# Determinants of Peer Selection*

Lukas Kiessling Jonas Radbruch Sebastian Schaube

April 3, 2020

**Abstract**

Peers influence behavior in many domains. We study whom individuals choose as peers and explore individual determinants of peer selection. Using data from a framed field experiment at secondary schools, we analyze how peer choices depend on relative performance, personality differences, and the presence of friendship ties. Our results document systematic patterns of peer choice: friendship is the most important determinant, albeit not the only one. Individuals exhibit homophily in personality and, on average, prefer similar peers who perform slightly better. Our results help to rationalize models of differential and nonlinear peer effects and to understand reference group formation.

**Keywords:** Peer Effects, Peer Selection, Social Comparison, Reference Points

**JEL-Codes:** C93, D01, D03, J24, L23

*Lukas Kiessling: Max Planck Institute for Research on Collective Goods, lkiessling@coll.mpg.de; Jonas Radbruch: University of Bonn and IZA, radbruch@iza.org; Sebastian Schaube: University of Bonn, sebastian.schaube@uni-bonn.de. We thank Lorenz Goette, Ingo Isphording, Sebastian Kube, Pia Pinger, and Matthias Sutter for helpful feedback and comments. We also thank the schools and students that participated in the experiments. This research was undertaken while all authors were at the University of Bonn. We did not obtain an IRB approval for this project because at the time of the experiment there was no IRB at the University of Bonn's Department of Economics, where all three authors were employed when conducting the study. However, we would like to stress that the schools' headmasters approved the study, written parental consent was required for students to take part in the study, and participation was voluntary. Moreover, the experiment is in line with the requirements of the BonnEconLab. The data stem from an experiment registered in the AEA RCT Registry as AEARCTR-0005562. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (projects A01 and A02) is gratefully acknowledged.

# 1 Introduction

Peer effects have been documented across many different environments: skills of classmates influence grades at school (Sacerdote, 2011), and co-workers affect their own performance both in highly stylized settings (Falk and Ichino, 2006) and workplaces (Bandiera, Barankay, and Rasul, 2010; Mas and Moretti, 2009). Peers, therefore, constitute a crucial determinant of our performance. But who are these people whom we select as peers? Answering this question is important for the design and successful implementation of policies that exploit social comparisons and peer effects more generally in educational contexts or firms (e.g., Carrell, Sacerdote, and West, 2013; Kőszegi, 2014; Roels and Su, 2014). However, we do not know much about the underlying process of peer selection. In particular, several sometimes conflicting determinants of peer selection are conceivable. If high-performing peers serve as a reference point, they can motivate individuals to exert more effort (e.g., Abeler et al., 2011; Koch and Nafziger, 2011). Others choose peers to compete with them or they select specific friends, as they might make a task more enjoyable (Park, 2019).

In this paper, we study the selection of peers and link these choices to three potential determinants. Specifically, we measure with whom individuals want to interact and analyze the extent to which these preferences for specific peers depend on (i) relative performance, (ii) personality differences, and (iii) the presence of friendship ties.[1] By studying these determinants, we can quantify the magnitudes of performance and social aspects in the peer selection process, and examine their relationship. In this paper, we therefore highlight the role of individual characteristics for reference group formation. In doing so, we provide a microfoundation for models of differential and nonlinear peer effects.

In order to study the selection of peers, we use data from a framed field experiment with over 600 students aged 12 to 16. In the experiment, students took part in two running tasks, first alone, then simultaneously with a peer. Between the two runs, we collected two different types of preferences for peers, which were subsequently used to form pairs for the second run. More specifically, we elicited students' preferences for peers by allowing them to name up to six classmates with whom they would like to be paired (*name-based preferences*), or choose their peer's relative performance (*performance-based preferences*). Furthermore, we elicited personality measures and the social network within each class.

[1] We differentiate between friends and peers as two distinct, albeit related, concepts. While friends can be peers, not all friends have to be peers across all situations.

This setup has four crucial features to analyze the selection of peers. First, the classroom environment enabled students to state meaningful preferences for known peers (*name-based preferences*) allowing for social aspects. Second, using a running task yields direct measures of performance. This allows us to measure preferences over the relative performance of peers (*performance-based preferences*), creating a preference measure for peers that abstracts from social considerations. Third, our analysis relies on preference measures and not just realized peer selection, which might be influenced by the limited availability of peers. Fourth, by focusing on a single peer in the second run, we circumvent issues associated with multiple reference points (Kahneman, 1992), as students interact with one peer only.

Our analysis proceeds in three steps. First, we describe the heterogeneity in both preference measures. We find that friendship ties play a crucial role. About 80% of the three most-preferred name-based peers are friends. Nonetheless, this number declines to less than 50% when considering the fifth- or sixth-ranked peer. Moreover, we observe that students on average prefer slightly faster peers (0.20 SD in terms of performance in the first run). Yet, this masks large heterogeneities in performance-based preferences: Approximately half of the students want to interact with similar (slightly faster or slower) students. The other half prefer peers who differ in their relative performance by more than one second.

In a second step, we study the determinants of peer selection based on names, factors driving an individual's decision whom to choose as a peer.We estimate the extent to which peer selection patterns can be explained by differences in past performance, differences in personality, and the presence of friendship ties. Our results show that all three dimensions matter, although friendship ties are the most important determinant. If two students are friends, this increases their nomination probability by 39 percentage points. Nonetheless, we find substantial homophily in past performance as well as personality showing that students select peers who are similar to them. A one standard deviation difference in past performance (difference in personality) reduces the probability of selecting a given classmate by approximately 8.8 percentage points (14.1 pp.), which corresponds to 0.42 ranks (0.75 ranks). Importantly, these homophily effects hold even conditional on friendship ties highlighting that friendship alone does not help to explain our results. In other words, students select those friends as peers who are close to them with respect to personality and performance. In addition, we uncover heterogeneous selection patterns across subgroups. We show that the importance of these dimensions differs between males and females as well as high- and low-ability students. In particular, male subjects exhibit stronger homophily in performance and personality than female subjects, and friendship ties as relatively more important for low rather than high ability students.

In a third step, we explore the relationship between performance- and name-based preferences. Our results show that when students select peers based on names, they try to target their preferred relative performance level. This demonstrates that subjects nominate similarly performing peers not only due to homophily, but also due to preferences over relative performance. Still, the social dimensions of peer selection remain unaffected, which highlights the multidimensionality of preferences for peers.

This paper relates to the rich literature on peer effects. Although their importance is undisputed, evidence on whom people select as peers remains scarce. Yet, already Manski (1993, p. 536) noted that the "*informed specification of reference groups is a necessary prelude to [the] analysis of social effects*". This implies that studies on peer effects have to take a stance on who constitutes a reference or peer group, thus specifying who exerts potential peer effects. For example, it is common to specify the set of classmates or co-workers as reference groups on an ad-hoc basis. However, only parts of these groups may constitute relevant peers, and misspecification thereof attenuates peer effect estimates due to measurement error (Cornelissen, Dustmann, and Schönberg, 2017; Dube, Giuliano, and Leonard, 2019). In order to circumvent this problem and to accommodate different peer definitions, a growing body of literature estimates peer effects for different groups separately, differentiating between gender (Beugnot et al., 2019; Black, Devereux, and Salvanes, 2013; Hoxby, 2000; Lavy and Schlosser, 2011) or allowing friends and non-friends to exert different peer effects (Aral and Nicolaides, 2017; Bandiera, Barankay, and Rasul, 2010). We document that friendship is the most important determinant for peer selection, thereby validating the use of friends as a proxy for peers. Moreover, our results show that people exhibit systematic peer choice patterns: they prefer peers of similar performance and with similar personality traits. Hence, only a subset of people may serve as peers and affect subsequent behavior, and as we show in the present paper, this subset may differ by an individual's characteristics such as gender or ability. This motivates the separate estimation of peer effects for different subgroups and demographic characteristics, i.e., differential peer effects, as their effects on behavior may differ.[2] Relatedly, an individual's impact may differ across the ability distribution: it might be large on classmates or co-workers with similar abilities, whereas for others with vastly different ability levels the effect might be small. Nonlinear peer effects implicitly incorporate these patterns of peer selection, since they allow different individuals (in terms of

[2]In principle, differential peer effects can be due to (i) only some individuals being relevant peers, (ii) only some individuals exerting peer effects, or (iii) a combination of both.

their ability) to exert different effects (Burke and Sass, 2013; Mas and Moretti, 2009; Tan and Netessine, forthcoming).

Individuals often self-select into workplaces or organizations based on institutional characteristics or individual traits. For example, employees select into workplaces based on the latter's characteristics (e.g., incentive schemes, Dohmen and Falk, 2010; Niederle and Vesterlund, 2007), students and their parents choose schools based on the academic performance of the school (Burgess et al., 2014), and individuals sort into occupations and organizations based on individual traits (e.g., prosociality, Carpenter and Myers, 2010; Friebel, Kosfeld, and Thielmann, 2019). We advance this literature by studying the process of peer selection within organizations or social groups. In a similar vein, Cicala, Fryer, and Spenkuch (2018) study how students choose peer groups by sorting into specific tasks based on their comparative advantage. Our approach differs and links peer selection to social and non-social determinants to investigate how individuals weight these. By this, our paper adds to a growing literature modeling the selection of friends and the formation of social networks (for an overview, see Graham, 2015; McPherson, Smith-Lovin, and Cook, 2001). However, we deliberately differentiate between friends and peers as two distinct, albeit related, concepts. In our setting, we define peers as those individuals a person wants to interact with in a subsequent task. Friendship ties may be one factor that determines whether to choose a specific individual as a peer. Although it is quantitatively the most important factor in the peer selection process, it is neither a necessary nor sufficient indicator for actual peer choices. Reversely, a given set of peers, i.e., the set of people one repeatedly interacts with, can influence the formation of social networks (Girard, Hett, and Schunk, 2015; Marmaros and Sacerdote, 2006). Methodologically, our analysis follows a framework similar to Girard, Hett, and Schunk (2015), who study friendship formation at a university and find homophily in several personality traits and economic preferences. By contrast, we focus on peer selections within established social networks and allow – among other factors – friendship ties to affect these. We therefore move beyond previous work that takes peers as given.

In our experiment, we induce peer effects by allowing for mutual social comparisons. Thus, we deliberately focus on a specific, highly relevant mechanism with importance for adolescents' educational investments (Bursztyn and Jensen, 2015) as well as their skill formation (Pagani, Comi, and Origo, forthcoming). Furthermore, such social comparisons are also relevant for other populations, environments, and outcomes: they may harm or benefit effort provision and work performance (Ashraf, Bandiera, and Lee, 2014; Bradler et al., 2016; Cohn et al., 2014; Kosfeld and Neckermann, 2011), increase working hours (Rosaz, Slonim, and Villeval, 2016), reduce job satisfaction (Card et al., 2012), change consumption patterns (Kuhn et al., 2011), or negatively affect happiness and overall well-being (Clark and Senik, 2010). Common across these examples is that the resulting behavior is not driven by monetary but non-monetary incentives. In our experiment, we also employ non-monetary incentives (grades for a physical education lesson and intrinsic motivation) in combination with possible comparisons to others' behavior. As we show in a companion paper, these social comparisons indeed induce additional incentives and raise performance in our task (Kiessling, Radbruch, and Schaube, 2020).[3]

In contrast to these studies documenting the consequences of social comparisons, the present paper explicitly analyzes individual preferences for these comparisons and thus relates to a small literature studying whom people compare to: while some studies find that people compare themselves to friends, co-workers, or neighbors (Clark and Senik, 2010; Knight, Song, and Gunatilaka, 2009), others focus on comparisons along performance levels (Falk and Knell, 2004) or with one's own past (Senik, 2009). By contrast, ours is the first study to combine preferences along a social dimension with information about preferences for relative performance and detailed personality measures. By doing so, our results also inform the literature on the specification and formation of reference points (for an overview, see O'Donoghue and Sprenger, 2018). Selected peers may serve as an "aspiration level" or goal that constitutes a reference point, as introduced by Kahneman and Tversky (1979) and used similarly, for example, by Brookins, Goerg, and Kube (2017) and Koch and Nafziger (2011). Some studies (e.g., Cerulli-Harms, Goette, and Sprenger, 2019; Schwerter, 2019) debate the nature and the location of reference points. In this paper, we argue that such social reference points can arise endogenously through peer selection.

The remainder of the paper is structured as follows. The next section presents the data and describes our sample. Section 3 documents two kinds of preferences for peers, based on relative performance and names. We analyze the determinants of peer selection in section 4 and conclude. Finally, section 5 discusses our results and concludes.

## 2 Data

In most environments, it is difficult to observe with whom people compare their own performance. This is especially challenging when there is not only a single peer

[3]Moreover, framed field experiments with similar tasks have been used to examine related phenomena such as competitiveness (Gneezy, Niederle, and Rustichini, 2003), favoritism (Belot and van de Ven, 2011), or discrimination (Rao, 2019).

available as an objective standard, but rather when several peers are observed at the same time. Additionally, the selection of peers may not only be based on preferences over some target performance; rather, it is potentially based on a much broader set of peers' characteristics.

In this paper, we use the dataset of a framed field experiment studying the self-selection of peers (Kiessling, Radbruch, and Schaube, 2020) to overcome these difficulties. The experiment elicited preferences for peers in a sample of over 600 students and thus allows us to study the peer selection process. In addition to these preferences, the experiment elicited the social network and several personal characteristics.

## 2.1 Experiment

The experiment was embedded into physical education classes in German secondary schools. Subjects participated in two "suicide runs", each comprising a series of short sprints along the lines of a volleyball court: alone at the beginning of the experiment, and then with a single peer at the end.[4] This means that during the second run two students performed the task simultaneously, although their times were recorded individually. To exclude any audience effects, no other classmates were present during the first or second run. For the second run, we randomly assigned classes to one of three treatment conditions, which implemented different peer assignment rules: random assignment of peers, self-selection based on names, or self-selection based on relative performance. We informed the students of all three treatments at the beginning of the experiment, but revealed the implemented treatment only directly before the second run took place.

Between the two runs, we elicited the students' preferences for peers based on names and relative performance, which we describe in detail in the following section. To implement self-selected peers, we used the elicited preferences for peers and implemented a "stable roommate" algorithm proposed by Irving (1985). This algorithm has the property that it matches students in pairs that are stable and it is a (weakly) dominant strategy for subjects to report their preferences truthfully.[5] Students were matched within their own gender only.

[4]The exact task was to sprint and turn at every line of the volleyball court. Subjects had to line up at the baseline, from where they started running to the first line of the court (6 meters). After touching this line, they returned to the baseline again, touching the line on arrival. The next sprint took the students to the middle of the court (9 meters), the third to the second attack line (12 meters), and the final sprint to the opposite baseline (18 meters), each time returning back to the baseline. They finished by returning to the starting point. The total distance of this task was 90 meters.

[5]Students were not informed about the exact matching algorithm. Instead, they were truthfully told that it was in their best interest to reveal their true preferences.

All students from grades 7 to 10 (corresponding to age 12 to 16) of three German secondary schools were invited to participate in the experiment. Approximately two weeks prior to the experiment, teachers distributed parental consent forms. These forms contained a brief, very general description of the experiment. Only those students who had handed in the parental consent before the study took place participated in the study.

In order to incentivize students in both runs, we stressed that their performance in both runs would be reported to and graded by their teacher. However, teachers received no information about the pairings from the second run. Hence, students were solely Incentivized by their own performance. After the experiment, we reported the individual times to teachers for grading, and only then students learned about their own time (i.e., there was no feedback during the experiment). As physical education classes often lack explicitly measured student performances, teachers were very happy to receive the recorded times. Moreover, students themselves were intrinsically motivated and expressed a strong interest in their individual time. Between the two runs, students filled out a questionnaire eliciting their preferences for peers according to the two dimensions indicated above. In addition, we elicited sociodemographic characteristics, several personal characteristics, and the social network within each class. In the following, we describe each of these survey elements in more detail. We present the experimental instructions and protocols in Appendix F.

## 2.2 Preference Elicitation

The survey elicited two distinct measures for peer preferences, which were used to implement self-selected peers in the experiment. First, we elicited preferences that abstract from social considerations and focus on preferences over relative performance (*performance-based preferences*). Second, we asked for preferences for those settings in which social information is available (*name-based preferences*). Students were aware that their preferences would be used with a positive probability to determine their peers in the second run and that it was in their best interest to reveal their true preferences. These preferences were elicited for the whole sample and independent of the treatment itself, as the treatment was only assigned after the survey took place. Note that these preferences are revealed, rather than stated, preferences. In particular, there was a positive probability that these preferences would be taken into account due to the random assignment of treatments after the survey.

We first discuss how we elicited preferences for peers based on relative performance. The survey presented subjects with ten categories comprising one-second intervals starting from (4, 5] seconds slower than their own performance in the first run, to (0, 1] seconds slower and (0, 1] seconds faster up to (4, 5] seconds faster. Figure 1 shows the elicitation screen. Subjects indicated the relative performance interval from which they would prefer a peer for the second run, irrespective of the potential peer's identity. This made it very hard for the students to target specific classmates or to base their decision on any characteristics besides the relative performance.[6] In the first row of the table, subjects indicated their most-preferred time interval and thereby the peer's relative performance. In the second row, they indicated their second most-preferred interval, and so forth. The preference for peers based on relative performance corresponds to the highest-ranked time interval. We asked students to rank their seven most-preferred time intervals and therefore elicited a partial ranking of potential peers for performance-based preferences. Naturally, each time interval could only be chosen once, but it potentially included several peers. Similarly, some intervals might have been empty.[7]

Figure 1: Screenshot of the Survey Question on Performance-based Peer Preferences

| | 4-5 seconds slower | 3-4 seconds slower | 2-3 seconds slower | 1-2 seconds slower | 0-1 seconds slower | Own time | 0-1 seconds faster | 1-2 seconds faster | 2-3 seconds faster | 3-4 seconds faster | 4-5 seconds faster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1st Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |
| **2nd Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |
| **3rd Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |
| **4th Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |
| **5th Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |
| **6th Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |
| **7th Preference** | ○ | ○ | ○ | ○ | ○ | X | ○ | ○ | ○ | ○ | ○ |

**Notes:** The figure presents a screenshot of the survey module eliciting the preferences over relative performance. In particular, it elicits a partial ranking of ten categories of relative ability ranging from 4 to 5 seconds slower to 4 to 5 seconds faster.

[6]A potential concern would be that students target those time categories, which they expected to contain their friends. Reassuringly, under the performance-based matching students were paired with peers from their name-based preferences about as often as under the random matching, making such potential targeting behavior at least very ineffective.

[7]This implies the most preferred time interval might not have existed for all participants, especially at the top or bottom of the preference distribution. These students were then matched with a peer from a lower ranked preference. However, more than 50% of all students were paired with a peer from their first three preferences.

The second preference measure elicited preferences for situations in which selection could be based on the identity of the peer (*name-based preferences*), i.e., subjects could condition their decision on all known characteristics of their peers. We asked each student to state his or her six most-preferred peers from the same gender within their class. These classmates had to be ranked, creating a partial ranking of their peers. When subjects nominated a student, they were asked to indicate their belief about the relative performance of the person. The belief elicitation was similar to that of the performance-based preferences described above: subjects had to indicate their belief about the performance of the potential peer in the first run, using the same ten intervals and the same layout as above.

## 2.3 Personal Characteristics and Social Network

The survey also included several measures of personality traits and economic preferences: the Big Five inventory, as used in the youth questionnaire of the German Socioeconomic Panel (Weinhardt and Schupp, 2011), a measure of the locus of control (Rotter, 1966), competitiveness[8], general risk attitude (Dohmen, Falk, et al., 2011), and a short version of the INCOM scale for social comparison (Gibbons and Buunk, 1999; Schneider and Schupp, 2011). For each multiple-item scale, we extracted one underlying factor with a mean of zero and a standard deviation of one.

At the end of the survey, we elicited the social network of the class.[9] The elicitation asked every student to name up to six friends in their class. Due to this constraint, we focus on undirected links in our main analysis and explore different definitions allowing for different friendship intensities (e.g., directed or reciprocal links) in robustness checks. We define that friendship ties exist between person $i$ and $j$ if $j$ was either nominated by student $i$ as a friend, or $j$ herself nominated $i$ as a friend. This means that students can have more than six friends if they were nominated by participants who they did not nominate themselves.[10]

[8]Rather than using tournament entry decisions as measures of competitiveness, we introduced a continuous measure based on a student's agreement to four items on a seven-point Likert scale. The statements were: (i) "I am a person who likes to compete with others", (ii) "I am a person who gets motivated through competition", (iii) "I am a person who performs better when competing with somebody", and (iv) "I am a person who feels uncomfortable in competitive situations" (reversely coded). We then extracted a single principal component factor from those four items.

[9]As preferences were elicited as the first part of the survey, this ordering induced the maximum possible time lag between the two elicitations. This makes potential spillovers between these two measures unlikely.

[10]About 79% of the students nominated six friends. Thus, we were concerned that a maximum of six friends might be restrictive and accordingly define friendships based on undirected rather than directed links. As mentioned before, we relax this definition in robustness checks.

## 2.4 Summary Statistics

We present summary statistics of our sample in Table 1. Overall, we have preference measures and the information about the social network for 619 individuals from 39 classes. The grades range from grades 7 to 10 (aged 12 to 16), and 66% of the students were female.[11] Overall, 73% of all students in a class participated in the experiment.[12] The average class size was about 26 and students had approximately seven friends on average, with 80% of those friends being from a student's own gender. On average, females took 27.57 seconds to finish the first run, which did not vary by age. Male performance improved with age: while the average time of males in grade 7 was 25.33 seconds, it improved to 23.21 seconds in grade 10.

Table 1: Summary Statistics

| | 7th grade | 8th grade | 9th grade | 10th grade | Total |
|---|---|---|---|---|---|
| *Sociodemographic Variables* | | | | | |
| Age | 12.77 | 13.80 | 14.76 | 15.82 | 14.51 |
| | (0.48) | (0.45) | (0.39) | (0.53) | (1.22) |
| Female | 0.60 | 0.61 | 0.66 | 0.73 | 0.66 |
| | (0.49) | (0.49) | (0.47) | (0.45) | (0.47) |
| Number of friends | 6.93 | 7.18 | 7.01 | 6.50 | 6.86 |
| | (1.35) | (1.75) | (1.57) | (1.70) | (1.63) |
| Share of friends of own gender | 0.84 | 0.75 | 0.85 | 0.75 | 0.80 |
| | (0.19) | (0.24) | (0.20) | (0.26) | (0.23) |
| *Times (in sec)* | | | | | |
| Time 1 (Females) | 28.03 | 27.06 | 27.32 | 27.81 | 27.57 |
| | (2.75) | (2.06) | (2.28) | (2.71) | (2.50) |
| Time 1 (Males) | 25.33 | 24.18 | 23.60 | 23.21 | 24.04 |
| | (1.93) | (2.02) | (1.82) | (2.11) | (2.11) |
| *Class-level Variables* | | | | | |
| # Students in class | 25.54 | 25.97 | 26.29 | 25.01 | 25.68 |
| | (2.71) | (1.96) | (2.56) | (3.17) | (2.74) |
| Share of participating students | 0.75 | 0.69 | 0.77 | 0.71 | 0.73 |
| | (0.11) | (0.14) | (0.16) | (0.13) | (0.14) |
| # Classes | 7 | 8 | 11 | 13 | 39 |
| Individuals | 123 | 122 | 179 | 195 | 619 |

[11] These classes are from three Germany secondary schools from the highest track, preparing students for university entry after grade 12 (*Gymnasium*).

[12] Only those students who submitted parental consent forms prior to the experiment, who did not choose to abstain from the study (nobody did), and who were not absent from the physical education lesson took part in the study. Since students did not know the exact date when the study took place, we do not have any concerns about study-related absences from the classes.

# 3 Preferences for Peers

In this section, we describe two types of preferences for peers: first, students could select their most-preferred relative performance (*performance-based preference*); and second, students could select their preferred peers based on names (*name-based preferences*), allowing students to condition their peer choice on all characteristics known to them. These two distinct preference measures allow us to describe the students' peer selection, i.e., whom they prefer as peers.

## 3.1 Performance-based Preferences

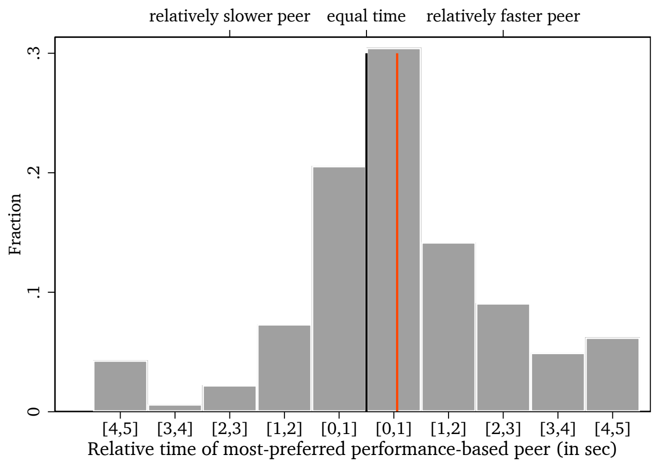
As described in section 2, we elicited a partial ranking over ten categories, with each category corresponding to a one-second time interval of relative performance. Figure 2 presents the preferences for the relative performance of peers. First, turning to the distribution of the most-preferred relative performance (Figure 2a), we find that students prefer performances from the entire possible set. Some students prefer peers who are 4 to 5 seconds slower, whereas others prefer peers who are up to 4 to 5 seconds faster than their own performance. Second, around half of the students prefer similar performing peers, i.e., their most-preferred peer has a performance within one second of their own performance in the first run. Finally, the majority of students prefers faster peers: the median of the distribution lies in the category with slightly faster peers, and on average students prefer peers who were .56 seconds faster in the first run, corresponding to .20 SD in terms of performances in the first run. Figure 2b shows the relationship of the first performance-based preference with the second and third ones. We observe that the second and third preferences are centered around the first performance-based preference.[13] Moreover, Appendix Figures A.2a and A.2b reveal that the distributions across genders is similar, with males preferring somewhat faster peers than females: while males prefer peers who are .90 seconds faster (.31SD in terms of performances in the first run), females select peers who are .38 seconds (.13SD) faster.

In general, these preferences partially support the conjecture of Festinger (1954, p. 121) that people compare themselves with others who are "close to [their] own

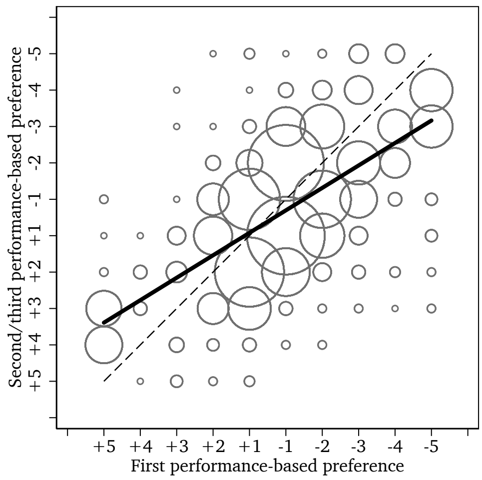
[13]In Appendix Figures A.1a and A.1b, we present the distributions of the second- and third-highest ranked interval. While the probability mass in these histograms is shifted away from an individual's own performance, this is simply an artifact of the limited number of categories, as can be seen in Figure 2b. The categories in which students preferred a much faster or much slower peer as the first preference naturally show a different pattern due to censoring. This explains why we do not find a perfect relationship with a slope of 1. When estimating a Tobit model accounting for censoring at the lower and upper limit, the regression coefficient on the second preferences is .97, with a standard error of .05, and we cannot reject that the coefficient equals unity.

Figure 2: Preferences for Relative Performance

(a) Distribution of Preferences




(b) Relationship among Preferences




**Notes:** Figure (a) presents a histogram of students' preferences over relative performance. The intervals are one-second intervals of relative performances in the first run. Vertical lines indicate own performance (black; equals zero by definition) and mean preference (red; where we used the midpoint of each interval to calculate the mean). Figure (b) presents the relationship of the first performance-based preference and the second/third preference to show that there is no additional information in performance-based preferences when taking more than the first preference into account.

ability" and are in line with evidence from other disciplines noting tendencies to engage in upward comparisons (e.g., Huguet et al., 2001). Nonetheless, this does not hold for all of our subjects. In particular, there is a sizable share of students preferring peers who do differ in ability.

## 3.2 Name-based Preferences

The second set of preferences allows students to state their preferences by selecting peers from a list of their classmates' names. In contrast to performance-based preferences, students can take into account all information known to them when selecting their preferred peer.

Table 2: Share of Selected Peers who are Friends

| Name-based preference | 1st | 2nd | 3rd | 4th | 5th | 6th | Average |
|---|---|---|---|---|---|---|---|
| Share of peers being friends | 0.89 | 0.79 | 0.73 | 0.60 | 0.49 | 0.41 | 0.65 |

**Notes:** This table presents the share of nominated peers for each of the six name-based preferences elicited in the survey who are friends.

We start by presenting the share of selected peers who are also friends of an individual in Table 2. While 89% of all individuals select a friend as their most-preferred peer, this number decreases by about 10 percentage points for each of the following ranks. One might think that this pattern is partially driven by the fact that students do not have a sufficient number of friends of the same gender in the class who they can select. Yet, our data shows that students have on average about seven friends, of which 78% are of their own gender, implying that students on average have 5.3 same-sex friends whom they could select (see Table 1). Thus, students predominately consider their friends as peers. However, they do not solely choose their peers based on friendship ties. Some students seem to avoid some of their friends in favor of other class members. Moreover, even if they select their friends as peers, the question remains how they decide conditional on the social network, which we analyze in the following.

# 4 Determinants of Peer Selection

To explore more formally the underlying determinants of peer selection, we analyze how the three fundamental dimensions – performance, personality, and friendship – relate to peer nominations. For our analysis, we use an empirical model similar to the approach by Girard, Hett, and Schunk (2015), who studied the determinants of link formation in social networks. We take the social network as given and study peer selection on this existing social network. In our main analysis, we focus on the extensive margin, i.e., whom to select as a peer, but consider the intensive margin or ranking of peers in robustness checks. In a second step, we study heterogeneities in those preferences by gender as well as ability, and investigate at the role of one determinant – the preferences for a specific relative performance – in more detail. In particular, we analyze the extent to which students target their performance-based preferences when selecting peers based on names.

## 4.1 Empirical Strategy

In order to analyze the determinants of peer selection in a structured way, we proceed in two steps. First, we analyze the extensive margin of peer selection. Let $y_{ij}$ equal one if individual $i$ nominates individual $j$ and zero otherwise. The dataset therefore contains one observation for each possible nomination within a group. In our main analysis, we define a person to be selected as a peer if this person is part of the first three nominated name-based peers. More specifically, a person is nominated, if she is one of the three students who somebody would be most willing to be paired with in the second run.[14] We want to understand the extent to which $i$'s nomination of $j$ depends on three determinants: (i) differences in terms of performance in the first run ($\Delta^t(t_i, t_j)$), (ii) differences in personality ($\Delta^p(p_i, p_j)$), and (iii) the presence of friendship ties ($F_{ij}$). Additionally, we allow for individual-level heterogeneity in terms of observed and unobserved characteristics by including either individual characteristics ($\Omega_{ij} = \lambda X_i + \pi X_j$) or individual-level fixed effects ($\Omega_{ij} = \nu_i + \nu_j$) as well as some idiosyncratic shock ($\epsilon_{ij}$) for each nomination. Our main specification is therefore given by:

$$y_{ij} = \underbrace{\alpha\Delta^t(t_i, t_j)}_{\substack{\text{Differences} \\ \text{in performance}}} + \underbrace{\beta\Delta^p(p_i, p_j)}_{\substack{\text{Differences} \\ \text{in personality}}} + \underbrace{\gamma F_{ij}}_{\substack{\text{Friendship} \\ \text{ties}}} + \underbrace{\Omega_{ij}}_{\substack{\text{Controls for} \\ \text{heterogeneity}}} + \epsilon_{ij} \tag{1}$$

In our application, we measure differences in terms of the Euclidean distance of the respective characteristic. Hence, similarity in terms of past performance is measured by the absolute distance $\Delta^t(t_i, t_j) = |t_i - t_j|$. In order to measure the difference in personality, we combine the set of standardized personality measures elicited in the survey (Big Five, locus of control, competitiveness, attitudes to engage in social comparisons, and risk attitudes) to define the distance $\Delta^p(p_i, p_j) = \sqrt{\sum_k (p_{ik} - p_{jk})^2}$

[14] Accordingly, we define $y_{ij} = 1$ if and only if $j$ is nominated in $i$'s first three name-based preferences and $y_{ij} = 0$ otherwise. Given that groups were normally not very large and – as shown in Kiessling, Radbruch, and Schaube (2020) – 81% of students were matched with one of their first three preferences in the name-based matching, we consider those individuals as the most important ones. In the Appendix, we relax this definition and consider different cut-offs. Appendix Table B.4 presents the results and shows that our results are qualitatively and quantitatively similar when choosing a different cutoff.

with $k$ indexing different personality measures.[15] Therefore, the coefficients $\alpha$ and $\beta$ can be interpreted as the influence of differences in past performance and personality on the likelihood of nominating someone as a peer. Negative coefficients ($\alpha < 0$, $\beta < 0$) provide evidence of homophily, namely the tendency of individuals to select others with similar characteristics (McPherson, Smith-Lovin, and Cook, 2001). Similarly, positive coefficients ($\alpha > 0$, $\beta > 0$) support heterophily, namely the tendency to avoid others who are similar.

In a series of robustness checks, we also study the intensive margin of peer selection. We adopt the same specification as for the extensive margin (equation (1)) with a crucial modification: We change the dependent variable to be $j$'s rank in $i$'s preferences. Hence, we can study how individuals rank those peers that were selected in a first step. For this, we define $y_{ij}$ to the rank that individual $i$ assigns individual $j$ in the nomination process. The highest-ranked peer receives a score of 6, and this score decreases by one with each rank in the preferences.[16]

## 4.2 Whom Do Individuals Choose as Peers?

We begin our analysis by studying the extensive margin of peer selection, i.e., whom individuals select as peers. Figure 3 provides first evidence of systematic peer selection patterns. Figure 3a shows that, as the difference in initial performance between two individuals increases, the likelihood of nominating the other as a peer decreases. Similarly, Figure 3b shows a similar trend for differences in personality. Taken at face value, these relationships point towards homophily in both performance and personality. Yet, these associations could be driven by a common underlying factor (e.g., friendship ties) and potentially measure the same effect.

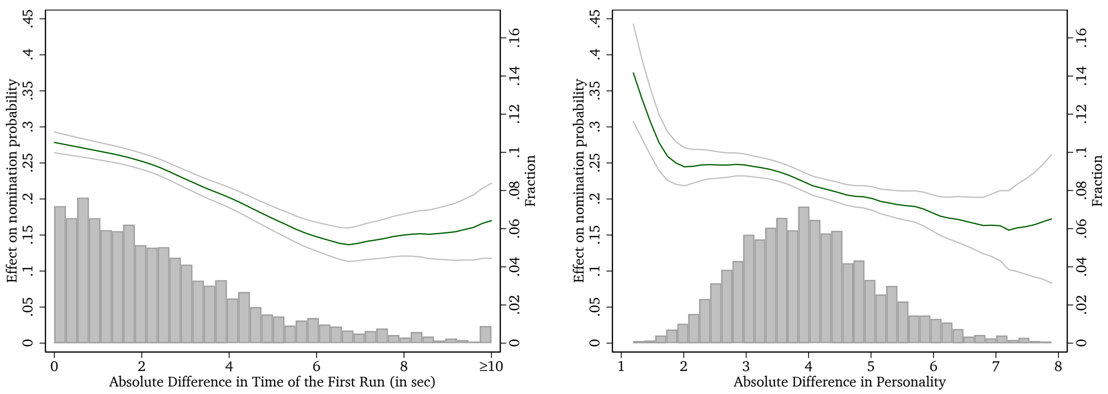
In order to disentangle the contribution of different factors in the peer selection process, Panel A of Table 3 presents a more structured analysis of the extensive margin. In particular, we estimate equation (1) using own and peer characteristics, as well as class fixed effects in column (1), as well as individual and peer fixed effects in column (2). The results show that friendship ties are the most important determinant of peer selection. If two students are friends, this increases the nomination probability

[15]In robustness checks, we allow each of these personality measures to enter separately to explore what is driving the estimated effects. The advantage of the index is that it reduces the degrees of freedom and yields a single coefficient, which makes the impact of personality easily comparable to absolute differences in performances.

[16]The exact score does not matter for our estimates, as the level is taken out by individual fixed effects. For the interpretation of the results, it is important to note that there is a difference of one between the scores. In the main robustness checks, we restrict attention to all individuals who have been nominated as peers, but also present other specifications, in which we estimate censored models on the whole sample.

Figure 3: Descriptive Patterns in the Peer Selection Process

(a) Performance Differences (b) Personality Differences




**Notes:** These figures present local linear regressions of peer nominations on (a) absolute differences in initial performance and (b) absolute differences in personality including 95% confidence intervals. The underlying histograms show the distribution of the respective regressor.

by 38 percentage points. However, we also find evidence of homophily in terms of both performances in the first run, as well as personality. According to the estimates in column (2), a one-second difference in past performance or a difference of one standard deviation in personality reduces the probability of nominating a person by 3-4 percentage points.

While these effects initially seem modest compared to the effect of friendship ties, it is necessary to take into account the underlying distributions of these variables. Conditional on friendship ties, increasing the absolute difference of performances in the first run by one standard deviation (2.10 sec) reduces the nomination probability by 6.3 percentage points. Similarly, increasing the difference in personality by one standard deviation reduces nomination probability by 4.5 percentage points.[17] Moreover, comparing columns (1) and (2) reveals that controlling for unobserved individual-level heterogeneity is important. Individual fixed effects allow us to capture this heterogeneity and thus controls for, e.g., the popularity of students, which is otherwise unmeasured.

These homophily results imply that high-ability students tend to select other high-ability students as peers, whereas low-ability students are more likely to select other low-ability students. Similarly, individuals in our sample tend to select peers who have a similar personality as themselves. Taken together this suggests that students' peer choices lead to segregated groups, which could result in different peer effects.

[17]Appendix Table B.1 presents summary statistics of the absolute differences in these characteristics.

Our analysis deliberately conditions on friendship ties to study the peer selection process in existing social networks. Students already formed friendship ties in the class and select for example only certain friends. Moreover, there is evidence that friendship nominations are driven by homophily themselves (e.g., Girard, Hett, and Schunk, 2015; Selfhout et al., 2010). Hence, to study how much homophily is captured by friendship ties, we exclusively focus on homophily in past performance and personality and omit friendship ties in column (3) of Table 3. Note that this specification is also informative for many settings, in which data on the underlying social network may not be available. On average, both homophily terms increase by about 50%. This indicates that about a third of observed homophily in settings without friendship data may be mediated by friendship ties.[18]

In a next exercise, we want to understand the relationship between the different determinants of peer selection better and analyze their interactions in column (4) of Table 3. This helps to understand whether homophily is more pronounced for friends. We find that differences in performance and personality do not interact and seem to be independent as the resulting coefficient on the interaction term is close to zero and precisely estimated. Although interacting friendship ties with absolute differences in personality yields a negative coefficient – suggesting stronger homophily in personality among friends – this effect is insignificant at conventional levels. Interestingly, we find that existing friendship ties increase the importance of differences in past performance. The homophily in performance among friends almost doubles from 3 percentage points to 5.4 percentage points for a one-second difference in initial performance.

We present additional support for these results in column (5). Here, we restrict the sample to the set of friends and thus ask whether homophily carries over to selection among friends. Our estimates on the two homophily terms for past performance and personality remain significant and increase in magnitude. Even conditional on being part of someone's social network, students therefore seem only to select those friends as peers who share similar characteristics.

As a last step, we provide several robustness checks for our key results presented and alleviate concerns with respect to the analysis presented here.

[18]Relatedly, we observe that the absolute differences in past performance and personality are smaller for the set of friends relative to the full set of potential peers as shown in Appendix Table B.1, indicating some degree of homophily in friendship nominations.

Table 3: Extensive Margin of Peer Selection

| | Student $i$ selected $j$ as a peer | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Abs. Diff. in Time of First Run | -0.016*** | -0.030*** | -0.042*** | -0.030*** | -0.058*** |
| | (0.003) | (0.005) | (0.006) | (0.009) | (0.012) |
| Abs. Diff. in Personality | -0.017*** | -0.040*** | -0.067*** | -0.040*** | -0.092*** |
| | (0.004) | (0.009) | (0.011) | (0.009) | (0.024) |
| Friendship Indicator | 0.381*** | 0.392*** | | 0.515*** | |
| | (0.014) | (0.017) | | (0.049) | |
| Abs. Diff. in Time of First Run × Abs. Diff. in Personality | | | | 0.002 | |
| | | | | (0.002) | |
| Abs. Diff. in Time of First Run × Friendship Indicator | | | | -0.024*** | |
| | | | | (0.007) | |
| Abs. Diff. in Personality × Friendship Indicator | | | | -0.016 | |
| | | | | (0.011) | |
| Controls for heterogeneity | Characteristics | Fixed effects | Fixed effects | Fixed effects | Fixed effects |
| Sample | All | All | All | All | Friends only |
| Observations | 6654 | 6646 | 6646 | 6646 | 2872 |
| Individuals | 612 | 612 | 612 | 612 | 612 |
| $R^2$ | 0.26 | 0.37 | 0.21 | 0.37 | 0.37 |

**Notes:** This table presents the results from the extensive margin analysis using a linear probability model according to equation (1) with an indicator of being nominated as one of the three most-preferred name-based peers as the dependent variable. Column (5) restricts the sample to the set of friends. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

**How do different definitions of friendship ties affect our results?** In our main specification of Table 3, we define friendship ties as undirected. As an alternative, we consider directed and reciprocal friendships in Appendix Table B.2. The coefficient on the friendship indicator increases when using those alternative definitions, which arguably measure more intense friendships, although coefficients on absolute differences in performance and personality remain unaffected. This alleviates the concern that the homophily terms in peer selection are mere artifacts of different friendship intensities.

**Which facets contribute to the homophily in personality?** In Appendix Table B.3, we break down the aggregated personality measure by allowing each of the personality measures to enter the model separately. We find that the effect mainly stems from homophily in agreeableness, tendencies to engage in social comparisons, and – to a lesser extent – competitiveness. Importantly, the coefficients on absolute differences in performances of the first run and the presence of friendship ties remain constant, indicating that the aggregation to a single distance measure for personality traits does not seem to be restrictive.

**Different definitions of the outcome variable.** We elicited a ranking of peers, but for our analysis, we defined an indicator for being selected as a peer if a classmate was selected as one of the three most preferred peers. Since students on average had on average the choice between 10.8 potential peers, we think that choosing a cutoff of approximately 30% is sensible. Yet, we agree that this cutoff is set arbitrarily. To check the sensitivity of our results to this definition, we replicate Table 3 and define an indicator for being selected at all, i.e., being among the six most preferred peers, as the outcome. Appendix Table B.4 demonstrates that our results do not seem to be driven by the choice of the exact cutoff and remain qualitatively similar when using this alternative definition as the estimates remain similar.

**Studying the intensive margin of peer selection.** Although the extensive margin analysis highlights whom students consider as peers, it reveals little about their relative importance. Since peers had to be ranked explicitly, we can exploit this information to explore what makes a peer relatively more important. In contrast to the previous analysis, we now exploit the full richness of the elicited preferences. In particular, we can move beyond a simple binary outcome measure (i.e., whether an individual selected somebody as a peer or not) and study how individuals rank peers who they nominated in a first step. We present the corresponding results in

Appendix C. In Appendix Table C.1, we present analyses analogous to those in Table 3, but use the rank in an individual's preferences as the dependent variable. We find similar determinants for the ranking of peers as for the extensive margin: on average, friends are ranked 1.71 ranks higher than non-friends and students exhibit homophily in performance in the first run, as well as in their personality. In particular, we find that the rank of a peer decreases by 0.18 ranks for each one-second time difference and by 0.27 ranks for each one standard deviation difference in personalities.

**Further robustness checks on the intensive margin.** Appendix C also reports analogous robustness checks for the intensive margin as for the extensive margin. In particular, we show that the same personality measures drive homophily in personality also drive the homophily on the intensive margin[19], using beliefs about the relative performance of others rather than absolute difference in actual performance does not change our conclusions[20], and estimating censored Tobit models retaining the whole sample or ordered logit models to relax the assumption of constant distances between ranks do not alter our conclusions as shown in Appendix Table C.3.

Taken together, these results and robustness checks provide evidence for systematic peer selection patterns. Even conditional on friendship ties, students select similar individuals in terms of their past performance and personality as peers.

## 4.3 Heterogeneities in Peer Selections

While the previous section has documented robust evidence of homophily in the peer selection process, different groups may choose peers differently. In order to predict the effects of different policies, such as assigning students into classrooms or workers into teams, it is important to understand whether peer selection patterns differ across

[19]Column (2) of Table B.3 in the Appendix splits up the aggregated personality measure. Similar to the extensive margin, we observe that agreeableness and the extent of engaging in social comparisons underlie the observed homophily in personality. More specifically, a one standard deviation larger difference in agreeableness or social comparison attitudes is associated with a decrease of 0.25 and 0.16 ranks, respectively.

[20]Note that we only elicited beliefs about relative performance for those students who were nominated as peers. Hence, we can only conduct this robustness check for the intensive margin and not for the extensive one. Nonetheless, as our results in Appendix Table C.1 reveal, our conclusions change neither in a qualitative nor quantitative sense when beliefs rather than actual performances are used. In fact, Appendix D shows that beliefs and actual relative performance are strongly related to each other and validates their consistency. For this, we lever a second belief elicitation over the relative performance of the peer in the first run that was elicited just before the second run took place. This second belief measure and the one used in the elicitation of name-based preferences are indeed highly correlated, indicating that the beliefs are meaningful.

observable characteristics. Hence, we now shed light on the underlying heterogeneity across subgroups.

We present heterogeneities by gender and initial performance in Table 4 analogous to the results in column (2) of Table3. Columns (1) and (2) split the sample by gender and reveal some profound differences in the peer selection behavior of males and females. In particular, we find that males exhibit significantly stronger homophily in past performance (p-value = 0.089) as well as personality (p-value = 0.002). By contrast, females seem to emphasize the presence of friendship ties more, although the difference is not statistically significant (p-value = 0.257).

In columns (4) and (5), we check for heterogeneities in ability. More specifically, we perform a median split of times in the first run within each gender and grade, and estimate equation (1) separately for both groups. Hence, we define ability for each age group and gender separately to make sure that we are not capturing age or gender effects. We find that the effect of friendship ties is more pronounced for slower students (p-value = 0.004), while faster students show larger homophily effects in personality (p-value = 0.067).

We perform analogous analyses for the intensive margin in Appendix Table C.4. Similar to our main results, we find qualitatively similar heterogeneous patterns on the intensive margin as on the extensive one.

Our results highlight differential peer selection patterns across different subgroups. Such heterogeneous patterns have to be taken into account when designing peer or group assignment policies. Moreover, differences in peer selection criteria help to understand why peer effects work differently across different groups: if high-ability students exhibit strong homophily in their peer selection, they will tend to select more similar students as peers. By contrast, low-ability students choose their friends as peers, who may have low or high ability.

Table 4: Heterogeneities on the Extensive Margin of Peer Nominations

| | Student *i* selected *j* as a peer | | | | | |
|---|---|---|---|---|---|---|
| | (1) Males | (2) Females | (3) p-value | (4) Low Abil. | (5) High Abil. | (6) p-value |
| Abs. Diff. in Time of First Run | -0.057*** | -0.027*** | 0.089 | -0.022** | -0.042** | 0.768 |
| | (0.016) | (0.005) | | (0.010) | (0.016) | |
| Abs. Diff. in Personality | -0.105*** | -0.025*** | 0.002 | -0.027** | -0.047*** | 0.067 |
| | (0.020) | (0.009) | | (0.012) | (0.010) | |
| Friendship Indicator | 0.348*** | 0.400*** | 0.257 | 0.434*** | 0.358*** | 0.004 |
| | (0.039) | (0.020) | | (0.021) | (0.022) | |
| Controls for heterogeneity | Fixed effects | Fixed effects | | Fixed effects | Fixed effects | |
| Observations | 1408 | 5238 | | 3303 | 3244 | |
| Individuals | 207 | 405 | | 308 | 301 | |
| $R^2$ | 0.39 | 0.35 | | 0.44 | 0.43 | |

**Notes:** This table replicates column (2) of Table 3 for different subsamples. More specifically, it presents results from a linear probability model according to equation (1) with an indicator of being nominated as one of the three most-preferred name-based peers as the dependent variable. Columns (1) and (2) analyze male and female subsamples, whereas columns (4) and (5) focus on high and low ability, defined according to the gender- and grade-specific median performance in the first run. Columns (3) and (6) present p-values of tests of equality between the two preceding columns. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

## 4.4 Targeting of Preferred Relative Performances

Finally, we examine the role of the preferred relative performances for the peer selection process and examine the relationship between the two sets of preferences. More specifically, we analyze the extent to which students target a relative performance level in the name-based selection process. In Appendix E, we provide graphical evidence on the relation between the preferred relative performance and the selected peers. We observe that both sets of preferences are positively associated with each other, but not perfectly related. Our preferred explanation for this imperfect relation is the fact that preferences for peers are multidimensional. They do not stem from a single factor, but rather are determined by the interplay of several factors.[21] Therefore, we ask whether students target peers with certain performance levels similar to their own, as indicated by the homophily documented in the previous section, or whether they try to target their preferred relative performance when selecting peers based on names.

To illustrate the notion that preferences for peers are indeed multidimensional, we enrich our previous model. In particular, we include the absolute deviation of a name-based peer's performance from the most-preferred performance in the peer selection model in equation (1). Table 5 presents the results of this exercise analogous to Table 3. We observe that the estimated homophily in past performance is much smaller than documented above. Instead, there is a sizable effect of targeting one's preferred relative performance, with highly significant coefficients ranging between 1.0 and 4.9 percentage points. At the same time, the point estimate for differences in performance remains negative in all specifications and significant in some. Together, these two effects are similar in size to the homophily in performance reported in Table 3. Thus, students mainly select individuals who are close to their most-preferred performance (targeting of specific relative performances), but they also select peers who are close to their own performance (homophily in performances). Importantly, the other coefficients on friendship ties and personality differences remain unaffected

[21]A second possible explanation is that the true relation is indeed perfect and measurement error attenuates this association. Subsequently, given a true coefficient of unity, the estimated coefficients correspond to the attenuation factor $\lambda$. Using the relationship $\lambda = 1/(1 + s)$, with $s$ being the noise-to-signal ratio (Cameron and Trivedi, 2005, p. 903f.), we can calculate $s$. Based on the estimates in Table E.1, in which we regress the preferred relative performance on a student's belief about the relative performance of her most-preferred peer, we obtain a coefficient $\hat{\beta} = 0.44$, implying $s = 1.27$. This ratio exceeds one, implying that the beliefs would need to contain more noise components than actual information. We thus conclude from this that measurement error alone is unlikely to be the sole cause for the imperfect relationship.

Table 5: Targeting of Preferred Relative Performances

| | Student $i$ selected $j$ as a peer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Abs. Diff. in Time of First Run | -0.007* | -0.012* | -0.016 |
| | (0.004) | (0.006) | (0.014) |
| Abs. Diff. from Perf.-based Preference | -0.010*** | -0.022*** | -0.049*** |
| | (0.003) | (0.006) | (0.011) |
| Abs. Diff. in Personality | -0.017*** | -0.039*** | -0.094*** |
| | (0.004) | (0.009) | (0.023) |
| Friendship Indicator | 0.381*** | 0.392*** | |
| | (0.014) | (0.017) | |
| Controls for heterogeneity | Characteristics | Fixed effects | Fixed effects |
| Sample | All | All | Friends only |
| Observations | 6654 | 6646 | 2872 |
| Individuals | 612 | 612 | 612 |
| $R^2$ | 0.26 | 0.37 | 0.37 |

**Notes:** This table presents the results of the linear probability model according to equation (1) using an indicator of being nominated as one of the three most-preferred name-based peers as the dependent variable and the absolute deviation from the most-preferred relative performance as an additional explanatory variable. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

by the inclusion of the preference for relative performance.[22,23] This highlights that the previous results are not a mere artifact of a preference for a specific relative performance; rather, it provides evidence that additional social dimensions are important for the peer selection process beyond mere reference points in performance.

These results stress that preferences over a peer's relative performance play a crucial role when selecting peers. While the most-preferred relative performance and the performance of the selected peer are strongly related, these measures do not coincide perfectly; rather, individuals also take into account other dimensions such as the peers' similarity in terms of past performance and personality, as well as existing friendship ties. By selecting peers based on their names, students can therefore condition on a richer information set. This suggests that social comparisons incorporate classical conceptualizations of reference points for effort provision (for an overview of the literature and the specification and formation of classical reference points, see O'Donoghue and Sprenger, 2018), but they also depend on social factors.

[22]Similar to the estimates previously presented, we split the personality index in its components in Appendix Table E.3 and these results also remain the same. That is, the personality effects are mainly driven by agreeableness, tendencies to engage in social comparisons, and competitiveness.

[23]In Appendix Table C.5, we replicate these findings for the intensive margin of peer selection. Appendix Table C.3 reports further robustness checks on the intensive margin using censored models (Tobit) and ordered logits. The results are similar to the ones reported here.

## 5 Discussion and Conclusion

Whom do individuals choose as peers? Answering this question is crucial to understand how peer effects work and how to design policies leveraging them. We use data from a framed field experiment and study preferences for peers to shed light on this issue. We find that individuals choose their peers predominantly, but not exclusively, along their social network. Friendship ties drive peer selections, but our sample also exhibits significant homophily in terms of individuals' performance and personality. Interestingly, among friends, similarity in performance becomes even more important for peer selection. In combination with students targeting their preferred relative times when selecting peers, this indicates that the selection of the peer is determined by a desire to compare one's performance with or even compete against that peer and not just about sharing the task or work, as it might be the case in other situations. While males choose more similar peers than females, low-performing individuals emphasize friendships more than their high-performing counterparts. By eliciting the desired relative performance of a peer, we find that most prefer peers with slightly higher but similar performance, which is in line with findings in social sciences (e.g., Blanton et al., 1999; Huguet et al., 2001). When selecting peers, individuals target a specific relative performance. This suggests that peer selection is therefore based on homophily in personality, friendship ties, and a desired performance level.

Our results have important implications for estimating peer effects, designing mechanisms with social preferences, and policy interventions. First, if friends are more likely to be chosen as peers, this could give rise to relatively larger peer effects of friends compared to non-friends. Similarly, if individuals choose peers with specific performances, these preferences may result in those peers exerting stronger effects than others from the same group. The evidence presented in this paper therefore provides a rationale for estimating models of differential (in terms of gender and friends) or nonlinear peer effects (in terms of own and peer ability).

Second, by demonstrating to whom individuals compare their performance, we inform theories of reference group formation. These insights, in turn, can be used to predict the effect of reorganizations and incentive contracts in a theoretically-disciplined manner (Ederer and Patacconi, 2010; Kőszegi, 2014). Finally, by using reassignment policies, teachers or managers influence the set of people from whom one can choose peers. On the one hand, these policies can have unintended consequences if subgroups emerge (Carrell, Sacerdote, and West, 2013). On the other hand, policy-makers who are aware of such preferences for peers can provide suitable peers by manipulating the pool of potential peers and hence indirectly affect peer selection.

The preferences for peers analyzed in this paper and their link to personal characteristics might be specific to situations with competitive components where only one's own performance matters. Other peers might be selected in cooperative settings or when different mechanisms underlie peer effects. Our results are not meant to be directly transferred to these settings. Rather, we demonstrate that the heterogeneity in social reference points and peer selection is based on systematic patterns of past performance, personality, and friendship ties. These determinants are also likely to matter in other settings. Given the importance of peers for a variety of outcomes, it is imperative to study peer preferences in a broad rang of settings in future research.

At the same time, our results open avenues for new interventions and research projects: If some peers exert positive effects on performance, can we encourage individuals to select into specific peer groups that help them to unfold their full potential? Relatedly, are students aware how their peers affect their own performance? Both of these issues raise the question whether preferences for peers would change if we provided individuals with information about peer effects or even "nudged" people to select specific peers. Our results are therefore a first step towards understanding the different aspects underlying peer choices. Future research on the interaction of personality, selection into environments, and the influence of peers is needed to improve our understanding of social comparison processes, the endogenous formation of peer groups, as well as their long-term consequences.

# References


Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman (2011). "Reference Points and Effort Provision". In: *American Economic Review* 101.2, pp. 470–492.

Aral, Sinan and Christos Nicolaides (2017). "Exercise Contagion in a Global Social Network". In: *Nature Communications* 8.14753.

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee (2014). "Awards unbundled: Evidence from a natural field experiment". In: *Journal of Economic Behavior and Organization* 100, pp. 44–63.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2010). "Social Incentives in the Workplace". In: *Review of Economic Studies* 77.2, pp. 417–458.

Belot, Michèle and Jeroen van de Ven (2011). "Friendships and Favouritism on the Schoolground – A Framed Field Experiment". In: *Economic Journal* 121.557, pp. 1228–1251.

Beugnot, Julie, Bernard Fortin, Guy Lacroix, and Marie Claire Villeval (2019). "Gender and peer effects on performance in social networks". In: *European Economic Review* 113, pp. 207–224.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2013). "Under Pressure? The Effect of Peers on Outcomes of Young Adults". In: *Journal of Labor Economics* 31.1, pp. 119–153.

Blanton, Hart, Bram P. Buunk, Frederick X. Gibbons, and Hans Kuyper (1999). "When Better-than-Others Compare Upward: Choice of Comparison and Comparative Evaluation as Independent Predictors of Academic Performance". In: *Journal of Personality and Social Psychology* 76.3, pp. 420–430.

Bradler, Christiane, Robert Dur, Susanne Neckermann, and Arjan Non (2016). "Employee Recognition and Performance: A Field Experiment". In: *Management Science* 62.11, pp. 3085–3099.

Brookins, Philip, Sebastian Goerg, and Sebastian Kube (2017). "Self-chosen goals, incentives, and effort".

Burgess, Simon, Ellen Greaves, Anna Vignoles, and Deborah Wilson (2014). "What Parents Want: School Preferences and School Choice". In: *Economic Journal* 125.587, pp. 1262–1289.

Burke, Mary A. and Tim R. Sass (2013). "Classroom Peer Effects and Student Achievement". In: *Journal of Labor Economics* 31.1, pp. 51–82.

Bursztyn, Leonardo and Robert Jensen (2015). "How Does Peer Pressure Affect Educational Investments?" In: *Quarterly Journal of Economics* 130.3, pp. 1329–1367.

Cameron, Colin and Pravin Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez (2012). "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction". In: *American Economic Review* 102.6, pp. 2981–3003.

Carpenter, Jeffrey and Caitlin Myers (2010). "Why volunteer? Evidence on the role of altruism, image, and incentives". In: *Journal of Public Economics* 94.11, pp. 911–920.

Carrell, Scott, Bruce Sacerdote, and James West (2013). "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation". In: *Econometrica* 81.3, pp. 855–882.

Cerulli-Harms, Annette, Lorenz Goette, and Charles Sprenger (2019). "Randomizing Endowments: An Experimental Study of Rational Expectations and Reference-Dependent Preferences". In: *American Economic Journal: Microeconomics* 11.1, pp. 185–207.

Cicala, Steve, Roland Fryer, and Jörg Spenkuch (2018). "Self-Selection and Comparative Advantage in Social Interactions". In: *Journal of the European Economic Association* 16.4.

Clark, Andrew and Claudia Senik (2010). "Who Compares to Whom? The Anatomy of Income Comparisons in Europe". In: *Economic Journal* 120.544, pp. 573–594.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frédéric Schneider (2014). "Social Comparison and Effort Provision: Evidence from a Field Experiment". In: *Journal of the European Economic Association* 12.4, pp. 877–898.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg (2017). "Peer Effects in the Workplace". In: *American Economic Review* 107.2, pp. 425–456.

Dohmen, Thomas and Armin Falk (2010). "You Get What You Pay For: Incentives and Selection in the Education System*". In: *Economic Journal* 120.546, F256–F271.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences". In: *Journal of the European Economic Association* 9.3, pp. 522–550.

Dube, Arindrajit, Laura Giuliano, and Jonathan Leonard (2019). "Fairness and frictions: The impact of unequal raises on quit behavior". In: *American Economic Review* 109.2, pp. 620–63.

Ederer, Florian and Andrea Patacconi (2010). "Interpersonal Comparison, Status and Ambition in Organizations". In: *Journal of Economic Behavior and Organization* 75.2, pp. 348–363.

Falk, Armin and Andrea Ichino (2006). "Clean Evidence on Peer Effects". In: *Journal of Labor Economics* 24.1, pp. 39–57.

Falk, Armin and Markus Knell (2004). "Choosing the Joneses: Endogenous Goals and Reference Standards". In: *Scandinavian Journal of Economics* 106.3, pp. 417–435.

Festinger, Leon (1954). “A Theory of Social Comparison Processes”. In: *Human Relations* 7.2, pp. 117–140.

Friebel, Guido, Michael Kosfeld, and Gerd Thielmann (2019). “Trust the Police? Self-Selection of Motivated Agents into the German Police Force”. In: *American Economic Journal: Microeconomics* 11.4, pp. 59–78.

Gibbons, Frederick and Bram Buunk (1999). “Individual Differences in Social Comparison: Development of a Scale of Social Comparison Orientation.” In: *Journal of Personality and Social Psychology* 76.1, pp. 129–147.

Girard, Yann, Florian Hett, and Daniel Schunk (2015). “How Individual Characteristics Shape the Structure of Social Networks”. In: *Journal of Economic Behavior and Organization* 115. Behavioral Economics of Education, pp. 197–216.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini (2003). “Performance in Competitive Environments: Gender Differences”. In: *Quarterly Journal of Economics* 118.3, pp. 1049–1074.

Graham, Bryan S. (2015). “Methods of Identification in Social Networks”. In: *Annual Review of Economics* 7.1, pp. 465–485.

Hoxby, Caroline (2000). “Peer Effects in the Classroom: Learning from Gender and Race Variation”.

Huguet, Pascal, Florence Dumas, Jean M. Monteil, and Nicolas Genestoux (2001). “Social Comparison Choices in the Classroom: Further Evidence for Students’ upward Comparison Tendency and its Beneficial Impact on Performance”. In: *European Journal of Social Psychology* 31.5, pp. 557–578.

Irving, Robert (1985). “An Efficient Algorithm for the “Stable Roommates” Problem”. In: *Journal of Algorithms* 6.4, pp. 577–595.

Kahneman, Daniel (1992). “Reference points, anchors, norms, and mixed feelings”. In: *Organizational Behavior and Human Decision Processes* 51.2, pp. 296–312.

Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision under Risk”. In: *Econometrica* 47.2, pp. 263–292.

Kiessling, Lukas, Jonas Radbruch, and Sebastian Schaube (2020). “Self-selection of peers and performance”.

Knight, John, Lina Song, and Ramani Gunatilaka (2009). “Subjective well-being and its determinants in rural China”. In: *China Economic Review* 20.4, pp. 635–649.

Koch, Alexander K. and Julia Nafziger (2011). “Self-regulation through Goal Setting”. In: *Scandinavian Journal of Economics* 113.1, pp. 212–227.

Kosfeld, Michael and Susanne Neckermann (2011). “Getting More Work for Nothing? Symbolic Awards and Worker Performance”. In: *American Economic Journal: Microeconomics* 3.3, pp. 86–99.

Kőszegi, Botond (2014). “Behavioral Contract Theory”. In: *Journal of Economic Literature* 52.4, pp. 1075–1118.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn (2011). "The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery". In: *American Economic Review* 101.5, pp. 2226–2247.

Lavy, Victor and Analía Schlosser (2011). "Mechanisms and Impacts of Gender Peer Effects at School". In: *American Economic Journal: Applied Economics* 3.2, pp. 1–33.

Manski, Charles (1993). "Identification of Endogenous Social Effects: The Reflection Problem". In: *Review of Economic Studies* 60.3, pp. 531–542.

Marmaros, David and Bruce Sacerdote (2006). "How Do Friendships Form?" In: *Quarterly Journal of Economics* 121.1, pp. 79–119.

Mas, Alexandre and Enrico Moretti (2009). "Peers at Work". In: *American Economic Review* 99.1, pp. 112–145.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook (2001). "Birds of a Feather: Homophily in Social Networks". In: *Annual Review of Sociology* 27.1, pp. 415–444.

Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy Away from Competition? Do Men Compete too much?" In: *Quarterly Journal of Economics* 122.3, pp. 1067–1101.

O'Donoghue, Ted and Charles Sprenger (2018). "Reference-dependent preferences". In: *Handbook of Behavioral Economics: Foundations and Applications 1*. Elsevier Amsterdam.

Pagani, Laura, Simona Comi, and Federica Origo (forthcoming). "The Effect of School Rank on Personality Traits". In: *Journal of Human Resources*.

Park, Sangyoon (2019). "Socializing at Work: Evidence from a Field Experiment with Manufacturing Workers". In: *American Economic Journal: Applied Economics* 11.3, pp. 424–455.

Rao, Gautam (2019). "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools". In: *American Economic Review* 109.3, pp. 774–809.

Roels, Guillaume and Xuanming Su (2014). "Optimal Design of Social Comparison Effects: Setting Reference Groups and Reference Points". In: *Management Science* 60.3, pp. 606–627.

Rosaz, Julie, Robert Slonim, and Marie Claire Villeval (2016). "Quitting and peer effects at work". In: *Labour Economics* 39, pp. 55–67.

Rotter, Julian B. (1966). "Generalized Expectancies for Internal Versus External Control of Reinforcement". In: *Psychological Monographs: General and Applied* 80.1, pp. 1–28.

Sacerdote, Bruce (2011). "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In: ed. by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Vol. 3. Handbook of the Economics of Education. Elsevier, pp. 249–277.

Schneider, Simone and Jürgen Schupp (2011). “The Social Comparison Scale: Testing the Validity, Reliability, and Applicability of the IOWA-Netherlands Comparison Orientation Measure (INCOM) on the German Population”. In: *DIW Data Documentation*.

Schwerter, Frederik (2019). “Social Reference Points and Risk Taking”.

Selfhout, Maarten, William Burk, Susan Branje, Jaap Denissen, Marcel Van Aken, and Wim Meeus (2010). “Emerging late adolescent friendship networks and Big Five personality traits: A social network approach”. In: *Journal of Personality* 78.2, pp. 509–538.

Senik, Claudia (2009). “Direct evidence on income comparisons and their welfare effects”. In: *Journal of Economic Behavior and Organization* 72.1, pp. 408–424.

Tan, Tom Fangyun and Serguei Netessine (forthcoming). “When You Work with a Superman, Will You Also Fly? An Empirical Study of the Impact of Coworkers on Performance”. In: *Management Science*.

Weinhardt, Michael and Jürgen Schupp (2011). “Multi-Itemskalen im SOEP Jugendfragebogen”. In: *DIW Data Documentation*.

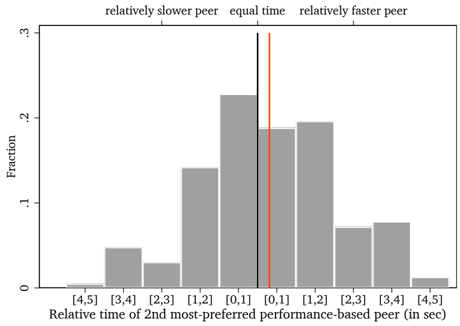
# Appendix – For Online Publication

# A Additional Material for Performance-based Preferences

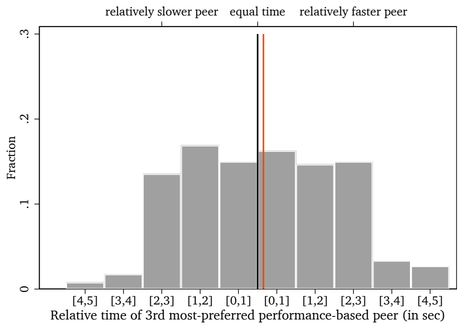
Figure A.1 presents the distribution of the second- and third-most preferred relative performance. We observe that these are also centered around the [0,1] second faster category, but show some different pattern. Nonetheless, as reported in section 3.1, the differences in the distribution are due to targeting the most-preferred relative performance. We thus restrict our attention to the first preference only.

Figure A.1: Distribution of Second and Third Performance-based Peer Preferences

(a) Second Performance-based Preference




(b) Third Performance-based Preference

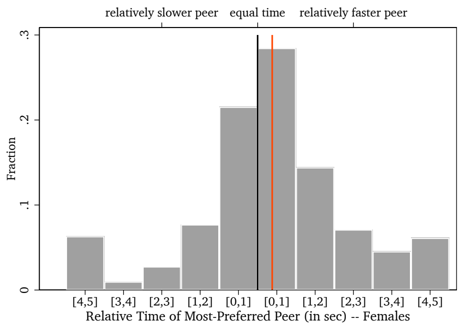



**Notes:** Figure (a) presents a histograms of students' preferences over relative performance. The intervals used here and in the survey are one-second intervals of relative performances in the first run. Vertical lines indicate own performances (black; equals zero by definition) and mean preference (red; where we used the mean of each interval to calculate the mean). Figure (b) presents the relationship of the first performance-based preference and the second/third preference.

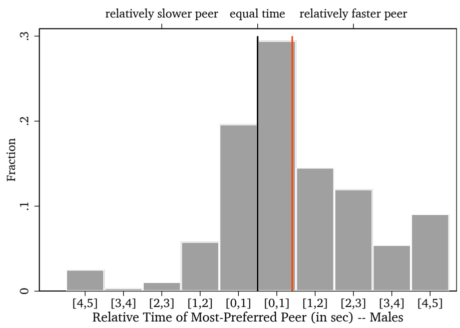
In Figures A.2a and A.2b, we present gender splits of the most-preferred relative performance. While both distributions are relatively similar, males prefer somewhat faster peers than females. On average, females prefer peers being .38 seconds faster, whereas males prefer peers being .90 seconds faster. These correspond to 13 and 31% of a standard deviation in performances of the first run.

Figure A.2: Distribution of Performance-based Peer Preferences by Gender

(a) Females




(b) Males




**Notes:** Figure (a) presents a histograms of students' preferences over relative performance. The intervals used here and in the survey are one-second intervals of relative performances in the first run. Vertical lines indicate own performance (black; equals zero by definition) and mean preference (red; where we used the mean of each interval to calculate the mean). Figure (b) presents the relationship of the first performance-based preference and the second/third preference.

# B Additional Material for Peer Selection Analysis

This appendix provides descriptive statistics and robustness checks for the analysis of the peer selection process. Table B.1 provides summary statistics for the variables used in the analysis. Table B.2 uses alternative definitions of friendship to show that our results are robust with respect to the exact definition. Table B.3 splits up the aggregate measure of personality and includes all dimensions separately. In Table B.4 Panel A we consider someone to be nominated if he is nominated at all, i.e., among the first six most-preferred peers, and zero otherwise. Similarly, Panel (B) estimates a Tobit specification using all potential peers, where we only observe the ranking for six most-preferred peers, and is censored otherwise.

Table B.1: Distribution of Absolute Differences

| | Absolute differences | | | | |
|---|---|---|---|---|---|
| | Mean | SD | 25th perc. | 50th perc. | 75th perc. |
| *Full sample* | | | | | |
| Abs. Diff. in Time of First Run | 2.55 | 2.10 | 0.93 | 2.06 | 3.59 |
| Friendship Indicator | 0.46 | 0.50 | 0.00 | 0.00 | 1.00 |
| Abs. Diff. in Personality | 3.99 | 1.12 | 3.20 | 3.92 | 4.66 |
| Abs. Diff. in Agreeableness | 1.13 | 0.85 | 0.44 | 0.96 | 1.65 |
| Abs. Diff. in Conscientiousness | 1.12 | 0.84 | 0.45 | 0.97 | 1.62 |
| Abs. Diff. in Extraversion | 1.13 | 0.84 | 0.45 | 0.95 | 1.66 |
| Abs. Diff. in Openness | 1.11 | 0.87 | 0.42 | 0.93 | 1.60 |
| Abs. Diff. in Neuroticism | 1.06 | 0.78 | 0.43 | 0.91 | 1.53 |
| Abs. Diff. in Locus of Control | 1.09 | 0.83 | 0.43 | 0.90 | 1.58 |
| Abs. Diff. in Social Comparison | 1.09 | 0.82 | 0.43 | 0.92 | 1.59 |
| Abs. Diff. in Competitiveness | 1.07 | 0.78 | 0.44 | 0.91 | 1.58 |
| Abs. Diff. in Risk Preferences | 1.12 | 0.87 | 0.45 | 0.90 | 1.79 |
| *For friends only* | | | | | |
| Abs. Diff. in Time of First Run | 2.33 | 1.95 | 0.84 | 1.87 | 3.25 |
| Abs. Diff. in Personality | 3.89 | 1.10 | 3.13 | 3.80 | 4.50 |

**Notes:** This table presents summary statistics for absolute differences in several characteristics. The upper panel considers all characteristics for the whole sample, while the lower panel restricts the characteristics to friends only.

Table B.2: Robustness Checks: Alternative Definitions of Friendship Ties

| | Student $i$ selected $j$ as a peer | | |
|---|---|---|---|
| | (1) Undirected | (2) Directed | (3) Reciprocal |
| Abs. Diff. in Time of First Run | -0.030*** | -0.029*** | -0.029*** |
| | (0.005) | (0.005) | (0.005) |
| Abs. Diff. in Personality | -0.040*** | -0.035*** | -0.032*** |
| | (0.009) | (0.008) | (0.007) |
| Friendship Indicator | 0.392*** | 0.454*** | 0.507*** |
| | (0.017) | (0.015) | (0.017) |
| Controls for heterogeneity | Fixed effects | Fixed effects | Fixed effects |
| Observations | 6646 | 6646 | 6646 |
| Individuals | 612 | 612 | 612 |
| $R^2$ | 0.37 | 0.41 | 0.42 |

**Notes:** This table presents the results from the extensive margin analysis using a linear probability model according to equation (1) with an indicator of being nominated as one of the three most-preferred name-based peers as the dependent variable for varying definitions of friendship ties. Column (1) uses undirected friendships as in the main text, column (2) defines friendship ties as directed, while column (3) only considers reciprocal links. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

Table B.3: Robustness Checks: Splitting up Personality Index

| | (A) Peer Nominated | (B) Peer Nomination Ranking |
|---|---|---|
| | (1) | (2) |
| Abs. Diff. in Time of First Run | -0.029*** | -0.173*** |
| | (0.005) | (0.032) |
| Abs. Diff. in Agreeableness | -0.033*** | -0.249*** |
| | (0.009) | (0.054) |
| Abs. Diff. in Conscientiousness | -0.002 | -0.054 |
| | (0.008) | (0.053) |
| Abs. Diff. in Extraversion | -0.015 | -0.104 |
| | (0.010) | (0.063) |
| Abs. Diff. in Openness | -0.004 | 0.025 |
| | (0.010) | (0.069) |
| Abs. Diff. in Neuroticism | -0.014 | -0.141* |
| | (0.009) | (0.083) |
| Abs. Diff. in Locus of Control | -0.003 | -0.056 |
| | (0.009) | (0.079) |
| Abs. Diff. in Social Comparison | -0.022*** | -0.163*** |
| | (0.007) | (0.055) |
| Abs. Diff. in Competitiveness | -0.018* | -0.072 |
| | (0.009) | (0.059) |
| Abs. Diff. in Risk Preferences | -0.005 | 0.006 |
| | (0.009) | (0.086) |
| Friendship Indicator | 0.393*** | 1.716*** |
| | (0.017) | (0.113) |
| Controls for heterogeneity | Fixed effects | Fixed effects |
| Sample | All | All |
| Observations | 6646 | 2756 |
| Individuals | 612 | 612 |
| $R^2$ | 0.37 | 0.40 |

**Notes:** Panel A presents the results from the extensive margin analysis using a linear probability model according to equation (1) with an indicator of being nominated as one of the three most-preferred name-based peers as the dependent variable, but in which we allow for each personality measure to enter separately. Panel B presents analogous results of the intensive margin using the ranking among those who are nominated as peers. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

Table B.4: Alternative Definition of Outcome Variable

| | Student $i$ nominated $j$ as a peer | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Abs. Diff. in Time of First Run | -0.017*** | -0.032*** | -0.046*** | -0.030*** | -0.037*** |
| | (0.004) | (0.006) | (0.007) | (0.009) | (0.009) |
| Abs. Diff. in Personality | -0.016*** | -0.023*** | -0.058*** | -0.040*** | -0.059*** |
| | (0.005) | (0.007) | (0.010) | (0.009) | (0.017) |
| Friendship Indicator | 0.495*** | 0.507*** | | 0.515*** | |
| | (0.022) | (0.024) | | (0.049) | |
| Abs. Diff. in Time of First Run × Abs. Diff. in Personality | | | | 0.002 | |
| | | | | (0.002) | |
| Abs. Diff. in Time of First Run × Friendship Indicator | | | | -0.024*** | |
| | | | | (0.007) | |
| Abs. Diff. in Personality × Friendship Indicator | | | | -0.016 | |
| | | | | (0.011) | |
| Controls for heterogeneity | Characteristics | Fixed effects | Fixed effects | Fixed effects | Fixed effects |
| Sample | All | All | All | All | Friends only |
| Observations | 6654 | 6646 | 6646 | 6646 | 2872 |
| Individuals | 612 | 612 | 612 | 612 | 612 |
| $R^2$ | 0.35 | 0.46 | 0.27 | 0.37 | 0.44 |

**Notes:** This table presents the results from the extensive margin analysis using a linear probability model according to equation (1) with an indicator of being nominated as one of the sixth most-preferred name-based peers (i.e., whether a person is nominated as a peer at all) as the dependent variable. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

# C Intensive Margin of Peer Selection

We replicate our analysis using the intensive margin of peer selection rather than the extensive margin. To do so, we adopt our main specification (1) but use ranks in an individual's preferences as the dependent variable. These ranks range from 6 for the most-preferred peer to 1 (0) for the lowest, but still preferred peer (for unranked peers).

Appendix Table C.1 replicates our main results of Table 3. We relax our the definition of friendship ties from undirected links to directed or reciprocal links in Appendix Table C.2. Column (2) of Table B.3 in Appendix B presents results for different facets of our aggregate personality factor, while Appendix Table C.3 presents results from Tobit models (including all same-gender classmates as potential peers) and ordered logit models (to relax the assumption that the difference between the most-preferred and second most-preferred peer is the same as the difference between the second and the third one). We study heterogeneous peer selection patterns at the intensive margin in Appendix Table C.4. Finally, Appendix Table C.5 presents our results from the horserace between homophily in past performance and targeting of performance-based preferences in the peer selection process.

Across all of these analyses, we find that our results from the main text using the extensive margin of peer selection are qualitatively similar.

Table C.1: Intensive Margin of Peer Selection

| | Student *j*'s rank in *i*'s preferences | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Abs. Diff. in Time of First Run | -0.090*** | -0.178*** | -0.202*** | |
| | (0.013) | (0.035) | (0.037) | |
| Abs. Diff. in Beliefs about Times in First Run | | | | -0.184*** |
| | | | | (0.051) |
| Abs. Diff. in Personality | -0.088*** | -0.270*** | -0.358*** | -0.261*** |
| | (0.019) | (0.073) | (0.081) | (0.079) |
| Friendship Indicator | 2.246*** | 1.710*** | | 1.756*** |
| | (0.070) | (0.115) | | (0.118) |
| Controls for heterogeneity | Characteristics | Fixed effects | Fixed effects | Fixed effects |
| Sample | All | All | All | Beliefs |
| Observations | 6654 | 2756 | 2756 | 2756 |
| Individuals | 612 | 612 | 612 | 612 |
| $R^2$ | 0.36 | 0.40 | 0.30 | 0.39 |

**Notes:** This table presents results of the intensive margin, using the ranking among those who are nominated as peers. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

Table C.2: Robustness Checks: Alternative Definitions of Friendship Ties

| | Student $j$'s rank in $i$'s preferences | | |
|---|---|---|---|
| | (1) Undirected | (2) Directed | (3) Reciprocal |
| Abs. Diff. in Time of First Run | -0.178*** | -0.176*** | -0.170*** |
| | (0.035) | (0.033) | (0.033) |
| Abs. Diff. in Personality | -0.270*** | -0.260*** | -0.222*** |
| | (0.073) | (0.069) | (0.065) |
| Friendship Indicator | 1.710*** | 1.704*** | 1.817*** |
| | (0.115) | (0.106) | (0.102) |
| Controls for heterogeneity | Fixed effects | Fixed effects | Fixed effects |
| Observations | 2756 | 2756 | 2756 |
| Individuals | 612 | 612 | 612 |
| $R^2$ | 0.40 | 0.42 | 0.45 |

**Notes:** This table presents the results from the intensive margin analysis using a specification based on equation (1) with individual $j$'s rank in $i$'s preferences as the dependent variable for varying definitions of friendship ties. Column (1) uses undirected friendships as in the main text, column (2) defines friendship ties as directed, while column (3) only considers reciprocal links. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

Table C.3: Censored Tobit and Ordered Logit Models

| | Student *j*'s rank in *i*'s preferences | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Tobit | Tobit | OLogit | OLogit |
| Abs. Diff. in Time of First Run | -0.145*** | -0.173** | -0.085*** | -0.045 |
| | (0.042) | (0.070) | (0.022) | (0.029) |
| Abs. Diff. from Perf.-based Preference | | -0.235*** | | -0.049*** |
| | | (0.056) | | (0.018) |
| Abs. Diff. in Personality | -0.077* | -0.358*** | -0.091*** | -0.093*** |
| | (0.044) | (0.081) | (0.029) | (0.029) |
| Friendship Indicator | 4.981*** | 4.863*** | 1.515*** | 1.505*** |
| | (0.311) | (0.267) | (0.107) | (0.108) |
| Controls for heterogeneity | Fixed effects | Fixed effects | Characteristics | Characteristics |
| Sample | All | All | All | All |
| Observations | 6654 | 6654 | 2811 | 2811 |
| Individuals | 612 | 612 | 612 | 612 |

**Notes:** This table presents the results from the intensive margin analysis using a censored Tobit model (column (1)-(2)) and ordered logit models (columns (3)-(4)). To ease interpretation, we coded the highest-ranked peer as 6, and this number decreases by one for each subsequent rank. We code all students who were not nominated as part of the six most-preferred peers as 0. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

Table C.4: Heterogeneities on the Intensive Margin of Peer Nominations

| | Student *j*'s rank in *i*'s preferences | | | | | |
|---|---|---|---|---|---|---|
| | (1) Males | (2) Females | (3) p-value | (4) Low Abil. | (5) High Abil. | (6) p-value |
| Abs. Diff. in Time of First Run | -0.162 | -0.180*** | 0.884 | -0.195*** | -0.028 | 0.765 |
| | (0.116) | (0.036) | | (0.060) | (0.112) | |
| Abs. Diff. in Personality | -0.486*** | -0.200** | 0.063 | -0.204* | -0.304** | 0.778 |
| | (0.119) | (0.092) | | (0.113) | (0.128) | |
| Friendship Indicator | 1.441*** | 1.776*** | 0.234 | 2.174*** | 1.436*** | 0.012 |
| | (0.245) | (0.134) | | (0.140) | (0.202) | |
| Controls for heterogeneity | Fixed effects | Fixed effects | | Fixed effects | Fixed effects | |
| Observations | 777 | 1979 | | 1260 | 1230 | |
| Individuals | 207 | 405 | | 308 | 301 | |
| $R^2$ | 0.46 | 0.38 | | 0.51 | 0.50 | |

**Notes:** This table replicates Panel B of Table 3 for different subsamples. More specifically, it presents results from the intensive margin analysis using the ranking among those who are nominated as peers, in which better rankings correspond to higher values of the dependent variable (6: highest, 1: lowest). Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

Table C.5: Targeting of Preferred Relative Performances

| | Student *j*'s rank in *i*'s preferences | |
|---|---|---|
| | (1) | (2) |
| Abs. Diff. in Time of First Run | -0.052 | |
| | (0.048) | |
| Abs. Diff. in Beliefs about Times in First Run | | -0.149*** |
| | | (0.051) |
| Abs. Diff. from Perf.-based Preference | -0.150*** | -0.175*** |
| | (0.043) | (0.029) |
| Abs. Diff. in Personality | -0.269*** | -0.260*** |
| | (0.072) | (0.076) |
| Friendship Indicator | 1.705*** | 1.722*** |
| | (0.117) | (0.121) |
| Controls for heterogeneity | Fixed effects | Fixed effects |
| Sample | All | Beliefs |
| Observations | 2756 | 2756 |
| Individuals | 612 | 612 |
| $R^2$ | 0.40 | 0.41 |

**Notes:** This table presents analogous results to Table 5 for the intensive margin. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

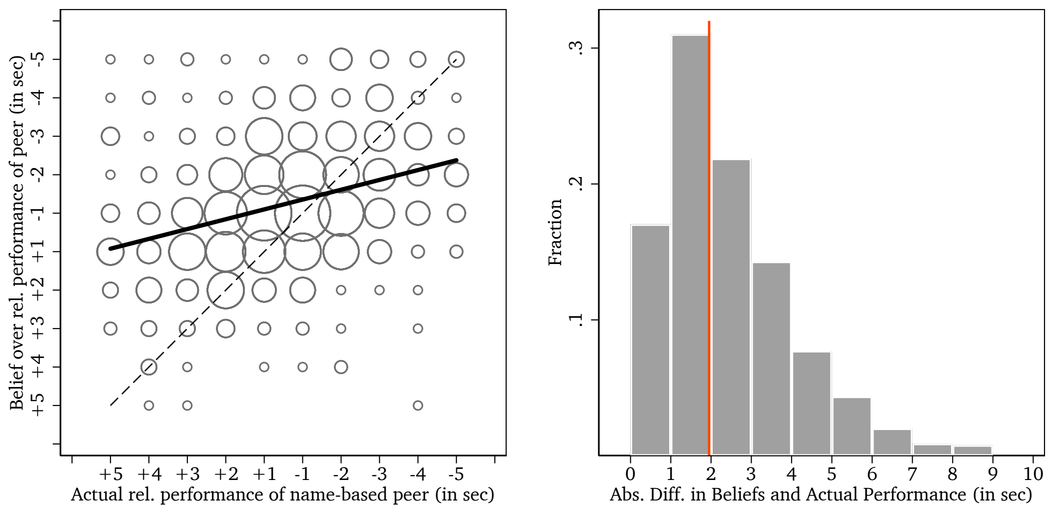
# D Relationship of Beliefs and Actual Performance

In this section, we first describe the relationship between beliefs and actual performance. Afterwards, we provide evidence that the beliefs are meaningful, which is consistent over time by leveraging a second measurement of the same belief.

Beliefs about relative performance and actual relative performance do not necessarily coincide. We therefore check how these two relate to each other. Figure D.1a presents a scatter plot of the belief about relative performance of name-based peers and their actual relative performance. We observe that, although the relationship is not perfect, these two are significantly related, as is confirmed by the corresponding regressions in Table D.1. Figure D.1b displays the absolute differences between the beliefs and the actual relative performance. On average, these two have an absolute difference of 1.95 seconds.

Figure D.1: Relationship of Beliefs and Actual Performance

(a) Relationship (b) Absolute Difference




**Notes:** Figure (a) presents the relationship beliefs about and actual relative performance of the name-based peers. The corresponding regression is presented in Table D.1. Figure (b) presents a histogram of the absolute difference in beliefs and actual performance. The vertical line in (b) indicates mean absolute difference (red; where we used the mean of each interval to calculate the mean). The intervals used here and in the survey are one-second intervals of relative performances in the first run.

Moreover, we are interested whether the beliefs capture pure noise or whether they are constant over time. To check for consistency of the beliefs, we lever a second (binary) belief, elicited right before the second run, and compare it to the beliefs elicited as part of the name-based preferences. The first two columns of Table D.2 use the continuous measure of beliefs about relative performance, as elicited in the name-based preferences, as the dependent variable. The second set of columns uses

Table D.1: Relationship between Beliefs about and Actual Relative Performance

| | (a) Peer's relative time (continuous) | | (b) Peer is faster (binary) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Relative time of most-preferred name-based peer | 0.25*** | 0.24*** | | |
| | (0.04) | (0.04) | | |
| Preferred name-based peer is faster | | | 0.27*** | 0.25*** |
| | | | (0.05) | (0.05) |
| Personality | No | Yes | No | Yes |
| Class FEs, Gender, Age | Yes | Yes | Yes | Yes |
| Individuals | 566 | 562 | 566 | 562 |
| $R^2$ | .21 | .23 | .16 | .17 |

**Notes:** This table presents least squares regressions using a peer's relative performance according to the beliefs of the name-based preferences as the dependent variable. Figure D.1 presents the results graphically. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

a binary version of this indicating whether the student believed that the peer has been faster or slower. The sample is restricted to those students with peers who are nominated somewhere in the name-based preferences (i.e., for whom we have beliefs) and that are matched as a peer in the second run (i.e., only for those for whom we have a second belief measure). This naturally oversampled observations in Name. We thus check whether the pattern differs depending on the treatment. As can be seen, the two measures are significantly related with a correlation of .58. Moreover, this correlation does not significantly vary with the assigned treatment.

Table D.2: Consistency of Beliefs

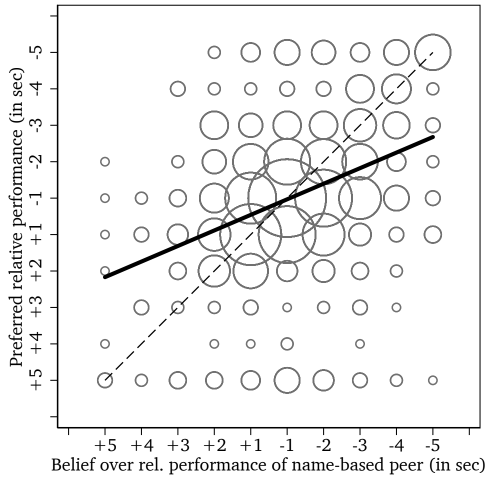
| | (a) Cont. belief | | (b) Binary belief | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Believe peer is faster | 1.96*** | | 0.58*** | |
| | (0.23) | | (0.05) | |
| RANDOM × Believe peer is faster | | 2.00*** | | 0.53*** |
| | | (0.27) | | (0.06) |
| NAME × Believe peer is faster | | 1.92*** | | 0.59*** |
| | | (0.23) | | (0.05) |
| PERFORMANCE × Believe peer is faster | | 2.01*** | | 0.58*** |
| | | (0.23) | | (0.05) |
| N | 345 | 345 | 345 | 345 |
| $R^2$ | .26 | .27 | .3 | .31 |

**Notes:** This table presents least squares regressions using the beliefs about the peer's performance, as elicited in the name-based preferences, as the dependent variable. The sample is restricted to those subjects with peers who are nominated in the name-based preferences and are actually matched for the second run, for which we have elicited a second (binary) belief measure. 89 observations are from students in RANDOM, 180 from NAME, and 87 from PERFORMANCE. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

# E Additional Material for Relationship of Preferences

Figure E.1 and Table E.1 provide a first view on the relation of performance- and name-based preferences. More specifically, we associate the preferred relative performances of each individual with beliefs about the relative performance of their peers nominated in the name-based preferences. We observe a positive relationship between the two measures, as shown in Figure E.1.

Figure E.1: Relationship of Performance- and Name-based Preferences for Peers




**Notes:** The figures present the relationship between performance- and name-based preferences using beliefs about peer's performance. Corresponding regressions are presented in Table E.1.

Table E.1 quantifies this relationship: if students select a peer who they believe is one second faster, this is associated with an increase in the relative performance in the performance-based preference by .44 seconds on average (columns (1) and (2)). Similarly, we observe a significant positive relationship between binary indicators of believing that the most-preferred name-based peer is faster and choosing a faster peer in the performance-based preference in columns (3) and (4). Nonetheless, the relationship between name- and performance-based preferences is not perfect, as would be the case if the preferences over relative performance were the only determinants of name-based preferences. If this were the case, we should observe regression coefficients of unity.

One potential explanation for the imperfect relationship between performance- and name-based preferences is measurement error. Here, we show that measurement error is unlikely to explain the imperfect association alone. Assume that we have classical measurement error and the true coefficient corresponds to one ($\beta = 1$), then by the standard attenuation bias formula (Cameron and Trivedi, 2005, p. 903f.), we

Table E.1: Relationship between Preferences based on Names and Relative Performance

| | Peer preference over rel. perf. | | | |
|---|---|---|---|---|
| | (a) Continuous | | (b) Binary | |
| | (1) | (2) | (3) | (4) |
| Belief over peer's rel. perf. | 0.44*** | 0.44*** | 0.29*** | 0.29*** |
| | (0.06) | (0.06) | (0.04) | (0.04) |
| Personality | No | Yes | No | Yes |
| Class FEs, Gender, Age | Yes | Yes | Yes | Yes |
| Individuals | 627 | 623 | 582 | 578 |
| $R^2$ | .25 | .28 | .17 | .2 |

**Notes:** This table presents least squares regressions using a peer's relative performance in one-second intervals or an indicator for preferring a faster peer according to the performance-based preferences as the dependent variable. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level. Figure E.1 presents the results graphically.

have that if $x^* = x + v$ with $v$ being a mean-zero error with variance $\sigma_v^2$,

$$p \lim \hat{\beta} = \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_v^2}\beta = \lambda\beta = \lambda \tag{2}$$

as $\beta = 1$ and where $\lambda$ is the attenuation factor.[1] Thus, the regression coefficients in Table E.2 correspond to the attenuation factors that would be needed for a perfect relationship. For a more intuitive interpretation, we rewrite the factor in terms of the noise-to-signal ratio $s$ such that $\lambda = 1/(1 + s)$. The noise-to-signal ratio tells us how much noise relative to signals the data should have if the true relationship is given by $\beta = 1$. We reproduce Table E.1 here and additionally present the corresponding noise-to-signal ratios of each coefficient below the corresponding regressions. We find that all ratios exceed one, which implies that the measurements would need to have more noise components than actual information. We thus conclude that measurement error alone cannot explain the imperfect relationship.

[1]For the multivariate case the formula is slightly different, but the basic idea remains the same.

Table E.2: Relationship between Performance- and Name-based Preferences

| | (a) Peer's relative time (continuous) | | (b) Peer is faster (binary) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A: Using name-based beliefs* | | | | |
| Belief over peer's performance | 0.44*** | 0.44*** | | |
| | (0.06) | (0.06) | | |
| Belief over peer's performance (0/1) | | | 0.29*** | 0.29*** |
| | | | (0.04) | (0.04) |
| Personality | No | Yes | No | Yes |
| Class FEs, Gender, Age | Yes | Yes | Yes | Yes |
| Individuals | 627 | 623 | 627 | 623 |
| $R^2$ | .25 | .27 | .17 | .2 |
| Noise-to-signal ratio for $\beta = 1$ | 1.3 | 1.3 | 2.5 | 2.5 |
| *Panel B: Using name-based actual perf.* | | | | |
| Relative time of peer | 0.10*** | 0.09*** | | |
| | (0.03) | (0.03) | | |
| Peer is faster (0/1) | | | 0.04 | 0.03 |
| | | | (0.04) | (0.04) |
| Personality | No | Yes | No | Yes |
| Class FEs, Gender, Age | Yes | Yes | Yes | Yes |
| Individuals | 566 | 562 | 566 | 562 |
| $R^2$ | .11 | .13 | .095 | .12 |
| Noise-to-signal ratio for $\beta = 1$ | 9.2 | 10 | 26 | 28 |

**Notes:** This table presents least squares regressions using a peer's relative performance in one-second intervals or an indicator for preferring a faster peer according to the performance-based preferences as the dependent variable. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level. The reported signal-to-noise ratio describes the extend of measurement error needed if the true relationship is actually perfect (i.e., $\beta = 1$) rather than imperfect ($\beta < 1$). Accordingly, a noise-to-signal ratio larger than one indicates more noise than signal, equal to one corresponds to as much signal as noise and less than one more signal than noise. Figure E.1 presents the results graphically.

Table E.3: Robustness Checks: Splitting up Personality Index

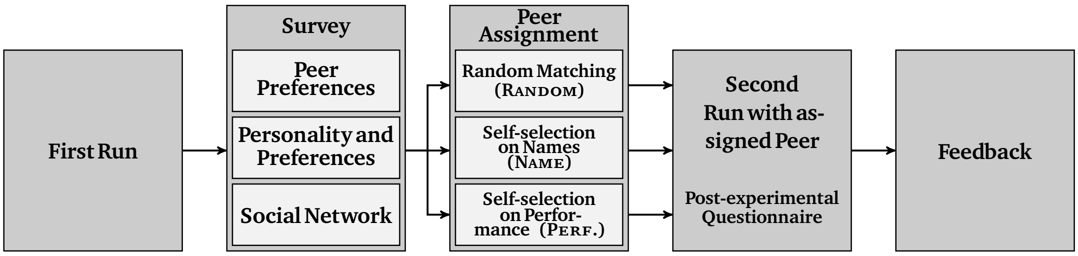
| | (A) Peer Nominated | (B) Peer Nomination Ranking |
|---|---|---|
| | (1) | (2) |
| Abs. Diff. in Time of First Run | -0.011* | -0.052 |
| | (0.006) | (0.045) |
| Abs. Diff. from Perf.-based Preference | -0.022*** | -0.145*** |
| | (0.006) | (0.042) |
| Abs. Diff. in Agreeableness | -0.034*** | -0.242*** |
| | (0.009) | (0.057) |
| Abs. Diff. in Conscientiousness | -0.002 | -0.055 |
| | (0.008) | (0.052) |
| Abs. Diff. in Extraversion | -0.015 | -0.103 |
| | (0.010) | (0.063) |
| Abs. Diff. in Openness | -0.003 | 0.033 |
| | (0.010) | (0.069) |
| Abs. Diff. in Neuroticism | -0.014 | -0.136 |
| | (0.009) | (0.082) |
| Abs. Diff. in Locus of Control | -0.003 | -0.055 |
| | (0.009) | (0.078) |
| Abs. Diff. in Social Comparison | -0.021*** | -0.161*** |
| | (0.007) | (0.054) |
| Abs. Diff. in Competitiveness | -0.018* | -0.072 |
| | (0.010) | (0.057) |
| Abs. Diff. in Risk Preferences | -0.005 | -0.006 |
| | (0.009) | (0.085) |
| Friendship Indicator | 0.394*** | 1.711*** |
| | (0.017) | (0.114) |
| Controls for heterogeneity | Fixed effects | Fixed effects |
| Sample | All | All |
| Observations | 6646 | 2756 |
| Individuals | 612 | 612 |
| $R^2$ | 0.37 | 0.41 |

**Notes:** Panel A presents the results from the extensive margin analysis using a linear probability model according equation (1) with an indicator of being nominated as one of the three most-preferred name-based peers as the dependent variable, but in which we allow for each personality measure to enter separately and add absolute deviations of the most-preferred relative performance as an additional regressor. Panel B presents analogous results of the intensive margin using the ranking among those who are nominated as peers. Standard errors are shown in parentheses and clustered at the class level. *, **, and *** denote significance at the 10, 5, and 1 percent level.

# F Experimental Instructions and Protocol

The following figure illustrates our experimental design. Below, we present translations of the German instructions for the experiment.

Figure F.1: Experimental design




**Notes:** This figure illustrates the experimental design. Peer assignment rules (RANDOM, NAME, PERFORMANCE) are randomly assigned on a classroom level. For this paper, we use running times from the first run and preferences for peers as well as personality traits measured in the survey. The consequences of different peer assignment rules is analyzed in Kiessling, Radbruch, and Schaube (2020).

### Introduction to the Experiment

Welcome, everyone, to today's physical education session. As you might have already noticed, today's session is going to be different. As you already know, you will take part in a scientific study. For that purpose, you have received a parental consent form and handed it back to your teacher. If you have not handed it back to your teacher, you will not take part in the study.

The study is going to be conducted by the three of us: Lukas Kiessling, Sebastian Schaube, and I am Jonas Radbruch. If you have any questions throughout the study, you can address us at any point in time.

The study comprises several parts. For the first part, we would like you to do a running task called suicide runs. My colleague will shortly demonstrate this exercise. *(The following verbal explanation was accompanied by a physical demonstration of the exercise.)*

You start at the baseline of the volleyball court and run to this first line. You touch it with your hand and run back to the baseline. You touch the baseline with your hand and run to the next line. Touch it again, back to the baseline; touch it, and then to the third line, back to the baseline, to the fourth line and then you return to the baseline.

Each of you will run alone, and the goal is to be as fast as possible. After this run, we will hand you a computer to fill out a survey.

After all of you have ran and filled out the survey, you will run for a second time. This time, you will run at the same time as another student. During the survey we will ask you – among other questions – with whom you would like to run. You will receive detailed information about this later on.

The goal during both runs is to be as fast as possible. We will record your running times and hand it to your teacher. Your teacher will grade your performance during both runs.

Before we start with the study, we would like to remind you again that your participation is voluntary. If anyone does not want to take part in the study, then please inform us now.

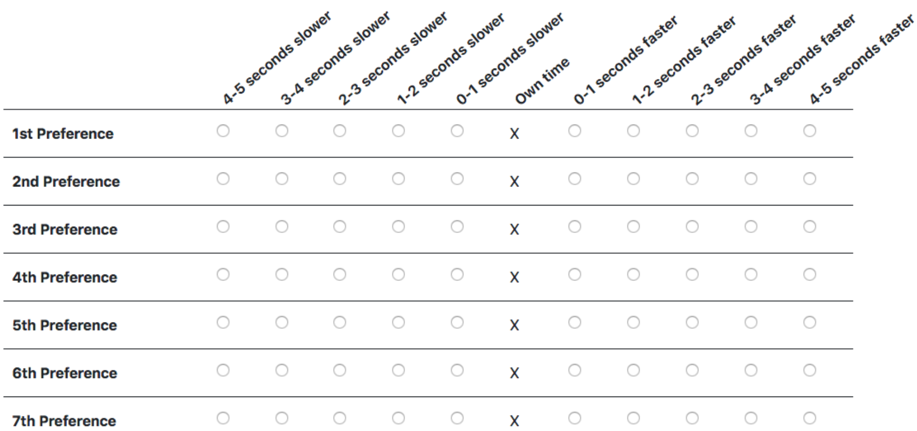
Do you have any further questions? If this is not the case, please start with the warm-up, before we start with the experiment.

*(Introduction was followed by short warm-up by students. After a short warm-up, all students were asked to leave the gym and wait in an accompanying the hallway until they were called into the gym to take part in the first run. We asked students whether they had understood the task and, if necessary, we explained the task again. Directly afterwards, they were asked to leave the gym and were led to a different room. There we asked them to complete the survey on a computer we handed them.)*

### Screenshots of the Preference Elicitation During the Survey

*(The following two screenshots, Figures F.2 and F.3, display translated elicitation screens for performance- and name-based preferences for peers.)*

Figure F.2: Performance-based preferences



| | 4-5 seconds slower | 3-4 seconds slower | 2-3 seconds slower | 1-2 seconds slower | 0-1 seconds slower | Own time | 0-1 seconds faster | 1-2 seconds faster | 2-3 seconds faster | 3-4 seconds faster | 4-5 seconds faster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1st Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 2nd Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 3rd Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 4th Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 5th Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 6th Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 7th Preference | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |

Figure F.3: Name-based preferences

| | ID of running mate | 4-5 seconds faster | 3-4 seconds faster | 2-3 seconds faster | 1-2 seconds faster | 0-1 seconds faster | Own time | 0-1 seconds slower | 1-2 seconds slower | 2-3 seconds slower | 3-4 seconds slower | 4-5 seconds slower |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1st Preference | --------- | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 2th Preference | --------- | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 3rd Preference | --------- | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 4th Preference | --------- | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 5th Preference | --------- | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |
| 6th Preference | --------- | ○ | ○ | ○ | ○ | ○ | x | ○ | ○ | ○ | ○ | ○ |

**Introduction to the Second Run for the Whole Class**

*(Class was gathered for announcement)*
We will shortly start with the second run. For this purpose, a partner has been selected for you. In your class, the partner has been selected *randomly [based on your indication how fast you want your partner to be] [based on the classmates you nominated]*. We would like to remind you that the objective is to be as fast as possible and it is only about your own time. Your teacher will receive a list with your performance, but no information about the pairs.
*(The list with pairs was read out aloud to the students, and students were accompanied to the waiting zone. Students were called into the gym one pair after the other. In the gym, they were led to separate, but adjacent tracks. Each student was accompanied by one experimenter, who recorded their time as well their responses to four additional questions.)*

**Individual Introduction Directly Before the Second Run**

The two of you will now run simultaneously. Your partner has been selected randomly *[based on your indication how fast you want your partner to be] [based on the classmates you nominated].*
*(We then asked each subject to assess their relative performance in the first run.)* Please guess which of you two was faster during the first run?

### Post-run questionnaire after the second run

*(Directly after a pair participated in the second run, we asked each of the two subjects the following three questions in private.)*

(1) How much fun did you have during the second run? Please rate this on a scale from 1 – no fun at all – to 5 – a lot of fun.

(2) If you were to run again, would you prefer to run alone or with a partner?)

(3) How much pressure did you feel from your partner during the second run? Please rate this on a scale from 1 – no pressure at all – to 5 – a lot of pressure.