

If I can do it, so can you! Peer Effects on Perseverance*

Berno Buechel^a, Lydia Mechtenberg^b, and Julia Petersen^b

March 16, 2018

Abstract

Successful performance – be it in school, at the job, or in sports activities – requires perseverance, i.e., persistent work on a demanding task. We investigate in a controlled laboratory experiment how an individual’s social environment affects perseverance. We find evidence for two kinds of peer effects: being observed by a peer can postpone the decision to give up, while observing a peer can be informative and have more sustainable effects. In particular, we show that successful peers affect perseverance positively if they communicate their success in a motivating way and negatively otherwise. Our experimental results suggest that peers affect perseverance indirectly, via influencing self-confidence.

Keywords: Perseverance; Peer Effects; Experiment

JEL-Classification: C91; D90; J24

^a University of Fribourg, Department of Economics, Bd. de Pérolles 90, 1700 Fribourg, Switzerland. Email: berno.buechel@unifr.ch. Web: www.berno.info.

^b University of Hamburg, Department of Economics, Von-Melle-Park 5, 20148 Hamburg, Germany. Email: lydia.mechtenberg@wiso.uni-hamburg.de, petersen.julia@t-online.de. Tel.: +49-30-25491-446.

* This paper is a substantial revision of an older version that circulated under the title “Peer Effects and Perserverance.” Special thanks go to the WISO lab crew: Olaf Bock (leader), Thais M. Hamasaki, Jan Papemeier, Fenja Pauls, and Merlinde Claudia Tews; and to Willi Flemming and Lorenz Rumberger for supreme research assistance.

1 Introduction

1.1 Motivation

Some people are more successful than others. Be it in school, at the job, or in sports activities, there are tasks which require painful effort at some point. While many people let themselves get distracted or fully give up, others stay focused and keep working on the successful completion, i.e. they *persevere*. Clearly, while perseverance is to some extent a matter of personality, it is also, and probably more importantly, shaped by the social environment, which includes peers, teachers, parents, and coaches. Personal interactions can have two very distinct effects on individual perseverance.¹

On the one hand, others who know about the task at hand and interact with the individual after the task can induce social pressure. This is the typical view that peer effects work through *being observed* by peers. On the other hand, *observing peers* before the own task can affect an individual's ability to persevere. More experienced peers, as well as mentors and coaches, may foreshadow the value of perseverance and affect the confidence of the one aspiring to successfully complete a task.² It is important to understand the conditions for positive and negative peer effects on perseverance because they can have severe consequences. Applications range from students' educational success, over performance of sports athletes, to success within organizations.

In this paper, we test both categories of peer effects on perseverance in a controlled laboratory experiment. We find that both types of peer effects are present in some way. Peer pressure to justify the own performance after the task induces participants to try for a longer period before giving up. Communication with a successful peer before the task, on the other hand, can either induce participants to give up immediately or to persevere, depending on the content of the communication. This last result means that a successful peer can have a crucial effect on the perseverance of a less experienced individual, while unsuccessful peers are rather ignored, according to our data.

¹ In a field study with children, Alan et al. (2016) provide evidence for the view that perseverance is not a fixed personality trait but can be fostered if children learn to acknowledge the value of grit for achievement. The authors do not consider peer effects.

² As an illustration of the two categories of peer effects, we may consider a leisure time runner who aspires to complete a marathon. Joining a group of runners might have a positive effect for the ability to persevere for multiple reasons. Committing to regularly practicing together and having announced the own plan to participate in a certain marathon to the group clearly fall into the category of peer effects through peer pressure. Learning about the challenges that running a marathon implies and increasing the own self-confidence by observing that others have succeeded to finish a marathon clearly falls into to the category of peer effects through learning. Yet, to choose a group of runners, it is important to know which effect is at work, under which conditions they are working, and when there are actually negative peer effects.

We provide one of the first experimental studies on how peers influence individual perseverance by their own success or failure.³ We find that the effect depends on the way peers communicate about the difficulty of the task.

1.2 Two Categories of Peer Effects

The vast majority of the studies on peer effects suggest that being observed by peers can enhance performance, e.g. due to peer pressure.⁴ By imposing some social cost on a person whom they observe giving in to a temptation, they can help that person not to give in. An example for such an effect in the educational context is a group of students who do all their work in the library and who would feel ashamed if watched by the others surfing the internet instead of writing their term paper.⁵

The second category of peer effects on perseverance is elaborated in the theory of Battaglini, Bénabou and Tirole (2005), hereafter: BBT, which provides us with our main hypotheses. They stipulate that observing one's peers – rather than being observed by them – can affect own performance. The reason is that the behavior of others can be informative about how much the task at hand would challenge one's own ability to persevere.

In BBT, individuals have imperfect knowledge about their ability to persevere on a task (that may be anything from not drinking alcohol to preparing for an exam or working on a paper).⁶ But they know that levels of this ability are positively correlated within their peer group. Hence, observing how peers react to temptations provides individuals with additional information about their own level of ability. This, in turn, affects their *self-confidence* and

³ The only other study we are aware of is the one by Gerhards and Gravert (2018).

⁴ See, e.g. Hoxby (2000), Hoxby and Weingarth (2005), Zimmerman (2003), Falk and Ichino (2006), Mas and Moretti (2009), Calvo-Armengol, Patacchini and Zenou (2009) who combine theory and empirics to investigate peer effects in an educational setting, and Rosaz et al. (2016) who investigate being observed and simultaneously observing a peer in the lab. Like our paper, Rosaz et al. (2016) implement a cognitive task and allow for communication. However, their focus is not on separating the effects of being observed and of observing a peer. Moreover, communication plays a different role in their setting (reducing social distance between peers).

⁵ This can also be a conscious strategy of individuals with a lack of self-control. Considering peers as commitment devices is an interesting variation of the general idea that commitment devices can help overcome self-control problems. See Ariely and Wertenbroch (2002), Houser et al. (2010), and Burger et al. (2011). While perseverance necessitates self-control, self-control does not imply perseverance, since perseverance is determined by many other factors, such as self-knowledge, self-confidence and incentives. Our paper studies perseverance rather than self-control. For theoretical studies on self-control problems, see, e.g., Strotz (1955), Thaler and Shefrin (1981), Ainslie and Haslam (1992), Laibson (1997), O'Donoghue and Rabin (1999), Gul and Pesendorfer (2001), Bénabou and Pycia (2002), and Fudenberg and Levine (2006). Experimental contributions on time preferences include Ariely and Wertenbroch (2002), Andersen et al. (2008), Benhabib, Bisin and Schotter (2008) and Andreoni and Sprenger (2012).

⁶ BBT equate this ability with self-control. However, the interpretation could easily be generalized to a composite ability, including self-control, but also physical and mental strength.

consequently the degree to which they themselves exhibit perseverance in the future.⁷ Intuitively, observing how peers can handle similar challenges to their ability to persevere can be encouraging or discouraging (“if he can do it, then so can I” or “if not even he can do it, then I do not even have to try”). Hence, the theory of BBT is complementary to the vast majority of the literature on peer effects which assumes that being observed by – rather than observing – one’s peers affects own performance.

1.3 Our Contribution

In a large lab experiment with overlapping generations, we address the question whether observing or being observed affects own perseverance and performance. Hence, we test the major claims that BBT and the empirical literature on peer effects make about causality against each other. For this purpose, we match subjects who face a challenge to perseverance with subjects who faced the same challenge before and whose success or failure is observed by the former. We let the matched pairs talk to each other in bilateral free text chats about the challenge. Some of the subjects who chat *after* the task anticipate being observed by and chatting with the peer, and some do not. The former setting is the *reputational treatment* in which *being observed* by peers may levy an additional cost on giving up. Subjects who chat *before* the task can be encouraged or discouraged to persevere later on by *observing* an experienced peer. We find that knowing in advance that one has to justify one’s own performance toward one’s peers prevents giving up immediately but does not necessarily have long-lasting effects on perseverance. Hence, there seems to be a reputational effect as stipulated by the empirical literature on peer effects. However, we also find effects on those unpracticed peers who observe how their advanced peers explain their own success. More specifically, both motivating and de-motivating messages from the advanced peer affect own perseverance if and only if the peer had been successful in her own task. A peer’s success or failure in itself, unmediated by communication, has no visible effect on perseverance in our data. Our findings are consistent with BBT and the empirical test of their theory in Battaglini, Diaz, and Patacchini (2015).

In our lab experiment, communication is the channel through which information that is assumed to be public in BBT is made available. Communication addresses in particular information on how, according to the experienced peer’s belief, the subjective difficulty of the task at hand – or, alternatively, the ability to persevere on it – is distributed in the population.

⁷ Relatedly, Bénabou and Tirole (2002, 2003) and Mechtenberg (2009) theoretically study effects of communication on self-confidence of subjects with incomplete self-knowledge. These studies find both positive and negative effects on self-confidence, i.e. learning from a peer can be both encouraging and discouraging.

Hence, our lab experiment combines the approach of BBT with a communication channel similar to the one discussed in Bénabou and Tirole (2002, 2003) and Mechtenberg (2009). The presence of communication channels is also the main feature of our experimental design that essentially distinguishes our lab experiment from the one conducted by Gerhards and Gravert (2018). As opposed to them, we do not find that observing one’s peers – without the mediation of communication – has any effect on perseverance; in our lab experiment, it is the interaction between listening to communication and observing success that is effective.

2 Experimental Design

We first describe the task environment and then turn to the different treatments and the overlapping-generations structure.

2.1 The Task Environment

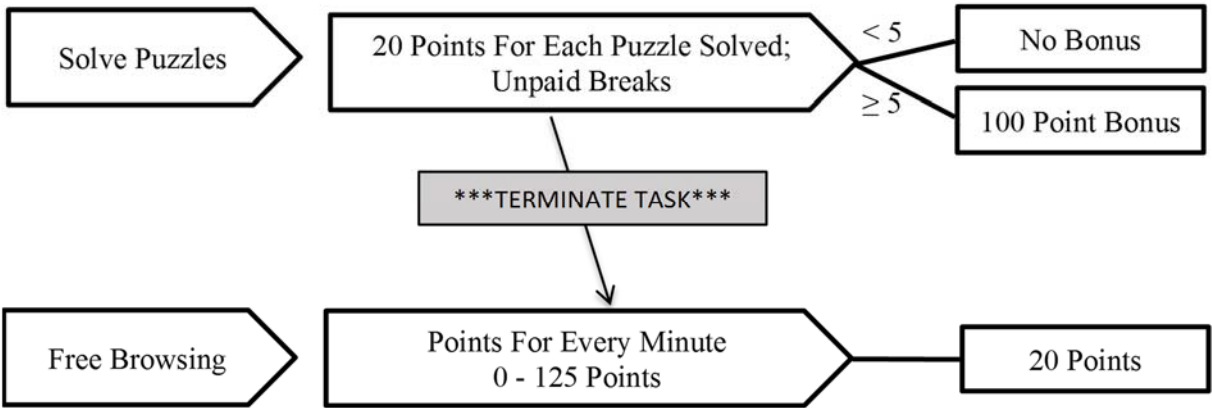
To measure perseverance, we implemented the combination of a real-effort task with access to the internet (see, e.g., Corgnet et al. 2015). In order to create the necessary variation in perseverance, we chose a real-effort task that is (1) depleting, to require an effort to persevere, (2) meaningless, to reduce average intrinsic motivation, and (3) novel to our subject pool, to exclude experience from previous real-effort experiments. Subjects had to identify, among a number of meaningless strings of letters (like abCOCAcaZAgbCBZ) any such strings that contain the letter combination “abc” (in all combinations of small and capital letters) *if* that is followed by a vowel. (For such tasks and their depleting effect, see Baumeister et al., 1998, and many others as surveyed in Hagger, Wood, & Stiff, 2010.)

The task was divided into ten subtasks, each consisting of 99 strings displayed on the computer screen, with five input fields for the five strings that had to be identified. A subtask was counted as correctly solved if at least four of the five strings were identified correctly. The entire task was counted as correctly solved if five out of the ten subtasks were solved. We chose these standards (after piloting the real-effort task) to allow approximately one third of our subjects to solve the entire task correctly. For the correctly solved task, a bonus of 100 points was paid. Additional points were paid for each subtask that was correctly solved. The difficulty of the task was increasing in that more and more letters a, b, and c appeared in the strings, while still only five strings had to be identified. The points awarded for correct subtasks were fixed to 20. Subjects had 50 minutes to solve the task, but they could also give up on it whenever they wished. Those who wished to give up could click the “terminate task”

button always visible on the task screen and spend the remainder of their time browsing the internet.

Browsing the internet is a common diversion at the work-place and hence a realistic obstacle to perseverance. However, it is to be expected that the laboratory setting activates the social norm to exclusively concentrate on the experimental task. Moreover, access to the internet is typically blocked in lab experiments to enhance control. Therefore, granting access to the internet may well create a demand effect; subjects might deduce from the design that browsing the internet would be encoded as lack of perseverance. Thus, simply granting access to the internet might not be sufficient to induce any usage of it. To prevent the demand effect and to overcome the social norm of “no surfing in the lab”, we paid a small bonus of 20 points for terminating the real-effort task and paid “surfing points” for browsing the internet afterwards, thereby providing a possible rationale for giving up the task. Browsing the internet in minute t yielded $t/10$ points to someone who terminated the task. For instance, if a subject clicked the “terminate-task” button in minute 12, she earned 1.2 points in minute 12; 1.3 points in minute 13, and so on. Hence, who only browsed the internet and did not even start working on the task received $125+20$ points in total. Generally, clicking the “terminate task” button at time T yielded $125 - \frac{1}{20}T^2 + 20$ points in total. 100 points correspond to 5 Euros. The payoff structure of the task environment is depicted in Figure 1. The respective amount of points possible to gain for continuing the task and for terminating it and surfing for the rest of the time was continuously announced on the experimental screen in z-tree.

Figure 1: Overview of Payoff Structure



The payoff structure is such that the following holds in virtually any case: If one solves at least five of the ten subtasks correctly, i.e., if one is eligible for the task bonus, then it pays off to continue on the task until the end, that is, to never click the “terminate task” button. If one solves less than five of the ten puzzles correctly, then it pays off to quit the task, that is, to click the “terminate task” button. In this case, the best choice is to terminate right after the

beginning of the task or right after solving at most one subtask. We did *not* communicate during the task whether the subject had reached the threshold for the bonus (this would strongly determine performance). Also, we took great care to explain the payoff structure and the choice between continuing and terminating the task in the instructions, using Figure 1. Subjects were given an example of a subtask in the instructions for them to assess their own confidence in their performance. Also, in the course of their work at the task we elicited their beliefs about whether or not they solved the last subtask correctly. Hence, we could validate that their termination decision was optimal, given their beliefs, which is true for more than 90% of the participants. (More precisely, of those who believed that they had solved five subtasks or more correctly, no one gave up. Of the remaining participants, most (84%) gave up at some point.)

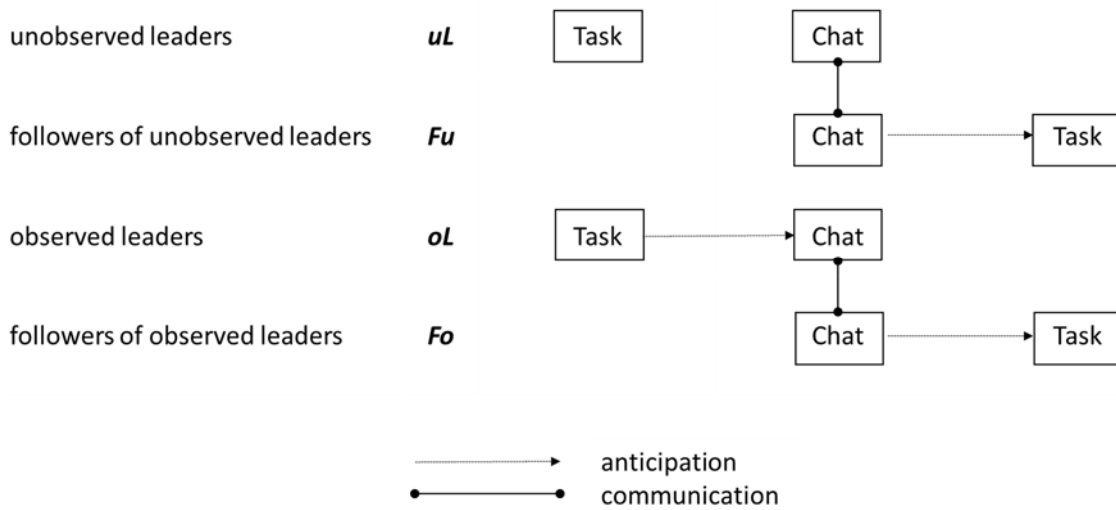
To familiarize subjects with the website that they could browse after task termination, we explicitly stated in the instructions that taking occasional breaks from the task and surfing during the breaks are allowed. However, we did not pay any points for browsing the internet during a break. Breaks could be scheduled freely by clicking the break button and switching to a specific website created by us for a deliberate amount of time.⁸ Moreover, prior to each subtask, subjects had to go on a website to search a password with which they could unlock the next subtask in z-tree. The website offered a broad variety of pictures and articles in many categories such as politics, sports, diets, celebrities, cartoons, and news; and during the unlocking of each subtask, subjects were made aware of the possible diversions offered on the website.

2.2 Treatments

We conducted four different treatments in a between-subjects design, two *leader* treatments and two *follower* treatments. We now use the “leader” and “follower” terminology to easily refer to a more experienced peer versus an unexperienced peer. Subjects were invited in subsequent cohorts, with a leader cohort preceding a follower cohort. A cohort comprised twelve subjects, half as many as were cubicles in the lab. All treatments had two parts, a *task part* and a *chat part*. The task part was spent in the task environment as described above. In the chat part, subjects in a follower treatment were randomly matched one-to-one with the subjects of the preceding leader treatment and could chat with their matched partner for three minutes. The treatment structure is depicted in Figure 2.

⁸ Note that breaks need not signify lack of perseverance but might be part of an optimal working schedule since the task is depleting.

Figure 2: Treatments: unobserved leaders (uL), followers of unobserved leaders (Fu), observed leaders (oL), and followers of observed leaders (Fo)



THE CHAT PART A follower matched with a leader from the precedent cohort in the first (chat) part of a follower treatment learned the failure or success and the give up decision of their matched partner, i.e., whether or not their matched partner had earned the bonus of 100 points and whether or not she had terminated the task. This was communicated to the followers on their computer screen and remained visible there throughout the next (task) part. In addition, followers could communicate with the matched leader via a chat window during the chat part. The chat opportunity lasted for three minutes. Chat format was free text; there were no prohibitions, except that anonymity must be kept. We asked the leaders from the preceding cohort to address the following three questions: 1) *Please estimate the fraction of participants in percent that manage the task, i.e. who reached the bonus of 100 points.* 2) *Does it pay off to strive for the bonus or should the task be quit immediately?* 3) *Please, explain your recommendation.* Question 1, in particular, was asked to measure the *similarity* that a leader assumed between herself and the average person, regarding performance on the task – a variable of central importance in BBT. A successful (unsuccessful) leader who believes more than half of the other subjects to be successful (unsuccessful) as well deems herself rather *similar* to the average person; a successful (unsuccessful) leader who believes that no other is successful (unsuccessful) as well deems herself *exceptional*. We assume that a follower – who has no experience with the task – tends to perceive herself as an average person with regard to the task environment. Thus, the answer that her matched leader gives to question 1 provides her with information about how similar she herself is to her matched leader. Since BBT implies that only similar peers have effects on subjects' perseverance, we

expect the answers given by the leaders to question 1 to set the conditions for any effects of their success or failure on the perseverance of the matched followers.

LEADER TREATMENTS In the treatment “*observed leaders*” (*short: oL*), subjects are leaders: They first do the task part and then the chat part. During the task part they know the content of the second part, i.e., that they will be matched with followers (*after* carrying out the task), that these followers will learn their success or failure and their perseverance, and that they will chat for three minutes. By contrast, in the other leader treatment “*unobserved leaders*” (*short: uL*), they do not anticipate the details of the chat part since instructions for that are only given after completion of the task part. Hence, subjects do not anticipate that they will be matched with followers of the subsequent cohort and that information about their own perseverance and performance will be provided to their matched partners.

FOLLOWER TREATMENTS In the “*followers of observed leaders*” treatment (*short: Fo*) subjects are followers: They first do the chat part and then the task part. They are matched with leaders from *oL* (observed leaders) in the chat part. In the other follower treatment “*followers of unobserved leaders*” (*short: Fu*) subjects are matched with leaders from *uL* (unobserved leaders) during the chat part.

2.3 Implementation

In total, we have $N=336$ participants which amounts to 84 subjects per treatment.⁹ Cohorts are overlapping: always two consecutive cohorts were chatting with each other after the first of the two had finished the task and before the second started with it. Hence, for logistic reasons we conducted multiple sequences of sessions per day. We distributed these sequences of sessions across different days and made sure that treatments do not differ much in the days and time slots at which they were run. The experiment was run at the WISO-laboratory of the University of Hamburg. A different room outside the computer lab was reserved for reading of instructions and a subsequent quiz to guarantee understanding of the decision and payoff structure.¹⁰ Another room outside the lab was reserved for payment (in cash and in private). Communication other than required in the chat part of the experiment was prohibited, and compliance with this rule was monitored by the staff of the lab throughout the entire experiment. In the time span between the end of the task and leaving (in treatments *Fo* and *Fu*) or chatting with the subsequent cohort (in treatments *oL* and *uL*), subjects answered a questionnaire. This questionnaire elicited age and sex of the subject and other control

⁹ There was only one case of a system failure within a regular session, which reduces the number of observations in *Fo* to 83.

¹⁰ The instructions and the quiz are provided as an online appendix.

variables. The chat data were analyzed following the procedure in Brandts and Cooper (2007). A research assistant unaware of the purpose of the study and the different treatments did the transcription and coding. The reliability of the coding was checked and confirmed by having another, equally unaware research assistant, independently coding the same chat protocols. The coding of the three questions that had to be addressed in the chat into answer categories was straight-forward.

3 Results

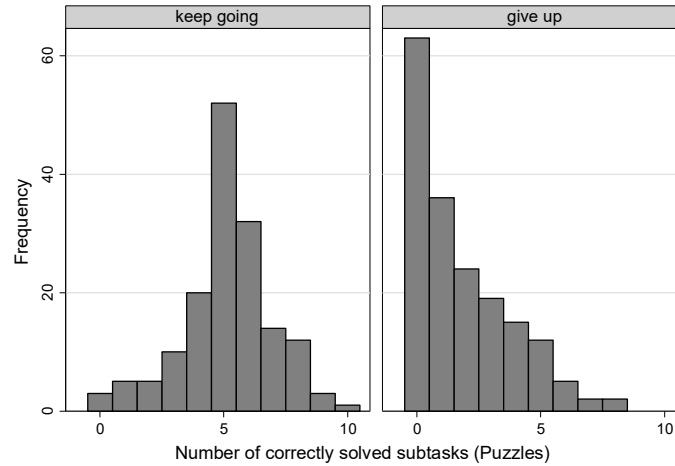
Before we turn to the effects of observing peers and to the effects of being observed by peers, we first provide some summary statistics for the main variables.

3.1 Some Summary Statistics

The main outcome variables are the number of subtasks solved correctly (*Puzzles*) and the dummy whether the bonus is reached or not (*Bonus*). Table 1 reports summary statistics of these main outcome variables and also reports the payoff earned in EUR (Payoff) and the minutes participants stayed in the task before clicking the “terminate task” button (MinutesTask). The average number of correctly solved puzzles is below four, but still roughly every third participant (34%) reached the bonus of 100 points, which requires to solve five or more. There was indeed one person who solved 10 puzzles correctly, each time believed that the puzzle is solved correctly, and received the bonus of 100 points. (This leads to the maximal number of points of $10 \cdot 20 + 10 \cdot 1 + 100 = 310$, which yields the maximal payoff of EUR $15.50 + 8.00 = 23.50$.) The average payoff was around 16 EUR.

In Table 1 there are separate outcome statistics for those participants who gave up, i.e. clicked the “terminate task” button at some point, and those who stayed in the task, reported in the last two columns. Separating these two groups, we observe that the bonus was reached by 73% of those who stayed in the task. The histogram below shows that those who stayed in the task solved typically five or six puzzles correctly. Very few managed to solve nine or ten tasks. Most of those who gave up solved none or less than three tasks. Table 2 reports the timing of giving up, i.e. the minutes before clicking the “terminate task” button, separated by the number of puzzles solved. Those who solved three or four puzzles gave up after 44 of 50 minutes, which indicates that they strived for the bonus until they realized that they would not reach it. Participants who solved two puzzles or less gave up earlier.

Figure 3: Histogram of Puzzles by Giving Up Decision



Since each participant was asked after each subtask whether she believed that she had solved it correctly, we can analyze the participants' beliefs about the number of tasks they had solved correctly. The correlation of this belief with the actual amount solved correctly (*Puzzles*) is 0.88. Looking at the difference between these two variables, it turns out that a majority of participants, 72%, has exactly accurate beliefs, while 24% overestimate the amount, and 4% underestimate it. Hence, beliefs are quite accurate, but there is a bias consistent with overconfidence. Most overconfident participants overestimate the amount solved by one or two.

We now first analyze the effects of observing peers and communication, i.e., we focus on the follower treatments. Afterwards, we turn to the effects of being observed, i.e., we focus on the observed leader (*oL*) treatment.

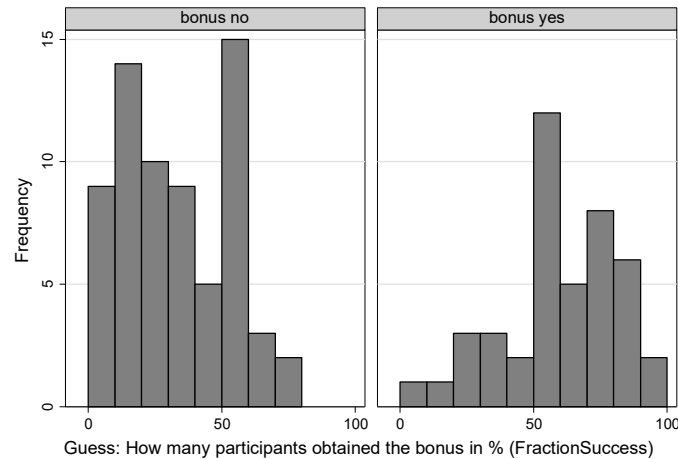
3.2 Effects of Observing Peers and Communication

Consider the follower treatments, *Fo* and *Fu*. From the perspective of the followers, no treatment variable changes between these two treatments. Hence, when analyzing the followers' reaction to what they observe from their matched leaders, we pool the data from *Fo* and *Fu*.

The main aspect observed by a follower is whether her matched leader did or did not receive a bonus (captured by the dummy variable *BonusPeer*). This piece of information is highlighted during the chat and on each subtask in the follower treatments. Moreover, in the chat the leaders are asked to provide a guess about the fraction of participants that was able to receive the bonus. The answers range from 0% to 100%, and we call this continuous variable *FractionSuccess*. As the histograms in Figure 4 show, leaders who do not receive the bonus more often guess that the number of participants receiving the bonus is small, but there is still

a considerable variation across leaders such that our subjects in the follower treatments get very different signals about the difficulty of the task and their personal likelihood of being able to receive the bonus.

Figure 4: Histogram of *FractionSuccess*, i.e. the leader’s estimate concerning how many participants manage to reach the bonus, differentiated by whether the leader has received the bonus herself (right panel) or not (left panel)



FractionSuccess provides information on the leader’s belief about how likely it is that the follower can succeed. Combined with the leader’s success or failure, *FractionSuccess* provides information on the peer’s belief on how similar she is to the general population and hence to the follower. If it follows from *FractionSuccess* that most will succeed or fail like the peer, this signals high similarity. Moreover, in the chat leaders gave explicit recommendations whether to give up or to strive for the bonus, and we elicited how difficult the task is described as. These two variables *Recommendation* and *Difficulty* are highly correlated with the variable *FractionSuccess*, but coarser measures since they are binary, respectively ordinal.

In the following, we report the results of regressing the main outcome variables (*Puzzles* and *Bonus*) on *BonusPeer* and *FractionSuccess*, which are the two most important variables of observing and communicating with a peer. They cover the objective information provided by the experimenter that the peer (leader) has or has not received the bonus, as well as the subjective message of how hard it is to reach the bonus. Since it must be expected that these two variables interact with each other, we also include the interaction term. The outcome variable *Puzzles* is a count variable which we estimate by a Poisson regression. The outcome variable *Bonus* is a binary variable, for which we use a logistic regression.

We first illustrate the results by plotting the predicted outcome values by different levels of *FractionSuccess* for both subjects with matched leaders who received a bonus and subjects with matched leaders who did not. We obtain the predicted number of puzzles and the

predicted probability of a bonus for values of *FractionSuccess* that range from 0 to 100 with increments of 5.

Figure 5: Marginsplots of Puzzles and Bonus by *FractionSuccess* and *BonusPeer*

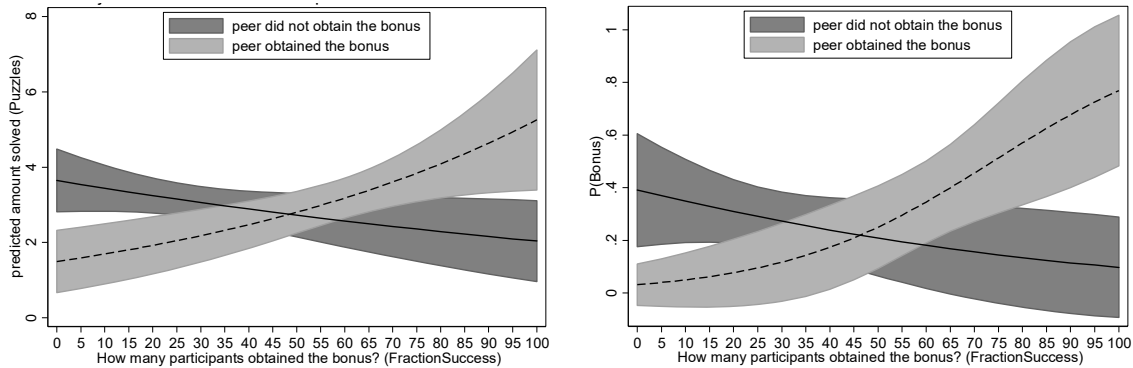


Figure 5 shows how the outcome variables are increasing in *FractionSuccess* for a matched leader who received a bonus and rather decreasing in *FractionSuccess* for a matched leader who did not. The graphs indicate that the best performance is predicted for those who observe that their matched leader obtained the bonus *and* who receive the message that a very high fraction is believed to be able to reach the bonus. Lowest performance is obtained if the matched leader received a bonus but reports her belief that only very few are able to do so. It seems that a message from a peer who has been successful can be both particularly encouraging and particularly discouraging.

To investigate this interaction effect and to test for significance of the difference between having a peer with bonus and a peer without bonus, we run a sequence of regressions each time modifying the variable *FractionSuccess* by subtracting a constant from it. More precisely, we generate the variables *FractionSuccess*₀, *FractionSuccess*₁₀, *FractionSuccess*₂₀, ..., *FractionSuccess*₁₀₀, with *FractionSuccess*_k defined as *FractionSuccess* minus *k*. The results of these regressions are all identical apart from the coefficient of *BonusPeer*. Consider the interaction term, which is the product of the binary variable *BonusPeer* and the continuous variable *FractionSuccess*_k. When the variable *FractionSuccess*_k is zero, the interaction term is always zero such that the coefficient of *BonusPeer* picks up the whole effect. We use this fact to provide an interpretation for the effects. All coefficients of *BonusPeer*, which are the only coefficients that differ from the ones in Table 3, are reported in Table 4. Hence, the entries in column FS0 of Table 4 correspond to the columns (1'), (2'), and (3) of Table 3.

First, we observe that significance levels are very similar across model specifications. Thus, the results seem to be robust. Second, the coefficients are significant for low values and for high values of *FractionSuccess* but not in between. Moreover, the sign of the effects changes

from low to high values. We provide an interpretation for each outcome variable, starting with *Puzzles*.

EFFECTS ON PUZZLES Pure coefficients of the Poisson model cannot be interpreted easily due to the log link function. Hence, we use the incidence rate ratio (IRR) for interpretation. (The significance levels remain roughly the same comparing the robust and the IRR – option). At the lower end of the distribution, IRR coefficients are significant for *FractionSuccess* in the range from zero to thirty. For example, the IRR at *FractionSuccess20* = 0,592 (see column ‘FS20’) is the estimated rate ratio comparing whether the matched leader received the bonus, i.e., was successful, or not, while holding the message *FractionSuccess* at the level of 20%. Those whose matched leader was successful are expected to have an incidence rate for *Puzzles* 0,592 times that of subjects whose peer was not successful, which is a decrease of 41%. Thus, a low *FractionSuccess* is bad news, in particular if the matched leader has received the bonus; it seems to transport the message “I did it but it was so hard, you should not try it”. Low *FractionSuccess* of a matched leader who did not receive the bonus is in principle bad news as well, but much less so. The statement that only few reached a bonus herself does not seem highly credible if it comes from a person who did not manage to receive the bonus. Maybe subjects believe that their unsuccessful peer did not even try and only wants to justify her own behavior.

The IRR-coefficients at the upper end are significant for *FractionSuccess* in the range from 70 to 100. For instance, holding *FractionSuccess* at a high level of 80%, i.e., considering *FractionSuccess80*, the IRR coefficient is 1.786 (column ‘FS80’): Those who receive the news that their matched leader has received the bonus are expected to have an incidence rate for *Puzzles* 1,786 times that of subjects that receive the opposite news, which corresponds to an increase of almost 79%. This is the good-news effect; it seems to transport the message “I could do it *and* I am similar to the average person, hence so can you”. It seems encouraging to hear that many of the subjects are able to receive the bonus from a person who did it as well.

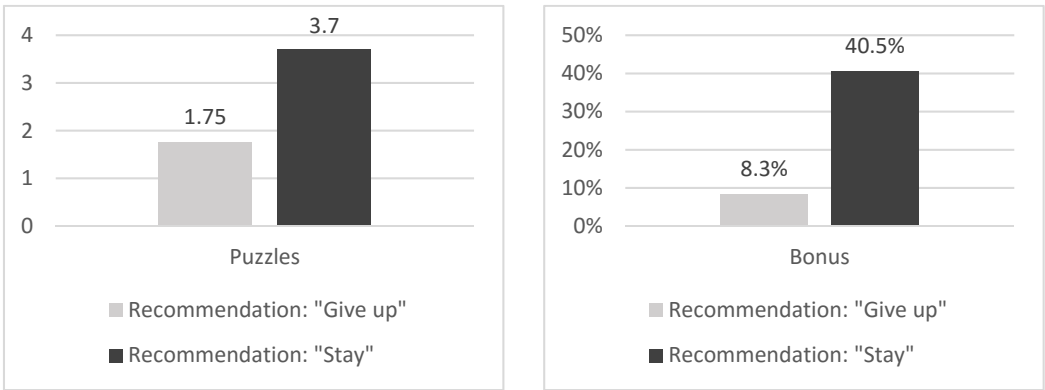
EFFECTS ON BONUS and PAYOFF Turning to the outcome variables *Bonus* and *Payoff*, we observe that the significance levels for the different *FractionSuccess* values are very similar to Poisson estimates of the outcome variable *Puzzles*. Effects are significant at the upper and the lower end of *FractionSuccess* with the same direction of the effect for each outcome variable. For instance, if the peer received a bonus and reported that only 20% were successful (i.e. at *FractionSuccess20*) *Payoff* decreases by EUR 2.01. At the other end of the spectrum, *Payoff* if one’s matched leader received a bonus and reported that 80% were successful increases by EUR 2.38, while the average payoff is EUR 15.73. Hence, there is a

sizeable effect, not only on the number of puzzles solved and on the likelihood to receive a bonus, but also on *Payoff*.

ROBUSTNESS These results remain valid when we include variables from the questionnaire. As a test of robustness, we rerun the regressions including all variables from the questionnaire that survived the stepwise reduction procedures. The magnitude and significance levels are virtually the same.¹¹

As another test of robustness we exchange the variable for the message *FractionSuccess* with the binary variable *Recommendation*, which is one if the matched leader recommends to stay in the task and zero if she recommends to give up at some point.¹² Table 5 reports the main outcome variables in dependence of whether the matched leader has recommended to give up or to stay in the task. The table gives an indication that the recommendation to stay has generally a positive effect on the outcome. This strongly indicates that it is indeed perseverance which is mainly required to succeed in the task. It also confirms the strong interaction between the content of the chat and the success of the peer. Recommendations of peers who received the bonus heavily affect outcome variables. In particular, the most encouraging leaders are those who have received the bonus and recommend striving for the bonus. Recommendation to stay in the task leads to significant increase of *Puzzles* (Mann Whitney U test, $p < 0.05$) and a significant increase of the fraction of bonus reached (Fisher exact test, $p < 0.05$), given that the matched leader has obtained the bonus.

Figure 6: Recommendation from a peer who has reached the bonus positively affects both the number of correctly solved subtasks (*Puzzles*, left panel, $p < 0.05$) and whether the bonus has been reached (*Bonus*, right panel, $p < 0.05$).



Summarizing we find both a good news effect and a bad news effect. Thereby, the signal

¹¹ The magnitudes are even a bit higher and the p -values are even a bit smaller when additional variables are included. These results can be requested from the authors.

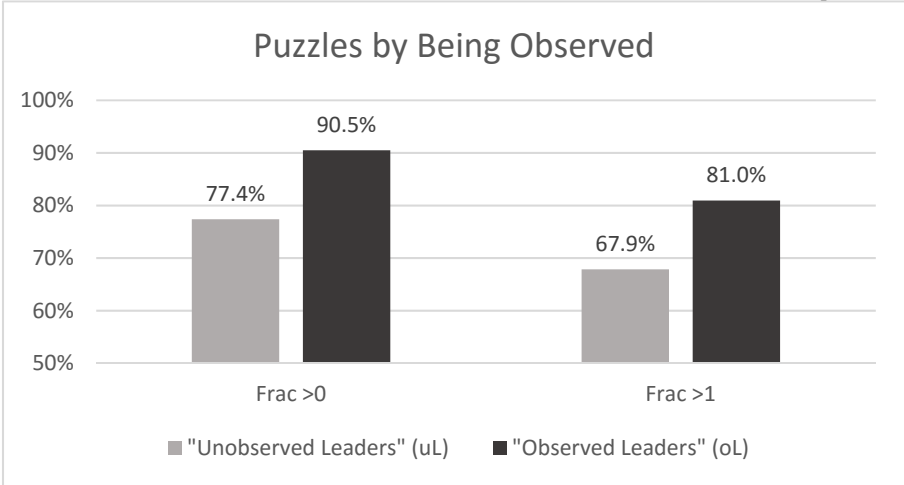
¹² The coding of the chat for the recommendation entails three categories: “stay in the task”, “work first on the task and quit later”, and “quit immediately”. The binary variable *Recommendation* simply distinguishes between the first and the two latter categories.

about the peer’s performance alone is not sufficient since it is put in relation to what the peer communicates. Good news is obtained from a successful peer who communicates in an encouraging way. Bad news are obtained from a successful peer who communicates in discouraging ways.

3.3 Effects of Reputation

We now compare the observed leader treatment (*oL*), in which subjects know that their performance will be shown to a peer and that they are going to chat with this peer, to the other treatments. Table 6 reports summary statistics of the variable *Puzzles* for *oL* and for all the other treatments.¹³ Two-sample Mann-Whitney U (MWU) tests reveal: *Puzzles* is significantly higher under *oL* than under both follower treatments ($p < 0.1$), and significantly higher under *oL* than under all other treatments pooled ($p < 0.05$). The effect of being observed on *Puzzles* is not so strong at the mean or the median as it is at the lower end of the distribution. While the null hypothesis is not rejected for *oL* vs. *uL* (the unobserved leader treatment) when using the MWU test (with $p = 0.264$), it is rejected when considering the fraction of participants who solved at least one, respectively at least two, puzzles. Fisher exact tests reveal that a fraction of subjects who solved at least one (or at least two) puzzles correctly is significantly higher under *oL* than under *each* of the three other treatments ($p < 0.05$), as well as higher under *oL* than under all other treatments pooled, of course. Hence, subjects who know that their result will be observed by a peer and that they have to justify themselves in a chat, tend to solve more puzzles on average and less often solve zero or only one puzzle.

Figure 7: Fraction who solved at least one subtask, respectively at least two subtasks, is higher in the observed leader treatment than in the unobserved leader treatment (Fisher exact tests: $p < 0.05$).



¹³ For the follower treatments, we have seen that there are positive and negative effects, depending on the peer. These effects may well cancel each other out such that no treatment effects may be apparent when comparing the follower treatments with the *uL* treatment.

It seems that the fact that there will be an observer to whom the subject has to justify her performance motivates the participants to try the puzzles at least for a while. If subjects manage to solve five puzzles this behavior pays off with a bonus of 100 points. Otherwise, this is not so clear because at the same time the subjects could have earned points for surfing in the internet. In particular, subjects who solve two, three or four puzzles can neither accumulate much points in the surfing mode nor receive the bonus of 100 points. Correspondingly, we do not find that the observed leader treatment oL significantly increases the likelihood of receiving the bonus (not reported here). In the oL treatment the fraction is 36% and in the other treatments 33%. Hence, it is not surprising that the difference of the observed leader treatment oL to the other treatments in the number of puzzles solved does not materialize in higher payoffs. The average payoffs are 15.81 for oL and 15.70 EUR for all other treatments (these numbers are not repeated in the appendix).

To summarize, we find that knowing in advance that one has to justify one's own perseverance toward one's peers prevents giving up immediately. Hence, there seems to be a reputational effect as stipulated by the empirical literature on peer effects. Our overall findings therefore provide evidence both for the direction of causality assumed by the empirical literature on peer effects and the direction of causality suggested by BBT.

4 Discussion and Conclusion

In our lab experiment, we investigated the two causal channels through which peers might influence individual perseverance. One channel ("being observed by peers") is stipulated by the empirical literature on peer effects, the other channel ("observing peers") is stipulated in a theoretical model (BBT). Importantly, our findings on the interaction of communication and observed success corroborate the approach of BBT who argue that some assumed similarity between one's own and one's peers' ability to persevere is necessary for one's own perseverance to be affected by the peer's success or failure. In our lab experiment, such beliefs are formed through communication when the advanced peers express their beliefs about how widespread their own success or failure may be among the experimental subjects (and, thus, how similar they believe themselves to be to the average student). Our results confirm the prediction of BBT that a peer's success enhances own perseverance the more, the more likely it is that her type – i.e., her ability to persevere – is similar to one's own: Only

successful peers who expressed the belief that more than half of the experimental subjects were successful, too, had a significant stimulating effect on the perseverance of the subjects matched with them. Intuitively, their success could be interpreted as the message “Since I am an average person, and I did it, you can do it, too”. However, our experiment also shows that, unpredicted by BBT, subjects can be affected negatively by their peer’s success if they are told that their peer believes herself to be the exception rather than the rule. In this case, BBT would predict no effect rather than a negative effect. Psychologically, we believe our finding to be intuitive: The success of an – according to herself – exceptionally perseverant peer may be interpreted as the message “I did it, but since I am an exceptional person, you do not even have to try”. Hence, our experiment is a first step to bridge the gap between BBT and the literature on how communication affects self-confidence. It confirms the important role that Bénabou and Tirole (2002, 2003) and Mechtenberg (2009) ascribe to the way advanced peers communicate with beginners whose self-confidence is still under formation; and it shows that – and how – one and the same fact (the success of an advanced peer) can be turned into a motivating or a de-motivating message.

Surprisingly, we find that a peer’s failure, even if justified by this peer herself as likely to be widespread among the experimental subjects, has no significant effect on own perseverance. At first glance, this seems contrary to BBT, who would predict a negative effect. However, BBT assume *objective public* information about the similarity between levels of the ability to persevere among peers whereas we, in our lab experiment, implement *subjective private information* (through peer-to-peer communication) about such similarities. Hence, it seems that in our lab experiment, only successful peers are taken to be credible in their communication about how similar they believe themselves to be to the average person. Therefore, the results of our lab experiment can be used to test the predictions of BBT about the effects of successful (i.e., highly perseverant) peers on own perseverance but may reveal nothing regarding BBT’s predictions about the effects of unsuccessful peers.

We conclude that both channels of causality can be active. Consequently, they might even complement each other and lead to a virtuous circle between the ability to persevere and positive peer effects.

References

- Ainslie, G., and N. Haslam (1992): Hyperbolic Discounting. In: Choice over Time, ed. by G. Loewenstein and J. Elster, 57-92. Russell Sage Foundation, New York.
- Alan, S., Boneva, T., & Ertac, S. (2016). Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. HCEO Working Paper.
- Ameriks, J., Caplin, A., Leahy, J., and T. Tyler (2007): Measuring Self-Control Problems. *American Economic Review* 97(3), 966-972.
- Andersen, S., Harrison, G., Lau, M., and E. Rutström (2008): Eliciting Risk and Time Preferences. *Econometrica* 76(3), 583-618.
- Andreoni, J., and C. Sprenger (2012): Risk Preferences are not Time Preferences. *American Economic Review* 102 (7), 3333-3356.
- Ariely, D., and K. Wertenbroch (2002): Procastination, Deadlines, and Performance: Self-Control by Pre-Commitment. *Psychological Science* 13(3), 219-224.
- Battaglini, M., Bénabou, R., and J. Tirole (2005): Self-Control in Peer Groups. *Journal of Economic Theory* 123, 105-134.
- Battaglini, M., Díaz, C., and E. Patacchini (2017). Self-Control and Peer Groups: An Empirical Analysis. *Journal of Economic Behavior & Organization* 134, 240-254.
- Bénabou, R., and M. Pycia (2002): Dynamic Inconsistency and Self-Control: a Planner-Doer Interpretation. *Economics Letters* 77 (3), 419-424.
- Bénabou, R., and J. Tirole (2002): Self-Confidence and Personal Motivation. *Quarterly Journal of Economics* 117 (3), 871-916.
- Bénabou, R., and J. Tirole (2003): Intrinsic and Extrinsic Motivation. *Review of Economic Studies* 70(3), 489-520.
- Benhabib, J., Bisin, A., and A. Schotter (2010): Present Bias, Quasi Hyperbolic Discounting, and Fixed Costs. *Games and Economic Behaviour* 69(2), 205-223.
- Brandts, J., and D. J. Cooper (2007). It's what you say, not what you pay: An experimental study of manager-employee relationships in overcoming coordination failure. *Journal of the European Economic Association* 5 (6), 1223-1268.
- Burger, N., Charness, G., and J. Lynham (2011): Field and Online Experiments on Self-Control. *Journal of Economic Behavior and Organization* 77(3), 393-404.
- Calvo-Armengol, A., Patacchini, E., and Y. Zenou (2009): Peer Effects and Social Networks in Education. *Review of Economic Studies* 76, 1239-1267.

Cameron, A., and P. Trivedi (2005): *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.

Corgnet, B., Hernán-González, R., and E. Schniter (2015). Why real leisure really matters: Incentive effects on real effort in the laboratory. *Experimental Economics*, 18(2), 284-301.

Falk, A., and A. Ichino (2006): Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24(1), 39-58.

Fudenberg, D., and D. Levine (2006): A Dual-Self Model of Impulse Control. *American Economic Review* 96(5), 1449-1476.

Gerhards, L., & C. Gravert (2018). Because of you I did not give up – How peers affect perseverance. Mimeo, University of Hamburg.

Gul, F., and W. Pesendorfer (2001): Temptation and Self-Control. *Econometrica* 69(6), 1403-1435.

Houser, D., Schunk, D., Winter, J., and E. Xiao (2010): Temptation and Commitment in the Laboratory. IEW Working Papers 488, Institute for Empirical Research in Economics – University of Zurich.

Hoxby, C. (2000): Peer Effects in the Classroom: Learning from Gender and Race Variation. NBER Working Paper 7867.

Hoxby, C., and G. Weingarth (2005): Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects. Mimeo, Harvard University.

Laibson, D. (1997): Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics* 112(2), 443-478.

Mas, A., and E. Moretti (2009): Peers at Work. *American Economic Review* 99(1), 112-145.

Mechtenberg, Lydia (2009): Cheap Talk in the Classroom: How Biased Grading at School Explains Gender Differences in Achievements, Career Choices and Wages. *The Review of Economic Studies* 76(4), 1431-459.

O'Donoghue, T., and M. Rabin (1999): Doing it now or later. *American Economic Review* 93(2), 186-191.

Rosaz, J., Slonim, R., & Villeval, M. C. (2016). Quitting and Peer Effects at Work. *Labour Economics*, 39, 55-67.

Strotz, R. (1955): Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies* 23(2), 165-180.

Thaler, R., and H. Shefrin (1981): An Economic Theory of Self-control. *Journal of Political Economy* 89(2), 392-406.

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

Appendix: Tables

Table 1: Summary Statistics of Lab Experiment

variable	N	mean	sd	min	max	p50	mean if kept going (N=157)	mean if given up (N=178)
Puzzles	335	3.37	2.52	0	10	4	5.15	1.80
Bonus	335	0.34	0.47	0	1	0	0.73	0.00
MinutesTask	335	39.55	14.73	0.05	49.98	49.55	49.98	30.34
Payoff	335	15.73	3.19	8.10	23.50	15.20	17.12	14.5

Notes: Calculation is based on data from laboratory experiment conducted in 2015. The variable *Puzzles* is a count variable capturing the number of subtasks solved correctly. The variable *Bonus* is binary and equal to 1 if the bonus was reached; zero otherwise. The variable *MinutesTask* captures the time elapsed until the “terminate task” button is pushed. When it is never pushed, then *MinutesTask* is fifty, which is measured by the computer as 49.98. Variable *Payoff* captures the experimental payout in EUR. “Given up” means that the “terminate task” button is pushed at some point, “kept going” means that it was never pushed.

Table 2: MinutesTask by Puzzles if Given Up

Puzzles	0	1	2	3	4	5	6	7	8	Total
Mean	17.29	27.83	36.88	43.74	43.91	45.01	41.66	46.41	46.90	30.34
Std. Dev.	12.05	10.91	10.68	7.21	7.31	4.38	5.09	0.11	4.03	15.08
Freq.	63	36	24	19	15	12	5	2	2	178

Notes: The variable *MinutesTask* captures the time elapsed until the “terminate task” button is pushed. This table only reports the observations in which this button was pushed at some point (“given up”). *Puzzles* is a count variable capturing the number of subtasks solved correctly.

Table 3: Basic Regressions

	(1)	(1')	(2)	(2')	(3)
	Poisson Puzzles	IRR Puzzles	Logit Bonus	Odd Ratios Bonus	OLS Payoff
BonusPeer	-0.892* (0.476)	0.410*** 0.125	-2.956** (1.358)	0.052** 0.071	-3.586** (1.455)
FractionSuccess	-0.006 (0.005)	0.994 0.004	-0.018 (0.013)	0.982 0.014	-0.025 (0.018)
FractionSuccess x BP	0.018** (0.008)	1.019*** 0.006	0.064*** (0.024)	1.066*** 0.026	0.075*** (0.027)
constant	1.294*** (0.157)		-0.444 (0.443)		16.367*** (0.657)
N	110	110	110	110	110
χ^2	6.282	11.594	7.566	8.869	
P	0.099	0.009	0.056	0.031	0.060
Pseudo R^2	0.021	0.021	0.065	0.065	
F					2.543
R^2					0.058

Notes: Dependent variables are *Puzzles* in models (1) and (1'), *Bonus* in models (2) and (2'), and *Payoff* in model (3). Robust standard errors are reported in parentheses. Variable *FractionSuccess* is the fraction of participants who receive the bonus as estimated by the matched peer of the observed subject. The variable *BonusPeer* is equal to 1 if the matched peer has received the bonus and zero otherwise. "FractionSuccess x BP" stands for the interaction term between *FractionSuccess* and *BonusPeer*. *, **, *** indicate significance at the 10%, 5% and 1% level.

Table 4: Effect of BonusPeer at Different Levels of FractionSuccess

	FS0	FS10	FS20	FS30	FS40	FS50	FS60	FS70	FS80	FS90	FS100
Puzzles											
Poiss. IRR	0.410*** 0.125	0.493*** 0.127	0.592** 0.127	0.712* 0.127	0.856 0.132	1.028 0.153	1.236 0.201	1.486** 0.286	1.786** 0.414	2.147*** 0.594	2.581*** 0.838
Bonus											
Logit OR	0.052** 0.071	0.098** 0.114	0.186* 0.178	0.353 0.277	0.668 0.441	1.263 0.777	2.391 1.586	4.524* 3.565	8.561** 8.22	16.202** 18.772	30.661** 42.07
Payoff											
OLS	-3.586** (1.455)	-2.841** (1.223)	-2.095** (1.011)	-1.35 (0.836)	-0.605 (0.724)	0.14 (0.707)	0.885 (0.79)	1.631* (0.947)	2.376** (1.149)	3.121** (1.375)	3.866** (1.616)

Notes: Dependent variables are *Puzzles* in the first row, *Bonus* in the second row, and *Payoff* the third row. The columns stand for the effect of the bonus of the matched peer when her estimated fraction of subjects who reach the bonus is 0, 10, ..., 100. As detailed in the main text, we obtain these effects by using the variable *FractionSuccess_k*, which is defined as *FractionSuccess* minus constant *k*, and considering the coefficient for *BonusPeer*. The coefficients in the first row report the incident rate ratios (IRR) of the Poisson regressions that correspond to model (1') in Table 3. The coefficients in the second row report the odd ratios of the logit regressions that correspond to model (2') in Table 3. The coefficients in the third row report the marginal effects of the ordinary least squares (OLS) regressions that correspond to model (3) in Table 3. Robust standard errors are reported in brackets. *, **, *** indicate significance at the 10%, 5% and 1% level.

Table 5: Outcome Variables by Recommendation and Bonus of the Matched Peer

Bonus of Peer Recommendation by Peer	Any		No Bonus		Bonus	
	Give Up	Stay	Give Up	Stay	Give Up	Stay
N	68	53	56	16	12	37
Puzzles	2.81	3.47	2.94	3.04	1.75	3.7
Bonus	26.47%	37.74%	30.36%	31.25%	8.33%	40.54%

Notes: Means of outcome variables *Puzzles* and *Bonus*. Columns distinguish cases by whether the matched peer has reached the bonus and by the peer’s recommendation. The coding of the chat for the recommendation entails three categories: “stay in the task”, “work first on the task and quit later”, and “quit immediately”. The binary variable *Recommendation* simply distinguishes between the first (“Stay”) and the two latter categories (“Give Up”).

Table 6: Puzzles by Reputation

Puzzles	N	mean	Sd	p50	Frac >0	Frac >1
“observed leaders” (oL)	84	3.83	2.37	4	90.5%	81.0%
Others (Fo, uL, Fu)	251	3.22	2.56	3	76.9%	63.8%
“followers of observed leaders” (Fo)	83	3.13	2.54	3	79.5%	63.9%
“unobserved leaders” (uL)	84	3.38	2.54	4	77.4%	67.9%
“followers of unobserved leaders” (Fu)	84	3.13	2.62	4	73.8%	59.5%
Total	335	3.37	2.52	4	80.3%	68.1%

Notes: The variable *Puzzles* captures the number of subtasks solved correctly. The columns “Frac > 0” and “Frac > 1” report the fraction of observations that solved more than zero, respectively more than one, subtasks correctly. Rows distinguish observations by treatments.