Teamwork revisited: Social preferences and knowledge acquisition in the field

Frauke von Bieberstein, Jonas Gehrlein, Anna Güntner*

Abstract

Combining a lab-in-the-field experiment with field data, we study the effect of social preferences on performance in a modified teamwork setting, where public good production constitutes a prerequisite for individual performance but is not a goal in itself. Examples of such modified team settings are knowledge sharing, peer coaching, and helping on the job – all highly relevant topics for organizations today. As opposed to a standard public good setting, we find that teams with conditional cooperators are not more successful jointly or individually. In contrast, selfish individuals tend to perform better individually, without generating negative externalities for their team partners.

Keywords: teamwork, social preferences, knowledge sharing, helping on the job, lab-in-the-field experiment

JEL classifications: D90, M12, M21

*University of Bern, Institute for Organization and Human Ressource Management, Engehaldenstr. 4, 3012 Bern, Switzerland; corresponding author: Frauke von Bieberstein, frauke.vonbieberstein@iop.unibe.ch, +41 31 631 5464; Jonas Gehrlein: jonas.gehrlein@iop.unibe.ch; Anna Güntner: anna_guentner@mckinsey.com

1 Introduction

Teamwork is key for organizations today (Lazear and Shaw, 2007, Delarue et al., 2008). From an economic point of view, teamwork is a public good, where individuals jointly manage an asset (the team production) from which all of them benefit, but that requires costly and unobservable individual resources to build or to sustain (Alchian and Demsetz, 1972, Holmstrom, 1982, Rob and Zemsky, 2002). Group members' social preferences strongly influence contributions of individual resources to the public good. Production is more successful when the group counts conditional cooperators – individuals who behave reciprocally – among its members (Fischbacher et al., 2001). This has been shown in the lab, where the common workhorse for these studies is a repeated public good game (Isaac et al., 1984), as well as in the field, where for instance groups with a larger share of conditional cooperators are more successful in forest commons management (Rustagi et al., 2010).

Given the above findings, it seems that firms and organizations relying on teamwork should try to hire as many conditional cooperators as possible. However, in practice a widespread type of teamwork exists that is not fully captured by a public good setting: exchanges among workers with the objective to share knowledge and to help each other. This type of teamwork requires costly and unobserved individual effort but is often unincentivized because it does not directly produce tangible team output. Instead, teamwork in this modified sense creates a favorable environment for individual capability building where *in-*

dividual performance *subsequent* to team production is measured and incentivized. One situation of this kind is the exchange of best practice (internal knowledge) in firms (Tsai, 2001, Szulanski et al., 2016) – a crucial factor for generating and continuously renewing competitive advantage (Szulanski, 1996). The modified team setting also covers helping, peer coaching, and cross-functional thinking – all important topics for organizations today.

In this article we examine whether the positive effects of prosocial behaviors like conditional cooperation carry over to the modified team setting where individual output is measured and incentivized. On the one hand, prosociality could still be favorable given that on-the-job helping and knowledge sharing requires an investment of time and effort by the person who helps, without the certainty of receiving help back in the future. Helping thus resembles a public good situation. On the other hand, compared to classical teamwork, it is not clear that maximum help is always optimal in our modified team setting, because helping the team partner can take away valuable resources from one's own production. Due to their sense for reciprocity, conditional cooperators might be at risk of helping too much (even if help is inefficient), if the team partner has been helpful in the past. In contrast, a free rider might choose more strategically when and how much to help.

We study the role of social preferences in such a modified setting, where team production constitutes a prerequisite for subsequent quantifiable individual performance. We are able to control for individual ability and to obtain estimates for the intensity of group exchanges. The subjects of our study are mathematics freshmen enrolled in a challenging mandatory course of the curriculum. Learner-learner exchanges constitute a core component of knowledge acquisition according to the educational standards in mathematics (Ernest, 2010) - an idea communicated to the student by the faculty. In order profit from positive effects of learner-learner exchanges (Eisenkopf, 2010), students in our study are required to hand in weekly homework assignments in pairs. We measure students' social preferences in a public good game (lab-in-the-field), using the standard procedure from Fischbacher et al. (2001). In addition, we measure initial mathematical ability and collect data on team performance (time spent on the task jointly and team members' self-reported levels of satisfaction with joint work). Our main outcome variable is individual performance in the final exam.

The public good game is especially suited to illustrate the situation our student subjects encounter during the work on the assignments: Since these assignments are of considerable difficulty, team production is important to obtain the necessary skills to solve the assignments and to ultimately perform well individually in the exam. The public good is reflected in situations where one student already obtained the necessary skill to handle a specific problem but the other did not. As each team receives the grade on an assignment collectively, there is no incentive to spend additional time to help a struggling team member instead of just writing down and handing in the solution. However, over time, if helping is efficient, mutual help will maximize the skills of both team partners. Thus, taken together the public good in our setting is the sum of individual skills obtained from teamwork.

We find that teams with conditional cooperators are not more successful jointly or individually in the modified setting. In contrast, free riders have a significantly better individual performance in the final exam than other social types, controlling for individual ability. Furthermore, in our study free riders do not generate externalities, negative or positive, on their partners' individual performance. The most likely explanation for our findings is that, contrary to classical teamwork, maximum help may not be socially optimal in our modified setting. A student might spend so much effort on helping her partner, that there might not be enough time left to proceed oneself to the next level of understanding. This is especially true for conditional cooperators, who might feel compelled to help their team partner a lot if the team partner has been helpful in the past, even if helping is inefficient. In contrast, a free rider may be especially good at deciding when and how much to help in order to secure the team partners' help in the future without losing too much time.

2 Related literature

For our modified team setting, the economics literature has mostly examined the optimal task structure of specialization versus teamwork (Itoh, 1991) and optimal incentives for helping on the job (Drago and Garvey, 1998, Siemsen et al., 2007). Helping behavior in teams has been examined when agents have career concerns (Auriol et al., 2002), helping and sabotaging has received special attention in literature on tournament incentives (Lazear, 1989, Kräkel, 2005). All of these models assume rational agents who are only interested in their own economic well-being. Helping in this context mostly emerges due to incentive schemes that are based on joint production.

The management literature, on the other hand, has devoted substantial interest to the modified team setting. For instance, Flynn (2003) examines the effect of helping behavior on the job on individual productivity. He finds that employees who maintain an equitable balance between helping and being helped have the highest individual productivity. Also, the vast literature on organizational citizenship focuses specifically on behaviors that are not part of one's narrow job description, but benefit the organization as a whole (Organ, 1988, Podsakoff et al., 2000, DiPaola and Tschannen-Moran, 2014).

The effect of social preferences on contributions in the public good setting has been studied extensively, mostly in the lab (Fischbacher et al., 2001, Fischbacher and Gächter, 2010, Chaudhuri, 2011). A common finding is that groups that count conditional cooperators among its members produce higher contribution levels. There are several reasons for this finding. First, conditional cooperators contribute to the common project if they perceive that others are contributing as well. Second, they motivate selfish members, who rationally expect contributions from conditional cooperators in return, to participate more in the common project, at least in the first periods of a repeated game. Finally, the effects are amplified when punishment opportunities are given, as conditional cooperators negatively reciprocate free riding, punishing it even at a personal cost (Fehr and Fischbacher, 2004).

There is ample evidence that social preferences are related to behavior in real life. Less selfish individuals are more likely to donate to charity (Benz and Meier, 2008) and to participate in crowd-sourcing, such as Wikipedia (Algan et al., 2013). Gneezy et al. (2015) find that fishermen who must rely on teamwork due to environmental factors show more pro-social preferences across a range of experimental games than their neighbors who work individually. A detailed overview of literature on the generalizations of social behavior from the lab to the field, including its limitations, is given in Burks et al. (2016). This study shows that US truck drivers' social preferences are related to their behavior towards peers, but not towards experimenters, with whom they have less social connections.

Conditional cooperation (reciprocity) is particularly relevant to determine real-world outcomes. For example, cross-country skiers contribute more to the preparation of tracks if they believe that others do so, too (Heldt, 2005). Students' donations to a scholarship correlate with their beliefs about others' donations (Frey and Meier, 2004). Reciprocity measured as second mover behavior in a trust game (trustworthiness) is related to sales people's choice of selling strategy and success (Essl et al., 2018). Finally, according to a large scale survey with employees in Germany, reciprocity influences tax morale (Frey and Torgler, 2007) and effort exerted at work (Dohmen et al., 2009).

3 Study design and data collection

The data was collected during the winter semester 2015/2016 in the course *Analysis I* at a major university in Germany. The 10 ECTS course is mandatory for freshmen and is the main course in mathematics in the first semester.[1] The course consists of two lectures per week as well as a weekly tutorial in which take-home assignments are discussed. Class attendance is not compulsory. Exam failure rates in mathematics typically lie at about 30 - 50% with 40 - 60% of students dropping out of university during the first few semesters. Mathematics is thus generally a challenging university subject and *Analysis I* is the first demanding course in mathematics.

To encourage learner-learner exchange, students are given weekly take-home assignments to solve in teams of two. The problem sets are designed to be difficult to the degree that most students cannot successfully solve them on their own. Accordingly, the instructor makes students aware of the importance of joint work as a prerequisite for a successful exam performance.

Each assignment is graded per team, meaning that both team members receive the same amount of points. At least 50% of the maximum attainable points must be obtained to gain exam admission. In our data set no team failed to receive exam admission by at least the last assignment and since solutions could always be copied from other teams, the team output is arguably not externally incentivized. The required time, however, to reach the necessary points varies considerably. This

[1] The curriculum of the first semester consists of 31 ECTS with a total of 19 ECTS in mathematics and 12 ECTS in related subjects.

opens the possibility to judge the quality of team performance by the number of weeks needed to gain admission to the exam.

We conducted a lab-in-the-field experiment with the students in order to assess their social preferences. In addition, we were given access to pseudo-anonymized data of individual exam grades, team composition and scores obtained on homework assignments during the semester.[2] Figure 1 illustrates the timing of events and data collection.

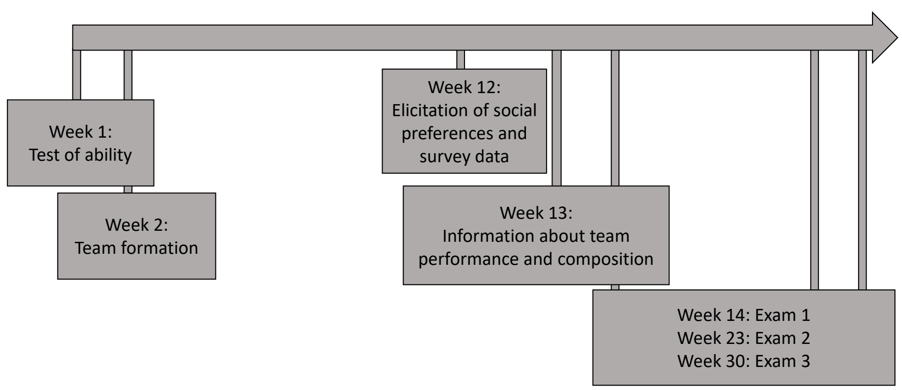

Figure 1: Timing of the study

In the first lecture, students took a test of high school math. The grade on this test was used as the measure of mathematical ability. Controlling for ability is important to avoid endogeneity in later analy-

[2]The students agreed that their matriculation number would be matched with their player number and stored physically at the university where the study was conducted. The list (or key) was used only once at the end of the semester to match the experimental data with the respective course data. Students furthermore were ensured that the key would be erased by the end of the study.

ses, as students with higher ability are more likely to earn better course grades.

In the second lecture students were instructed by the lecturer to form teams of two. These teams had to turn in the homework assignments together. Due to fairness concerns and as requested by the instructor, we had to keep the team formation procedure unchanged compared to previous semesters. Thus, we were not allowed to manipulate the formation of teams. However, teams formed at the beginning of the semester and students had to sign-up immediately after class with their partner. This reduces the probability that students selected their partners strategically. In our survey, 93% of students stated that they had not known the team partner before their studies.

In week 12 of the semester, all students present in the lecture participated in a classroom paper-pen public good experiment following the procedure in Fischbacher et al. (2001). In total 111 students out of the 158 enrolled students (roughly 70%) were present in this lecture and also participated in the initial test of ability. All students present chose to take part in the study.[3] Participants decided on an unconditional contribution and a full (conditional) contribution table. Every participant was endowed with €5 (10 points). Integer contributions between 0 and 10 points were allowed. In order to capture the context of the take-home assignments, experimental groups consisted of two members. The efficiency factor of $\alpha = 0.7$ guaranteed a sufficient balance between individual and collective utility. To minimize

[3] One student refused to share his/her pseudo-anonymized exam data and one failed to fill out the questionnaire. Both students' answers were excluded from the analysis.

social concerns towards their peers, students were informed that their decisions would be matched with another student from a different university.[4] After the experiment, students filled in a survey, stating how many hours per week they generally spend on the take-home assignment working individually and in the team measured in 5-hour intervals and how satisfied they were with the team collaboration, their partners' competence and engagement. In addition, students entered their demographic data and indicated when they had first met their team partner. Finally, students signed a consent form allowing us to use their pseudo-anonymized data in the study.

There were three exam dates in weeks 14, 23, and 30, respectively. Students were free to register for one of these three dates (in case of failure they could retake one of the subsequent exams) or they could even postpone the exam to the next semester. All tests were designed to be equally difficult. Table 1 summarizes the measures used in the study.

Table 1: The dependent and independent variables of the study

	Variable	**Measure**
Dependent	Individual performance	Points earned in the exam (min. 0; max. 60)
	Self-stated satisfaction with teamwork	Self-stated satisfaction with partner's engagement, competence and team cooperation in general
	Team performance	Number of take-home assignments required to achieve the admission to the exam (min. 6; max. 13)
Indep.	Ability	Test of mathematical skills in the first lecture
	Social preferences	Experimental measures from public good game
	Individual effort, team effort	Self-reported hours spent on the assignments weekly individually and jointly with the team partner

[4] We collected the matched data two weeks later and organized payments shortly after.

158 students were enrolled in *Analysis I*. For 111 students, we have the measure of ability from the first lecture and the behavior in the lab-in-the-field experiment.[5] For 85 of these students, we additionally have the experimental measures of their team partner. Both types of observations receive attention in our main analysis. 76 observations contain individual performance, ability and experimental measures of *both* team members. This enables us to assess the quality of teamwork by perceived competence, engagement and overall satisfaction of a team member, while controlling for ability.

4 Research hypotheses

We assume a simple knowledge acquisition function where knowledge k_i for individual i is a function of *i's* ability a_i and the effort that i spends on homework assignments. This effort can be either invested in individual problem solving ps_i or in helping the team partner h_i. Both types of effort increase own knowledge, however, problem solving is more effective individually while helping also benefits the team partner's knowledge. Finally, parts of the total available time T that each individual has at its disposition can also be spend completely unproductively on leasure l_i. Productive effort is thus given as

$$ps_i + h_i = T - l_i. \tag{1}$$

Taken together, the knowledge acquisition function looks as follows:

[5] Note that attendance in the lectures is not mandatory.

$$k_i = a_i + ps_i + \alpha \sum_{j=1}^{2} h_j, \qquad (2)$$

where α is the Marginal Per Capita Return (MPCR) of the public good, indicating the efficiency factor of helping each other. In common public goods with $n = 2$ individuals it is usually assumed that $0.5 < \alpha < 1$, indicating that contributions are individually detrimental (because effort would be more productively spend on ps_i) but collectively efficient.

We start with standard predictions regarding the effect of ability and effort on performance ($\frac{\partial k_i}{\partial a_i} > 0$, $\frac{\partial k_i}{\partial ps_i} > 0$, $\frac{\partial k_i}{\partial h_i} > 0$).

H1: Ability and effort positively influence performance in the final exam.

- Subjects with higher ability have a higher individual performance.
- Effort devoted to teamwork and individual work increases individual performance.

Now we turn to the hypothesis regarding the focus of our study - the effect of social preferences on performance in the modified team setting where team-production is not a goal in itself but a prerequisite for individual performance. In our setting, helping becomes relevant when one team partner knows how to approach a problem and the other does not. This situation resembles a public good game: Helping the team partner is individually costly because it takes away valuable resources from own production ($\alpha < 1$), but if both partners help each other the total knowledge of the team members increases ($n\alpha > 1$). It has been shown that production of a public good is more successful when the group counts conditional cooperators (Fehr and Fischbacher, 2004, Rustagi et al., 2010), because conditional cooperators contribute if they think the other group members are contributing as well and they are willing to punish deviators even at a personal cost. In our case of teamwork, punishment could mean withholding help for the team partner in the future or it could mean fewer social interactions with the team partner.

Contrary to the public good setting, in our setting individual performance can be measured and is the ultimate goal of the principal (the instructor). Given the above argumentation, conditional cooperators and their team partners invest more effort into helping each other (higher h_1 and h_2). Assuming that helping is efficient ($\alpha > 0.5$), conditional cooperators and their team partners should thus be more successful in the final exam compared to other social types.

H2: Individual performance.

- Conditional cooperators' individual performance is higher than that of other social types, controlling for ability.

- Conditional cooperators' partners' individual performance is higher than that of other social types, controlling for ability.

Finally, we predict that this positive effect on individual performance also translates into a higher satisfaction with teamwork itself.

Hypothesis 3 refers to literature about intrinsic satisfaction of cooperation between workers. Several studies show both theoretically and empirically, that workers value team production in a cooperative environment (Kosfeld and von Siemens (2011), Rabin (1993), Hamilton et al. (2003)).

H3: Effect of social preferences on the quality of teamwork.

- Individuals whose team partner is a conditional cooperator are more satisfied with their partners' engagement, competence, and with the overall collaboration than individuals paired with a person of any other social type.

The following section contains analyses evaluating these hypotheses.

5 Results

Figure 2 shows the distribution of social preferences and ability among the 111 subjects. 48.7% ($N = 54$) are classified as conditional cooperators following the definition by Fischbacher et al. (2001). 17.1% ($N = 19$) contribute 0 in every conditional decision and are classified as free riders. Around 17.1% ($N = 19$) show a hump-shaped contribution pattern. 2 subjects (1.8%) altruistically contributed all their endowment independently of the decision of the counterpart. For 15.3% ($N = 17$) social preferences cannot be classified according to standard definitions.

Ability is coded discretely between 1 and 5, 1 being the lowest and 5 the highest grade. A Kruskal-Wallis rank test[6] cannot reject the hypothesis that the distribution of ability is not different across social types ($\chi^2_4 = 3.050$, $p = 0.55$), i.e., social preferences do not explain differences in ability.

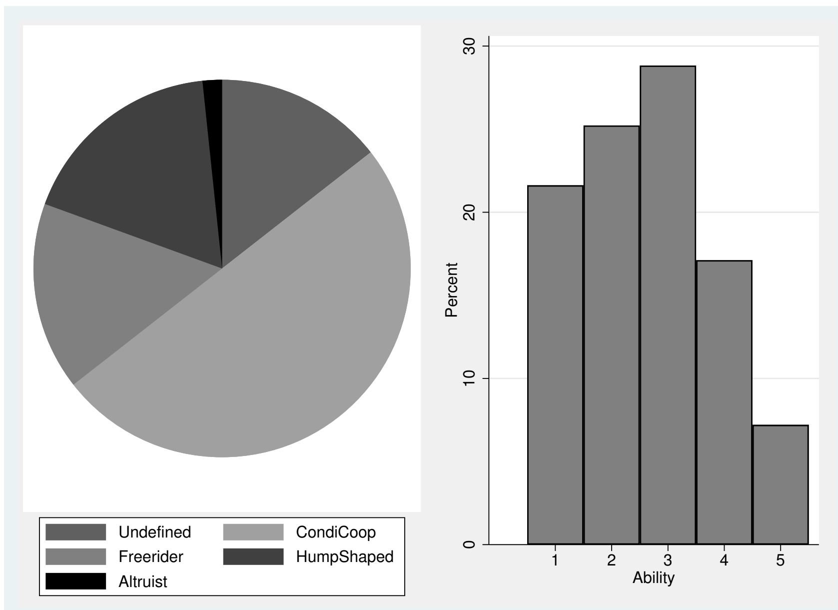

Figure 2: Distribution of ability and social preferences among subjects. Classification of social preferences according to Fischbacher et al. (2001)

Table 2 summarizes the number of students participating in each exam, the average number of points for those who passed (our main depended variable), and the percentage of students who failed the exam.

[6] All tests are two-sided.

Table 2: Results of the exams. Of the 158 students 8 did not sign up for any exam and 44 failed, of which 43 students took the exam again. In the column *Total* each student is counted once. Exams are designed to be of comparable difficulty. Differences in failure rate can be explained by selection effects.

	Exam I	Exam II	Exam III	Total
No. of students	8	127	58	150
Average points by students who pass	38.8	36.1	31.0	34.9
Failure rate	37.5%	41.7%	53.4%	29.3%

Table 3 contains the main regression analyses of this study related to hypotheses 1 and 2. The dependent variable is individual performance measured by the amount of points obtained in the exam, distributed between 0 and 60, with 30 points as the minimum amount necessary to pass.[7] In model (1) the independent variables are ability and the self-reported number of weekly hours spent on the home work individually as well as working together with the team partner.[8] For both effort variables the category “less than 5 hours per week” serves as reference. Model (2) considers the effect of the student’s social type and model (3) the effect of the team partner’s social type. In all cases the undefined social types serve as the reference category.[9] Changing the reference type to all non-freerider yield qualitatively similiar results.

[7] We use tobit regressions to fit the data to the lower and upper bounds as given by the exam. Coefficients and significance are qualitatively the same if we use an OLS regression instead.

[8] Table 5 in the Appendix shows the same regression with combined effort levels. Results remain qualitatively the same with more effort implying a higher performance on the exam.

[9] Two students classified as altruists are also part of the reference category. Regressions, in which these two observations are part of a separate category or in which these are dropped yield similar results to those presented here.

Table 3: Tobit regressions on the exam score; robust standard errors in parentheses; lower and upper bounds in tobit 0 and 60 respectively.

	(1) Effort&Ability	(2) (1)+Type	(3) (2)+Partner's type
Ability	3.703***	3.794***	4.039***
	(0.985)	(1.007)	(1.091)
Team effort: 5 - 10 hrs	-0.426	-0.679	3.948
	(2.586)	(2.603)	(2.724)
Team effort: > 10 hrs	6.386	7.265*	9.047**
	(3.910)	(3.666)	(3.905)
Indiv. effort: 5 - 10 hrs	9.596***	10.08***	9.876***
	(2.824)	(2.616)	(3.379)
Indiv. effort > 10 hrs	-1.815	-0.396	7.783
	(4.849)	(4.978)	(5.102)
Free rider		9.793***	6.925**
		(3.267)	(3.474)
Conditional cooperator		4.249	3.058
		(3.427)	(3.385)
Hump-shaped		3.449	4.832
		(3.874)	(4.361)
Partner free rider			0.373
			(4.992)
Partner conditional cooperator			-1.580
			(4.015)
Partner hump-shaped			4.903
			(4.160)
Constant	17.33***	12.58***	10.95*
	(3.284)	(4.182)	(5.561)
Observations	111	111	85
Pseudo R^2	0.027	0.034	0.037

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Our first hypothesis predicts that the main output – individual performance – depends on one's ability and the effort invested in solving the assignments. Table 3 provides evidence in support of this hypothesis. The coefficient of *Ability* is significant in all three models. For example, according to model (1) a student who obtains a one point higher grade in the ability test in the first week receives around additional 3.7 points in the final exam. Regarding effort, it can be seen that individual effort is already useful at a lower level, while working together pays off in grades only after students spend more than 10 hours per week working together. This is a reasonable result, because individual task solving capacities are attained after a short phase in the beginning, whereas teams need to spend time with social interactions and overcoming communication problems at first. Afterwards, teamwork becomes extremely effective, with team members earning 7 to 9 additional points individually in the exam. Table 5 in the Appendix combines both types of effort for each student.[10] Here, in all specifications more effort is associated with a higher individual performance. Evidence can be summarized as:

Result 1: Ability and effort, both individual and joint, increase individual performance.

Our main hypothesis 2 examines individual performance of social types. According to models (2) and (3) in Table 3, conditional cooperators do not perform significantly better than other social types. In addition, in model (3) the coefficient of *Partner conditional cooperator* is negative and insignificant, meaning that working together with a conditional cooperator does not lead to higher individual grades.

[10]Different specifications such as controlling for the exam date leave the results qualitatively the same.

In contrast, free riders significantly outperform other social types, earning on average around 7 - 10 more points in the exam.[11] Next, we examine whether the success of free riders comes at the expense of their team partners. In model (3) the coefficient of *Partner free rider* is not significant, meaning that being on the team with a free rider does not lead to earning lower individual grades. Incorporating the partner's ability in Models (1) - (3) does not change the results, while the respective coefficient is small and insignificant in every model.[12] The main result of the study can be summarized as follows:

Result 2: Conditional cooperators do not perform significantly better in the exam compared to other social types. In addition, conditional cooperator's team partners also do not perform significantly better. In contrast, free riders produce significantly higher individual output than other social types and do not generate externalities for their partners. These results hold controlling for ability.

The third hypothesis concerns the quality of teamwork which is a precondition for individual success. Team partners' self-reported satisfaction with joint work serves as proxy for this evaluation. The satisfaction with the partners' engagement, competence, and with the overall collaboration is measured on a scale between 0 ("fully unsatisfied") to 3 ("fully satisfied"). Table 4 presents the regression results for each of these assessments.

[11] Results are similar if we pool all types that are not free riders in the reference group.
[12] Results available on request.

Table 4: Tobit regression on team partner's perceived competence, engagement and overall satisfaction; lower and upper bounds 0 "fully unsatisfied" to 3 "fully satisfied"

	(1) Competence	(2) Engagement	(3) Satisfaction
Partner conditional cooperator	0.742	0.256	-0.326
	(0.494)	(1.038)	(0.208)
Partner free rider	1.280**	-0.058	-0.120
	(0.540)	(1.054)	(0.305)
Partner hump-shaped	0.066	-0.777	-0.054
	(0.598)	(1.089)	(0.327)
Partner's ability	0.475***	0.502	-0.083
	(0.175)	(0.324)	(0.093)
Ability of person who assesses	-0.131	-0.309	0.135
	(0.160)	(0.291)	(0.110)
Constant	1.258***	2.971***	1.930***
	(0.456)	(1.091)	(0.288)
Observations	76	76	76
Pseudo R^2	0.061	0.027	0.017

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The team partner's assessment of conditional cooperators' competence is positive but not significant. In contrast, partners are significantly more satisfied with free riders' competence compared to the reference category of unclassified individuals ($p < 0.05$). This result holds when controlling for ability, i.e., ability is not the driving factor of the assessment. A student's ability itself is a significant factor of satisfaction with his or her competence as assessed by the teammate ($p < 0.01$). Regarding the assessment of engagement and overall satisfaction, there are no significant differences across the partner's social types. In line with these results, Table 6 in the Appendix further examines team effectiveness as the number of assignments required to achieve exam admission. It indicates that group composition in terms of social preferences does not affect team production. This suggests that conditional cooperators again fail to thrive in this modified team setting. Furthermore, it strengthens the result that free riders do not exert negative externalities on their team members.

Summing up the evidence of Tables 4 and 6, the following is stated:

Result 3: Free riders are, independent of their ability, perceived as more competent by their team partners. There is no such effect for conditional cooperators or other social types. The assessment of team partner's engagement and the overall collaboration does not depend on the partner's social type. Additionally, the team's composition with respect to social types does not influence team effectiveness.

6 Discussion and conclusion

The present study tests the effect of social preferences on performance in a modified setting. In this setting, collective work is not externally incentivized, but necessary for production of valuable individual output. This framework applies to a number of important situations in the organizational context, such as internal knowledge sharing, helping on the job, or peer coaching. The study context – a lab-in-the-field experiment with university students combined with field data – allows to reliably measure individual output (final grades), controlling for ability.

Our initial hypotheses regarding a higher achievement of conditional cooperators with respect to individual performance and quality of teamwork must be rejected. We do not find a higher performance in the exam for conditional cooperators or their team partners. Given that teamwork is a prerequisite for individual performance in our setting, it is thus not surprising that conditional cooperators are also not perceived as more engaged or competent by their team partners and their teams do not need less time to earn exam admission. In contrast, free riders perform significantly better in the final exam compared to other social types. As an additional result, in our study free riders do not generate externalities, negative or positive, on their partners' individual performance. Teams that include free riders do not need significantly more time to achieve joint production targets and team partners are significantly more satisfied with free riders' competence than with the competence of other social types, even after controlling for ability.

Overall, based on our data we conclude that when team exchanges are just part of a bigger picture, ultimately resulting in individual production, conditional cooperators' presence has no positive effect and free riders' presence does not bear a negative effect on team production, as it does in classical teamwork. On the contrary, free riders seem to be more successful in producing individual output without hurting others' performance or the joint performance of the team. A possible reason for the success of free riders could be that they have a tendency towards rationality, which is an advantage for math studies. This explanation is, however, unlikely to be decisive in our case, as free riders do not perform better in the initial test and we control for ability in our regression.

In our view, the most likely reason for our findings is that, contrary to classical teamwork, maximum help may not be socially optimal in our modified setting. A student might spend so much effort on helping her partner, that there might not be enough time left to proceed oneself to the next level of understanding. This is especially true for conditional cooperators, who might feel compelled to help their team partner a lot if the team partner has been helpful in the past. In contrast, a free rider may be especially good at deciding when and how much to help in order to secure the team partners' help in the future without losing too much time. This behavior could even result in free riders being perceived as particularly competent, if they sometimes explain their knowledge in detail and sometimes just state the solution to a problem. In contrast, conditional cooperators might feel compelled to help even if this is inefficient if the team partner has been helpful in the past. In particular, in the case of helping, teamwork might sometimes be efficient and sometimes inefficient (formally, sometimes $\alpha > 0.5$ and sometimes $\alpha < 0.5$). For instance, if one team member absolutely does not understand a solution, it might take inefficiently long to explain just the basics of the solution to this person. While a conditional cooperator might still feel compelled to spend this time on helping, a freerider might choose more strategically when and how much to help.

By showing that free riders may work more efficiently in a modified team setting, the study provides clear managerial implications. First, hiring people with selfish social preferences for jobs that involve the modified team setting may prove useful based on the results of the study. Also, for firms who like to capture the benefits of teamwork but at the same time aim to reduce the problem of free riding, it might be possible in some settings to transform teamwork in a modified team setting by adding an individually measurable output. That way, organizations would not spend resources to screen out free riders at the stage of recruitment or give up on team output lost due to free riding, but strategically embed workers based on their social preference into suitable team environments.

In the given setting, several factors beyond social preferences and ability can influence performance. For instance, the degree of dynamic inconsistency (impatience) may vary among students leading some of them to abandon coursework during the semester in favor of leisure and partying despite the wish to earn a strong grade (Augenblick et al., 2015). In addition, the level of trust towards the course instructor and his claim that teamwork is important, may lead students to take the homework assignments more or less seriously. We chose not to control for these factors for two reasons. First, because we are not aware of studies that show that they are related to social preferences. Second, we aimed to keep data collection simple to realize in a classroom environment during a lecture.

The study, of course, has several limitations. The analysis uses self-reported data on the team partner's assessment and on effort invested, which might be biased systematically by different social types. In addition, we were not able to use random assignment to teams. Although

93% of students claimed that they had not known the team partner before their studies, there might still be unobserved factors influencing team composition that we are not able to control for. Finally, since our study focused on only one course of the curriculum, we are not able to observe the effect of partnering with a free rider in a multitasking context, e.g., taking multiple courses into account. It may well be that students in teams with free riders must invest more effort to study for the selected challenging mandatory course and that their performance in other courses suffers. This is of course possible, however, unlikely, as reported time spent working with team was not lower for teams with free riders and free riders' engagement was not reported lower that the engagement of of subjects with other social preferences.

Future research on the modified team setting should use longitudinal data to study how the effects of social preference on performance develop over time and also take into account multitasking, as described in the previous paragraph. In addition, it would be interesting to study social preferences in the modified team setting with employees of a company where helping and knowledge sharing are important prerequisites for success. Finally, it would be interesting to conduct a laboratory experiment where in the first part, subjects' social types are elicited using the standard procedure from Fischbacher et al. (2001), and in the second part subjects play a repeated public good game where contributions are sometimes inefficient but still benefit the other player. According to our results, we would predict that conditional cooperators choose to contribute in the second part even in inefficient situations if the other player contributed in these situations in the past.

References

Alchian, A. A. and Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, 62(5):777–795.

Algan, Y., Benkler, Y., Morell, M. F., and Hergueux, J. (2013). Cooperation in a peer production economy experimental evidence from wikipedia. In *Workshop on Information Systems and Economics*, pages 1–31, Milan, Italy.

Augenblick, N., Niederle, M., and Sprenger, C. (2015). Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics*, 130(3):1067–1115.

Auriol, E., Friebel, G., and Pechlivanos, L. (2002). Career concerns in teams. *Journal of Labor Economics*, 20(2):289–307.

Benz, M. and Meier, S. (2008). Do people behave in experiments as in the field?–evidence from donations. *Experimental Economics*, 11(3):268–281.

Burks, S. V., Nosenzo, D., Anderson, J., Bombyk, M., Ganzhorn, D., Götte, L., and Rustichini, A. (2016). Lab measures of other-regarding preferences can predict some related on-the-job behavior: Evidence from a large scale field experiment. Working paper.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14(1):47–83.

Delarue, A., Van Hootegem, G., Procter, S., and Burridge, M. (2008). Teamworking and organizational performance: a review of survey-based research. *International Journal of Management Reviews*, 10(2):127–148.

DiPaola, M. and Tschannen-Moran, M. (2014). Organizational citizenship behavior in schools and its relationship to school climate. *Journal of School Leadership*, 11(5):424–447.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioural outcomes. *The Economic Journal*, 119(536):592–612.

Drago, R. and Garvey, G. T. (1998). Incentives for helping on the job: Theory and evidence. *Journal of labor Economics*, 16(1):1–25.

Eisenkopf, G. (2010). Peer effects, motivation, and learning. *Economics of Education Review*, 29(3):364–374.

Ernest, P. (2010). Reflections on theories of learning. In Bharath, S. and English, L., editors, *Theories of Mathematics Education*, Advances in Mathematics Education, pages 39–48. Springer.

Essl, A. M., von Bieberstein, F., Kosfeld, M., and Kröll, M. (2018). Sales performance and social preferences. Working paper, University of Bern.

Fehr, E. and Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, 8(4):185–190.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1):541–556.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.

Flynn, F. J. (2003). How much should I give and how often? The effects of generosity and frequency of favor exchange on social status and productivity. *Academy of Management Journal*, 46(5):539–553.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review*, 94(5):1717–1722.

Frey, B. S. and Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics*, 35(1):136–159.

Gneezy, U., Leibbrandt, A., and List, J. A. (2015). Ode to the sea: Workplace organizations and norms of cooperation. *The Economic Journal*, 126(595):1856–1883.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, 111(3):465–497.

Heldt, T. (2005). Conditional cooperation in the field: Cross-country skiers' behavior in Sweden. Working paper, Department of Economics and Society, Dalarna University, Sweden.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, 13(2):324–340.

Isaac, R. M., Walker, J. M., and Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice*, 43(2):113–149.

Itoh, H. (1991). Incentives to help in multi-agent situations. *Econometrica: Journal of the Econometric Society*, pages 611–636.

Kosfeld, M. and von Siemens, F. A. (2011). Competition, cooperation, and corporate culture. *The RAND Journal of Economics*, 42(1):23–43.

Kräkel, M. (2005). Helping and sabotaging in tournaments. *International Game Theory Review*, 7(2):211–228.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economy*, 97(3):561–580.

Lazear, E. P. and Shaw, K. L. (2007). Personnel economics: The economist's view of human resources. *Journal of Economic Perspectives*, 21(4):91–114.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington Books/DC Heath and Com.

Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., and Bachrach, D. G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management*, 26(3):513–563.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302.

Rob, R. and Zemsky, P. (2002). Social capital, corporate culture, and incentive intensity. *RAND Journal of Economics*, pages 243–257.

Rustagi, D., Engel, S., and Kosfeld, M. (2010). Conditional cooperation and costly monitoring explain success in forest commons management. *Science*, 330(6006):961–965.

Siemsen, E., Balasubramanian, S., and Roth, A. V. (2007). Incentives that induce task-related effort, helping, and knowledge sharing in workgroups. *Management Science*, 53(10):1533–1550.

Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal*, 17(S2):27–43.

Szulanski, G., Ringov, D., and Jensen, R. J. (2016). Overcoming stickiness: How the timing of knowledge transfer methods affects transfer difficulty. *Organization Science*, 27(2):304–322.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal*, 44(5):996–1004.

A Supplementary analysis

Table 5: Tobit regressions on the final score; lower and upper bounds in tobit 0 and 60 respectively. Table contains new categories for combined effort.

	(1) Effort&Ability	(2) (1)+Type	(3) (2)+Partner's type
Ability	4.160***	4.237***	4.397***
	(0.960)	(1.000)	(1.082)
Individual & team effort: 10 - 15 hrs	2.770	2.172	7.133**
	(3.054)	(3.010)	(3.484)
Individual & team effort: 15 - 20 hrs	6.750*	6.918*	8.271*
	(3.544)	(3.531)	(4.768)
Individual & team effort: > 20 hours	6.758	8.380	16.73***
	(6.123)	(5.924)	(4.367)
Free rider		8.929**	5.871
		(3.677)	(3.591)
Conditional cooperator		3.248	3.268
		(3.633)	(3.084)
Hump-shaped		0.934	3.664
		(3.844)	(4.280)
Partner free rider			-0.594
			(4.586)
Partner conditional cooperator			-3.437
			(3.255)
Partner hump-shaped			3.110
			(3.464)
Constant	15.70***	12.42***	11.50**
	(3.279)	(4.239)	(5.233)
Observations	111	111	85
Pseudo R^2	0.018	0.024	0.050

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6 investigates the impact of social preferences on team performance measured by the timing of admission to the exam. We note that this measure is not perfect since there are no incentives attached to early completion. Nonetheless, on average strong teams should earn the required amount of points earlier. The variance in the data (10 assignments on average needed to earn exam admission with a standard deviation of 1.5; minimum 6 assignments required) indicates that it can be used as a proxy for team performance.

Table 6: Tobit regression on number of assignments required to reach admission to exam. Negative coefficients mean earlier admission. Bounds are 6 and 13.

	Number of assignments
Ability: Both medium	-0.268
	(0.979)
Ability: Both high	-2.020**
	(0.806)
Ability: Low and high	0.05
	(0.668)
Ability: Low and medium	-0.207
	(0.722)
Ability: Medium and high	-0.591
	(1.160)
Individual & team effort: 10 - 15 hrs	-0.378
	(0.501)
Individual & team effort: 15 - 20 hrs	-1.644*
	(0.874)
Individual & team effort: > 20 hrs	-1.516
	(1.049)
At least one conditional cooperator	-0.072
	(1.242)
Both conditional cooperator	0.133
	(1.396)
At least one free rider and no conditional cooperator	-0.057
	(1.398)
Constant	10.590***
	(1.629)
Observations	38
Pseudo R^2	0.096

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$