UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 659

Because of you I did not give up - How peers affect perseverance

Leoni Gerhards and Christina Gravert

June 2016

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

Because of you I did not give up - How peers affect perseverance*

Leonie Gerhards† [1] and Christina Gravert‡ [2]

[1]University of Hamburg
[2]University of Gothenburg

May 2016

Abstract

Various empirical papers have shown that peers affect productivity and behavior in the workplace. However, the mechanisms through which peers influence each other are still largely unknown. In this laboratory experiment we study a situation in which individuals might look at their peers' behavior to motivate themselves to endure in a task that requires perseverance. We test the impact of unidirectional peer effects under individual monetary incentives, controlling for ability and tactics. We find that peers significantly increase their observers' perseverance, while knowing about being observed does not significantly affect behavior. In a second experiment we investigate the motives to self-select into the role of an observing or an observant subject and what kind of peers individuals deliberately choose. Our findings provide first insights on the perception of peer situations by individuals and new empirical evidence on how peer groups emerge.

Keywords: grit, perseverance, laboratory experiment, peer effects, real effort
JEL codes: C91, D03, M50, J24

*We are grateful for helpful comments from seminar and conference participants at the Universities of Aarhus, Amsterdam, Copenhagen, Gothenburg, Hamburg, Lüneburg, Munich, Nottingham, Trier, the Copenhagen Network of Experimental Economics, the Meeting of German Economists Abroad 2014, the Nordic Conference in Behavioral and Experimental Economics 2015, the Colloquium on Personnel Economics 2016 and in particular Florian Englmaier, Gerd Mühlheußer and Uwe Sunde. This research was supported by funding from the Danish Council for Independent Research — Social Sciences (Det Frie Forskningsråd — Samfund og Erhverv), grant number 12-124835 and from the Swedish Research Council (Vetenskapsrådet) (2009-33472-67049-56).

†Department of Economics, University of Hamburg, Von-Melle-Park 5, 20146-Hamburg, Germany. Tel.: +49 40 42838 5573. Email: leonie.gerhards@wiso.uni-hamburg.de

‡Department of Economics, University of Gothenburg, Vasagatan 1, S-41124 Gothenburg, Sweden; christina.gravert@economics.gu.se, Tel.: +46 (0)31 786 5252

1 Introduction

Peers affect productivity and behavior in the workplace. Both theoretical papers (for instance Kandel & Lazear (1992)) as well as several empirical studies (such as the papers by Falk & Ichino (2006), Mas & Moretti (2009), Bandiera et al. (2010) and Baeker & Mechtel (2013)) show that individuals are influenced by others working besides them. However, the exact channels through which individuals influence each other are still largely unexplored (Charness & Kuhn 2011). Peer effects could, for example, stem from imitation of behavior (Bandura & Walters 1963, Mobius & Rosenblat 2014), an increase in knowledge about the task (Banerjee 1992, Guryan et al. 2009), competitive preferences (O'Keeffe et al. 1984), guilt and shame (Kandel & Lazear 1992), a desire for conformity (Bernheim 1994, Bellemare et al. 2010), self-perception in comparison to others (Deci & Ryan 1985) or the pure joy of working with others (Bandiera et al. 2010).

In this paper we apply a novel experimental design to study one isolated channel of peer effects: We analyze if and how peers affect each others' perseverance on a task. By focusing on perseverance, our paper also contributes to the emerging economic literature on non-cognitive skills. Recent studies in the fields of psychology and economics have shown that perseverance over a longer period of time – an attribute also known as grit – is an important determinant for various life outcomes. It has been shown that grit elicited from questionnaires is a reliable predictor for pupils' success in the national spelling bee, retention rates of West Point Military Academy cadets and Grade Point Averages (GPA) among Ivy League undergraduates (Duckworth et al. 2007, Eskreis-Winkler et al. 2014). Gerhards & Gravert (2015) introduce the first experimental measure of grit in a short-term laboratory real effort task and show that more gritty behavior leads to higher earnings in the experiment irrespective of subjects' initial ability. In the present paper we extend this work and combine the elicitation of grit with a peer setting in order to study how gritty behavior is influenced by the information about a peer's perseverance. This is an important question since up to date there has been no consensus on whether perseverance in the face of challenges is an inherent trait or a malleable attribute that can be learned (Alan et al. 2015). If it is an inherent trait then we should not find any peer effects in perseverance. If however, it is something that can be learned or at least stimulated by a peer, then we should see a difference in perseverance between an isolated and a peer situation.

We base our research design on the following idea: While workers might not know their co-workers' initial abilities and cannot assess the quality of the output before the termination of a project, they often have a good estimate of how much effort their colleagues exert and how long they remain working on the task at hand. Consider for instance graduate students working in the same office: They can observe on a daily basis when their peers come and go and whether they are working on their research projects or browsing the internet. But different fields of research and methods might make it difficult to compare the peers' progresses and outputs to their own. Nevertheless, observing the other students working longer and more diligently likely affects a student's perseverance on his or her own project.

There exists a wide range of other situations, for instance in sports, social and educational

contexts, in which perseverance plays an important role and diligent comrades, moreover, affect their peers. From a research perspective, however, these field settings might be confounded by several factors that cannot be controlled for. Examples are (self-)selection based on homophilistic preferences (which can cause spurious correlations (Manski 1993)), "reflection problems" of peers simultaneously influencing each other or complementarities between individuals in case their work tasks or payoffs are interrelated. We therefore decided to conduct a controlled laboratory experiment that excludes these and other confounding factors. Our subjects are randomly assigned to treatments and work on a word play task in which they solve anagrams on an individual piece rate basis. As a way to avoid working hard, subjects can choose to solve easier anagrams or to skip individual hard anagrams. Both avoidance behaviors come at an explicit cost. We compare the subjects' performance and amount of avoidance behavior when working by themselves ("Baseline Treatment") to a peer situation ("Random Matching Treatment") where half of the subjects are randomly assigned the role of what we call an "Observer" who is informed about his or her randomly matched "Peer's" avoidance behavior. The Peers know that they are observed by another subject, but do not receive any further information about their Observer.

We find, first, that Observers (compared to subjects from the Baseline Treatment) significantly reduce their avoidance behavior if they are presented with information about their Peers' respective behavior. Peers' and Baseline Treatment subjects' avoidance behavior, on the other hand, does not differ significantly. Second, personal attributes such as ability and self-reported grit play less of a role in explaining the avoidance behavior of subjects from the Random Matching Treatment than in the Baseline Treatment. And third, the Observers' inclination to switch to easy anagrams is significantly positively correlated with their Peer's switching behavior. Interestingly, this effect is even stronger if the Peer is of similar or worse ability.

Taken as a whole, we conclude that staying perseverant on a task does not only depend on individuals' own grittiness, but hinges on the peer group composition at hand. In the conclusion we provide a range of domains in which we expect these peer effects in perseverance to affect peoples' performance and discuss resulting policy implications

Related literature: A handful of other papers have looked at peer effects and performance in laboratory settings. The first three papers that we will briefly describe below consider peer effects in subjects' performance and output in general. The forth paper focuses on peer effects in perseverance in particular.[1]

Georganas et al. (2015) randomly ask their subjects to work on Gill & Prowse's (2012) slider task and assign them into three groups. Subjects in the first group are observed by another subject, those from the second group observe another subject and subjects in the third group work

[1]Furthermore, a number of field experiments study the effect of related personal attributes such as self-control on real-life outcomes (see for instance Buechel et al. (2014) who link self-control to academic success and Kaur et al. (2010) who link it to success at work). Also, a number of laboratory experiments investigate the determinants of shirking in real-effort tasks (see for instance Corgnet et al. (2015) on diligence in summing up numbers or Bucciol et al. (2013) on the ability to withstand temptation in a counting task) and the effects of being observed by group members (see for example Charness et al. (2007) and Andreoni & Bernheim (2009)).

in isolation. The authors find no overall peer effects on performance in the first two groups when subjects work under individual piece rate incentives. They conjecture that this could be explained by a significant and dominant learning effect that conceals potential treatment effects.[2] Beugnot et al. (2013) test whether the strength of peer effects depends on the actual presence of a peer. In one of their treatments subjects are informed about a peer's outcome who has worked on the same real effort task (a math task) in an earlier session. In a second treatment subjects are continuously informed about the performance of a peer who is simultaneously working in the same room. In both treatments the authors find significant peer effects for men, but not for women. They conclude that the observed effects are likely driven by rivalry between peers, which they find to be stronger for men than for women. Van Veldhuizen et al. (2015) set out to replicate the main features and findings of Mas & Moretti's (2009) seminal field study in the lab. Groups of four students work in a team to solve a predetermined number of mathematical calculations. By design, the subjects' payoffs do not depend on the other participants, but their workload does. Depending on their role in the team, subjects can either observe others, are observed by others, do both things simultaneously or cannot do either. In contrast to Mas & Moretti (2009), the authors find no difference between being observed and observing a coworker. They do, however, discover heterogeneous peer effects in productivity. When subjects are aware of the general productivity level of their peer, some subjects reciprocate a fast co-worker, while others free-ride on the faster co-worker's effort.

Lastly, similar to our study, Bonein & Denant-Bomont (2015) study individuals' perseverance – more precisely their self-control – in the presence of a peer. In their experiment subjects work on the slider task and can commit themselves to an output level *ex ante*. The authors find that subjects choose higher penalties for not reaching their goal if they know that their output will be reported to another participant with whom they interact in a later stage of the experiment. However, they do not find any effects on increased perseverance on the actual task. Besides the apparent differences between their and our experimental setups, it should moreover be noted that Bonein & Denant-Bomont (2015) measure perseverance as a zero/one decision. Once a subject decides to give in to temptation he or she could not return to the experiment. In a sense, their perseverance measure is hence more strict than our measures of giving up in the form of skipping or switching to easy anagrams, which can happen multiple times during an experimental session. Furthermore, compared to the studies by Georganas et al. (2015), Beugnot et al. (2013) and van Veldhuizen et al. (2015), our design allows to analyze peer effects in perseverance in isolation, keeping biasing learning effects, competitive motives and interdependecies between subjects at a minimum. First, we restrict our analysis of treatment effects to the final part of the experimental sessions when further learning can largely be ruled out. Second, we inform Observers only about their Peers' avoidance behavior and not about their performance or output. And third, we consider performance in individual tasks only.

The remainder of the paper proceeds as follows: Section 2 presents a simple model of peer

[2]In a team incentive treatment they find that observed participants increase their productivity a bit faster than the control group. However, this group catches up to them within 3 minutes of the experiment. Possibly, being observed leads to slightly faster learning in this experiment.

effects in perseverance. In section 3 we turn to the description of our experimental design. We then discuss the results from our Baseline and Random Matching Treatment. In the last part of the paper, we widen our focus to the conscious decision of selecting peers. First, we analyze which personal characteristics predict a preference for being observed (i.e. self-selection into the Peer's role) rather than observing someone else (i.e. self-selection into the Observer's role). Second, we investigate which type of Peer Observers choose to observe given information about his previous performance. We provide a detailed description of the experimental design and a discussion of our findings from this additional "Chosen Matching Treatment" in section 4. The paper concludes in section 5.

2 Conceptual framework

We use the following simple 2-person model that is based on Ichino & Maggi (2000) to illustrate how peers can affect each others' perseverance on our experimental work task. Consider person i who chooses a level of perseverance $P_i > 0$ in order to maximize his utility $U_i = G(P_i, \theta_i) - C(P_i, P_j)$. The first term of the utility function captures person i's gain from staying perseverant on hard anagrams. It depends both on P_i and θ_i. As we will discuss in detail below, in our experiment we will measure P_i indirectly as i's decision not to skip individual hard anagrams and not to switch to easy anagrams. The individual preference parameter θ_i could for instance reflect task ability or grit. We assume the gain term to be concave in perseverance ($\partial G/\partial P_i > 0$ and $\partial^2 G/\partial P_i^2 < 0$) and that higher values of θ_i indicate a person with a higher marginal gain from perseverance P_i ($\partial^2 G/\partial P_i \partial \theta_i > 0$). The possibility of peer effects is introduced with the second term. Person i incurs costs that depend both on his own perseverance P_i and on his peer's perseverance P_j. If $\partial^2 C/\partial P_i \partial P_j \leq 0$, i's costs of perseverance decreases in the peer's perseverance P_j. That is, it may be costly for person i not to keep up with person j's perseverance. In the absence of peer effects, on the contrary, we have $\partial^2 C/\partial P_i \partial P_j = 0$.

Each person chooses P_i to maximize his individual utility U_i. Assuming, for simplicity, a linear relationship, we can express the optimal level of perseverance as $P_i = \theta_i + \beta P_j$ where the weighting factor β measures the strength of peer effects. Positive peer effects are reflected in $\beta > 0$, $\beta < 0$ represent negative peer effects and $\beta = 0$ implies absence of peer effects.

As we will describe in detail below, in our experiment only Observers from the Random Matching Treatment are informed about another person's perseverance. Therefore only the Observers' perseverance will be a function of θ_i and P_j as described above and given by $P_i^O = \theta_i + \beta \theta_j$. On the contrary, Peers from the Random Matching Treatment as well as subjects from the Baseline Treatment will choose their optimal level of perseverance only based on their individual preferences: $P_i^I = \theta_i$. Obviously, $P_i^O > P_i^I$ for $\beta > 0$. That is, if positive peer effects exists, Observers should stay more perseverant on the experimental task than Peers and subjects in the Baseline Treatment.

Furthermore, it should be noted that in the above model P_i^O does not depend on the Peer's level of perseverance P_j. That is, the mere presence of the Peer should improve the Observer's

perseverance compared to that of subjects from the Baseline Treatment. The socio-psychological literature discusses this effect as the so-called social facilitation paradigm (Zajonc 1965).

3 Peer effects

3.1 Experimental design

During the experimental sessions subjects work on an anagram word play task in which they have to rearrange the letters of English words to form new ones. As an example consider the word "top" that can be rearranged to "pot" and "opt". We accept all possible anagrams that can be build from a word as a correct solution and offer two levels of difficulty. "Easy anagrams" consist of 3 to 4 letters, "hard anagrams" comprise 5 to 7 letters.[3] As discussed in detail in Gerhards & Gravert (2015) the anagram task is perceived as rather challenging and does not favor any of the sexes in the sense that men's and women's performance does not differ significantly.[4] Irrespective of treatment, all subjects are presented with the same anagrams in the same order.

The experiment starts with a five minute practice round in which subjects are only asked to solve easy anagrams. Performance in this part is not monetarily incentivized. It allows subjects to familiarize themselves with the experimental task.

In the main part of the experiment subjects are paid according to their performance and are free to choose between solving hard and easy anagrams. We consider working on hard anagrams as the main task, working on easy anagrams serves as an outside option. This is reflected in both the framing of the instructions and in the monetary incentives attached to them. For each correctly solved hard anagram subjects earn DKK 5.00. For each solved easy anagram subjects earn DKK 0.50. If subjects decide to switch to easy anagrams at the beginning of a new round, they have to bear a cost of DKK 3.00. The total cost of switching to easy anagrams for one round hence consists of the explicit switching cost and the implicit cost of reduced earning opportunities.

Each anagram is presented for up to 90 seconds. If a subject enters a correct solution, a new anagram is displayed immediately. If a subject does not manage to solve an anagram within the given time frame, a new anagram is generated free of cost. When working on hard anagrams, subjects are, moreover, given the opportunity to "skip" individual anagrams. Then they do not have to wait until the end of the 90 seconds, but a new anagram is generated immediately. Just like switching to easy anagrams, also each skip comes at a cost of DKK 3.00. In the following analysis, we consider switching to easy anagrams and excessive skipping of hard anagrams as manifestations of non-perseverance.

Table 1 summarizes the monetary incentives. We chose this incentive structure to make switching to easy anagrams clearly monetarily unattractive. Subjects should be able to see without any

[3] An example for a hard anagram is the word "teachers" that can for instance be rearranged to "hectares".

[4] The median "challenge" assessment of participants from the Baseline Treatment was 4 out of 5 points. Moreover, there are neither significant gender difference in the number of easy anagrams solved in the practice period (Mann-Whitney ranksum test result: $p = 0.84$), nor in the number of hard anagrams solved per hard rounds in the first ($p = 0.68$) or second part of the experiment ($p = 0.33$) in the Baseline Treatment.

Table 1: Monetary incentives in DKK

Action	Payoff
Solving a hard anagram	5.00
Solving an easy anagram	0.50
Skipping a hard anagram	−3.00
Switching to easy anagrams	−3.00

DKK 1.00 corresponds to approximately EUR 0.13 or US-Dollar 0.18.

formal calculation that even for less able individuals switching to easy anagrams is not a monetarily optimal decision. Their choice can hence only be rationalized by a strong desire to avoid working hard.[5]

The main part of the experiment is divided into two identical parts of 30 minutes each. Part 1 mainly serves us to measure subjects' tactics on the task. It is defined as the subjects' average earnings in rounds in which they work on hard anagrams and reflects their ability to find their payoff maximizing mixture of solving and skipping hard anagrams: Some subjects might obtain a given payoff by skipping and solving many anagrams, while others generate the same payoff by skipping and solving fewer anagrams. Part 2 varies across treatments and is hence the working period of interest in our data analysis.

Both parts are partitioned into ten rounds of three minutes. In the first round of each part subjects have to work on hard anagrams. We chose this set-up in order to stress the default character of working on hard anagrams and to make sure that the subjects get to know the level of difficulty of hard anagrams. At the beginning of each of the following nine rounds subjects can choose to "stay with the hard anagrams" or to "switch to the easy anagrams" for the coming three minutes.[6] We hence allow subjects to go back to solving hard anagrams after having switched to easy anagrams in the previous round (and vice versa).

In both part 1 and part 2, an information box on the computer screen continuously informs subjects about how many hard anagrams they have already skipped since the beginning of the current part and how many hard anagrams they have skipped in the current round. In rounds in which subjects work on easy anagrams, no such information is displayed.

At the end of part 1 subjects receive feedback about their productivity on hard anagrams and payoffs in that part. Furthermore, we ask them without previous announcement to make an incentivized guess which performance quintile they belong to.[7] In particular, we ask them to make this guess based on their number of correctly solved hard anagrams. We reward the correct answer with DKK 20.00. After subjects enter their guess they are informed about their actual performance

[5]Indeed, for only 2 out of our 152 subjects from the Baseline and Random Matching Treatment we find that their lowest round earnings on easy anagrams in part 1 lay above their highest round earnings on hard anagrams (-3.00 vs. -2.50 and -1.00 vs. 0).

[6]See https://www.dropbox.com/s/qq838eaioo4yc3w/InstructionsAndScreenshots.pdf?dl=0 for a screen shot of this decision stage and examples of other screens that were displayed during the experimental sessions.

[7]In one of our sessions only 6 subjects participated. Therefore we asked them to guess which performance tertile they belonged to instead.

quintile and part 2 starts.

Treatments: As mentioned above, the treatments are implemented in part 2 of the experimental sessions. In the Baseline Treatment, the instructions presented in the beginning of part 2 simply ask the subjects to continue working on the anagram task, as they did before in part 1.[8] Subjects in the Random Matching Treatment are presented a new set of instructions on their computer screens. These inform them that half of the subjects in their session will be randomly assigned the role of an Observer (in order to prevent framing effects called "Person A") and the other half will assume the role of a Peer (called "Person B"). Always one Observer and one Peer are randomly matched for part 2. Both are presented with the same anagrams in the same order.

Observers learn their Peer's performance quintile from part 1. Moreover and more importantly, Observers are informed about their Peer's skipping and switching behavior during part 2: When Observers work on hard anagrams their computer screens display the total number of hard anagrams the matched Peer has skipped since the beginning of that part and whether the Peer is working on hard or easy anagrams in the current round. This information is updated at the beginning of each round, that is, every three minutes. To prevent the emergence of so called "rat races" we do not provide them with further information on the number of correctly solved hard or easy anagrams. Peers receive no information about their Observer's performance, but are informed about the type of information that the Observer receives about them.

Questionnaire: After having finished working on the real-effort task we ask the subjects to fill out a short questionnaire that comprises two parts. In the task-specific part we ask them how they perceived working on the task. In the survey part we elicit a number of non-cognitive skills and personality traits through non-incenitivized survey questions. In particular, we administer the Short-Grit-Scale (Duckworth & Quinn 2009) which comprises eight items that are ranked on a 5 point scale and measure the tendency to sustain effort and maintaining interests over a long period of time. The full list of items is presented in Appendix A.

A valid concern when running questionnaires at the end of an experiment is that the behavior during and the outcome of the experiment might influence the answers given in the questionnaire. To address this point, we randomize the order of the experiment and the survey part of the questionnaire in the Baseline Treatment at the session level. The task-related questions are always asked at the end of the experiment. As discussed in Gerhards & Gravert (2015), the sequence of real effort task and survey questions neither significantly affects the answers given in the questionnaire, nor the performance on the task.

Procedures: We ran 16 experimental sessions in spring and summer 2014 at the Cognition and Behavior (Cobe) Lab at Aarhus University. Subjects were recruited via the laboratory's online recruiting website from a subject pool of mostly undergraduate students from all faculties. In total

[8]Note that the Baseline Treatment presented in this paper corresponds to the data set that was use to analyze further aspects in the companion paper Gerhards & Gravert (2015).

152 subjects participated, out of which 69 (45 percent) were female. We observe 62 subjects in the Baseline Treatment and 90 subjects in the Random Matching Treatment (45 Observers, 45 Peers). Most of the subjects in our sample are Danish undergraduates. The participants had on average 9 years of English at school.

At the beginning of each session consent forms and detailed instructions about the experimental task were distributed to the subjects.[9] The subjects had 10 minutes to read the instructions. Afterwards the experimenter asked if there were any questions or if anyone needed more time to read. When all subjects denied, the experiment (programmed in z-Tree (Fischbacher 2007)) started and all further instructions were provided on the computer screens. At the end of each session either part 1 or part 2 was randomly selected for payments. Average earnings for the 90 minutes sessions amounted to DKK 120.00, including the reward for a correct guess about the subject's own performance rank and a DKK 40.00 show-up fee. The payments were directly transferred to the subjects' bank accounts.

3.2 Results

In the following statistical analysis we will consider the four main outcome variables from part 2: subjects' round earnings from working on hard or easy anagrams, their number of solved hard anagrams per round in which they worked on hard, their number of skipped hard anagrams and their tendency to switch to easy anagrams. One can well consider the first variable a consolidated measure of the three latter ones: The number of solved hard anagrams increases subjects' earnings, while skipping and switching reduces them.

Figure 1 gives a first overview of how the means of our outcome variables evolve over the ten rounds in the Baseline and Random Matching Treatments. In general, the graphs paint a rather positive picture of peer effects: The upper left graph suggests that Peers and Observers in the Random Matching Treatment realize higher mean earnings than their counterparts from the Baseline Treatment who work in isolation. This seems to result from the fact that Random Matching subjects, on the one hand, solve more hard anagrams (see the upper right graph) and, on the other hand, display less avoidance behavior (see the two lower graphs). In particular Observers skip and Peers switch on average less often than subjects from the Baseline Treatment.

In the following regression analysis we are able to control for subject-specific abilities and tactics and go into more depth with regards to the effects of being in a peer situation. We begin by comparing Observers' and Peers' behavior to that of subjects from the Baseline Treatment. Subsequently, we move on to study potential heterogeneous treatment effects. Lastly, we will consider the question of how Observers are affected by specific attributes of their matched Peers.

[9] All instructions and screenshots of the experimental program can be downloaded from https://www.dropbox.com/s/qq838eaioo4yc3w/InstructionsAndScreenshots.pdf?dl=0

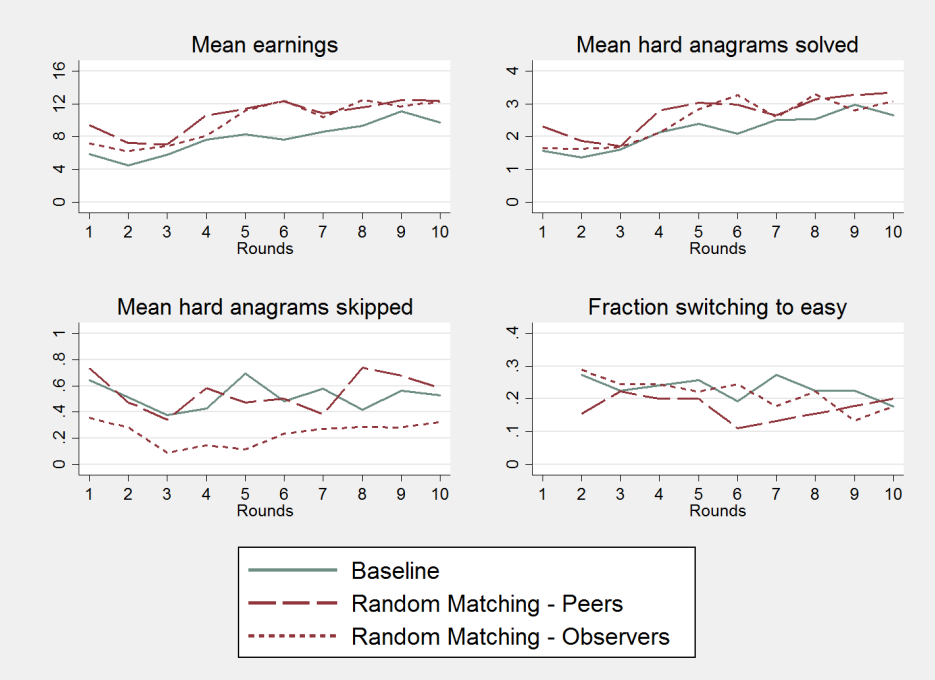

Figure 1: Outcome variables of interest across treatments

3.2.1 Observers react to their Peers, but Peers don't seem to care.

Using the pooled OLS models (1) to (3) and the pooled Probit Model (4) in Table 2, we test for general treatment effects and in particular role effects. We regress our four outcome variables on role dummies that take the value 1 for subjects in the Observer and Peer role, respectively. That is, we treat the subjects from the Baseline Treatment as reference group. Moreover, in all four specifications we control for subjects' tactics and task ability.

We define tactics as the average round earnings that subjects realize when working on hard anagrams in part 1. These earnings depend on the number of correctly solved and skipped anagrams per round worked on hard. They indicate the subjects' capability to find their individual payoff-optimal mix of skipping and solving anagrams, which may well vary across subjects. Some subjects might skip individual hard anagrams that they cannot solve in reasonable time, deliberately accepting the costs of DKK 3.00, in order to try their luck on the next word that might yield them a payoff of DKK 5.00. Other subjects might refrain from skipping entirely in order not to reduce their earnings. Our measure allows us to control for these individual tactics in a unified way.

Ability is defined as the number of solved easy anagrams in the unincentivized practice period. Ideally, we would have preferred to define ability analogously to the tactics measure, that is, as the average number of correctly solved anagrams when working on hard anagrams in part 1. However, this measure is by definition correlated with our tactics variable (Spearman's $\rho = .87$, $p < .01$) and would have lead to severe multicollinearity problems in our regressions. Therefore, we rely on performance in the practice period as our measure for ability. It is positively correlated with the

Table 2: Observers and Peers compared to Baseline Treatment subjects

	Earnings	Hard anagrams... ... solved	... skipped	Switch to easy
	(1)	(2)	(3)	(4)
Observer	-0.039	-0.236	-0.393***	0.067
	(0.829)	(0.213)	(0.143)	(0.234)
Peer	-0.267	-0.055	-0.094	-0.009
	(0.908)	(0.225)	(0.163)	(0.236)
Constant	7.380***	1.895***	0.729***	-1.064***
	(0.731)	(0.169)	(0.133)	(0.177)
Tactics	Yes	Yes	Yes	Yes
Ability	Yes	Yes	Yes	Yes
R^2	0.406	0.476	0.158	
Pseudo R^2				0.128
Independent observations	152	152	152	152
Observations	1520	1232	1232	1368

Regression 1: pooled OLS, dependent variable: Earnings; Regression 2: pooled OLS, dependent variable: Number of solved hard anagrams (given worked on hard); Regression 3: pooled OLS, dependent variable: Number of of skipped anagrams in hard; Regression 4: pooled Probit, dependent variable: Decision to switch to easy anagrams. Note that all continuous independent variables are standardized. Robust standard errors are clustered at the subject level and given in parentheses: * p<0.10 ** p<0.05 *** p<0.01. All regressions include round dummies.

average number of correctly solved hard anagrams ($\rho = .50$, $p < .01$) and comparably less strongly correlated with our tactics variable ($\rho = .43$, $p < .01$).

As becomes evident from Table 2, most of the role differences in Figure 1 are not statistically significant once one includes the necessary controls for subjects' tactics and ability. Subjects earnings, their number of solved hard anagrams and their propensity to switch to easy anagrams do not differ across roles and hence across treatments. The coefficient of the Peer dummy is in neither of the regressions significant. We do, however, find that Observers display less avoidance behavior by skipping significantly fewer hard anagrams than subjects from the Baseline Treatment. We summarize our findings as follows:

Result 1 *Being able to observe another subject significantly increases the Observers' perseverance when working on hard anagrams. In contrast, being aware of the fact that another subject observes one's actions does not significantly affect Peers' performance.*

This finding is in line with our model for positive peer effects from section 2. It is, however, in contrast to both Mas & Moretti (2009) who find stronger effects for Peers than for Observers, and to van Veldhuizen et al. (2015) who find that there is no difference in the behavior of Peers and Observers. A potential explanation lies in the fact that Observers and Peers in our study are not dependent on each other – neither in their workload, nor in their payoff. It thus seems highly

unlikely that Peers work hard out of guilt or fear of upsetting their team, which might be a driving factor in the other two experiments.

Another noteworthy aspect is the fact that we find significant overall peer effects in skipping, but not in switching. We suspect that the high explicit and implicit costs of switching to easy anagrams reduce subjects' tendency to switch in general, leaving less room for peers to have an effect. Nonetheless, we continue to consider peer effects in both avoidance behaviors – skipping and switching – in the following steps of our analysis, since the overall null effects might conceal noteworthy heterogeneous treatment effects.

3.2.2 Successful subjects perform even better in a peer situation

The question that arises from Result 1 is whether the average peer effect depends on specific attributes of the Observers. In particular, is the positive effect on perseverance driven by high skilled Observers who get more motivated by being in a peer situation – or rather by low skilled individuals who imitate their Peers' tactics? To answer this question we will consider interaction effects between individuals' task ability and their role as well as between their tactics on hard anagrams and their role.

Moreover, we will tackle the question of whether subjects can learn from each other to behave more perseverant conditional on their grit level that we elicit in the final questionnaire using Duckworth & Quinn's (2009) scale. This issue is of particular interest since Gerhards & Gravert (2015) find that self-reported grit is significantly correlated with a lower tendency to skip hard anagrams if subjects work in isolation. It is an open question whether this similarly holds in situations in which peers are present.

The regressions reported in Table 3 build on the regressions reported in Table 2. Our dependent variables are again earnings (Model (1)), the number of hard anagrams solved (Model (2)), the number of skipped hard anagrams (Model (3)) and the decision to switch to easy anagrams (Model (4)). We regress these outcome measures on role dummies for Observers and Peers (that is, subjects from the Baseline Treatment are again treated as the reference group to both roles), our measures of tactics and task ability and subjects' self-reported grit. Moreover, we include interaction terms of these personal attributes and the role dummies. All continuous variables are standardized in order to facilitate the interpretation of the main and interaction effects.

When it comes to earnings, the positive and significant tactics main effect in Model (1) reveals that a successful tactic in part 1 increases Baseline subjects' earnings in part 2 of the experiment. Interestingly, this positive tactics effect is even stronger for Observers, as indicated by the corresponding positive and highly significant interaction term. Similarly, Observers benefit comparably more from their general task ability than Baseline subjects. Self-reported grit influences neither Baseline subjects', nor Observers' earnings significantly.

The findings on performance on hard anagrams from Model (2) largely replicate our result from Model (1). Successful tactics on hard anagrams positively affect the Baseline subjects' amount of correctly solved anagrams. Moreover, this effect is even stronger for Observers and Peers. We

Table 3: Heterogeneous treatment effects

		Hard anagrams...		
	Earnings	... solved	... skipped	Switch to easy
	(1)	(2)	(3)	(4)
Observer	0.370	-0.253	-0.353***	0.248
	(0.760)	(0.190)	(0.128)	(0.307)
Peer	0.494	-0.003	-0.161	0.210
	(0.763)	(0.195)	(0.153)	(0.304)
Tactics	4.695***	0.931***	0.312**	-1.398***
	(0.707)	(0.208)	(0.147)	(0.409)
Observer × Tactics	3.617***	0.804***	-0.233	0.722
	(1.189)	(0.269)	(0.171)	(0.487)
Peer × Tactics	1.638	0.697**	0.186	0.875*
	(1.017)	(0.301)	(0.240)	(0.512)
Ability	-0.265	0.166	-0.017	0.985***
	(0.667)	(0.188)	(0.130)	(0.210)
Observer × Ability	2.792*	0.413	0.173	-1.021***
	(1.466)	(0.343)	(0.182)	(0.293)
Peer × Ability	0.249	0.000	0.217	-0.592*
	(0.930)	(0.249)	(0.205)	(0.315)
Self-reported grit	0.178	-0.215	-0.370***	-0.008
	(0.492)	(0.155)	(0.134)	(0.189)
Observer × Self-reported grit	-0.043	0.061	0.348**	-0.077
	(0.765)	(0.204)	(0.148)	(0.254)
Peer × Self-reported grit	-0.905	0.161	0.488***	-0.012
	(0.672)	(0.186)	(0.161)	(0.244)
Constant	6.833***	1.822***	0.722***	-1.282***
	(0.658)	(0.145)	(0.126)	(0.266)
R^2	0.445	0.511	0.252	
Pseudo R^2				0.166
Independent observations	152	152	152	152
Observations	1520	1232	1232	1368

Regression 1: pooled OLS, dependent variable: Earnings; Regression 2: pooled OLS, dependent variable: Number of solved hard anagrams (given worked on hard); Regression 3: pooled OLS, dependent variable: Number of of skipped anagrams in hard; Regression 4: pooled Probit, dependent variable: Decision to switch to easy anagrams. Note that all continuous independent variables are standardized. Robust standard errors are clustered at the subject level and given in parentheses: * p<0.10 ** p<0.05 *** p<0.01. All regressions include round dummies.

summarize our findings below:

Result 2 *Being able to observe a Peer performing the same task reinforces the generally positive effects of successful tactics (and task ability) on earnings through an increase in the number of correctly solved hard anagrams.*

This result suggests that in our experiment in particular high performers benefit from being in a peer situation. Model (3) yields further interesting results. First, it reproduces our finding from Table 2: Being in the Observer role generally reduces subjects' amount of skipping. The respective "Observer" main effect is highly significant and the size of the coefficient remains largely unaffected by the inclusion of further control variables (compare the corresponding coefficient in Model (3) of Table 2). Moreover, the positive and significant "Tactics" main effect suggests that a certain amount of skipping can indeed be explained by money-maximizing, strategical consideration. And lastly, we find that Baseline subjects skip significantly less the more gritty they are. Remarkably however, this effect is significantly less pronounced for Observers and Peers. Apparently, these subjects from the Random Matching treatment skip fewer hard anagrams irrespective of their initial level of grit. This indicates that own grit and being in a peer situation can act as substitutes when it comes to reducing avoidance behavior.

Model (4) provides similar insights in the sense that also the Random Matching subjects' inclination to switch depends less on their personal characteristics than the Baseline subjects': In particular, we find that the Baseline subjects' tendency to switch to easy anagrams depends negatively on their tactics on hard anagrams and positively on their task ability (remember, we defined the ability variable as performance on easy anagrams). Observers' – and to a lesser extent also Peers' – switching decisions, on the other hand, depend significantly less on their ability. Taken together, we summarize our findings from Model (3) and (4) as follows:

Result 3 *Subjects' abilities and attributes play less of a role in peer situations in which they can observe others or are aware of the fact that their performance is reported to a peer.*

Ultimately, our findings from Model (3) suggest that being in a peer situation can work as a substitute for own grit. A potential reason for this could be that Observers become more perseverant by imitating the behavior of their Peer. An alternative explanation could be that the mere existence of the Peer leads Observers to become more perseverant on the task. In the following we will disentangle these two effects – that is the information effect resulting from observing the Peer's actual behavior and the social facilitation effect caused by the mere presence of the Peer.

3.2.3 Similar able Peers affect Observers' performance most

In the regressions presented in Table 4 we focus on potential determinants of Observers' avoidance behavior in the Random Matching Treatment.[10] As described in section 3.1 above, in the beginning

[10]Note that we can use only 42 out of our 45 Observers in the regressions in Table 4 since three Observers switched to easy anagrams in every single round in which this was possible. They hence never received information about their Peers' avoidance behavior.

of part 2 we inform the Observers about their own and their Peers' performance quintile from part 1. Furthermore, in rounds in which Observers work on hard anagrams, the computer screen continuously informs them about their Peer's avoidance behavior in part 2. In particular, Observers are informed whether their Peer works on hard or easy anagrams in the present round and how many hard anagrams the Peer has skipped until the previous round.[11] In the pooled OLS Model (1) we estimate the effect of the Peer's skipping information on the Observers' amount of skipping. In the pooled Probit Model (2) we study how the Peer's switching decision in the last round affects the Observers' decision to switch to easy anagrams in the present round. Since these information effects might well depend on the similarity (or dissimilarity) of the Observer and Peer, we add a further dummy variable that takes the value one if the Peer was from a higher performance quintile (i.e. solved more hard anagrams correctly in part 1) than the Observer. We additionally interact this "Peer from higher quintile" dummy with our information variables.

Table 4: Determinants of Observers' avoidance behavior

	(1) Skips in hard	(2) Switch to easy
Peer from higher quintile	-0.071	0.770*
	(0.176)	(0.464)
Peer's skips until present round	-0.072	
	(0.043)	
Peer's skips × Peer from higher quintile	0.149	
	(0.099)	
Peer switched in previous round		1.122**
		(0.514)
Peer switched × Peer from higher quintile		-1.842***
		(0.652)
Constant	0.270*	-2.019***
	(0.150)	(0.595)
Tactics	Yes	Yes
Ability	Yes	Yes
R^2	0.186	
Pseudo R^2		0.151
Independent observations	42	42
Observations	317	280

Regression 1: pooled OLS, dependent variable: Number of of skipped anagrams in hard; Regression 2: pooled Probit, dependent variable: Decision to switch to easy anagrams. Note that all continuous independent variables are standardized. Robust standard errors are clustered at the subject level and given in parentheses: * p<0.10 ** p<0.05 *** p<0.01. All regressions include dummy variables for sub-periods.

Model (1) reveals that neither the main effect of Peer's skipping in the previous rounds, nor the main effect of Peer's relative ability is significant. Neither of the two variables is able to significantly explain the Observer's tendency to skip hard anagrams. This finding, together with the significantly

[11] The instructions emphasized that we classify the decision to switch to easy anagrams like a skip of a hard anagram. The latter figure therefore also includes the number of switches.

negative "Observer" main effect from Model (3) in Table 3, suggests that Observers reduce their amount of skipping compared to Baseline subjects due to a general social facilitation effect (Zajonc 1965). Apparently, any kind of information about the Peer leads the Observer to skip significantly less often. The actual behavior of the Peer plays less of a role. The insignificant interaction term "Peer's skips × Peer from higher quintile" moreover suggests that this holds irrespective of the Peer's relative ability.

However, when we consider the Observers' tendency to switch to easy anagrams, that is, the more fundamental form of avoidance behavior, Model (2) implies that the Peer's behavior actually does influence the Observer's behavior. First, the positive "Peer switched in previous round" main effect indicates that if a similar or less able Peer switched (did not switch) in the previous round, this significantly positively affects the Observer's decision (not) to switch. Second, as indicated by the negative coefficient of the interaction term "Peer switched × Peer from higher quintile", the Peer's switching decision is significantly less correlated with the Observers' decision if the Peer is from a higher performance quintile. Lastly, and in a similar vein, the positive and significant main effect of "Peer from higher quintile" suggests that observing a better Peer staying with hard anagrams leads the Observer to switch to easy more often. We summarize our findings from Table 4 as follows:

Result 4 *The Observers' tendency to skip is neither significantly affected by their matched Peers' relative task ability, nor by their actual skipping behavior. Observers do, however, imitate their Peer's switching behavior if they are matched to a similar or less able Peer. If the Peer is out of reach ability-wise, his presence and his actual behavior have less of an effect on the Observer's switching decisions.*

4 Self-selection into the Observer and Peer Roles

The findings from our treatments raised two further research questions. First, since the Peer's impact on their Observer differs depending on their relative abilities, we became interested in the question of what kind of Peer Observers would like to observe if they had a choice. If subjects were able to correctly anticipate our findings from Result 4, we would expect them to choose a Peer of similar or slightly worse ability. Also according to the social comparison model by Falk & Knell (2004), which is based on the social comparison theory by Festinger (1954), Observers should on average choose Peers that are similar to them. This inference is deduced from balancing two competing motivations: On the one hand, an upwards comparison can inspire individuals to work harder (self-improvement). On the other hand, a downwards comparison can make them feel better about themselves (self-enhancement). Thus, as one's own ability rises, the reference level for social comparison rises as well in order to stay close to one's own level. Battaglini et al. (2005) extend the ideas of Falk & Knell (2004) and consider social comparison in situations that require perseverance in particular. They predict that the ideal peer is someone who has a slightly worse self-control problem than oneself, because this makes the peer's successes more encouraging and his failures

less discouraging. Based on these theoretical models and our previous findings we conjecture that Observers in our experiment choose on average Peers who are equally or slightly less able than themselves.

Second, we investigate more generally which factors and personal characteristics are able to explain a desire to voluntarily self-select into the Observer and Peer role. The questionnaire data from the Random Matching Treatment provide a first hint. Since we exogenously and randomly assigned Observers and Peers to their roles in this treatment, this data allows us to cleanly identify potential differences in task perception depending on role. In particular, we asked all subjects from the Random Matching Treatment (1) how much they enjoyed working on the task, (2) how challenging, (3) how exhausting they perceived the task and, (4) how much effort they provided during the task. Answers were given on 5-point scales. At the aggregate, ranksum tests do not reveal significant differences between the answers provided by Observers and Peers ($p = .43$, $p = .25$, $p = .47$ and $p = .25$, respectively). But the picture changes if we restrict the group of Observers to those 22 subjects who were matched to a Peer from a higher performance quintile. These Observers report to having enjoyed the task significantly less ($p = .07$) and having perceived the task as a greater challenge ($p = .03$) than the 45 Peers. From this finding one could conjecture that subjects in the Chosen Matching Treatment will exhibit a tendency to avoid the Observer role – in particular if they expect to observe a Peer from a better performance quintile, for instance, due to their own rather low ability. However, one could also argue that subjects dislike the position in which they are observed by others and hence rather avoid the Peer role. The fact that only avoidance behavior and not positively connoted productivity is reported to the Observer might reinforce this effect. It is hence an empirical question which of the two effects prevails, i.e. which role is preferred on average.

In order to tackle the questions of peer selection and role choice, we designed a third treatment, the Chosen Matching Treatment, that we will describe in the following.

4.1 Experimental Design: Chosen Matching Treatment

In the summer and fall of 2014 we invited 90 additional subjects to the lab using the same recruiting procedure as the Baseline and Random Matching Treatment. During the experimental sessions the only difference to the Random Matching Treatment is the intervention that we introduce between part 1 and part 2: After guessing and being informed about their performance quintile in part 1 all subjects have to make two choices for part 2. First, we ask them to indicate whether they want to assume the role of an Observer or that of a Peer (labeled Person A and Person B, similar to the Random Matching Treatment). We inform subjects that the computer allocates the desired roles under the premise that within each performance quintile always one of the subjects assumes either role. This procedure is strategy-proof, as the subjects' best strategy is still to state their preferences truthfully, and yet it gives us the flexibility to allocate the experimental roles in an

efficient way.[12]

Subsequently, we ask the subjects to state from which performance quintile they want to observe a Peer in case they are assigned the role of an Observer. Using these procedures it is possible that more than one Observer is matched to one Peer. For all participants who actually assume the role of a Peer we elicit their belief on somebody choosing to observe them or not. However, at this stage the Peers are not informed whether and how many subjects observe them. Only after they finish working on the anagram task in the end of part 2 they receive information on whether their guesses about actually being observed were correct and whether they will earn additional DKK 20.00 for their correct guess.

4.2 Results: Chosen Matching Treatment

In the following, we first consider subjects' peer choices. That is, in the words of Falk & Knell (2004) we focus on the question of whether on average the self-enhancing (downward) or the self-improving (upward) comparison motivation determines subjects' peer choices.

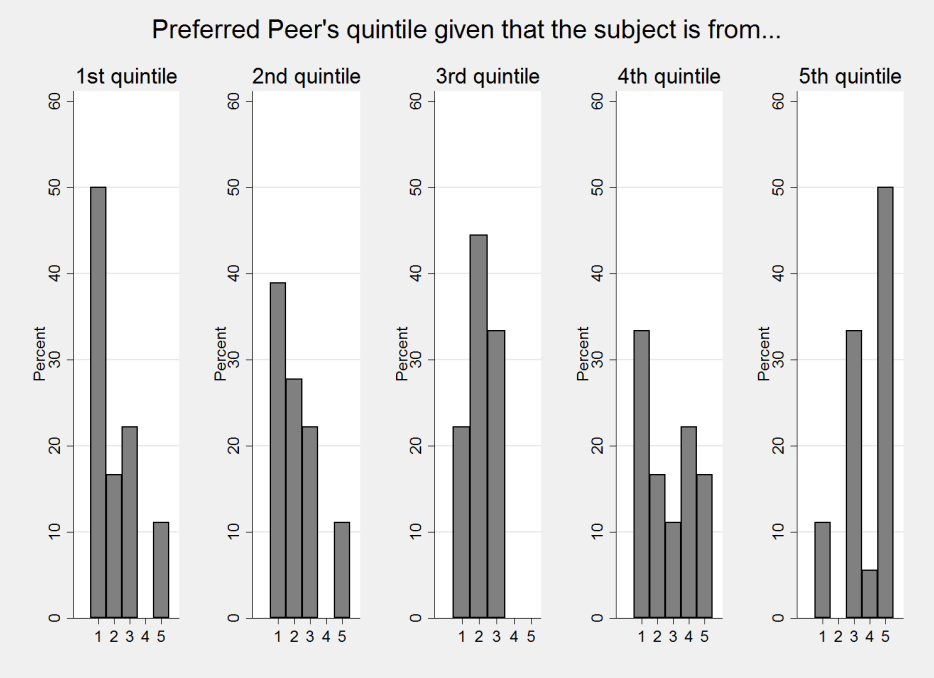

Figure 2: Choice of Peer quintile

Figure 2 shows for each of the subjects' performance quintile separately, their preferred Peer's performance quintile. As evident, in all five quintiles, at least fifty percent of the subjects prefer to be informed about the avoidance behavior of a Peer from their own or a (slightly) better performance

[12]In fact, 29 out of 90 subjects in that treatment were allocated their roles randomly: 6 out of 28 subjects who wanted to be an Observer had to assume the Peer role, 23 out of 62 subjects who stated a preference for the Peer role had to take on the role of an Observer. The remaining 61 subjects could assume the roles they initially stated as their preferred choice.

quintile. This suggests that, on average, the self-improvement comparison motivation seems to prevail in our experiment.

However, there is also a non-negligible fraction of subjects from the five quintiles that wishes to observe the avoidance behavior of a (slightly) worse peer, which is in line with Battaglini et al.'s (2005) prediction.

Next we consider subjects' role choices. Interestingly, we find that 62 out of the 90 subjects want to assume the Peer role, that is, they want to be the ones being observed. Using a regression analysis we aim at eliciting the determinants of this choice. In individual Probit models we regress the subjects' decision to become a Peer (= 1, vs. an Observer = 0) on their average performance on hard anagrams in part 1, their amount of skipping and switching in part 1, overconfidence (regarding their performance quintile in part 1) as well as on their self-reported levels of narcissism, grit, altruism and gender that we elicit in the questionnaire. It turns out that only two of these variables have significant effects on the desire to assume the Peer role. Subjects' average performance on hard anagrams is positively correlated with their wish to be observed, while subjects' inclination to switch to easy anagrams is negatively correlated with it. The corresponding regression Table B.1 is reported in Appendix B.

Still puzzled by the finding that about two thirds of the subjects wanted to take on the Peer role, we further consulted their statements in the final questionnaire where we asked them to explain the reasons why they selected a specific role. Interestingly, 41 out of the 62 would-be Peers stated having chosen that role in order to avoid any information that might distract them from their work. One of those subjects, for instance, stated: *"I did not want to observe anybody else, as I thought it would be annoying while working."*. These subjects were hence consciously self-selecting out of a situation in which they would receive information about a peer. Since we did not provide them with a neutral "No peer option", they could only self-select into the Peer role in our experiment.

Only three of the remaining 21 subjects who wanted to assume the Peer role stated that they wanted to do so in order to provide a good example for their Observer such that he or she could learn from observation. One of them for instance wrote that she, *"felt very confident in the experiment, so [she] figured that someone else, might benefit from seeing, what/how [she] was doing."*

Taken together, our results from the Chosen Matching Treatment indicate that individuals would rather like to avoid receiving information about a peer's performance in a cognitively demanding task such as ours. However, when being forced into an Observer position, they seem to be guided by the motive of self-improvement, suggesting that they are aware of the potentially motivating peer effects that we found in the Random Matching Treatment.

5 Conclusion

Although the literature on peer effects has been growing over the last couple of years, the precise mechanisms how peer effects arise are still a black box. In this paper we uncovered some new puzzle pieces that help us to better understand how individuals are affected by peers. In particular,

instead of simply measuring how peers influence produced output, we study how the presence of a peer affects subjects' behavior in a rather qualitative dimension, that is, how peers affect their co-workers' perseverance.

We measure perseverance in a word play task in which subjects can skip anagrams or switch to easier anagrams in order to avoid working on the default task solving hard anagrams. After working independently in part 1 of the experiment, subjects in our Random Matching Treatment are randomly assigned the role of an Observer or a Peer and randomly matched to each other.

We find that Observers' skipping behavior decreases significantly when they receive information about a Peer. This effect is independent of the Peer's actual behavior, which suggests a general social facilitation effect (Zajonc 1965). However, when only considering the switching behavior of Observers who are matched to a similarly or less able Peer, we find that these Observers react to the actual behavior of their Peer. In particular, Observers seem to act according to the motto "if he can do it, so can I": when their Peer stayed with hard anagrams in the previous round, the Observer has a higher propensity to stay with hard, too. If, conversely, the Peer is out of reach ability-wise, his presence and his actual behavior have less of an effect on the Observer's switching behavior. For Peers, knowing to be observed does not affect their behavior. On average, they do not behave significantly different than subjects from our Baseline Treatment who work in isolation.

The fact that all Observers, independent of their ability, increase their perseverance in the presence of a Peer has policy implications for diverse situations in which peer effects can be expected, including, but not limited to sport, social, work and educational contexts. The fact that the social facilitation effect alone increases a group's average perseverance allows for the formation of teams with varying ability and grit levels. In the organizational context superiors are well advised to arrange working conditions that permit workers to observe each other regardless of their individual tasks, in order to profit from perseverance-increasing peer effects. Examples being open plan offices or open factory work spaces. In a similar vein, our findings suggest that (university) students could benefit from studying in the library or study rooms instead of working individually at home. In the domain of sports, athletes might benefit from training with other athletes – even if they compete in a different sport – in order to reap motivational effects by observing peer athletes staying persistent in their training.

However, although our findings from the Random Matching Treatment suggest that peers increase their co-workers' perseverance, the results from our third treatment, the Chosen Matching Treatment, suggest that a significant part of the work-force might prefer to avoid the peer situation entirely. When given the choice between the Observer and Peer role, the majority of subjects prefers the Peer role. Subjects' survey responses suggest that they use this as a means to avoid being distracted by information about another subject's actions, even at the expense that their own avoidance behavior is observed by others. This finding warrants further study, as it shines new light on how to interpret peer effects. Individuals might prefer and even deliberately choose to work alongside a peer in a tedious task such as Falk & Ichino's (2006) folding letters into envelopes task. However, if they are confronted with a mentally more challenging task that requires a high

level of concentration, the findings from the present study indicate that they might prefer to work in isolation in order to not get distracted by peer information.

References

Alan, S., Boneva, T. & Ertac, S. (2015), 'Ever failed, try again, succeed better: Results from a randomized educational intervention on grit', *Working Paper No. 2015-009. Human Capital and Economic Opportunity Working Group* .

Andreoni, J. & Bernheim, B. D. (2009), 'Social image and the 5050 norm: A theoretical and experimental analysis of audience effects', *Econometrica* **77**(5), 1607–1636.

Baeker, A. & Mechtel, M. (2013), The role of task meaning on output in groups: Experimental evidence, IAAEU Discussion Papers 201308, Institute of Labour Law and Industrial Relations in the European Union (IAAEU).

Bandiera, O., Barankay, I. & Rasul, I. (2010), 'Social incentives in the workplace', *Review of Economic Studies* **77**(2), 417–458.

Bandura, A. & Walters, R. H. (1963), 'Social learning and personality development'.

Banerjee, A. V. (1992), 'A simple model of herd behavior', *The Quarterly Journal of Economics* **107**(3), 797–817.

Battaglini, M., Benabou, R. & Tirole, J. (2005), 'Self-control in peer groups', *Journal of Economic Theory* **123**(2), 105 – 134.

Bellemare, C., Lepage, P. & Shearer, B. (2010), 'Peer pressure, incentives, and gender: An experimental analysis of motivation in the workplace', *Labour Economics* **17**(1), 276 – 283.

Bernheim, B. D. (1994), 'A theory of conformity', *Journal of Political Economy* **102**(5), 841–877.

Beugnot, J., Fortin, B., Lacroix, G. & Villeval, M. C. (2013), Social networks and peer effects at work, IZA Discussion Papers 7521, Institute for the Study of Labor (IZA).

Bonein, A. & Denant-Bomont, L. (2015), 'Self-control, commitment and peer pressure: a laboratory experiment', *Experimental Economics* pp. 1–26.

Bucciol, A., Houser, D. & Piovesan, M. (2013), 'Temptation at work', *PLoS ONE* **8**(1), 1–5.

Buechel, B., Mechtenberg, L. & Petersen, J. (2014), Peer effects and students' self-control, MPRA Paper No. 53658, Munich Personal RePEc Archive.

Charness, G. & Kuhn, P. (2011), 'Lab labor: What can labor economists learn from the lab?', *Handbook of labor economics* **4**, 229–330.

Charness, G., Rigotti, L. & Rustichini, A. (2007), 'Individual behavior and group membership', *American Economic Review* **97**(4), 1340–1352.

Corgnet, B., Gonzalez, R. H. & Mateo, R. (2015), 'Cognitive reflection and the diligent worker: an experimental study of millennials', *PloS one* **10**(11), e0141243–.

Deci, E. L. & Ryan, R. M. (1985), *Intrinsic motivation and self-determination in human behavior*, Springer Science & Business Media.

Duckworth, A. L., Peterson, C., Matthews, M. D. & Kelly, D. R. (2007), 'Grit: Perseverance and passion for long-term goals', *Journal of Personality and Social Psychology* **92**(6), 1087–1101.

Duckworth, A. L. & Quinn, P. D. (2009), 'Development and validation of the short grit scale (grit-s)', *Journal of Personality Assessment* **91**(2), 166–174.

Eskreis-Winkler, L., Shulman, E. P., Beal, S. A. & Duckworth, A. L. (2014), 'The grit effect: predicting retention in the military, the workplace, school and marriage.', *Frontiers in Psychology* **5**, 36.

Falk, A. & Ichino, A. (2006), 'Clean evidence on peer effects', *Journal of Labor Economics* **24**(1), pp. 39–57.

Falk, A. & Knell, M. (2004), 'Choosing the joneses: Endogenous goals and reference standards', *Scandinavian Journal of Economics* **106**(3), 417–435.

Festinger, L. (1954), 'A theory of social comparison processes', *Human Relations* **7**(2), 117–140.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* (53658).

Georganas, S., Tonin, M. & Vlassopoulos, M. (2015), 'Peer pressure and productivity: The role of observing and being observed', *Journal of Economic Behavior and Organization* **117**(C), 223–232.

Gerhards, L. & Gravert, C. (2015), 'Grit trumps talent? an experimental approach', *Working Papers in Economics 629, Department of Economics, University of Gothenburg* .

Gill, D. & Prowse, V. (2012), 'A Structural Analysis of Disappointment Aversion in a Real Effort Competition', *American Economic Review* **102**(1), 469–503.

Guryan, J., Kroft, K. & Notowidigdo, M. J. (2009), 'Peer effects in the workplace: Evidence from random groupings in professional golf tournaments', *American Economic Journal: Applied Economics* **1**(4), 34–68.

Ichino, A. & Maggi, G. (2000), 'Work environment and individual background: Explaining regional shirking differentials in a large italian firm', *The Quarterly Journal of Economics* **115**(3), 1057–1090.

Kandel, E. & Lazear, E. P. (1992), 'Peer pressure and partnerships', *Journal of political Economy* pp. 801–817.

Kaur, S., Kremer, M. & Mullainathan, S. (2010), 'Self-control and the development of work arrangements', *American Economic Review* **100**(2), 624–28.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *The Review of Economic Studies* **60**(3), pp. 531–542.

Mas, A. & Moretti, E. (2009), 'Peers at work', *American Economic Review* **99**(1), 112–145.

Mobius, M. & Rosenblat, T. (2014), 'Social learning in economics', *Annual Review of Economics* **6**(1), 827–847.

O'Keeffe, M., Viscusi, W. K. & Zeckhauser, R. J. (1984), 'Economic Contests: Comparative Reward Schemes', *Journal of Labor Economics* **2**(1), 27–56.

van Veldhuizen, R., Oosterbeek, H. & Sonnemans, J. (2015), Peers at work: From the field to the lab, Tinbergen institute discussion papers 14-051/i, Tinbergen Institute.

Zajonc, R. B. (1965), 'Social facilitation', *Science* **149**(3681), 269–274.

A Questionnaire

Grit Scale: Scored from (1) ("Not at all like me") to (5) ("Very much like me"). Items 1, 3, 5 and 6 are reversely coded.

1. New ideas and projects sometimes distract me from previous ones.
2. Setbacks don't discourage me.
3. I have been obsessed with a certain idea or project for a short time but later lost interest.
4. I am a hard worker.
5. I often set a goal but later choose to pursue a different one.
6. I have difficulty maintaining my focus on projects that take more than a few months to complete.
7. I finish whatever I begin.
8. I am diligent.

B Additional tables

Table B.1: Chosen Matching Treatment: Determinants of Peer role choice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Solved hard anagrams in part 1	0.183*							
	(0.100)							
Skipped hard anagrams in part 1		0.019						
		(0.037)						
Switches to easy anagrams in part 1			−0.108**					
			(0.053)					
Self-reported narcissism				0.102				
				(0.528)				
Self-reported grit					−0.073			
					(0.234)			
Self-reported altruistic inclination						0.016		
						(0.052)		
Overconfident regarding performance quintile in part 1							−0.279	
							(0.294)	
Female								−0.111
								(0.280)
Constant	0.114	0.446***	0.810***	0.445	0.724	0.379	0.587***	0.541***
	(0.242)	(0.164)	(0.213)	(0.283)	(0.759)	(0.407)	(0.172)	(0.186)
Pseudo R^2	0.033	0.003	0.037	0.000	0.001	0.001	0.008	0.001
Observations	90	90	90	90	90	90	90	90

Probit regressions with robust standard errors that are given in parentheses: * p<0.10 ** p<0.05 *** p<0.01.