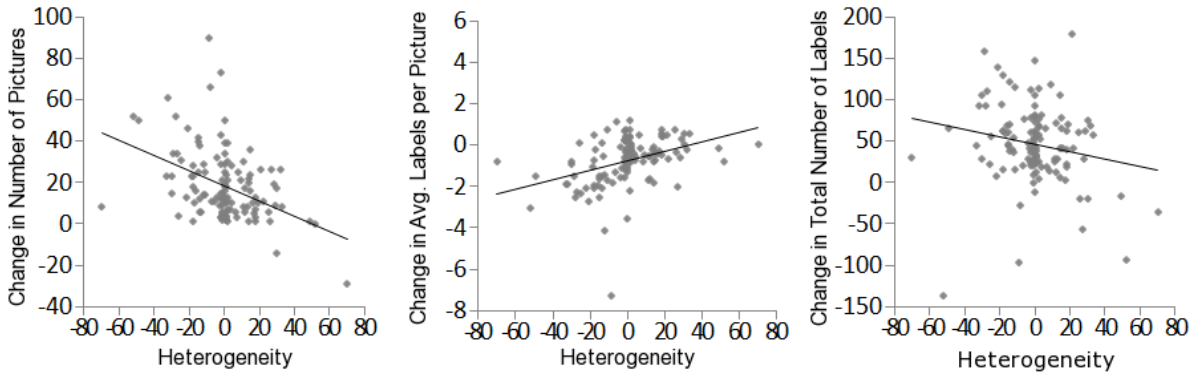


FIGURE 2. REACTION OF DIFFERENT WORKERS TO PEER PRESSURE

Notes: The figure shows the quantity, quality and productivity change for workers who provide more or less quantity in the no-feedback phase than their later co-worker. The solid line represents the least square regression slope.

The graphs in Figure 2 suggest a correlation of between heterogeneity and quantity, quality and productivity, which is confirmed by Spearman Rank Correlation tests. Quantity change is significantly negatively correlated to heterogeneity (spearman's rho=-0.37, $p<.01$) whereas quality is significantly positively correlated (spearman's

rho=0.47, p<.01). This means that workers who are worse than their (later) co-workers in quantity provision increase their quantity under peer pressure more than workers who are better, but this is achieved at the expense of decreasing quality significantly more. Concerning productivity, the correlation with heterogeneity is significantly negative (spearman's rho=-0.18, p<.05). This means that workers who were ahead their co-worker increase their productivity less strongly under peer pressure.

We again substantiate our findings by regression analyses, while at the same time controlling for absolute levels of individual quantity, quality and productivity in the no-feedback phase.[13] Moreover, we control for heterogeneity in social preferences by adding the score in the social value orientation (SVO) slider measure as independent variable. Tables 4 to 6 present the results. While the impact of heterogeneity on quantity and quality is highly significant across specifications, the influence of heterogeneity on productivity becomes (partly) insignificant, in particular when controlling for the initial quantity and quality. The less significant correlation for productivity is intuitively appealing, as opposing correlations of quantity and quality with respect to heterogeneity neutralize the correlation for productivity.

Comparing this finding with existing evidence on adjustment processes due to peer pressure, our results represent a refinement. For example, Mas and Moretti (2009) reveal a stronger influence of peer pressure on workers with below average productivity. However, in their paper productivity is the observable work dimension. Our results suggests that when work dimensions can be disentangled, productivity is no longer the most important variable for the adjustment processes, but rather the dimension in which co-workers compare themselves.

[13] Furthermore, controlling for these variables avoids conclusions that emerge due to "regressions to the mean".

TABLE 4. INFLUENCE OF HETEROGENEITY ON QUANTITY CHANGE IN THE PEER TREATMENT

	(1)	(2)	(3)	(4)	(5)
Heterogeneity	-0.36***	-0.38***	-0.43***	-0.53***	-0.41***
	(0.04)	(0.04)	(0.06)	(0.07)	(0.05)
SVO-slider-value		-0.33***	-0.30***	-0.32***	-0.32***
		(0.11)	(0.10)	(0.10)	(0.10)
Quality-NF			-1.57		
			(1.22)		
Quantity-NF				0.32***	
				(0.11)	
Productivity-NF					0.08***
					(0.24)
Observations	122	122	122	122	122
R-squared	0.1741	0.2241	0.2448	0.3113	0.3047

Notes: OLS Regressions with "difference in number of pictures between feedback and no-feedback phase" as dependent variable. Robust standard errors in brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01. "Heterogeneity" represents the difference between pictures described by the worker and her co-workers in the no feedback phase, "SVO-slider-value" represents the worker's SVO-slider score.

Another noteworthy observation in Table 4 is the significant relation between the SVO-slider-score and the quantity change due to peer pressure. Note that a smaller SVO-slider score indicates a more egoistic / competitive attitude. We find that the more competitive an individual is, the stronger she increases the quantity under peer pressure. Hence, competitive preferences seem to be a driver of peer effects. This fits nicely into previous findings about explanations for peer pressure, as Mas and Moretti (2009) already identified social pressure as a driver of peer effects. Competitiveness and the reaction to social pressure are both part of social comparison processes (Festinger 1954) and closely related.

TABLE 5. INFLUENCE OF HETEROGENEITY ON QUALITY CHANGE IN THE PEER TREATMENT

	(1)	(2)	(3)	(4)	(5)
Heterogeneity	0.02***	0.02***	0.02***	0.03***	0.02***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
SVO-slider-value		0.01	0.01	0.01	0.01
		(0.01)	(0.01)	(0.01)	(0.01)
Quality-NF			0.09		
			(0.08)		
Quantity-NF				-0.01	
				(0.01)	
Productivity-NF					0.00
					(0.01)
Observations	122	122	122	122	122
R-squared	0.15151	0.1547	0.1688	0.1649	0.1691

Notes: OLS Regressions with "difference in average number of labels per picture between feedback and no-feedback phase" as dependent variable. Robust standard errors in brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01.

TABLE 6. INFLUENCE OF HETEROGENEITY ON PRODUCTIVITY CHANGE IN THE PEER TREATMENT

	(1)	(2)	(3)	(4)	(5)
Heterogeneity	-0.45**	-0.45**	-0.22	-0.11	-0.48**
	(0.17)	(0.18)	(0.22)	(0.39)	(0.21)
SVO-slider-value		0.06	-0.07	0.03	0.07
		(0.38)	(0.35)	(0.37)	(0.37)
Quality-NF			7.17**		
			(3.37)		
Quantity-NF				-0.69	
				(0.51)	
Productivity-NF					0.07
					(0.11)
Observations	122	122	122	122	122
R-squared	0.0362	0.0364	0.0946	0.0913	0.0432

Notes: OLS Regressions with "difference in total number of labels between feedback and no-feedback phase" as dependent variable. Robust standard errors in brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01.

When thinking about possible adaptation processes of workers to peer pressure it is interesting to consider the self-assessment and self-reported stress level. Does an increase in quantity and productivity due to peer pressure also lead to an increase exhaustion and stress? We therefore turn to the self-evaluation of participants which they had to complete after each work phase. Table 7 summarizes how participants' self-evaluation changed from the no-feedback to the feedback work phase in both treatments.[14] Mann-Whitney-U tests detect no significant difference between treatments with respect to changes of subjective effort (p=0.83), stress (p=.63) and

[14] Recall that answers were given on a seven point scale.

exhaustion(p=.33).[15] This is remarkable, as one might expect higher values in the peer treatment due to the significantly stronger increase in productivity.

TABLE 7. MEAN CHANGE OF SUBJECTIVE EVALUATION

	Control	Peer
Effort	-0.20	-0.15
Stress	0.57	0.39
Exhaustion	0.61	0.90

5.3 Optimal Group Composition

In the previous section we only analyzed how individual behavior adjusts to peer pressure. For employers, it is of practical relevance how to compose co-workers to harness theses individual adjustment processes. As pointed out by Mas and Moretti (2009), an optimal group composition in terms of exploitation of peer pressure may decrease labor costs for a constant productivity level. In our setup, labor costs are constant, but for given labor costs, our experiment design allows to precisely investigate whether peer pressure has more positive effects in co-worker teams that are rather heterogeneous or homogeneous. To that end, we split co-worker teams in our peer treatment into two groups according to their absolute value of the "Heterogeneity" measure, i.e. the difference in quantity by the two co-workers in the no-feedback phase.[16] More specifically, the "homogenous" group (n=46) contains all workers in the 95% confidence interval around "Heterogeneity" mean of 0 which is [-3.5;+3.5]. The other, "heterogenous" group (n=76) thus contains individuals that rather dissimilar in their ex ante quantity provision.

[15] Additionally, ordered probit regressions to control for individual characteristics did not reveal any differences between treatments.

[16] We focus on the "Heterogeneity" measure as the previous section showed the prevalent importance of this variable on adjustments.

Figure 3 summarizes the changes of average quantity, quality and productivity between the no-feedback and feedback work-phase for heterogeneous as well as homogeneous group members.In the homogeneous groups, quantity increases by an average of 16.2 pictures, while in the heterogeneous groups the increase was 19.4 pictures. At the same time, quality decreases, in homogeneous groups by an average of 0.45 labels and in the heterogeneous groups by 0.93 labels per picture. Overall productivity increases on average by 48.9 labels in the homogeneous groupsand 45 labels in the heterogeneous groups. Results from Mann-Whitney U test reveal a significant difference between homogeneous and heterogeneous groups in terms of quality decrease ($p<.05$), while there is no significant difference for quantity and productivity ($p=0.2$ and $p=0.98$, respectively).

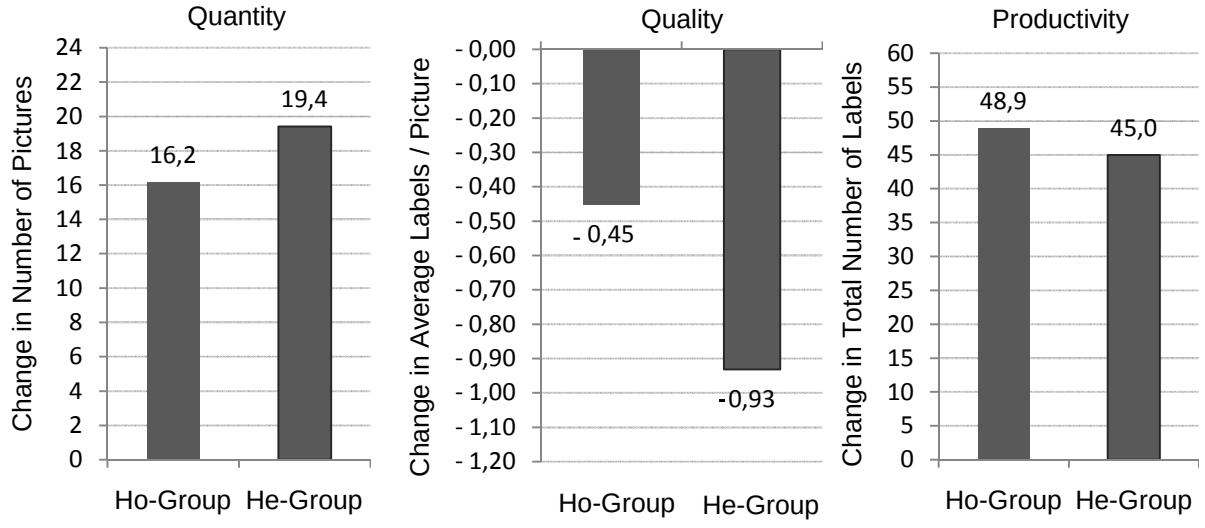


FIGURE 3. MEAN CHANGE FROM NO-FEEDBACK TO FEEDBACK PHASE FOR GROUPS

Notes: The figure shows, for both types of groups, the mean change of quantity (=number of described pictures), quality (= average labels per picture) and productivity (=total number of labels) from the no-feedback to the feedback work phase. In case of quality it reports the unweighted mean of subjects' mean number of labels.

We again substantiate our findings by regression analyses. The regressions in Tables 8 to 10 present the analysis of how team composition affects the change in quantity, quality and productivity in the peer treatment.[17] The independent variable "Hetero-Group" is a dummy variable that indicates that a worker is part of a heterogeneous team. Results confirm the significant effect of group composition on quality. Again, quantity and productivity seems to be unaffected by group composition. To check the robustness of these results, we extended the interval for defining the homogeneous groups up to an interval of [-10;+10] for the Heterogeneity variable. Results do not change. Hence, our data suggest that groups should be composed of rather homogenous workers to optimally harness peer effects. Although overall productivity does not seem to be affected much by the co-worker composition, the crowding out of quality by peer pressure is higher in heterogeneous teams. The reason for these results can easily be seen in Figures 4 to 6 in Appendix C. While Figure 2 in the previous chapter already indicates a stronger adjustment of individuals running behind, the Figure5 in Appendix C shows that this tendency is disproportionately high for the quality dimension, which is responsible forthe significant lower quality output in heterogeneous groups.

[17] In the previous regressions in Tables 1 to 6 we did not include "difference in average number of labels per picture" as independent variable when measuring effects on "difference in number of pictures" and vice versa. We did so, because these regressions include participants from the control treatment. Since the independent variables "control", "difference in average number of labels per picture" and "difference in number of pictures" are highly correlated we had to exclude the latter ones to avoid multicollinearity. Since the regressions in Table 8 to 10 do not contain participants from the control treatment, we include these variables.

TABLE 8. INFLUENCE OF GROUP COMPOSITION ON QUANTITY

	(1)	(2)	(3)	(4)	(5)
Hetero-Group	-1.81	-0.93	-1.92	-2.03	-1.69
	(2.69)	(2.55)	(2.37)	(2.32)	(2.33)
Change in Quality	-10.62***	-10.46***	-11.20***	-10.77***	10.37***
	(0.93)	(0.85)	(0.86)	(0.81)	(0.81)
SVO-slider-value		-0.25***	-0.20**	-0.23***	-0.24***
		(0.09)	(0.08)	(0.08)	(0.09)
Quality-NF			-1.76*		
			(0.89)		
Quantity-NF				0.11	
				(0.10)	
Productivity-NF					0.05***
					(0.02)
Observations	122	122	122	122	122
R-squared	0.5067	0.5353	0.5630	0.55	0.5667

Notes: OLS Regressions with "difference in number of pictures between feedback and no-feedback phase" as dependent variable. Robust standard errors in square brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01. "Hetero-Group" represents whether worker exchanged peer information with a worker providing differing quantity in the no feedback phase, "Change in Quality" represents the difference in average number of labels per pictures between the no feedback and the feedback phase.

TABLE 9. INFLUENCE OF GROUP COMPOSITION ON QUALITY

	(1)	(2)	(3)	(4)	(5)
Hetero-Group	-0.32**	-0.29*	-0.35***	-0.38***	-0.32**
	(0.15)	(0.15)	(0.13)	(0.14)	(0.15)
Change in Quantity	-0.05***	-0.05***	-0.05***	-0.05***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SVO-slider-value		-0.01	-0.00	-0.01	-0.01
		(0.01)	(0.01)	(0.01)	(0.01)
Quality-NF			-0.17***		
			(0.05)		
Quantity-NF				0.01*	
				(0.01)	
Productivity-NF					0.00
					(0.00)
Observations	122	122	122	122	122
R-squared	0.5227	0.5307	0.5937	0.5588	0.5438

Notes: OLS Regressions with "difference in number of pictures between feedback and no-feedback phase" as dependent variable. Robust standard errors in square brackets.Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01. "Change in Quantity" represents the difference in number of pictures between the no feedback and the feedback phase.

TABLE 10. INFLUENCE OF GROUP COMPOSITION ON PRODUCTIVITY

	(1)	(2)	(3)	(4)	(5)
Hetero-Group	-3.93	-4.35	-1.52	1.82	-5.07
	(8.76)	(8.99)	(8.86)	(8.54)	(8.43)
SVO-slider-value		0.13	-0.06	0.03	0.14
		(0.39)	(0.35)	(0.37)	(0.38)
Quality-NF			8.07***		
			(2.94)		
Quantity-NF				-0.76**	
				(0.32)	
Productivity-NF					0.05
					(0.09)
Observations	122	122	122	122	122
R-squared	0.0017	0.0027	0.0873	0.0901	0.0062

Notes: OLS Regressions with "difference in number of pictures between feedback and no-feedback phase" as dependent variable. Robust standard errors in square brackets. Standard errors are clustered on groups. A group is a co-worker team in the peer treatment and a single person in the control treatment. Level of significance: *p<0.1, **p<0.05, ***p<0.01.

6 Summary and discussion

In this paper, we study peer effects in multi-dimensional work tasks. So far, peer pressure has received attention in the economic literature as an alternative or additional work incentivation and researchers have been concerned about optimal group composition to best utilize peer effects for productivity increase. However, the research so far has exclusively focused on peer pressure in one-dimensional work tasks, which is a far abstraction from many work settings. We argue that similar to crowding out of intrinsic motivation by monetary incentives, peer pressure that affects effort in one dimension of the work tasks might at the same time crowd out effort in other dimensions that are not observable by peers.

First, we present a simple theoretical model to predict the influence of peer pressure in a multi-dimensional work task that rests on the combination of models on peer pressure

with models of incentives in multidimensional work tasks. In the model we assume that co-workers are only capable of comparing each other's output in one dimension. We then compare the case to the case without co-worker monitoring and peer pressure, respectively. With plausible assumptions on worker's effort cost function our model predicts that the introduction of peer pressure entails a (complete) crowding out of effort in the non-observable work dimension by effort in the observable one.

Second, we test the predictions in a laboratory experiment. In particular, we implement a two-dimensional work task in fixed wages environment that consists of separate quantity and quality dimensions. Peer pressure is introduced in the quantity dimension by exchanging information on the quantity output between two co-workers, while no information is exchanged about co-workers output in the quality dimension.

In accordance to the theoretical expectations, we observe a highly significant increase in quantity due to the introduction of peer pressure while quality highly significantly decreases. Thus, our results demonstrate that there is no invisible border preventing psychological incentives, such as peer effects, to have similar drawbacks as pecuniary incentives. However, aggregate effort rises when peer pressure is introduced.

Beside these core findings, our experimental data identifies competitiveness as a driver of peer effects and reveals that stress levels are unaffected by peer pressure despite the significant increase in productivity. In contrast, Falk and Dohmen (2011) find in their study that payment schemes that induce higher performance are accompanied by higher levels of subjective effort, stress and exhaustion. A possible explanation for the difference between their study and ours might be the smaller increase of productivity in our study. Peer pressure does not increase productivity to the same extent as monetary incentives do.Therefore, stress levels might not increase to the same extent.Alternatively, there might be differences in the subjective perception of the

interventions. Hence, even in case of similar productivity gains by pecuniary and psychological incentives subjective assessment would only change in the first case. However, independent of the explanation, peer pressure seems to be a mean to increase productivity without increasing stress, which might have interesting implications for employers: The satisfaction of workers does not change due to the introduction of peer pressure. This means that the possibly negative consequences that accrue from a change of the payment regime, as suggested by Falk and Dohmen (2010), less likely appear under peer pressure despite significant productivity gains.

Moreover, we analyze whether different types of workers are affected differently by peer pressure. We differentiate workers by the ex-ante distance to their later co-worker in the work dimension that is subject to peer pressure and find that weaker workers who lag behind show a stronger increase in quantity due to peer pressure, but at the same time a stronger decrease in quality than workers who are ahead. Previous literature of Falk and Ichino (2006), Mas and Moretti (2009) and Guryan et al. (2009) discusses whether workers' susceptibility to peer pressure depends on workers attributes such as ability. They conclude that low-skilled workers respond stronger to peer pressure. However, in our multi-dimension task, a slightly modified conclusion is indicated: While the absolute skill or output level of the worker seems to be of minor influence, the ex-ante difference to the co-worker in the later observable work dimension is a significant predictor for workers' susceptibility to peer pressure. Generally, the effect of peer pressure is stronger for those providing less output in the dimension later exposed to peer pressure. However, the relation is less clear for the increase of productivity: those who increase quantity stronger also decrease quality stronger, leading to a roughly zero net gain of productivity.

We also investigate how to optimally harness peer pressure by analyzing whether employers should create teams with workers who are (ex-ante) similar or different in the level of quantity they achieve. We find that neither quantity nor productivity is sensitive to the mix of workers within a team. However, the decrease in quality is significantly stronger when teams are composed of heterogeneous workers. Hence, our data recommends composing work teams of rather similar workers. Comparing this result with results of previous studies on optimal group composition (e.g. Falk and Ichino 2006, Mas and Moretti 2009 and partly Bandiera et al., 2010) suggests that optimal group composition depends on work environment. While the mentioned literature on one-dimensional work tasks recommends mixing highly diverse workers, we arrive at the exact opposite conclusion in multi-dimensional work tasks.[18]

Finally, even if we find several particularities of peer pressure, drawbacks are generally the same as for monetary incentives. Hence, our findings might explain why peer pressure (or psychological incentives in general) is rarely utilized in real work environments, despite the fact that workers are highly responsive to it.

[18] Furthermore, our results most probably represent a lower bound for quality decrease. When the effort in the dimension observable by peers indeed rises stronger in heterogeneous groups as found by previous research, decrease in the other dimension might be even stronger.

References

Akerlof, George A. 1982. "Labor Contracts as Partial Gift Exchange."*The Quarterly Journal of Economics*, 97, 543-569.

Baker, George, Robert Gibbons and Kevin J. Murphy.2002, "Relational Contracts and the Theory of the Firm."*The Quarterly Journal of Economics*, 117(1), 39-84.

Bandiera, Oriana, IwanBarankay and Imran Rasul.2010. *"Social Incentives in the Workplace." The Review of Economic Studies*, 77(2), 417-458.

Bertrand, Marianne, Erzo F. P. Luttmer and SendhilMullainathan. 2000. "Network Effects and Network Cultures." *The Quarterly Journal of Economics*, 115(3), 1019-1055.

Chan, Tat, Jia Li and Lamar Pierce. 2010. "Compensation and Peer Effects in Competing Sales Teams." *Working Paper.*

Chen, Yan, F. Maxwell Harper, Joseph Konstan and Sherry XinLi. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens." *American Economic Review*, 100(4), 1358–98.

Duflo, Esther and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics*, 118(3), 815-842.

Eriksson, Tor, Anders Poulsen and Marie Claire Villeval. 2009. "Feedback and Incentives: Experimental Evidence." *Labour Economics*, 16, 679-688.

Falk, Armin and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics*, 24(1), 39-57.

Falk, Armin and Thomas Dohmen. 2010. "You Get What You Pay For: Incentives and Selection in the Education System." *The Economic Journal*, 120, 256-271.

Falk, Armin and Thomas Dohmen. 2011. “Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender.” *The American Economic Review*, 101, 556-590.

Falk, Armin, UrsFischbacher and Simon Gächter. 2010. “Living in two Neighborhoods – Social Interaction Effects in the Laboratory.” *Economic Inquiry*, 48, 1 – 16.

Farrel, Jospeh and Carl Shapiro. 1990. “Horizontal Mergers: An Equilibrium Analysis.” *The American Economic Review*, 80(1), 107-126.

Fehr, Ernst, Simon Gächter and Georg Kirchsteiger.1997. „Reciprocity as a Contract Enforcement Device: Experimental Evidence.“ *Econometrica*, 65(4), 833-860.

Festinger, Leon. 1954. “A theory of social comparison processes.” *Human Relations*, 7, 117-140.

Fischbacher, Urs. 2007. “z-Tree: Zurich Toolbox for Ready-made Economic Experiments.” *Experimental Economics*, 10(2), 171-178.

Glaeser, Edward L., Bruce Sacerdote and José A. Scheinkman. 1996. “Crime and Social Interactions.” *The Quarterly Journal of Economics*, 111(2), 507-548

Gneezy, Uri and Aldo Rustichini. 2000. “A Fine is a Price.” *Journal of Legal Studies*, 29, 1-17.

Gould, Eric D. and Eyal Winter. 2009. “Interactions between Workers and the Technology of Production: Evidence from Professional Baseball.” *The Review of Economics and Statistics*, 91(1), 188-200.

Graham, Bryan S. 2008. “Identifying Social Interactions through Conditional Variance Restrictions.”*Econometrica*, 76(3), 643-60.

Greiner, Ben. 2004. "An Online Recruitment System for Economic Experiments." In: Kremer, Kurt, Macho, Volker (Eds.), Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen, 79-93.

Guryan, Jonathan, Kory Kroft and Matthew J. Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1(4), 34-68.

Holmstrom, Bengt. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *The Review of Economic Studies*, 66(1), 169-182.

Holmstrom, Bengt and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization*, 7, 24-52.

Horton, John J. 2010. "Employer Expectations, Peer Effects and Productivity: Evidence from a Series of Field Experiments." *Working Paper*.

Johnson, Ryan M., David H. Reiley and Juan Carlos Munoz. 2012. "The War of Fare: How Driver Compensation Affects Bus System Performance."*Working Paper*.

Kandel, Eugene and Edward P. Lazear. 1992. "Peer Pressure and Partnerships" *Journal of Political Economy*, 100(4), 801-817.

Kaur, Supreet, Michael Kremer and SendhilMullainathan. 2010. "Self Control and the Development of Work Arrangements." *The American Economic Review: Papers and Proceedings*, 100, 624-628.

Kreps, David M. 1997. "Intrinsic Motivation and Extrinsic Incentives."*The American Economic Review*, 87(2), 359-364.

Kroft, Kory. 2008. “Takeup, Social Multipliers and Optimal Social Insurance.” *Journal of Public Economics*, 92(3-4), 722-737.

MacLeod, W. Bentley and James M. Malcomson. 1989. “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment.” *Econometrica*, 57(2), 447-480.

MacLeod, W. Bentley and James M. Malcomson. 1998. “Motivation and Markets.” *The American Economic Review*, 88(3), 388-411.

Mas, Alexandre and Enrico Moretti. 2009. “Peers at Work.” *The American Economic Review*, 99(1), 112-145.

Murphy, Ryan O., Kurt A. Ackermann and Michel J. J. Handgraaf.2011. “Measuring Social Value Orientation.” *Judgment and Decision Making*, 6(8), 771-781.

Paarsch, Harry J. and Bruce Shearer. 2000. “Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records.” *International Economic Review*, 41(1), 59-92.

Pendergast, Canice. 1999. “The Provision of Incentives in Firms.” *Journal of Economic Literature*, 37(1), 7-63.

Sacerdote, Bruce. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *The Quarterly Journal of Economics*, 116(2), 681-704.

Zajonc, Robert B. 1965. “Social facilitation.”*Science*, 149, 269-274.

Zimmerman, David J. 2003. “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment.” *The Review of Economics and Statistics*, 85(1), 9-23.

Appendix A. Instructions (Translated from German)

General Part

Welcome to today's experiment! In this experiment you will be able to earn money. The amount you will earn depends on your decisions and the decisions of other participants.

This experiment consists of three parts. The instructions concerning the first part have been handed out to you with this sheet. The instructions concerning the second and third part will be handed out at the beginning of the respective part. Additionally to your payout from all three parts of the experiment you will receive a fixed payment of 2.50€ for your presence.

Please turn of your mobile phone and abstain from communicating with other participants. Please raise your hand if you have any questions concerning the experiment. We will come to you and answer your question.

All decisions made during the experiment as well as all payments will be kept anonymous.

First Part

In this part you will be presented six decision situations. In each of them you can choose, which amount of money you will receive and which amount of money another participant will receive.

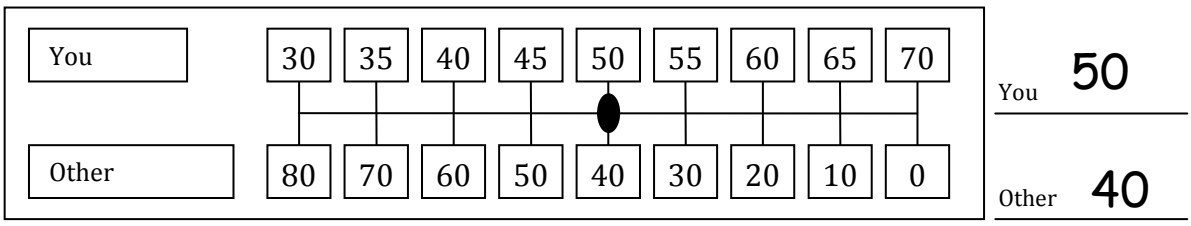
Please choose **one** distribution for each decision situation. All amounts of money in this part are denoted in **€-Cent**.

At the end of this part, one of the six decision situations will be chosen by chance. The decision you made in this situation will be paid out at the end of the experiment.

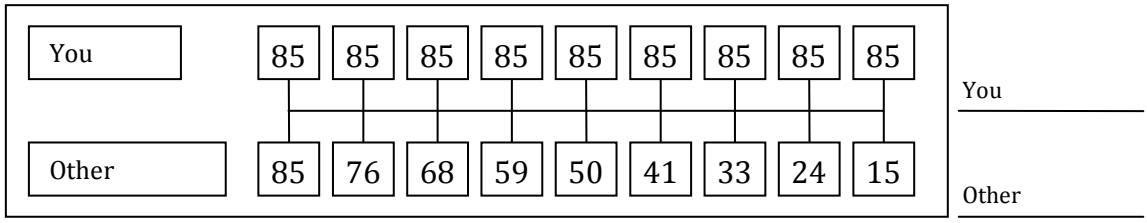
Please note that answers are neither right nor wrong in this part. You are exclusively asked to state your individual preferences. **Once you have made a decision, please mark the appropriate position on the center line and write the respective distribution of money onto the marks on the right**. With your decision you influence your individual payoff, as well as the payoff of the other participant. This participant will be referred to as "Other" hereafter.

Which other participant will receive the money from you will remain open for now.

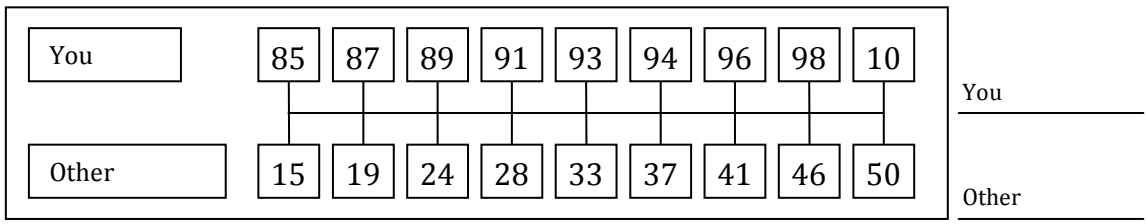
Example:



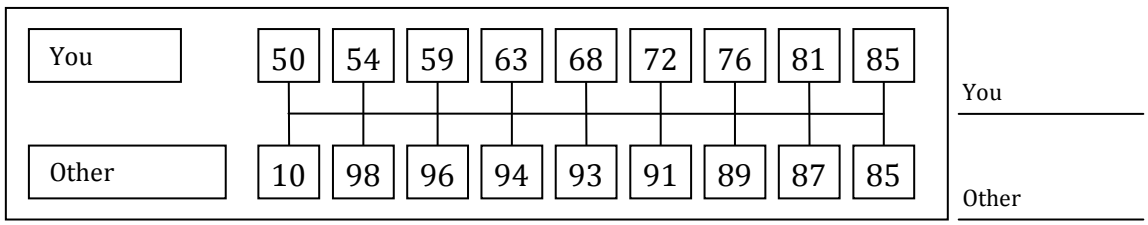
1)


You
85 85 85 85 85 85 85 85 85
Other
85 76 68 59 50 41 33 24 15
You
Other

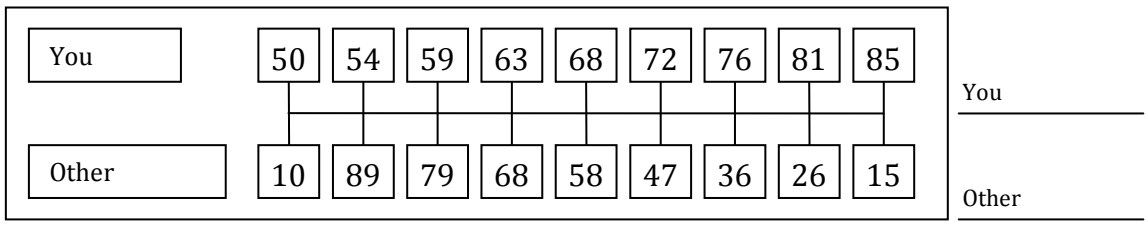
2)


You
85 87 89 91 93 94 96 98 10
Other
15 19 24 28 33 37 41 46 50
You
Other

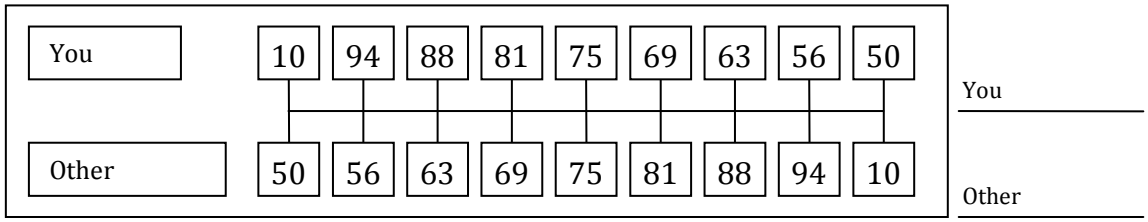
3)


You
50 54 59 63 68 72 76 81 85
Other
10 98 96 94 93 91 89 87 85
You
Other

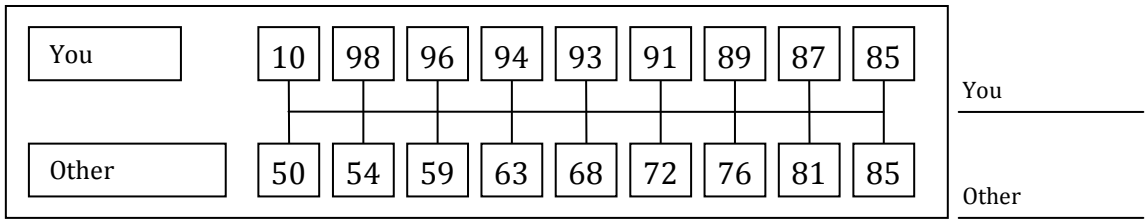
4)


You
50 54 59 63 68 72 76 81 85
Other
10 89 79 68 58 47 36 26 15
You
Other

5)


You
10 94 88 81 75 69 63 56 50
Other
50 56 63 69 75 81 88 94 10
You
Other

6)


You
10 98 96 94 93 91 89 87 85
Other
50 54 59 63 68 72 76 81 85
You
Other

Second Part

In this part of the experiment you will be presented photos on your screen. Your task will be to state attributes, with which you describe, what is depicted in the respective photo. An attribute can be located anywhere in the photo. E.g. also in the background or at the edge.

The described photos will be used for chair- and research-internal purposes.

You are asked to state at least one attribute per photo. You can state a maximum of 10 attributes, you see depicted in the respective photo.

If there are multiple kinds of one attribute to be seen in one photo, please state the respective attribute in **one row** with the respective count.

e.g. (correct):

1. **oneblackcow**
2. **threebrowncows**

instead of (wrong):

3. **blackcow**
4. **browncow**
5. **browncow**
6. **browncow**

Please state a **maximum of one** adjective per attribute:

e.g. one red rose

Concerning the input of the attributes:

Attributes have to be confirmed by pressing the "Enter" key. By doing so the attributes will be saved and an input field for a new attribute appears. If you do not want to enter a new attribute, please press the key "next photo".

If you request a new photo **before** pressing "Enter", the attribute will **not** be saved.

Before the actual start of this part, there will be a stage of testing. 3 photos will be presented to you, whereby you can familiarize yourself with the handling.

This part of the experiment will take 25 minutes. You will be paid 5 € for this part.

Please remember that communication with another participant is forbidden during this part of the experiment as well as during all other parts.

Third Part – Peer Treatment

In this part you will again be presented photos, and your task will again be to describe them.

The difference between this part and the former is that you will now be matched with another participant. This participant will be continuously informed about the number of photos described by you. You will also be continuously informed about the number of photos the other participant has described. You will find this information in the right top corner of your screen. Additionally, you will also see the number of photos you have described.

Before this part starts, you will be relocated so that you are sitting next to the participant who you are matched to.

This part of the experiment will take 25 minutes. You will be paid 5€ for this part.

Please remember that communication with another participant is forbidden during this part of the experiment as well as during all other parts.

Third Part – Control Treatment

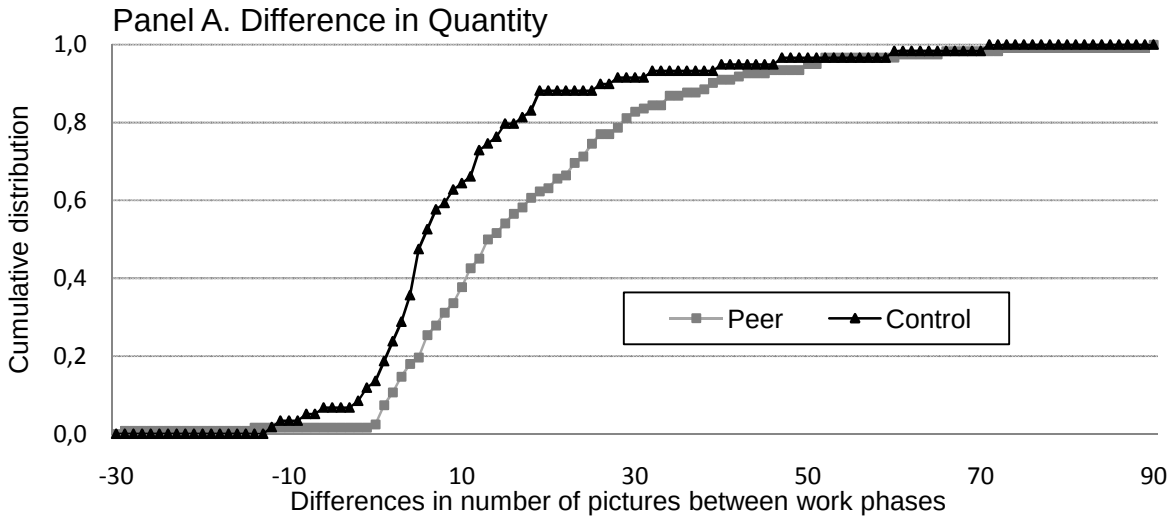
In this part you will again be presented photos, and your task will again be to describe them.

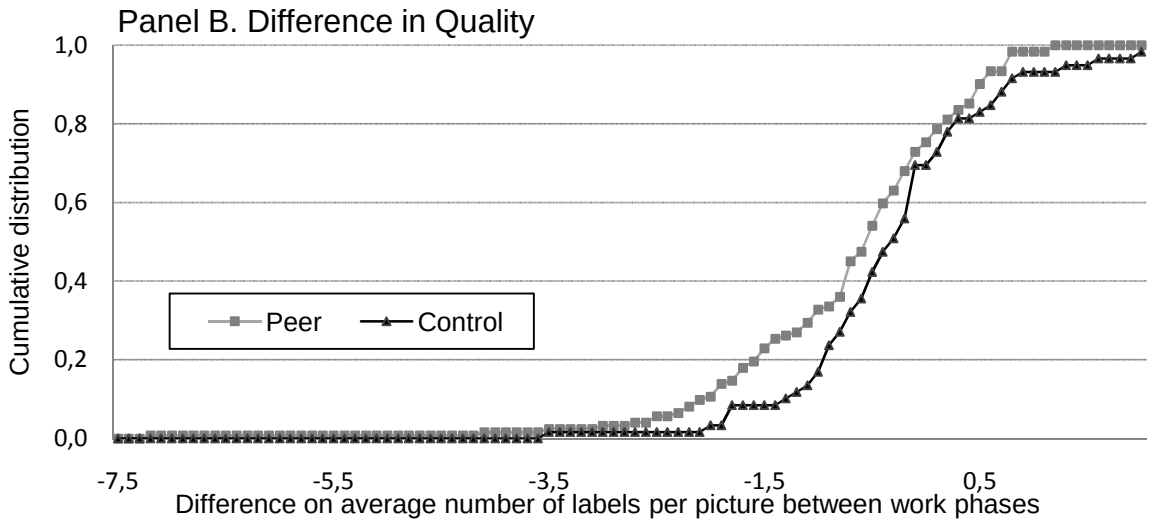
Before this part starts you will be relocated to a new cabin.

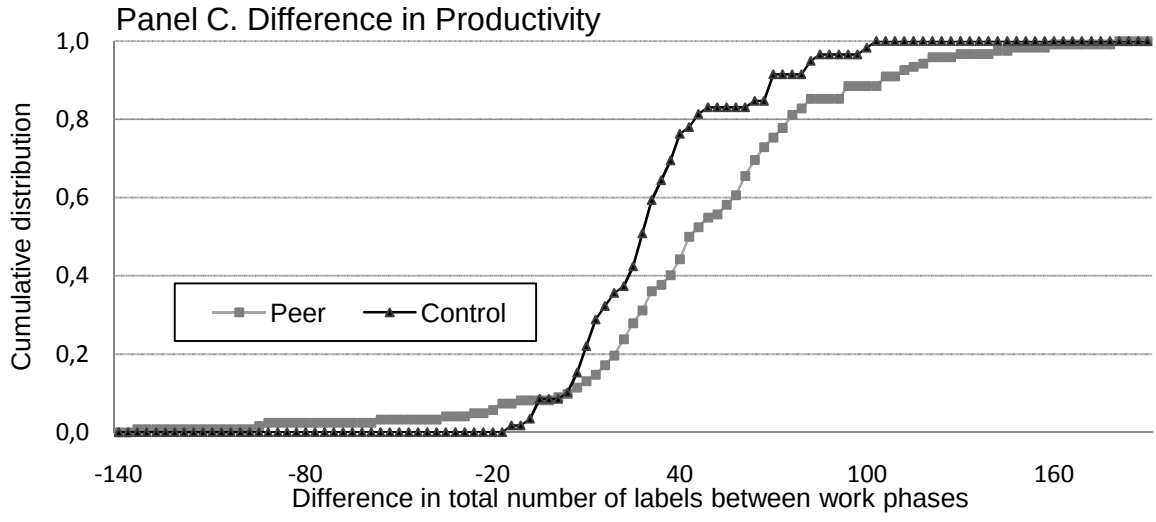
This part of the experiment will take 25 minutes. You will be paid 5€ for this part.

Please remember that communication with another participant is forbidden during this part of the experiment as well as during all other parts.

Appendix B. Cumulative distribution functions of changes in quality, quantity and productivity from the no-feedback to the feedback phase







Appendix C. Mean changes for different workers in different group types

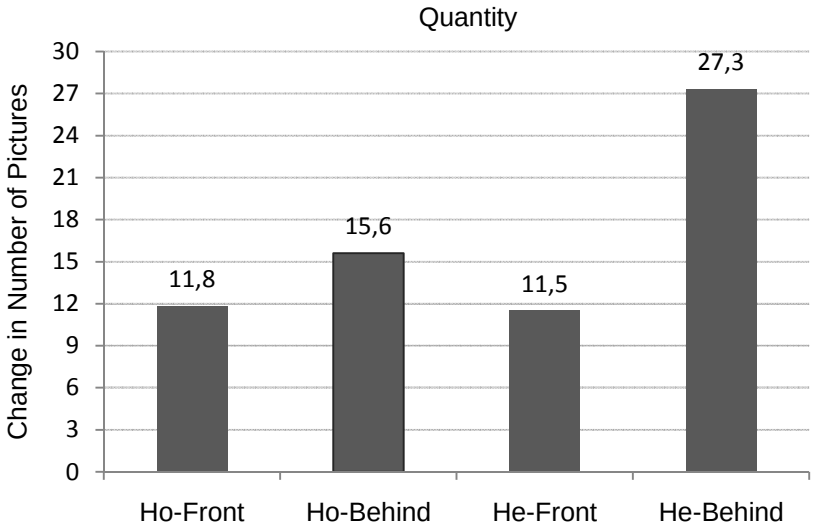


FIGURE 4. QUANTITY MEAN CHANGE FROM NO-FEEDBACK TO FEEDBACK PHASE

Notes: The figure shows for different worker types the mean change of quantity (=number of described pictures). "Ho-Front" represents workers with positive heterogeneity values in homogeneous groups, "Ho-Behind" represents workers with negative heteroneity values in homogeneous groups, "He-Front" represents workers with positive heterogeneity values in heterogeneous groups, "He-Behind" represents workers with negative heterogeneity values in heterogeneous groups.

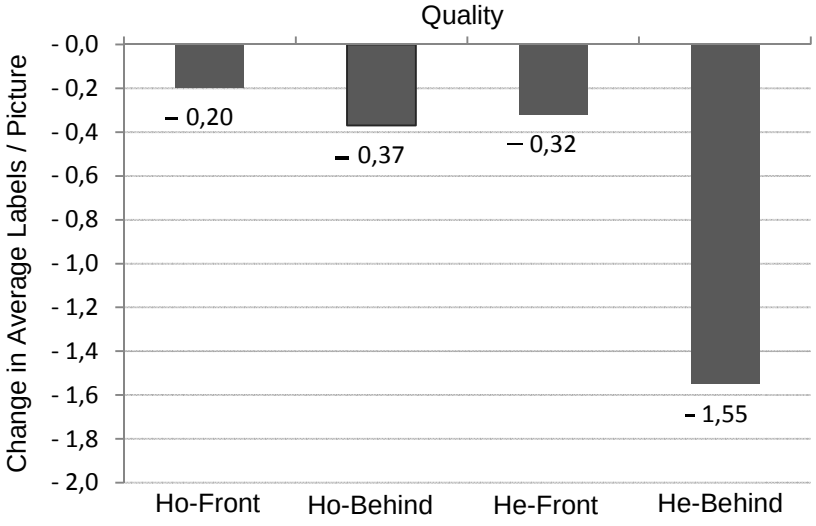


FIGURE 5. QUALITY MEAN CHANGE FROM NO-FEEDBACK TO FEEDBACK PHASE

Notes: The figure shows, for different worker types the mean change of quality (= average labels per picture).

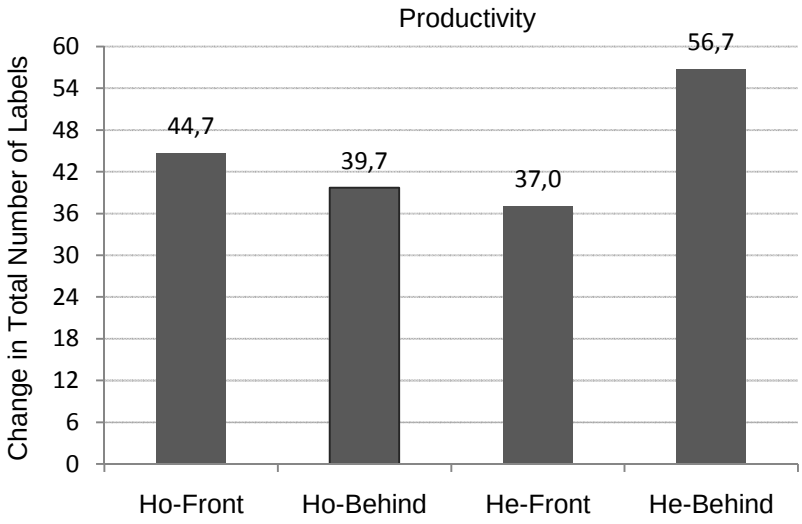


FIGURE 6.PRODUCTIVITY MEAN CHANGE FROM NO-FEEDBACK TO FEEDBACK PHASE

Notes: The figure shows, for different worker types the mean change of productivity (= total numbers of labels).