RATIONALITY & COMPETITION
CRC TRR 190

Interpersonal Preferences and Team Performance: The Role of Liking in Complex Problem Solving

Timm Opitz (MPI-IC)

Discussion Paper No. 492

February 09, 2024

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Interpersonal Preferences and Team Performance: The Role of Liking in Complex Problem Solving

Timm Opitz[†]

February 2024

Organizations increasingly rely on teams to solve complex problems. The ability of teams to work well together is critical to their success. I experimentally test whether team performance is affected by whether team members like each other. I find that teams in which partners like each other do not outperform teams in which partners dislike each other. However, teams in which one partner likes the other more than the other perform best. The performance differences result directly from changes in collaborative behavior when learning the team partner's interpersonal preferences, not indirectly from interacting with different individuals. Participants do not anticipate this pattern and expect to be most successful in a team where partners like each other. This provides insights into how teams should be optimally composed, when self-selection may be detrimental to performance, and what information about others' interpersonal preferences should be revealed.

Keywords: *Interpersonal Preferences, Teamwork, Liking, Complex Problem Solving Non-Routine Tasks*

JEL Codes: *C92, D23, D83, D91*

[†] Max Planck Institute for Innovation and Competition, Munich, Germany and Department of Economics, Ludwig-Maximilians-Universität (LMU), Munich, Germany; email: timm.opitz@ip.mpg.de

I would like to thank Vojtěch Bartoš, Florian Englmaier, Sebastian Erhardt, Svenja Friess, Dietmar Harhoff, David Heller, Elisabeth Hofmeister, Friederike Reichel, Cristina Rujan, Christoph Schwaiger, and seminar audiences at LMU Munich and Max Planck Institute for Innovation and Competition for their helpful thoughts and comments. I thank Margarita Gatsou for her excellent research assistance. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119), and the Max Planck Institute for Innovation and Competition is gratefully acknowledged. I thank the Munich Experimental Laboratory for Economic and Social Sciences (MELESSA) for providing laboratory resources. This experiment was pre-registered as AEARCTR-0011617 and received ethics approval at GfeW (approval number wxSdGh6x). Declaration of Interest: none.

1 Introduction

Teamwork is increasingly important in modern organizations. In addition, today's work tasks are more complex, more abstract, and more non-routine (Autor, Levy, & Murnane, 2003; Autor & Price, 2013). These tasks typically require workers to apply problem solving skills, rely on intuition, and be creative. Especially in complex tasks, teams have the potential to perform more efficiently than individuals due to skill complementarities (e.g., Ichniowski & Shaw, 2013). However, for these to materialize, teams must assign responsibilities, collaborate, communicate effectively, share information, and arrive at good solutions together (Deming, 2017). Thus, even "a team of experts does not necessarily make an expert team" (Salas, Reyes, & McDaniel, 2018).

Therefore, it is critical to understand how to make teams most effective when faced with complex problem solving. This requires an analysis of the environment in which teams operate, the skill composition of team members, as well as the collaboration within a team. Previous literature documents positive effects of incentive structures such as bonuses and tournaments on team performance (Englmaier et al., 2023a, 2023b), demonstrates the value of leadership (Englmaier et al., 2021), and studies optimal team composition based on individual characteristics in non-routine tasks (Hoogendoorn, Oosterbeek, & Van Praag, 2013; Hoogendoorn, Parker, & Van Praag, 2017; Hardt, Mayer, & Rincke, 2023). However, it remains challenging to predict how well a given team collaborates. Realizing the full potential of teams in non-routine tasks requires a deeper understanding of the determinants of successful team collaboration.

Interpersonal preferences of team members can play a crucial role for collaboration and team performance. Knowing that one is in a team with others who like them potentially increases satisfaction, facilitates helping behavior, or provides psychological safety to voice concerns and express opinions openly. All of these may contribute to a better *team spirit*[1], which is considered to be essential for team success and can explain why we often prefer to interact with those who also want to interact with us (Opitz & Schwaiger, 2023). In this paper, I examine whether teams perform better when team members like each other.

I conduct a laboratory experiment to test whether interpersonal preferences causally affect team performance in a non-routine task. The laboratory setting allows me to infer participants' interpersonal preferences and to manipulate information structures. My underlying hypothesis is that teams work better when the team members like each other. There can be two reasons for this. On the one hand, people who like each other may be more successful because of sorting. Given that perceived similarity has been shown to relate to interpersonal attraction (McWhirter & Jecker, 1967;

[1] *Team spirit* is defined as *The willingness [...] to work together and support[...] each other as part of a team* (New Oxford American Dictionary, 3rd ed.).

Montoya, Horton, & Kirchner, 2008), those who like each other may be more similar and therefore exert higher effort (e.g., Y. Chen & Li, 2009) or communicate more effectively. On the other hand, the feeling of being liked may causally change behavior. For example, people may adjust their behavior when they learn that their team members like them. They may provide higher effort, feel safer on the team, and ultimately become more productive and satisfied. In the experiment, I not only test whether those who like each other work better together, but also disentangle the underlying mechanisms by selectively revealing information about others' preferences.

The experiment consists of a team formation process and a non-routine team task that participants perform under one of two information structures. During the team formation process, participants indicate with whom they would like to interact in a payoff-irrelevant situation. They rank their potential partners on the basis of short personality profiles. Participants then perform a non-routine team task with a randomly assigned partner from those ranked before. This task requires participants to find a numerical solution to an abstract problem, often referred to as a *Guesstimation* task (Morgan, Neckermann, & Sisak, 2020).[2] Many companies use such tasks in job interviews because they demonstrate critical problem solving abilities and require coordination and communication –skills that have been identified to be key for the success in 21st century workplaces (Binkley et al., 2012; Aerlebaeck & Albarracin, 2019). In the baseline condition (*No-Info*), participants never learn how their partner ranked them. In the treatment condition (*Info*), participants receive information on how they are ranked by their matched partner before performing the *Guesstimation* task jointly. This allows me to compare whether teams where partners learn that they like each other perform better than teams who learn that the partners do not like each other. By contrasting behavior with the *No-Info* condition, I rule out the possibility that performance differences are due to people who like each other having *something* in common that might enhance performance. If this were the underlying mechanism, it would also operate when the partner's preferences are unknown. In this way, I establish that the disclosure of interpersonal preferences is the underlying mechanism.

I find that interpersonal preferences matter for performance in complex problem solving. While teams in which partners like each other perform similarly to teams in which partners dislike each other, teams in which one partner likes the other more than the other (*dissimilar liking*) perform best. Changes in collaborative behavior upon learning the preferences of the team partner are the causal mechanism underlying these performance differences. Participants do not anticipate this. Before the task, they expect to be more successful when being in a team where partners like each other due to higher effort provision. I present suggestive evidence that the performance differences are

[2] A classic example of a *Guesstimation* task (also referred to as a Fermi problem) is the question of the number of piano tuners in Chicago. The challenging thing about such problems is that individuals neither have direct empirical values from a similar problem, nor do they have the necessary data with which they could directly make a calculation.

driven by different communication patterns. In retrospect, participants evaluate the collaboration in a similar way depending on whether they were in a team where the partners liked each other, disliked each other, or had dissimilar preferences.

This paper contributes to three strands of literature. First, I add to an organizational literature on the self-selection of teams. Traditionally, management has been responsible for the composition of teams in firms. But there is an increasing development towards giving workers more flexibility in choosing their team partners –a situation that is the norm in other environments such as academic institutions or entrepreneurship.[3] Performance in self-selected teams can be higher because workers have better information on how to form teams effectively. On the other hand, workers may also form teams based on personal preferences that are not conducive to performance. Bandiera, Barankay, and Rasul (2013) illustrates this trade-off in a manual task by showing that team based incentives induce workers to team up with others of similar ability and forgoing the non-pecuniary benefit of working with their (lower-ability) friends. Through a structural model, Allocca (2023) shows higher performance of self-selected teams in a scientific organization. This is in line with findings of Boss et al. (2023), demonstrating the benefit of freely choosing team partners for a given pitch deck presentation task. Similarly, Fischer, Rilke, and Yurtoglu (2023) shows higher performance of self-selected teams in a task environment where the abilities of all team members play a significant role in the team production function. I contribute to this literature by showing that those who like each other do not necessarily perform better in complex problem solving, despite the expectation of higher team effort and better performance. This underscores the importance of understanding the dimensions on which individuals self-select into teams when assessing the effectiveness of giving them the freedom to do so.

Second, I relate to to the literature studying interpersonal preferences in teamwork. Gächter, Starmer, and Tufano (2023) demonstrates the importance of social relationships for team production. They find a strong positive association between group cohesion and performance in weak-link coordination games. This is consistent with evidence that social proximity can lead to higher levels of prosocial behavior, which is an important prerequisite for successful cooperation (Leider et al., 2009; Goeree et al., 2010). Opitz and Schwaiger (2023) use an approach similar to this paper in which interpersonal preferences are based on questionnaire responses within the experiment instead of relying on real-world friendship networks. They find higher cooperation in a Public Goods Game when participants are liked by their partner, and point to increased altruism and higher beliefs about the other's contributions as the underlying pathways.

[3]Some companies give employers complete flexibility in choosing projects and teams (e.g., Zappos), others allow employees to reshuffle teams in regular intervals (e.g., Microsoft), and still others have a long history of encouraging side-project time in self-selected teams (e.g., Google).

Third, the paper advances the literature on team performance in non-routine tasks. There is increasing interest in understanding the production function in complex work tasks through experimental approaches. Research questions range from the optimal team composition (Hoogendoorn, Oosterbeek, & Van Praag, 2013; Hoogendoorn, Parker, & Van Praag, 2017; Hardt, Mayer, & Rincke, 2023), to identifying team players (Weidmann & Deming, 2021), to effective incentive and governance structures (Morgan, Neckermann, & Sisak, 2020; Englmaier et al., 2021, 2023a, 2023b). The experimental approaches range from field experiments on entrepreneurship education programs and escape room games (e.g., Hoogendoorn, Oosterbeek, & Van Praag, 2013; Hoogendoorn, Parker, & Van Praag, 2017; Englmaier et al., 2023a) to laboratory experiments (Morgan, Neckermann, & Sisak, 2020) and online experiments (Hardt, Mayer, & Rincke, 2023). In this paper, I shed light on how interpersonal preferences affect collaboration of teams in a laboratory setting by employing a *Guesstimation* task, using chat communication.

The paper is structured as follows: Section 2 presents the experimental design, Section 3 outlines my hypotheses and results on interpersonal preferences in non-routine teamwork. Finally, I discuss and conclude in Section 4.

2 Experimental Design

Overview. The experiment consists of two parts. First, a team formation process. Second, a non-routine team task. During the team formation process, participants indicate with whom they want to interact in a payoff-irrelevant situation. This means that participants rank their four potential partners according to the desirability of interacting with them without this decision being tied to any potential payoff. This is how I operationalize *liking*. These interpersonal preferences are based on short profiles of the potential partners. The incentivized non-routine task is played with one randomly matched partner from those ranked. I analyze behavior under two information structures. In the baseline condition (*No-Info*), participants never know how much their partner likes them. In the treatment condition (*Info*), participants receive the information before the non-routine team task. I investigate whether performance in the non-routine task differs depending on how much team partners like each other, and I analyze different mechanisms of how this may affect team performance. The timeline of the experiment is visualized in Figure A1.

Team Formation. I form teams for the *Guesstimation* task in order to analyze whether performance differs depending on whether team partners like each other or not. I face two challenges in doing so. The first is to operationalize *liking*. Ideally, one needs to to elicit incentive-compatible

preferences about the desirability of interacting with each other, without these preferences being linked to the *Guesstimation* task. The second challenge is to create teams with different preference constellations. I intend to compare teams where both partners like each other, to teams where only one partner likes the other, to teams where both partners dislike the other. This can be achieved through randomization. By creating a payoff-irrelevant interaction, I solve both challenges.

Participants express their preferences for potential interaction partners based on seeing questionnaire responses of them. At the beginning of the experiment, participants complete a personality questionnaire. This contains 15 statements on personality traits, preferred leisure activities, and societal opinions that participants answer on a 4-point Likert scale (see Appendix A2). Intuitively, participants get an impression about each other based on seeing the responses to the questionnaire. They are then asked to rank their potential partners according to the desirability to interact with them, which I use as a proxy for *liking*. Participants have no further information about the payoff-irrelevant interaction the ranking refers to, so they do not know what the *interaction* will be about.[4] Their preferences therefore directly reflect which of their potential partners they enjoy interacting with the most and least.

Practically, participants are randomly assigned to groups of 8 players, whereby each group is evenly split into two *market sides*. Then the questionnaire responses to five randomly selected questions are shared between the participants across market sides. That is, each market side receives the answers to these questions from the four players on the other market side. Participants then rank the players based on their perception of the desirability of interacting with them.[5] The preference submission takes place via a strategy-proof mechanism (Abdulkadiroğlu & Sönmez, 1998). It is in the best interest of the participants to submit their true preferences.[6] This mechanism then determines the teams for the interaction.

Non-routine Team Task. The non-routine team task is a *Guesstimation* problem.[7] The goal is to estimate an unknown and highly unusual quantity by connecting a series of known or easily estimated quantities in a logical but non-routine way. Participants are paid for the accuracy of their response (Morgan, Neckermann, & Sisak, 2020). They are allowed to use a calculator, but are not given the possibility to search for additional information on the internet. To familiarize participants

[4]Ultimately, participants are asked to create three statements that they believe half of the student population would agree with and half would disagree with during this interaction.

[5]Participants have no additional information about each other such as gender, age, or socio-demographic characteristics. They only see the answers to the five questions.

[6]Participants are also informed that the indication of one's true preferences increases the likelihood of performing the interaction with this person. I use a random serial dictatorship mechanism. Thereby, participants are sorted and randomly listed in order. The first in line is then assigned their favorite partner, then it is the second's turn, and so on until the four two-player teams are formed in each group of 8 participants.

[7]See Appendix A3 for the set of problems.

with the task and to get a measure of individual ability, they first perform one *Guesstimation* task on their own.[8]

Participants work on the team task with one randomly selected team partner. The allocation mechanism randomly pairs individuals from each market side within the group of 8 players that ranked each other during the team formation stage. Thus, I create random teams out of players who indicated how much they *like* each other. Teams have 8 minutes to solve the joint task, and communicate via chat messages. For a team's answer to count, both team partners must enter the same solution. This requires close collaboration between the two partners.

Treatments. I assess whether interpersonal preferences affect performance in the non-routine team task. Participants perform the task under one of two information structures. They either learn how much their team partner likes them or they do not. That is, in one condition, participants learn the rank their partner placed them on for the interaction, in the other condition, they do not.

The treatment variation takes place precisely at the time of learning who the team partner is. In the experimental condition *No-Info*, participants learn who their team partner is. This includes a reminder of how they ranked their team partner, and how this partner answered the randomly chosen questions. In the experimental condition *Info*, participants additionally learn how their assigned partner ranked them. In this way, I reflect situations in which team partners like or dislike each other for reasons orthogonal to the specific task they perform. (here the *Guesstimation* task). Treatment is assigned at the group level. This means that either both team partners know about the other's preference, or neither partner knows.

Underlying Channels. I investigate the channels through which *liking* may affect task performance by capturing participants' ex-ante beliefs about their collaboration as well as their ex-post evaluation of it. Before the task, I assess whether participants expect to be more successful and enjoy the task more when they are in a team with partners who like each other. I distinguish between two motives for why participants may show a preference for working with someone they like and/or who likes them (Opitz & Schwaiger, 2023). On the one hand, a belief-based motive related to higher expected team performance. On the other hand, a preference-based motive about higher expected utility from doing the task itself. After the task, I ask participants about their collaboration experience and their willingness to continue working with their current partner, capturing both their success beliefs and their enjoyment of the task.

I elicit four ex-ante beliefs about the upcoming *Guesstimation* task once participants know who their partner is. First, they are asked about their motivation to provide high effort in the upcoming

[8]For the individual task, participants are given 5 minutes and can earn up to 5 Euro.

task on a scale from 0 to 10. Second, participants indicate their belief about their partner's effort. This belief is incentivized through the partner's self-reported effort on the previous question. If the respondent's assessment matches their partner's self-reported effort, they receive 2 Euro. If the assessment is not equal to the partner's report, they receive no payment. Third, participants provide their team performance beliefs. For this, I endow participants with 1 Euro. They can bet any amount (in increments of 10 Cents) on their own team's performance and keep the rest. If their team happens to be in the 25% best performing teams, the amount invested is quadrupled and paid out. If not, the investment is lost. In this way, they indirectly indicate their expected probability that their team will be among the 25% best performing teams.[9] Fourth, participants indicate their willingness to pay to perform an extra round of the task with fixed pay alone compared to with their partner.[10] This allows me to measure their expected task utility, the utility of doing the task with their partner regardless of any payoffs.

After the completion of the *Guesstimation* task and before any performance feedback, participants assess their collaboration experience on two different dimensions. First, they rate different aspects of their collaboration experience (see Appendix A2 for the questionnaire). Second, they decide whether they would like to perform an extra round of the task with performance-contingent payment with their team partner or alone. This binary decision captures both their (relative) performance beliefs as well as their task utility from doing the task with their team partner.

Of the two choices of whether to work alone or with your partner on an extra task, at most one will be implemented. There are three possible implementation procedures. First, participants do not do another round of the task (implemented with 90 percent probability). Second, participants do another round of the *Guesstimation* task with a fixed payoff. This is the scenario to which the ex-ante measure for expected "task utility" refers (5 percent probability). Third, participants do another round of the *Guesstimation* task with performance-contingent payoff. This is the scenario the ex-post measure for "task continuation" refers to (5 percent probability). In this way, I elicit incentive-compatible preferences over teamwork, but make sure that most participants do not have to do an extra round of the task. Given that there is an extra task, I follow the implementation procedure of Hardt, Mayer, and Rincke (2023): There is a 50 percent chance that participants get to work alone irrespective or their indicated preferences. Likewise, there is a 50 percent chance that

[9] I either pay out the beliefs about partner's effort or the beliefs about team performance with equal probability. I inform participants before the *Guesstimation* task which belief is drawn to be payoff-relevant.

[10] The multiple price list procedure to elicit the willingness to pay is displayed on Screen 10 in Appendix D.

their choice will be implemented. In particular, participants get to work together if both partners stated a preference for doing so.[11]

Experimental Procedures. The experiment was conducted at the Munich Experimental Laboratory for Economic and Social Sciences (MELESSA) in July 2023, with physical participation in the laboratory. In total, 240 student participants participated in the experiment across 13 sessions. The participants were recruited using the online system ORSEE (Greiner, 2015). The experiment was programmed with the software oTree (D. Chen, Schonger, & Wickens, 2016). Treatment was randomized within-session at the group level. On average, participants earned 19 Euro (including a show-up fee of 10 Euro). The experiment lasted about 55 minutes.

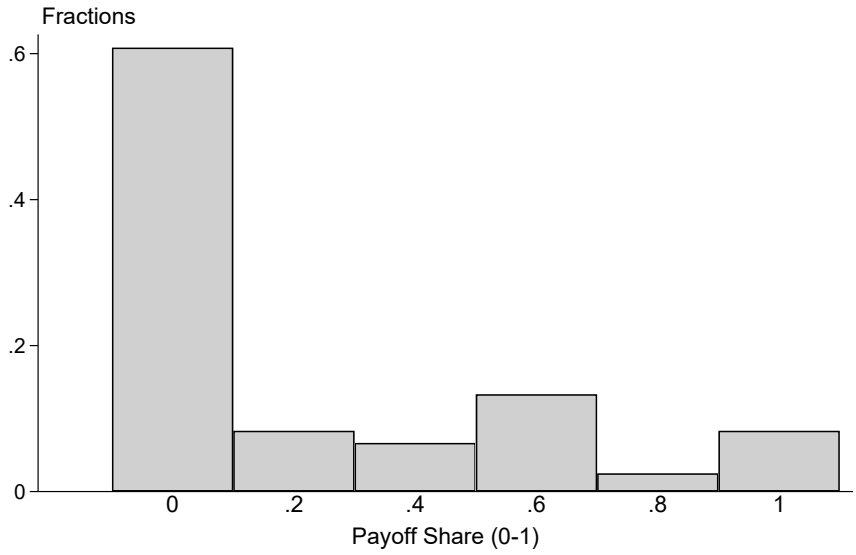
3 Main Results

3.1 Incentives and Task Performance

Participants could earn up to 10 Euro in the (collaborative) *Guesstimation* task. Following the incentive scheme of Morgan, Neckermann, and Sisak (2020), participants received the full payment if their answer deviated less than 10% from the true value. For deviations of up to 20% (40%/ 60%/ 80%) they received 8 Euro (6 Euro/ 4 Euro/ 2 Euro). For larger deviations or no response, they did not receive any payment.

Out of the 120 teams in the experiment, 72 submitted a valid answer within the given time. Those who did not submit a valid answer either failed to submit an answer within the time limit (33 teams), or gave different answers from each other (15 teams). From those who submitted a valid answer, around 65 percent achieved a positive payoff, while the rest of the teams submitted a solution that deviated by more than 80% from the true value. On average, participants earned 2.23 Euro from the collaborative *Guesstimation* task, Figure 1 shows the payoff distribution. The variable *Payoff Share* refers to the incentive scheme participants faced. For example, a value of 0.4 means that the participant received 4 out of the possible 10 Euro and implies that the answer deviated between 40% and 60% from the true value. Figure A.3 in Appendix C shows that performance across the five *Guesstimation* tasks was similar, ranging from averages of 0.19 to 0.24 in the payoff share.

[11] This requires first randomly selecting one of the rows in the multiple price list as the relevant one for the (ex-ante) assessment of expected task utility.



Notes. This figure displays the distributions of the payoff share in the collaborative *Guesstimation* task based on all 120 teams allocated to either the *Info* or *No-Info* treatment.

Figure 1: Payoff across Treatments

3.2 Estimation Strategy

I perform two different types of analyses to answer whether people who like each other perform better in complex problem solving and why. The first type of analysis examines behavior within the *Info* treatment. Conceptually, I compare teams where the partners like each other, to teams where the partners dislike each other, to teams where one partner likes the other more. This allows me to make general statements about which type of team performs best. However, this analysis does not disentangle whether the underlying mechanism is sorting or a causal effect of feeling liked by the team partner.

I analyze underlying mechanisms by comparing behavior across the treatments *Info* and *No-Info* in the second type of analysis. That is, I compare behavior of those who like each other (or not) across *Info* and *No-Info*. Thereby, I hold the underlying preferences constant, and isolate the effect of "knowing these preferences". For example, I can compare a team where both partners ranked each other first and knew it (in *Info*) with a team where both partners ranked each other first but did not know it (in *No-Info).* If people who like each other are, for example, more similar and this per se facilitates communication, these teams perform better even when participants do not know their partner's preferences (in *No-Info*). In contrast, finding that interpersonal preferences affect performance in *Info* but not in *No-Info* provides strong evidence for a causal effect of feeling liked and related changes in collaborative behavior.

For both types of analyses, I run two sets of regressions. In one set, I estimate the effect of the sum of both partners' preferences on the outcome.Preferences are a reverse-coded measure of the ranks they assign to each other. Setting a partner on the first and most favorable rank is converted to the highest possible preference, setting a partner on the last rank is converted to the lowest possible preference. In the other set, I add a measure of how dissimilar partners' preferences are. This allows me to analyze, for example, whether it makes a difference if both partners rank each other second, or if one partner ranks the other first and one partner ranks the other third in their preference list.

3.3 Liking and Task Performance

3.3.1 Liking within Treatment *Info*

Teams where partners like each other do not perform better than teams where partners dislike each other. Column (1) of Table 1 shows a negative but statistically insignificant impact of *joint liking* on performance. *Joint Liking* is defined as the sum of both partners' preference rankings. In Column (2) I add controls for team demographics and individual performance. It confirms a statistically insignificant but economically significant negative impact of about 18.5 percent on payoffs for each rank that one of the two partners ranked the other more favorably.

Teams with *dissimilar liking* perform better. I define *dissimilar liking* as the absolute difference between partners' preference rankings. The pre-registered analyses in Columns (3) & (4) indicate that teams where one partner likes the other more than vice versa are more successful. This also holds in Column (5) when adding fixed effects for the different *Guesstimation* problems. The coefficient for *joint liking* remains in the same magnitude when including *dissimilar liking* and only reaches marginal significance in Column (4). I summarize the main findings in Results 1 and 2.

Result 1 *Teams where partners like each other do not perform better than teams where partners dislike each other in Info.*

Result 2 *Teams where partners have dissimilar preferences perform significantly better in Info.*

In the next subsection, I analyze whether this pattern is driven by sorting. I show that partners who like (or dislike) one another do not perform differently per se. This is despite the fact that I observe homophily in the sense that those who are similar also like each other more (McPherson, Smith-Lovin, & Cook, 2001). The more similar a potential partner's answers are to one's own answers, the more favorable this partner is ranked ($r = 0.27$. $p <$0.001; Pearson's correlation coefficient). This implies that teams where both partners like each other are, on average, most similar

Table 1: Task Performance: Guesstimation [*Info*]

	Payoff Share [0-1]				
	(1)	(2)	(3)	(4)	(5)
Joint Liking (2-8)	-.032	-.037	-.035	-.043*	-.039
	[-.078,.013]	[-.085,.011]	[-.080,.010]	[-.089,.004]	[-.088,.010]
Dissimilarity Liking (0-3)			.063	.094**	.121**
			[-.018,.143]	[.008,.180]	[.024,.219]
Mean Dep. Var.	.20	.20	.20	.20	.20
Team Controls	No	Yes	No	Yes	Yes
Performance Controls	No	Yes	No	Yes	Yes
Problem FE	No	No	No	No	Yes
N	60	60	60	60	60

Notes. OLS Regressions.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The outcome variable is the fraction of the maximum possible payoff that teams achieved. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

in terms of their questionnaire responses. Theoretically, this can either improve performance by making it easier to communicate and share information, or it can be detrimental if similar partners have less complementary skills and information. Empirically, I show that sorting does not affect performance at large. Instead, the patterns in Results 1 & 2 stem from knowing how much their partner likes them.

3.3.2 Disentangling Selection and the Feeling of Being Liked

I show that the knowledge about the partner's preferences changes performance through comparing behavior across the treatments *Info* and *No-Info*. Table 2 corroborates the findings from analyzing the treatment *Info* in isolation. First, it shows that the preferences of team partners only matter in the *Info* treatment. If participants who like each other naturally worked better (or worse) together because they have similar communication patterns, thinking styles, or skills, this would also show up in *No-Info*.[12] Second, it confirms that participants who like each other do not perform better when knowing this information. The effect is consistently negative and statistically significant across Columns (1)-(3), showing that learning about each others' favorable preferences is detrimental to performance. Third, also this analysis supports that teams with dissimilar liking perform better in

[12]In Table A.1, I also analyze the treatment *No-Info* in isolation and consistently shows that this is not the case. Additionally, I show in Table A.15 that questionnaire responses and performance are largely uncorrelated. This means that participants cannot (objectively) identify who is good at the task individually (Column 1), whether someone is a good team partner in the task (Column 2), or whether someone is a good team partner given the available information about each others preferences (Column 3).

Info. Coefficients in Columns (2) & (3) are economically sizeable and statistically significant at the 10% level.

Table 2: Task Performance: Guesstimation [across Treatments]

	Payoff Share (0-1)		
	(1)	(2)	(3)
Joint Liking (2-8)	.034	.035	.036
	[-.012,.081]	[-.011,.081]	[-.011,.083]
Joint Liking X Info	-.068*	-.073**	-.072**
	[-.137,.001]	[-.142,-.005]	[-.143,-.001]
Info	.303*	.197	.170
	[-.050,.656]	[-.170,.564]	[-.218,.558]
Dissimilarity Liking (0-3)		-.040	-.038
		[-.145,.064]	[-.146,.070]
Dissimilarity Liking X Info		.124*	.134*
		[-.015,.263]	[-.011,.279]
Mean Dep. Var.	.23	.23	.23
Team Controls	Yes	Yes	Yes
Performance Controls	Yes	Yes	Yes
Problem FE	No	No	Yes
N	120	120	120

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The outcome variable is the fraction of the maximum possible payoff that teams achieved. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Joint Liking X Info* interacts this variable with a treatment indicator for *Info*. *Dissimilarity Liking* is the absolute difference between both partners' preferences, *Joint Liking X Info* captures the interaction with the treatment *Info*. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

Result 3 *Performance differences between teams with different interpersonal preferences are due to changes in behavior upon learning how partners like each other.*

I confirm these results in a robustness check where I correct for idiosyncratic mistakes when entering the final response. There are 12 teams that arguably coordinated on a common response, but one of the team members made a mistake when entering the solution. These mistakes are either typos or formatting errors. Table A.14 shows that the results on performance are qualitatively and quantitatively similar in this alternative specification where I assign them the payoff they would have received without the mistake.

Decision times provide additional support that performance differences are due to participants learning how much their partner likes them. I argue that behavior changes because the knowledge about interpersonal preferences changes collaboration. One alternative interpretation of the results is that participants become more interested in the characteristics of their partner when learning

how they were ranked. So they increase attention to their partner's questionnaire responses and adjust their beliefs and behavior due to more careful inspection.[13] Following this argument, I would expect participants in *Info* to spend significantly more time inspecting their partner's profile before proceeding to the task than participants in *No-Info*. However, participants in *Info* spend a similar amount of time on the screen where they learn who their partner is (24.8 seconds) compared to *No-Info* (22.5 seconds). This makes it unlikely that results are driven simply by increased attention to their partner's characteristics. Therefore it seems to be the knowledge about the partner's preferences that changes how well teams with the same underlying preferences perform. While some of these revelations increase performance, others decrease performance, resulting in very similar average performance across treatments (see Appendix B2 for a detailed analysis).

3.4 Underlying Channels

3.4.1 Ex-ante Beliefs and Preferences

I investigate whether participants' expectation about the upcoming task depend on how much they like their partner and how much their partner likes them in *Info*.[14] I separate beliefs about the efficacy of their upcoming interactions and preferences for working with their assigned partner.

First, participants expect their team to be more successful the more both team partners like each other. This is shown in Table 3. They especially expect to be significantly more successful with partners they like (Column 1). Together with the insignificant (but positive) point estimate on the preferences of their partner (Column 2), this translates into a significantly more positive evaluation of the team success when partners like each other. They also believe partners who like them to exert higher effort (Table A.4). In contrast, participants think that their own behavior will not be affected by either their own or their partner's preferences (Table A.2). In neither of these three assessments, participants consider the difference between their own and their partners' preferences to be meaningful.

Result 4 *Participants expect to be more more successful when being in a team where partners like each other. Whether or not their preferences align with their partner's preferences does not affect their expectations.*

Second, participants do not expect to receive a higher task utility from performing the *Guessti-mation* task in a team where partners like each other. They are equally willing to work alone in

[13] Holding such beliefs seems plausible, despite the fact that questionnaire responses are not correlated with actual performance as shown in Table A.15.

[14] In Appendix B3, I show how these measures compare between *Info* and *No-Info* and corroborate the findings of this Section.

Table 3: Team Success [*Info*]

	Expected Team Performance			
	(1)	(2)	(3)	(4)
Preference for partner (1-4)	.620***			
	[.177,1.064]			
Partner's preference (1-4)		.364		
		[-.099,.827]		
Joint Liking (2-8)			.447***	.442***
			[.147,.747]	[.140,.744]
Dissimilarity Liking (0-3)				.092
				[-.451,.636]
Mean Dep. Var.	7.358	7.358	7.358	7.358
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
N	120	120	120	120

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The outcome variable is the amount bet on the team being in the top 25%, with monetary values recoded on a scale of 0-10. The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Partner's preferences (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice.. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

a payoff-irrelevant task, irrespective of how much they like their partner (Table 4, Column 1) or how much their partner likes them (Column 2). They are not more prone to prefer working alone when there is a larger difference between their own and their partner's interpersonal preferences (Column 4).

Result 5 *Participants expect to derive a similar utility from performing the Guesstimation task with someone whom they like or who likes them. Whether or not their preferences align with their partner's preferences does not affect their expectations.*

Table 4: Task Utility [*Info*]

	Preference for Working Alone			
	(1)	(2)	(3)	(4)
Preference for partner (1-4)	-.002			
	[-.351,.347]			
Partner's preference (1-4)		.057		
		[-.299,.413]		
Joint Liking (2-8)			.024	.014
			[-.212,.261]	[-.223,.252]
Dissimilarity Liking (0-3)				.195
				[-.233,.623]
Mean Dep. Var.	5.475	5.475	5.475	5.475
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
N	120	120	120	120

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The outcome variable indicates the willingness to pay to perform another round of the task with a fixed payoff alone (as opposed to with their team partner). The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

3.4.2 Communication

Effective communication in sharing information, discussing approaches, and collaborating is key to success in the *Guesstimation* task. Because participants face limited time to solve the problem, there is a clear trade-off between working on parts of the task individually and communicating with their partner. I find strong evidence that extensive communication is not necessary for success. The correlation between the quantity of communication, measured by the total number of characters, and team performance is indeed negative in *No-Info* where preference revelation cannot affect communication patterns ($r = -0.344$; Pearson's correlation coefficient). Successful teams typically require less than 1000 characters to reach a solution. In contrast, less successful teams exchange up to 2000 characters. While this should not be taken as evidence that less communication is conducive to performance, it does show that coordination on a common solution can be achieved with little communication.[15]

Changes in collaboration behavior are the only channel through which the revelation of each other's preferences can have a causal effect on performance. Since collaboration only happens via chat, any changes *should* be observable in the (written) communication. Differences in

[15] Appendix Table A.13 shows in Column (1) that this also holds for the ex-post evaluation of the collaboration. How much teams communicate is not related to how positive participants evaluate the collaboration experience.

communication patterns may be due to changes in effort, shifting roles within the team, or changes in willingness to challenge the partner or compromise on a solution. Whatever the underlying reason for the differences in performance, it must operate through joint communication.

I find suggestive evidence that changes in communication patterns through the revelation of preference information drive the performance differences. Columns (1)-(3) of Table 5 consistently show that those who like each other communicate more in *Info*. Columns (1) & (2) also reveal that teams with *dissimilar liking* communicate less in *Info*, although the effect is imprecisely estimated, which is also the case when analyzing communication patterns in isolation in *Info* (Table A.8. Table A.9 shows no effects of the revelation of preferences on either communication asymmetry, turn-taking, or the time until the first message is sent.

Table 5: Team Communication [across Treatments]

	Total Length of Chat Messages		
	(1)	(2)	(3)
Joint Liking (2-8)	-38.819*	-39.233*	-42.086*
	[-84.267,6.628]	[-84.141,5.675]	[-87.848,3.677]
Joint Liking X Info	56.682*	61.796*	65.090*
	[-10.782,124.146]	[-5.048,128.639]	[-3.954,134.134]
Info	-350.430**	-259.216	-296.205
	[-696.427,-4.434]	[-617.546,99.114]	[-674.038,81.628]
Dissimilarity Liking (0-3)		-3.019	-5.015
		[-105.351,99.313]	[-110.085,100.054]
Dissimilarity Liking X Info		-94.906	-90.746
		[-230.384,40.572]	[-231.813,50.321]
Mean Dep. Var.	943.72	943.72	943.72
Team Controls	Yes	Yes	Yes
Performance Controls	Yes	Yes	Yes
Problem FE	No	No	Yes
N	120	120	120

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Joint Liking X Info* interacts this variable with a treatment indicator for *Info*. *Dissimilarity Liking* is the absolute difference between both partners' preferences, *Dissimilarity Liking X Info* captures the interaction with the treatment *Info*. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

Result 6 *Teams where partners know they like each other communicate slightly more, those who know about their misaligned preferences communicate slightly less.*

Inefficiency in communication seems to arise after the first *Guesstimate* is shared. Across and within conditions, different teams share their first *Guesstimate* at a similar point in time (Table A.10, Column 1), and this first *Guesstimate* is similarly accurate (Table A.10, Column 2). However, those

teams where partners like each other and know about this, then deviate inefficiently far away from this initial value. In contrast, those with *dissimilar liking* who know about this are more likely to submit final answers that are similar to the initial *Guesstimate*, which is conducive to performance (Table A.10, Column 3).

3.4.3 Ex-post evaluation

Participants do not evaluate their interaction ex-post differently depending on the interpersonal preferences of both team partners. Whether team partners liked each other or whether these preferences were misaligned does not determine how they judge the interaction retrospectively. The same is true in terms of their decision on whether to continue working with their team partner. In Table A.11, I show that neither psychological safety, the group climate, task, enjoyment, the willingness to compromise, nor the fear of disappointing others are affected by both partners' preferences.[16] Participants are also equally inclined to perform another identical round of a *Guesstimation* task with their partner, irrespective of their preferences. In *Info*, 60 percent want to continue working with their team partner. Table A.12 shows that the preference for working alone remains unaffected by whether team partners like each other or not, and whether their preferences are aligned or not. Taken together, these results provide suggestive evidence that participants did not experience greater utility from working with a partner who liked them, nor do they believe they were more successful with such a partner. In that sense, participants learned through the collaboration that their ex-ante expectations of higher team performance when liking each other were not met. At the same time, participants in teams with dissimilar preferences did not seem to experience collaboration more positively either, nor did they believe to be more successful and hence to continue with their partner.

Result 7 *Participants evaluate the collaboration in the Guesstimation task similarly, irrespective of whether they were in a team where partners liked each other, disliked each other, or had misaligned preferences.*

4 Discussion

In this study, I analyze how interpersonal preferences affect team performance in complex problem solving. To do so, I rely on a non-routine task that captures 21st century skills and is used frequently to assess the quality of job candidates. This *Guesstimation* task allows me to create a

[16] Columns (2)-(5) in Table A.13 show consistency in the responses between the different sets of questions, providing evidence that the null results are not driven by a lack of attention to these (non-incentivized) questions.

team task in which complementary knowledge and skills can come into play when participants communicate with each other effectively. Hence, I analyze whether interpersonal preferences change the effectiveness of teams in a task with high external validity.

I find that interpersonal preferences matter for effective teamwork. When individuals know each other's preferences, dissimilar preferences are conducive to increased team performance. That is, teams where one partner likes the other more perform better. However, teams where partners like each other do not perform better than teams where they do not. Team members do not anticipate this pattern. Before conducting the task, those in a team where partners like each other expect to be more successful in the task. I provide suggestive evidence that more effective communication (as opposed to more communication) drives the performance results and establish that the results are not due to homophily and sorting.

Relying on an experimental setting allows me to infer participants' preferences and manipulate information structures. Incentivizing truthful preference submission and tightly controlling information structures would be impossible in the field. This comes at a cost. First, I elicit preferences that are meaningful as they shift behavior (see also Opitz & Schwaiger, 2023), but are not as strong as they can be in real-word settings. A preference formed on the basis of five answers to questions is likely to be less strong and more malleable than a preference formed on the basis of hearing the opinions and observing the behavior of others. Second, the *Guesstimation* task abstracts from other dimensions where interpersonal preferences may matter in the work context. I focus on the performance aspect during the task, while working with friends may be especially important when it comes to fair attribution of responsibility, success, and failure after task completion (e.g., Jin et al., 2019).

This study extends the findings of Opitz and Schwaiger (2023), which shows the relevance of being liked for cooperation. In line with Opitz and Schwaiger (2023), I find that participants expect higher effort from a partner who likes them, and believe that this will translate into higher performance. This is the case even though I am studying a collaborative environment with aligned interests rather than a cooperative environment with conflicting individual interests, and I conceptualize interpersonal preferences slightly differently. In Opitz and Schwaiger (2023), preferences for team partners are based on the desirability to perform the cooperative game with them. In contrast, this study attempts to capture preferences that are not tied to the specific task – instead they refer to a payoff-irrelevant interaction. In that sense, this study is closer to a definition of *liking* than to *task-specific partner preferences*.[17] Despite the more favorable ex-ante beliefs, objective

[17]In this sense, the study relates more to papers that investigate the effects of friends at the workplace, including Bandiera, Barankay, and Rasul (2005) and Ashraf (2022).

performance is not better in teams where partners like each other. While in stylized one-shot interactions, these beliefs closely map into actions (and payoffs), the determinants of success in the collaborative problem solving environment of this study are more complex. This highlights the importance of better understanding the production function in such complex tasks and analyzing communication as a key component.

Overall, these findings raise important questions on how to make communication effective. Charness, Cooper, and Grossman (2020) shows that imposing communication costs can improve team performance on logical problems through decreasing the quantity and improving the quality of messages, Girotra, Terwiesch, and Ulrich (2010) shows the benefits from limiting communication temporarily. I also show that the quantity of communication does not necessarily translate into higher success. In addition, I find that the quantity of communication is largely unrelated to the self-reported collaboration experience of participants. While Girotra, Terwiesch, and Ulrich (2010) and Charness, Cooper, and Grossman (2020) change the nature of communication exogenously, my experimental condition holds everything constant except for the piece of information on how one was ranked by the team partner. This, in turn, influences how individuals communicate to combine their knowledge and solve the problem. Interestingly, participants did not report a better or worse collaboration experience depending on both partners' preferences.

The results suggest several important organizational implications. Self-selected teams would most likely have been detrimental for performance if preferences were either known or if individuals learned each other's preferences in the process of forming teams. Because teams are formed randomly in the experiment, I do not observe performance differences between the two information structures on average. Still, if participants had been allowed to form teams on their own, there had been more teams in which both partners liked each other –and these teams turned out to be less successful in the *Info* treatment. This has two underlying mechanism. First, the assignment would not be random anymore, take the preferences of both partners into account, and lead to higher fraction of individuals who work with their preferred partner.[18] Second, preferences themselves are likely to change because people take their partner's preference into account when forming a team, since they expect a partner who likes them to exert more effort. This implies that having teams that self-select based on how much team members like each other could be detrimental to overall performance without teams expecting this when they form. Teams where one or both partners know they are not necessarily each other's favorites perform better. In this sense, these findings contribute

[18]For example, a counterfactual assignment of participants to teams via the Deferred-Acceptance Algorithm (Gale & Shapley, 1962) instead of a random assignment would have increased the fraction of individuals who perform the *Guesstimation* task with their most preferred partner from 23 percent to 60 percent.

to a better understanding of how teams should be formed, and which preferences –if known– should be revealed.

References

Abdulkadiroğlu, A., & Sönmez, T. (1998). Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, *66*(3), 689–701.

Aerlebaeck, J. B., & Albarracin, L. (2019). The use and potential of fermi problems in the stem disciplines to support the development of twenty-first century competencies. *ZDM*, *51*(6), 979–990.

Allocca, A. (2023). 'No man is an island': An empirical study on team formation and performance. *CRC TRR 190 Discussion Paper No. 389*.

Ashraf, A. (2022). Performance ranks, conformity, and cooperation: Evidence from a sweater factory. *CESifo Working Paper No. 9591*.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*(4), 1279–1333.

Autor, D. H., & Price, B. (2013). The changing task composition of the us labor market: An update of autor, levy, and murnane (2003). *Working Paper*.

Bandiera, O., Barankay, I., & Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *Quarterly Journal of Economics*, *120*(3), 917–962.

Bandiera, O., Barankay, I., & Rasul, I. (2013). Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*, *11*(5), 1079–1114.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. *Assessment and teaching of 21st century skills*, 17–66.

Boss, V., Dahlander, L., Ihl, C., & Jayaraman, R. (2023). Organizing entrepreneurial teams: A field experiment on autonomy over choosing teams and ideas. *Organization Science*, *34*(6), 2097–2118.

Charness, G., Cooper, D. J., & Grossman, Z. (2020). Silence is golden: Team problem solving and communication costs. *Experimental Economics*, *23*, 668–693.

Chen, D., Schonger, M., & Wickens, C. (2016). Otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431–457.

De Dreu, C. K., Evers, A., Beersma, B., Kluwer, E. S., & Nauta, A. (2001). A theory-based measure of conflict management strategies in the workplace. *Journal of Organizational Behavior*, *22*(6), 645–668.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, *132*(4), 1593–1640.

Dimotakis, N., Davison, R. B., & Hollenbeck, J. R. (2012). Team structure and regulatory focus: The impact of regulatory fit on team dynamic. *Journal of Applied Psychology*, *97*(2), 421.

Edmondson, A. (2002). *Managing the risk of learning: Psychological safety in work teams*. Division of Research, Harvard Business School Cambridge, MA.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383.

Englmaier, F., Grimm, S., Grothe, D., Schindler, D., & Schudy, S. (2023a). The effect of incentives in non-routine analytical team tasks. *Journal of Political Economy (Forthcoming)*.

Englmaier, F., Grimm, S., Grothe, D., Schindler, D., & Schudy, S. (2023b). The efficacy of tournaments for non-routine team tasks. *Journal of Labor Economics (Forthcoming)*.

Englmaier, F., Grimm, S., Grothe, D., Schindler, D., & Schudy, S. (2021). The value of leadership: Evidence from a large-scale field experiment. *CESifo Working Paper No. 9273*.

Fischer, M., Rilke, R. M., & Yurtoglu, B. B. (2023). When, and why, do teams benefit from self-selection? *Experimental Economics*, *26*, 749–774.

Gächter, S., Starmer, C., & Tufano, F. (2023). Measuring "group cohesion" to reveal the power of social relationships in team production. *Review of Economics and Statistics (Forthcoming)*.

Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, *69*(1), 9–15.

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best idea. *Management Science*, *56*(4), 591–605.

Goeree, J. K., McConnell, M. A., Mitchell, T., Tromp, T., & Yariv, L. (2010). The 1/d law of giving. *American Economic Journal: Microeconomics*, *2*(1), 183–203.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.

Hardt, D., Mayer, L., & Rincke, J. (2023). Who does the talking here? The impact of gender composition on team interactions. *Working Paper*.

Hoogendoorn, S., Oosterbeek, H., & Van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, *59*(7), 1514–1528.

Hoogendoorn, S., Parker, S. C., & Van Praag, M. (2017). Smart or diverse start-up teams? Evidence from a field experiment. *Organization Science*, *28*(6), 1010–1028.

Ichniowski, C., & Shaw, K. (2013). Insider econometrics: A roadmap to estimating empirical models of organizational performance. In R. Gibbons & J. Roberts (Eds.), *The Handbook of Organizational Economics* (pp. 263–312). Princeton University Press.

Jin, G. Z., Jones, B., Lu, S. F., & Uzzi, B. (2019). The reverse Matthew effect: Consequences of retraction in scientific teams. *Review of Economics and Statistics*, *101*(3), 492–506.

Leider, S., Möbius, M. M., Rosenblat, T., & Do, Q.-A. (2009). Directed altruism and enforced reciprocity in social networks. *Quarterly Journal of Economics*, *124*(4), 1815–1851.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444.

McWhirter, R. M., & Jecker, J. D. (1967). Attitude similarity and inferred attraction. *Psychonomic Science*, *7*(6), 225–226.

Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? a meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships*, *25*(6), 889–922.

Morgan, J., Neckermann, S., & Sisak, D. (2020). Peer evaluation and team performance: An experiment on complex problem solving. *Working Paper*.

Opitz, T., & Schwaiger, C. (2023). Everyone likes to be liked: Experimental evidence from matching markets. *CRC TRR 190 Discussion Paper, No. 366*.

Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist*, *73*(4), 593.

Van Ginkel, W. P., & Van Knippenberg, D. (2008). Group information elaboration and group decision making: The role of shared task representations. *Organizational Behavior and Human Decision Processes*, *105*(1), 82–97.

Weidmann, B., & Deming, D. J. (2021). Team players: How social skills improve team performance. *Econometrica*, *89*(6), 2637–2657.

Appendices

A Design, Questionnaires & Materials

A1 Timeline of the Experiment

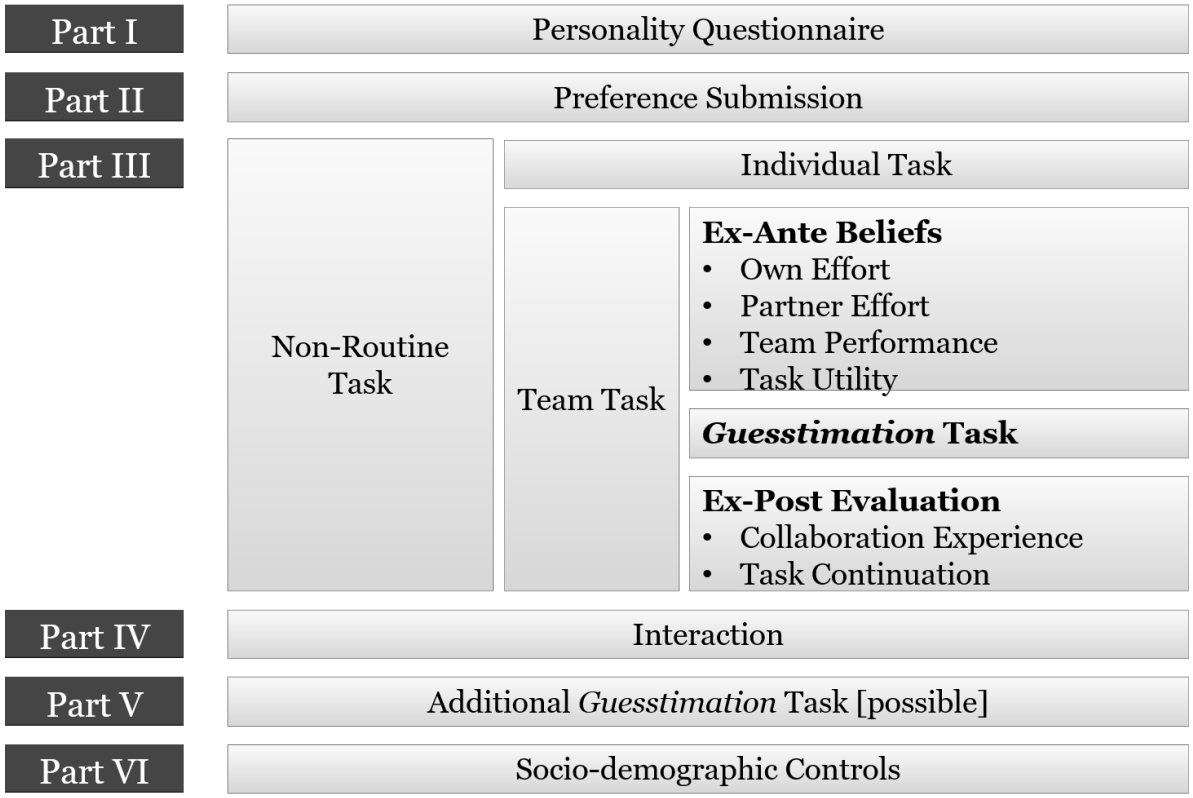


Figure A.1: Timeline of the experiment

A2 Personality Questionnaire

1. I would rather have a cat than a dog as a pet.
2. I prefer reading a book in the evening to watching a movie.
3. I prefer going to the beach on vacation to visiting a city.
4. I would rather spend an evening in a bar than partying in a club.
5. I prefer to living in a shared apartment than alone.
6. I am rather reserved and quiet.
7. I am easygoing, prone to laziness.
8. I am talented with my hands.
9. I often make decisions spontaneously and intuitively.
10. I tend to avoid conflict.
11. I support climate protests, even if they use means of civil disobedience (e.g., “Last Generation”).

12. I support quota regulations in the labor market for socially disadvantaged groups (e.g., for women or migrants).
13. There should be a requirement to wear a bicycle helmet.
14. The possession of marijuana should be legalized.
15. Unhealthy foods should be taxed more.

A3 *Guesstimation* tasks

In the experiment, each team faces one of the following *Guesstimation* tasks.

1. What is the route length of all streetcar lines in Germany?
2. How many businesses in the hairdressing trade (hair salons) were there in Germany in 2022?
3. How many dogs are there in Germany?
4. How much household waste was generated per inhabitant in Germany in 2021 (household waste+bulky waste + recyclables)?
5. How many passengers were transported by Deutsche Bahn ICEs in 2022?

A4 Collaboration Questionnaire

Psychological Safety (Edmondson, 2002, 1999)

- I felt like my team partner would judge me on the things that I said.
- I feel safe sharing my views with my team partner.

Task Satisfaction (Dimotakis, Davison, & Hollenbeck, 2012)

- I found real enjoyment in performing this task.

Group Climate (Van Ginkel & Van Knippenberg, 2008)

- – The atmosphere was good
  – Our collaboration was good

Willingness to Compromise (De Dreu et al., 2001)

- I tried to accommodate my teammate.
- I insisted we both give in a little.

Fear of Disappointing Others

- I did not want to disappoint my team partner.

B Additional Analyses

B1 Treatment *No-Info*: Participants do not know each others' preferences

There are no significant effects of team partners' preferences on team performance in *No-Info*. This confirm that not team characteristics are underlying the effects in the treatment *Info*. Also within *No-Info*, I compare teams with different preference constellations –with the only differences that participants were not aware of these preferences. Columns (1)-(4) of Table A.1 estimate the pre-registered specifications, Column (5) adds *Guesstimation* problem fixed effects. Results consistently show that there is no differential performance depending on the preferences of both partners. The positive point estimates for *joint liking* at least shows that there is no negative performance of sorting. While participants generally prefer to interact with those who are similar to them ($r = 0.119$, $p < 0.001$; Pearson's correlation coefficient based on the similarity of responses to five the questions that participants saw and *liking*), this does not translate into worse performance due to a lack in complementary skills or knowledge.

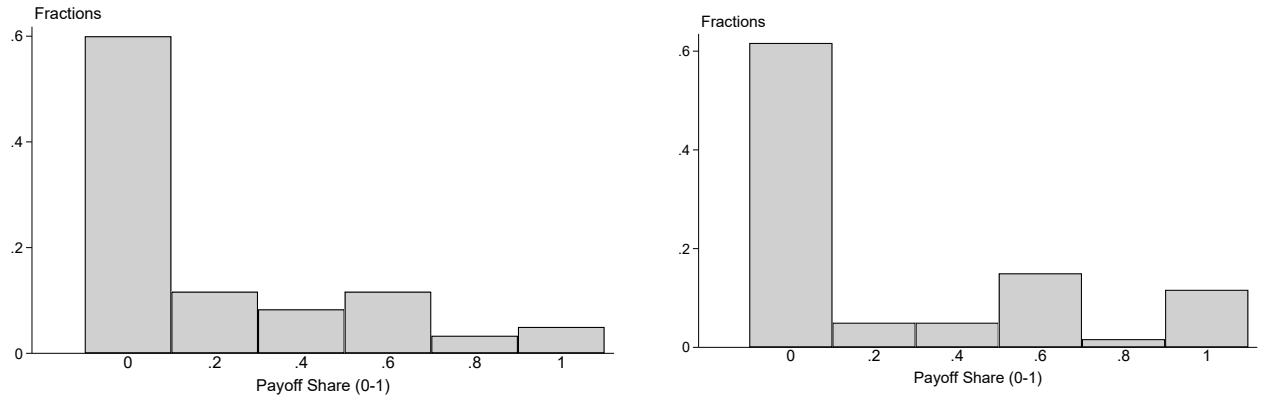
Table A.1: Task Performance: Guesstimation [*No-Info*]

	Payoff Share (0-1)				
	(1)	(2)	(3)	(4)	(5)
Joint Liking (2-8)	.038	.031	.039	.031	.032
	[-.011,.088]	[-.022,.083]	[-.011,.089]	[-.022,.084]	[-.023,.088]
Dissimilarity Liking (0-3)			-.029	-.028	-.017
			[-.142,.083]	[-.151,.095]	[-.148,.115]
Mean Dep. Var.	.25	.25	.25	.25	.25
Team Controls	No	Yes	No	Yes	Yes
Performance Controls	No	Yes	No	Yes	Yes
Problem FE	No	No	No	No	Yes
N	60	60	60	60	60

Notes. OLS Regressions.*** p <0.01, ** p <0.05, * p <0.1. The outcome variable is the fraction of the maximum possible payoff that teams achieved. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. *Dissimilarity Liking* is the difference between both partners' preferences. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants could face.

B2 Task Performance across Treatments

Task performance is similar across *Info* and *No-Info*. The average payoff share is 0.25 in *Info* and 0.20 in *No-Info* ($p = 0.765$; Mann-Whitney-U test (MWU)). Also both distributions (see Figure A.2) are not statistically significantly different from each other ($p = 0.985$; Kolmogorov–Smirnov test). Therefore, on average, there are no efficiency gains from knowing the preferences of one's partner before performing the *Guesstimation* task.



A) Payoff Shares in *Info*

B) Payoff Shares in *No-Info*

Notes. This figure displays distributions of payoff share in the joint *Guesstimation* task. Panel A shows the distributions in *Info*, Panel B in *No-Info*.

Figure A.2: Payoff Share across Treatments

B3 Ex-ante Beliefs and Preferences

Beliefs: Own Effort

Table A.2: Own Effort [*Info*]

	Expected Own Effort (0-10)			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	.131			
	[-.094,.356]			
Partner's Liking (1-4)		.060		
		[-.171,.291]		
Joint Liking (2-8)			.087	.091
			[-.066,.240]	[-.063,.245]
Dissimilarity Liking (0-3)				-.080
				[-.357,.197]
Mean Dep. Var.	8.692	8.692	8.692	8.692
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
N	120	120	120	120

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The outcome variable is self-reported motivation to exert high effort. The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

Table A.3: Own Effort [across Treatments]

	Expected Own Effort (0-10)			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	.093			
	[-.154,.339]			
Own Liking X Info	.042			
	[-.299,.382]			
Partner's Liking (1-4)		.077		
		[-.171,.324]		
Partner's Liking X Info		-.026		
		[-.368,.315]		
Joint Liking (2-8)			.056	.054
			[-.086,.199]	[-.090,.197]
Joint Liking X Info			.027	.033
			[-.186,.240]	[-.180,.247]
Dissimilarity Liking (0-3)				.109
				[-.209,.427]
Dissimilarity Liking X Info				-.188
				[-.614,.238]
Info	-.045	.120	-.074	.073
	[-.942,.853]	[-.780,1.021]	[-1.161,1.012]	[-1.069,1.216]
Mean Dep. Var.	8.65	8.65	8.65	8.65
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
Problem FE	No	No	No	No
N	240	240	240	240

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The outcome variable is self-reported motivation to exert high effort. The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Own Liking X Info* interacts this variable with a treatment indicator for *Info*; the other interactions in the table are defined accordingly. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on.*Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

Beliefs: Partner Effort

Table A.4: Partner Effort [*Info*]

	Expected Partner Effort			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	.199			
	[-.050,.448]			
Partner's Liking (1-4)		.299**		
		[.048,.550]		
Joint Liking (2-8)			.224***	.229***
			[.058,.390]	[.062,.396]
Dissimilarity Liking (0-3)				-.101
				[-.401,.200]
Mean Dep. Var.	8.40	8.40	8.40	8.40
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
N	120	120	120	120

Notes. OLS Regressions.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The outcome variable is the belief about the partner's motivation to exert high effort. The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Own Liking X Info* interacts this variable with a treatment indicator for *Info*; the other interactions in the table are defined accordingly. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on.*Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

Table A.5: Partner Effort [across Treatments]

	Expected Partner Effort (0-10)			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	.241*			
	[-.034,.516]			
Own Liking X Info	-.053			
	[-.432,.326]			
Partner's Liking (1-4)		.009		
		[-.266,.284]		
Partner's Liking X Info		.273		
		[-.107,.652]		
Joint Liking (2-8)			.083	.082
			[-.075,.241]	[-.077,.241]
Joint Liking X Info			.129	.137
			[-.107,.364]	[-.100,.374]
Dissimilarity Liking (0-3)				.057
				[-.295,.409]
Dissimilarity Liking X Info				-.186
				[-.658,.286]
Info	.034	-.742	-.712	-.549
	[-.965,1.033]	[-1.742,.258]	[-1.916,.492]	[-1.815,.718]
Mean Dep. Var.	8.43	8.43	8.43	8.43
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
Problem FE	No	No	No	No
N	240	240	240	240

Notes. OLS Regressions.*** $p<0.01$, **$p<0.05$, * $p<0.1$. The outcome variable is the belief about the partner's motivation to exert high effort. The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Own Liking X Info* interacts this variable with a treatment indicator for *Info*; the other interactions in the table are defined accordingly. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on.*Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

Beliefs: Team Performance

Table A.6: Team Success [across Treatments]

	Expected Team Performance			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	-.443*			
	[-.923,.037]			
Own Liking X Info	1.032***			
	[.371,1.694]			
Partner's Liking (1-4)		-.104		
		[-.591,.383]		
Partner's Liking X Info		.501		
		[-.170,1.172]		
Joint Liking (2-8)			-.185	-.176
			[-.462,.093]	[-.454,.102]
Joint Liking X Info			.629***	.615***
			[.216,1.042]	[.201,1.029]
Dissimilarity Liking (0-3)				-.400
				[-1.016,.215]
Dissimilarity Liking X Info				.503
				[-.323,1.328]
Info	-2.636***	-1.367	-3.186***	-3.543***
	[-4.381,-.892]	[-3.137,.402]	[-5.295,-1.076]	[-5.758,-1.329]
Mean Dep. Var.	7.43	7.43	7.43	7.43
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
Problem FE	No	No	No	No
N	240	240	240	240

Notes. OLS Regressions.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The outcome variable is the amount bet on the team being in the top 25%, with monetary values recoded on a scale of 0-10.The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Own Liking X Info* interacts this variable with a treatment indicator for *Info*; the other interactions in the table are defined accordingly. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on.*Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

Beliefs: Task Utility

Table A.7: Task Utility [across Treatments]

	Preference for Working Alone			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	.032			
	[-.337,.401]			
Own Liking X Info	-.047			
	[-.556,.462]			
Partner's Liking (1-4)		.151		
		[-.218,.519]		
Partner's Liking X Info		-.128		
		[-.636,.380]		
Joint Liking (2-8)			.061	.063
			[-.152,.275]	[-.151,.277]
Joint Liking X Info			-.058	-.074
			[-.375,.260]	[-.393,.244]
Dissimilarity Liking (0-3)				-.055
				[-.529,.418]
Dissimilarity Liking X Info				.352
				[-.282,.987]
Info	.374	.565	.536	.208
	[-.968,1.716]	[-.775,1.905]	[-1.087,2.159]	[-1.494,1.910]
Mean Dep. Var.	5.34	5.34	5.34	5.34
Team Controls	Yes	Yes	Yes	
Performance Controls	Yes	Yes	Yes	Yes
Problem FE	No	No	No	No
N	240	240	240	240

Notes. OLS Regressions.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The outcome variable indicates the willingness to pay to perform another round of the task with a fixed payoff alone (as opposed to with their team partner). The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Own Liking X Info* interacts this variable with a treatment indicator for *Info*; the other interactions in the table are defined accordingly. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on.*Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

B4 Communication

Table A.8: Team Communication [*Info*]

	Total Length of Chat Messages				
	(1)	(2)	(3)	(4)	(5)
Joint Liking (2-8)	18.3	12.9	21.3	17.8	12.6
	[-32.8,69.5]	[-38.3,64.1]	[-29.5,72.1]	[-32.6,68.2]	[-41.5,66.6]
Dissimilarity Liking (0-3)			-65.8	-85.5*	-77.8
			[-156.1,24.6]	[-178.9,7.9]	[-185.1,29.6]
Mean Dep. Var.	906.6	906.6	906.6	906.6	906.6
Team Controls	No	Yes	No	Yes	Yes
Performance Controls	No	Yes	No	Yes	Yes
Problem FE	No	No	No	No	Yes
N	60	60	60	60	60

Notes. OLS Regressions.*** p <0.01, ** p <0.05, * p <0.1. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

Table A.9: Team Communication [across Treatments]

	Asymmetric Communication	Turn-Taking	Time to First Message
	(1)	(2)	(3)
Joint Liking (2-8)	.001	-.432	-.240
	[-.012,.013]	[-1.473,.608]	[-3.793,3.313]
Joint Liking X Info	-.007	.502	-2.057
	[-.026,.011]	[-1.068,2.072]	[-7.418,3.303]
Dissimilarity Liking (0-3)	.003	-.401	2.544
	[-.025,.032]	[-2.789,1.988]	[-5.613,10.702]
Dissimilarity Liking X Info	.021	-.842	1.516
	[-.017,.058]	[-4.049,2.365]	[-9.437,12.468]
Info	-.010	-1.725	5.347
	[-.112,.091]	[-10.315,6.866]	[-23.988,34.681]
Mean Dep. Var.	.11	19.55	41.21
Team Controls	Yes	Yes	Yes
Performance Controls	Yes	Yes	Yes
Problem FE	Yes	Yes	Yes
N	120	120	120

Notes. OLS Regressions.*** p <0.01, ** p <0.05, * p <0.1. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Joint Liking X Info* interacts this variable with a treatment indicator for *Info*. *Dissimilarity Liking* is the absolute difference between both partners' preferences, *Dissimilarity Liking X Info* captures the interaction with the treatment *Info*. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

Table A.10: First Guesstimate [across Treatments]

	First Guesstimate [Time]	First Guesstimate [Payoffs]	Deviation from First Guesstimate
	(1)	(2)	(3)
Joint Liking (2-8)	.009	.014	-.015
	[-.025,.043]	[-.032,.060]	[-.084,.054]
Joint Liking X Info	.000	-.024	.113**
	[-.053,.053]	[-.095,.047]	[.006,.219]
Dissimilarity Liking (0-3)	-.029	-.054	.187**
	[-.108,.050]	[-.159,.052]	[.029,.346]
Dissimilarity Liking X Info	.014	.095	-.193*
	[-.093,.120]	[-.047,.238]	[-.407,.021]
Info	.042	.066	-.432
	[-.254,.339]	[-.332,.464]	[-1.029,.165]
Mean Dep. Var.	.51	.28	.51
Team Controls	Yes	Yes	Yes
Performance Controls	Yes	Yes	Yes
Problem FE	Yes	Yes	Yes
N	117	117	117

Notes. OLS Regressions.*** p <0.01, **p <0.05, * p <0.1. The values in square brackets represent the 95% confidence intervals. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Joint Liking X Info* interacts this variable with a treatment indicator for *Info*. *Dissimilarity Liking* is the absolute difference between both partners' preferences, *Dissimilarity Liking X Info* captures the interaction with the treatment *Info*. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

B5 Ex-post Evaluation

Collaboration Experience

Table A.11: Collaboration Experience [*Info*]

	Regression Coefficient [Confidence Interval]	
	Sum of partner's preferences	Preference Difference
Psychological Safety	-.019 [-.097,.059]	-.005 [-.145,.136]
Group Climate	.030 [-.079,.139]	-.036 [-.232,.161]
Task Enjoyment	.077 [-.058,.211]	.090 [-.151,.331]
Willingness to Compromise	.057 [-.032,.146]	-.101 [-.260,.059]
Fear of Disappointing Others	.020 [-.088,.128]	-.052 [-.246,.142]

Notes. OLS Regressions. *** p <0.01, ** p <0.05, * p <0.1. The values in square brackets represent the 95% confidence intervals. This table displays coefficients and confidence intervals from individual regressions, each controlling for age, gender, and individual performance.

Task Continuation

Table A.12: Task Continuation [*Info*]

	Preference for Working Alone			
	(1)	(2)	(3)	(4)
Own Liking (1-4)	.019			
	[-.059,.096]			
Partner's Liking (1-4)		-.010		
		[-.089,.069]		
Joint Liking (2-8)			.004	.004
			[-.049,.057]	[-.050,.057]
Dissimilarity Liking (0-3)				.007
				[-.088,.103]
Mean Dep. Var.	.40	.40	.40	.40
Demographic Controls	Yes	Yes	Yes	Yes
Performance Indiv. Task	Yes	Yes	Yes	Yes
N	120	120	120	120

Notes. OLS Regressions.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The outcome variable is a (binary) indicator of whether the participant would like to perform another round of the task alone (as opposed to with their team partner). The values in square brackets represent the 95% confidence intervals. *Own Liking (1-4)* takes the value of four if the partner was the first choice of the participant, three if the partner was the second choice, and so on. *Partner's Liking (1-4)* takes the value of four if the participant was the most preferred choice of their partner, three if the participant was the second most preferred choice, and so on. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Dissimilarity Liking* is the absolute difference between both partners' preferences. The *Demographic Controls* are the age and gender of the participant, *Performance Indiv. Task*, captures their performance in the individual *Guesstimation* task.

Correlation Table [*No-Info*]

Table A.13: Correlations of Message Length and Collaboration Experience

	Communication Quantity (1)	Psychological Safety (2)	Fear of Disappointing (3)	Task Enjoyment (4)	Willingness to Compromise (5)	Group Climate (6)
Communication Quantity	1.00					
Psychological Safety	0.01	1.00				
Fear of Disappointing	−0.02	−0.16 *	1.00			
Task Enjoyment	0.06	0.36***	−0.02	1.00		
Willingness to Compromise	0.18*	0.21**	−0.03	0.05	1.00	
Group Climate	0.12	0.40***	−0.07	0.56***	0.09	1.00

Notes. Pearson's Correlation Coefficients. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

B6 Adjusted Performance Measure

Table A.14: Task Performance: Guesstimation [across Treatments]

	Payoff Share (0-1)		
	(1)	(2)	(3)
Joint Liking (2-8)	.022	.023	.026
	[-.025,.069]	[-.023,.070]	[-.022,.073]
Joint Liking X Info	-.059*	-.066*	-.067*
	[-.129,.011]	[-.135,.003]	[-.139,.004]
Info	.253	.126	.137
	[-.105,.611]	[-.244,.496]	[-.254,.529]
Dissimilarity Liking (0-3)		-.054	-.049
		[-.160,.052]	[-.158,.060]
Dissimilarity Liking X Info		.150**	.151**
		[.010,.290]	[.005,.297]
Mean Dep. Var.	.25	.25	.25
Team Controls	Yes	Yes	Yes
Performance Controls	Yes	Yes	Yes
Problem FE	No	No	Yes
N	120	120	120

Notes. OLS Regressions.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The values in square brackets represent the 95% confidence intervals. The outcome variable is the fraction of the maximum possible payoff that teams would have achieved when correcting for typos or formatting errors team members made. *Joint Liking* is the sum of both partners' preferences. For example, a value of 8 means that both partners ranked the other as their first choice, a value of 2 means that both partners ranked the other as their last choice. *Joint Liking X Info* interacts this variable with a treatment indicator for *Info*. *Dissimilarity Liking* is the absolute difference between both partners' preferences, *Joint Liking X Info* captures the interaction with the treatment *Info*. The *Team Controls* are the age and gender of both team partners, *Performance Controls*, capture the individual task performance of each partner. *Problem FE* are fixed effects for the five different *Guesstimation* problems that participants face.

B7 Predictive Potential Questionnaire

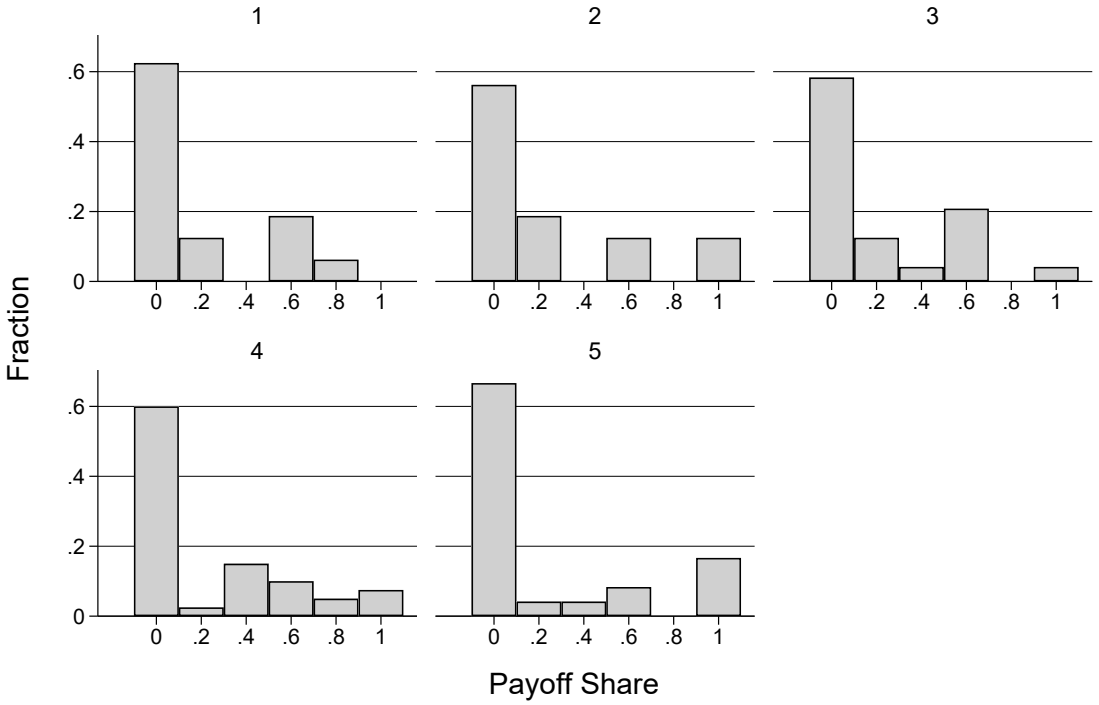
Predictive Potential Questionnaire

Table A.15: Predictive Potential Questionnaire

	Individual Performance	Joint Performance	
	All	All	Info
	(1)	(2)	(3)
Cat over Dog	-.000	-.003	.008
	[-.029,.029]	[-.039,.032]	[-.042,.059]
Book over Film	-.022	-.043	-.062
	[-.066,.022]	[-.097,.010]	[-.138,.013]
Beach over City	-.002	.026	.040
	[-.040,.036]	[-.020,.073]	[-.027,.107]
Bar over Club	-.010	.008	.016
	[-.049,.029]	[-.040,.055]	[-.051,.083]
Living Alone over Shared	.028	.024	.009
	[-.006,.061]	[-.018,.065]	[-.048,.067]
Reserved	.031	-.018	-.032
	[-.009,.071]	[-.067,.032]	[-.102,.037]
Lazy	-.031	-.038	-.034
	[-.074,.011]	[-.090,.014]	[-.112,.043]
Handy with Hands	.035*	-.037	.011
	[-.006,.076]	[-.088,.013]	[-.061,.084]
Spontaneous	-.005	.055**	.038
	[-.050,.039]	[.001,.110]	[-.039,.115]
Conflict Avoidant	.024	.038	-.022
	[-.018,.066]	[-.014,.090]	[-.090,.046]
Climate Civil Disobedience	.006	-.029	-.016
	[-.039,.051]	[-.084,.026]	[-.093,.060]
Quota Disadvantaged	-.020	.013	-.030
	[-.063,.022]	[-.039,.065]	[-.103,.042]
Bicycle Helmet Mandatory	-.000	.017	.066*
	[-.040,.040]	[-.031,.066]	[-.003,.135]
Legalize Marijuana	.007	.014	.019
	[-.029,.044]	[-.031,.058]	[-.044,.081]
Taxes Unhealthy Food	.008	-.014	.010
	[-.029,.046]	[-.060,.031]	[-.049,.069]
Observations	240	240	120

Notes. OLS Regressions. *** p <0.01, ** p <0.05, * p <0.1. The values in square brackets represent the 95% confidence intervals. Column (1) estimates the effect of questionnaire responses on the individual task performance. Columns (2) & (3) estimate the effect on the performance in the joint task. Column (2) includes all, participants, Column (3) only includes participants in the *Info* condition. For the wording of the questions, answered on a Likert scale from 1-4, see Appendix A2.

C Appendix Figures



Notes. This figure displays the payoff shares across the five different *Guesstimation* tasks. Task 1 asked about the length of all streetcar lines in Germany, Task 2 about the number of hair salons in Germany, Task 3 about the number of dogs in Germany, Task 4 about the yearly household waste of an inhabitant in Germany, and Task 5 about the number of passengers transported by Deutsche Bahn ICEs. All questions refer to numbers in either 2021, 2022, or 2023.

Figure A.3: Payoff Share across *Guesstimation* Problems

D Instructions

Appendix D includes the translated instructions of the experiment (from German). Treatment specific parts are shown in *italics* and the corresponding treatment is clearly indicated.

General (On-screen) Instructions (before Part I)

Welcome to the experiment and thank you for your participation!

Please do not speak from now on with any other participant.

Procedures

This experiment is designed to study economic decision-making. You can earn money by participating. The money you earn will be paid to you privately and in cash after the experiment.
The experiment lasts around 60 minutes and consists of five parts (I-V). At the beginning of every part, you receive detailed instructions. In addition, you will receive comprehension questions for some parts to help you understand how the experiment and the payoff conditions work. If you have questions after reading the instructions or during the experiment, please raise your hand or press the red button on your keyboard. The experimenter will then come to you and answer your questions privately.

Anonymity

The analysis of the experiment is anonymous; that is, we will never link your name with the data generated in the experiment. At the end of the experiment, you must sign a receipt for the receipt of the earnings. This receipt is only used for accounting and booking of the experiment money. No further personal data from the experiment will be passed on. Information collected during the experiment may be visible to other participants as the experiment progresses. You make all decisions anonymously, so no other participant can associate your decisions with you during the experiment.

Payment

For your participation in this experiment and the completion of a short questionnaire at the end of it you receive 10 €. In addition, you can achieve further payoffs during the experiment. During the experiment, you and the other participants will be asked to make a series of decisions. These can affect the payoffs for you, and potentially for other participants. Additionally, you can earn money by making correct assessments. You will receive detailed instructions about these decisions and how they affect the payoff.

Information on fold-out elements

On some places you will find green buttons. When you click on them, they unfold and contain information. Here you can see an example:

Sample Box

Blue buttons do not contain any information. When you click on them, the next page of the experiment will appear.

Part I

In the first part of the experiment, we ask you to truthfully fill out a questionnaire. This is a personality questionnaire, so there are no right or wrong answers.

You receive 2 € for the completion of the questionnaire.

Please answer the questions with the answer options:

- Does not apply
- Tends not to apply
- Tends to apply
- Applies

[The questionnaire items are shown in Section A2.]

Part II

Screen 1: Explanation Ranking

In the further course of the experiment, you will simultaneously interact with **one** other participant (your *team partner*). This **interaction will take about 5-10 minutes**.

Important: For that part of the experiment you will receive a payoff of 2 €.

There are **4 possible partners for the interaction.**

In the following, you will be able to indicate who you would prefer as a team partner. For this you will see a randomly selected part of your potential partners' answers from the questionnaire.

At the same time, your potential partners will see your answers (as well as the answers of other participants) to these questions.

After viewing the profiles, we ask you to **indicate a preference order**.

With this preference order **you indicate with whom of the participants you would prefer to interact**. Rank 1 means that you would most like to have this participant as your team partner.

Rank 2 means that you would second most like to have this participant as your team partner, and so on.

Important: You maximize your chances to interact with your preferred partner by indicating your true preferences.

This is intuitively the case, since the allocation mechanism

1. sorts all participants randomly,
2. gives the first participant his preferred team partner,
3. gives the second participant his preferred team partner from the remaining pool of possible partners,
4. and continues with this process until all participants have a partner.

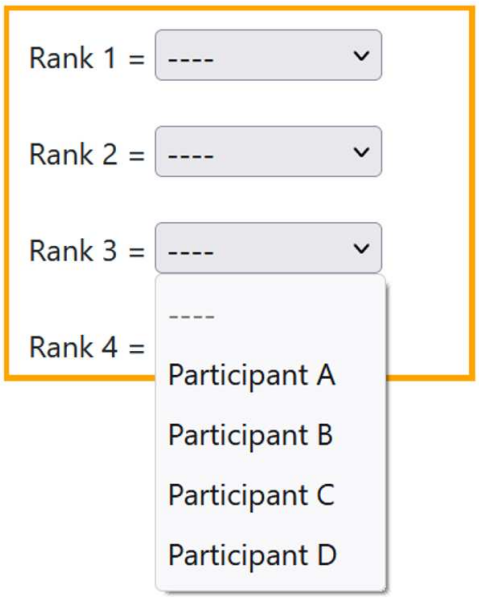
Screen 2: Choice of Team Partner

Sample screen:

Below you will find information from Part I about your potential team partners.

	Participant A	Participant B	Participant C	Participant D
Statement 1: I support quota regulations in the labor market for socially disadvantaged groups (e.g., for women or migrants).	Applies	Tends to apply	Tends not to apply	Tends not to apply
Statement 2: There should be a requirement to wear a bicycle helmet.	Does not apply	Applies	Does not apply	Does not apply
Statement 3: I would rather spend an evening in a bar than partying in a club.	Applies	Tends to apply	Does not apply	Applies
Statement 4: I am easygoing, prone to laziness.	Tends to apply	Tends to apply	Tends to apply	Does not apply
Statement 5: I prefer to live in a shared apartment than alone.	Does not apply	Does not apply	Tends to apply	Tends to apply

We now ask you to specify a preference order. With this preference order you indicate with which of the participants you would prefer to interact.



Information after the Choice of the Team Partner

Thank you for submitting your preferences. We will inform you later which participant you will work with in the interaction task in Part IV.

Before that, however, you will make some decisions and assessments in Part III.

Part III

Screen 1: Fermi-Problem, Instructions and Comprehension Questions

In Part III, you will be confronted with a problem that is also known as Fermi-Problem:

A Fermi-Problem is a **quantitative estimate** for a **problem** for which initially **little data is available**.

Therefore, you will **hardly be able to answer such a problem accurately at first go**. However, after **careful and structured thought**, you will be able to **converge to the correct answer**. This is called an **educated guess**.

The classic example of a Fermi-Problem is the question about the **number of piano tuners in Chicago**.

It is obvious that you do not know the exact solution to this question.

The best strategy is to **break the problem into parts** that can be worked on. This means that you need to make some assumptions in order to start a calculation based on them.

In this example, the following assumptions might be useful:

- Number of Chicago residents.
- Percentage of households with a piano.
- Frequency with which a piano is tuned.
- Duration to tune a piano.
- Weekly working hours of a piano tuner.

It is **not always easy to estimate the exact value at each step**.

It is often **easier and more reliable to estimate the upper and lower limits than the direct value**. How does one arrive at the best estimate based on the upper and lower limits?

It turns out that the average of upper and lower bounds is not the best estimate. The **geometric mean is a much better choice**. The geometric mean of two numbers is the square root of their product. For example, the geometric mean of 5 and 20 is 10, because $10 = \sqrt{5 * 20}$. Thus, if you are sure that Chicago's population is at least 1 million and at most 9 million, it makes sense to take $\sqrt{1 * 9} = 3$ million as an estimate. You can use the calcalator that is provided to you on the desk.

It is also often helpful to look at a problem from two different angles. Your confidence in your educated guess will increase if you arrive at a similar number another way.

Your **payoff will depend on the accuracy of your answer**. More detailed information on this is provided hereafter.

You have **limited time to solve the problem**. This will be shown to you on the screen.

Comprehension Questions

A Fermi problem refers to a question that many people can spontaneously answer correctly using their factual knowledge?

❍ Yes

❍ No

A Fermi problem is best solved by breaking it down into parts?

❍ Yes

❍ No

It makes sense to determine upper and lower limits for your estimates?

❍ Yes

❍ No

The arithmetic mean is preferable to the geometric mean in estimation?

❍ Yes

❍ No

Screen 2: Individual Fermi-Problem, Instructions

You will now be given a **first Fermi-Problem to solve on your own**.

You can **receive a maximum of 5 euros** for solving this problem.

- Your payoff is calculated as follows:
  - Your answer deviates by **less than 10% from the true value: 5 Euro**
  - Your answer deviates by **less than 20% from the true value: 4 Euro**
  - Your answer deviates by **less than 40% from the true value: 3 Euro**
  - Your answer deviates by **less than 60% from the true value: 2 Euro**
  - Your answer deviates by **less than 80% from the true value: 1 Euro**
  - Your answer deviates by **more than 80% from the true value: 0 Euro**

Time for solving the problem: 5 minutes

Screen 3: Individual Fermi-Problem

How many weddings were performed in Germany in July 2022?

[Info: You can drag the text field larger with the mouse at the bottom right corner.]

Logical Steps:

1.
2.
3.
4.
5.
…

Submit final response

Screen 4: Joint Fermi Problem, Instructions and Comprehension Questions

Explanation

In the following, you will solve **another Fermi-Problem**. This time you will work **together with a team partner**.

You have **8 minutes to solve it**. You can make **up to 10 euros**.

Important: Your answer only counts if **both team partners give the identical answer**.

Your team partner is **one of the four participants whose answers you have seen and based on which you have given a preference order for the interaction task**.

For the Fermi-problem, you were **randomly assigned one of the four participants**.

On the next page, you will learn who your team partner is for the Fermi-Problem.

In doing so, you will learn:

1. **How you ranked your team partner.**
2. ***How your team partner ranked you. [only in the treatment info]***
3. **What answers your team partner gave.**

Comprehension Questions

You have to enter exactly the same result as your team partner for it to count?

❍ Yes

❍ No

You have been randomly assigned your team partner?

❍ Yes

❍ No

You are about to find out how your team partner ranked you?

❍ Yes

❍ No

Screen 5: Joint Fermi Problem, Your Team Partner

Your submitted preference order:

Rank 1: Participant A

Rank 2: Participant C

Rank 3: Participant B

Rank 4: Participant D

- **Your assigned partner is: Participant C**

Below you will find the information from Part I about your team partner.

Information about your team partner

Continue

Screen 6: Joint Fermi Problem, Assessments

Before you perform the task with your team partner, **we will ask you for a few more assessments on the upcoming teamwork**.

You can earn extra money from two of the assessments. We will randomly select one of the two assessments and pay you the amount earned. You will get more detailed information about these decisions on the next pages.

With other decisions you can influence the course of the further experiment.

Screen 7: Joint Fermi Problem, Assessment 1: Your Motivation

Here we ask you for a subjective assessment of how motivated you are to make an effort in the following task with your team partner.

How motivated are you to make an effort in the following task (from 0 to 10)?

Screen 8: Joint Fermi Problem, Assessment 2: Motivation of your partner

Should you guess your team partner's answer to the previous question correctly, you will receive 2 €. If you estimate the information incorrectly, you will receive 0 €.

How much do you think your partner is motivated to do the following task (from 0 to 10)?

Screen 9: Joint Fermi Problem, Assessment 3: Success of your team

Now we ask you for an **estimation of the success of your team**. Again, you can receive money for a correct estimation.
For this estimation you have a **credit of 1 €, with which you can bet on the performance of your team**. You can **decide which part of the €1 you want to keep and which part you want to bet**.
The part you bet

- is multiplied by 4 if your team is in the top 25% of today's teams.
- is lost if your team is not in the top 25% of today's teams.

We illustrate this with two examples.

1. If you bet 1 € and your team is in the top 25% of today's teams, you receive 4 €. On the other hand, if you bet 1 € and your team is not in the top 25%, you receive 0 €.

2. if you bet 0,50 € and your team is among the best 25%, you receive 2,50 € (0,50 €*4 + 0,50 €). If you bet 0,50 € and your team is not in the top 25%, you receive 0,50 €.

How much do you want to bet (0 - 1 €)?

Screen 10: Joint Fermi Problem, Assessment 4: Alone or with partner in additional Fermi-Problem with fixed payoff?

It is **possible that in the further course of the experiment you will have to solve an additional Fermi-Problem** where you will **not be paid for the accuracy of your data**.

Instead, you will receive a **fixed payoff for this Fermi-Problem** that is not related to your performance.

The only thing we will check is whether you work conscientiously on the Fermi-Problem. The duration of the task will again be 5 minutes.

You can indicate whether you prefer to do this **task alone or with your current team partner**.

This indication is based on 11 choices, each of which requires you to choose Option A or Option B. Each of the rows in the table below represents one decision.

- At the end of the experiment, one of the 11 decisions is selected. The computer will randomly select (with equal probability) which decision will be chosen.
- If either you or your team partner chose option A in this decision, you will work the Fermi-Problem alone.
- If you both chose option B, a random mechanism will decide whether you will do the task individually or together again. Thus, your team partner will never know which decision you made (and you will never know which decision your team partner made).

Option A			**Option B**
Do the task alone and get 1.0€	❍	❍	Do the task together and get 2.0€
Do the task alone and get 1.1€	❍	❍	Do the task together and get 1.9€
Do the task alone and get 1.2€	❍	❍	Do the task together and get 1.8€
Do the task alone and get 1.3€	❍	❍	Do the task together and get 1.7€
Do the task alone and get 1.4€	❍	❍	Do the task together and get 1.6€
Do the task alone and get 1.5€	❍	❍	Do the task together and get 1.5€
Do the task alone and get 1.6€	❍	❍	Do the task together and get 1.4€
Do the task alone and get 1.7€	❍	❍	Do the task together and get 1.3€
Do the task alone and get 1.8€	❍	❍	Do the task together and get 1.2€
Do the task alone and get 1.9€	❍	❍	Do the task together and get 1.1€
Do the task alone and get 2.0€	❍	❍	Do the task together and get 1.0€

Screen 10: Joint Fermi Problem

Sample screen:

Remaining time: 4:48

Instructions

Your team partner (Information)

How many dogs are there in Germany?

Please indicate your **answer in millions** (as decimal character you can use both "," or ".").

Chatbox: Please communicate here with you team partner.

Send

Submit final response

Screen 11: Joint Fermi Problem, Cooperation Questionnaire

Please now indicate how much the following statements are true in relation to working with your team partner.

Please answer the questions with the answer options:

• Do not agree at all • Do not agree • Neutral • Agree • Agree fully

[The questionnaire items are shown in Section A4.]

Screen 12: Joint Fermi-Problem, Alone or with team partner in additional Fermi-Problem?

It is **possible that you will perform another round of the identical task later in the experiment**, where you will again be **paid for the accuracy of your data**.

You can **specify whether you prefer to perform this task alone or with your current team partner**.

- If either you or your team partner selected ”Alone”, you will work the Fermi-Problem alone.
- If you both chose ”Team Partner”, a random mechanism will decide whether you do the task individually or together again. Thus, your team partner will never know what decision you made (and you will never know what decision your team partner made).

Would you prefer to do another round of the task alone, or with your team partner?

❍ Alone

❍ Team partner

Part IV

Screen 1: Interaction Task, Instructions

- Your task is to create **3 questions** for the game **“I have never ...”**.
- The game works as follows:
  - A player starts a sentence with ”I have never...” and ends it with something he or she has never done before.
  - If one of the other players has done that thing before, he or she gets a point deducted.
- Please create **statements where you think about half of the players have done this thing**, but half have not.
- Think other players as “average Munich students”.
- Important: Again, both team partners have to enter the same answers. However, it is **not** crucial whether this is absolutely identical in wording

Your team partner

(based on the allocation mechanism)

Information

Screen 2: Interaction Task

[not displayed]

Part V

Screen 1: Additional task, Revelation

Chance has decided. You

1. do not solve an additional Fermi-Problem.
2. solve an additional Fermi-Problem with a fixed payment.
3. solve an additional Fermi-Problem with a performance-contingent payment.

Screen 2 (possible): Additional task

[not displayed]

Part VI

Screen 1: Socio-Demographic Questionnaire

Please provide the following statistical information.

- Age [integer]
- Gender [male; female; diverse]
- Field of study (faculty/major) [string]
- What language(s) is (are) your native language(s)? [string]
- What is your high school graduation grade? [number; 1-6]
- What is your high school graduation grade in mathematics? [number; 1-6]
- How many times have you participated in an economic laboratory study (including outside of this laboratory)? [0; 1-2; 3-5; 5+]