Uncovering Peer Effects in Social and Academic Skills

Román Andrés Zárate*

First version: November 2018
This version: October 2022

Abstract

This paper explores the impact of adolescent peers who are central in their social network on the formation of social skills and the academic performance of fellow students. I study a large-scale field experiment at selective public boarding schools in Peru that varies the type of peer defined by the median of social centrality and academic achievement. Students are assigned to (i) more socially central versus less socially central peers and (ii) higher-achieving versus lower-achieving peers. Peer effects are more pronounced for social skills than academic performance, and both vary by gender. While socially central peers lead boys to better social skills and improve their later-life outcomes, there are no effects for girls. Meanwhile, higher-achieving peers do not affect boys' academic performance but decrease girls' test scores. Gender differences in how beliefs about one's abilities respond to peer interactions explain both findings, revealing the importance of self-confidence in peer allocation policies.

*Department of Economics, University of Toronto. Email: ra.zarate@utoronto.ca. I am grateful to my advisors Esther Duflo, Parag Pathak, and Joshua Angrist for their guidance throughout this project. I also thank Daron Acemoglu, Abhijit Banerjee, and Frank Schilbach for their helpful suggestions. This paper would not have been possible without the support of the staff of the Ministry Education of Peru; I especially thank María Antonieta Alva and Cinthia Wicht for their collaboration and support, and Paola Encalada for her dedicated work at different stages of this project. I also thank Jose Carlos Chávez, Gerlad Eliers, Raul Panduro, and Paola Ubillus. Heather McCurdy provided generous help with the MIT IRB documentation. I am particularly grateful to Sydnee Caldwell, Joaquín Klot, Matt Lowe, Rob McMillan, and Francine Loza for their feedback. Claudia Allende, Andrei Bartra, Pía Basurto, Zach Brown, Alonso Bucarey, Adriana Camacho, Lorena Caro, Nicolás de Roux, Josh Dean, Juan Dubra, Leopoldo Fergusson, David Figlio, Chishio Furukawa, Isabel Hincapie, Nicolás Idrobo, Raji Jayaraman, Juan Galán, Arda Gitmez, Gabriel Kreindler, Stephanie Majerowicz, Santiago Melo, Mateo Montenegro, Ricardo Montero, Andrés Moya, Alan Olivi, Gautam Rao, Cory Smith, Vira Semenova, Román David Zárate, David Zarruk, and the participants at different conferences and seminars also provided useful comments and suggestions. This project is part of the MineduLab. Funding for this project was generously provided by the Weiss Family Fund and the NAEd/Spencer dissertation fellowship. The experiment was approved by the MIT IRB (ID 1702862092), and is registered at the AEA RCT Registry (ID 0002600) (Zárate, 2017). A previous version of this paper was circulated under the title "Social and Cognitive Peer Effects: Experimental Evidence from Selective High Schools in Peru".

1 Introduction

Adolescence is a crucial stage for developing personality and non-cognitive skills (Heckman and Mosso, 2014). Similar to academic skill formation, socialization in schools during adolescence can influence the accumulation of social skills for life. There is a growing appreciation of the importance of social skills in later life as recent empirical evidence documents that social skills are increasingly valued in the labor market (Deming, 2017; Weinberger, 2014). The number and type of friendships students have during adolescence can also bring long-term benefits to individuals; five more friends during this period raise wages as much as an additional year of schooling (Lleras-Muney et al., 2020). The existing literature has yet to provide causal evidence of the impact of socialization on social skills, alongside the traditional academic peer effects that students could experience at schools.

In this paper, I conduct a field experiment at selective boarding schools in Peru to study the consequences of having more socially central peers and higher-achieving peers on students' social, academic, and college outcomes. I depart from the typical random peer studies by implementing a two-by-two experimental design that manipulates two substantively interesting peer characteristics: (i) social centrality and (ii) academic achievement. I manipulate these peer characteristics by controlling the assignment to dorms in twenty-five boarding schools across all regions of Peru. By classifying students into types and randomizing students to their neighbor type in dormitories, the design guarantees both random and substantial variation in peers' social centrality and academic achievement. This experimental design overcomes concerns of weak variation that arise when estimating peer effects with random allocation to groups (Angrist, 2014).

The design considers two treatments and control arms: (i) more socially central versus less socially central peers and (ii) higher-achieving versus lower-achieving peers. To classify students as more or less socially central in the first treatment, I use the social network's eigenvector centrality[1] (henceforth "centrality"). Centrality is measured using a social network of students listing their friends, study partners, and preferred roommates. Students with a centrality above (below) the median are classified as more (less) socially central. For the second treatment, having higher-achieving versus lower-achieving peers, I use the median score of the admissions test for the boarding schools. I perform a stratified randomization of students to the two treatments that determine student-peer type combinations that vary the proportion of each peer type. Students' names are organized on a list based on these combinations. The schools use the lists to allocate students to specific beds in the dormitories, verifying that the neighbors' type by social centrality and academic achievement coincides with the randomly assigned treatments.

I estimate the impact of both treatments on students' social and academic outcomes. I consider three types of social outcomes:[2] (1) the total number of connections after the intervention, (2) psychological tests that measure social skills (see Appendix D for details), and (3) the number of peers who perceive the student as a leader or as a popular, friendly, or shy person. To measure academic outcomes, I use grades and standardized tests. To account for imperfect compliance between the assignment to treatments and actual neighbors in the dormitories, I exploit the experimental variation in a two-stage least-square (2SLS) model that uses the treatment assignment

[1]Eigenvector centrality measures a student's influence within his or her social network. High values indicate that a student is connected to many other individuals who also have high values.

[2]These outcomes follow the definition in Glaeser et al. (2002) of social capital: an individual's social skills and an individual's connections.

as an instrument for neighbors' characteristics.

The results show that peer effects vary by gender and skill. Being connected to a socially central peer improves a boy's social outcomes, as it affords them more connections and a better network position. Such boys also gain higher scores on psychological tests of social skills and are perceived by their peers to be more sociable. It is mainly the impact on boys with a low baseline centrality that drives these positive effects. The 2SLS model shows that a one-standard-deviation increase in a neighbor's centrality increases the number of connections by about 1.11 (p-value 0.028) for all boys and by 2.17 for less central boys (p-value 0.002). These results are robust to multiple checks, including randomization inference and multiple-hypotheses testing.

The positive effects of socially central peers on social skills persist in later life. Less central boys drop out less of the selective boarding schools and enroll more at better colleges after being assigned to more socially central neighbors. By contrast, being connected to more central peers does not affect girls' social outcomes. It also does not affect the academic performance of boys or girls.

The results show more pronounced peer effects on social outcomes than on academic performance. While peer effects on social outcomes are visible, dominated by gains in boys' social skills, academic peer effects are, on average, zero. If anything, neighboring a higher-achieving peer reduces the academic performance, especially math scores for lower-achieving girls. The 2SLS estimates show that, a one-standard-deviation increase in a neighbor's admission score reduces girls' math scores by 0.044σ (p-value 0.178) and reading scores by 0.120σ (p-value 0.006). For lower-achieving girls, the estimates reveal an even stronger negative treatment effect of -0.116σ (p-value 0.015) on math scores and -0.136σ (p-value 0.060) on reading scores.

I exploit a rich survey and social network information to assess the mechanisms driving the peer effects described above. In particular, I explore whether treatment effects on self-confidence are consistent with the effects on social and academic outcomes or whether the formation of friendships, as in Carrell et al. (2013), may drive the results in this paper.

The idea that peer interactions can affect self-confidence dates back to the "big fish, little pond" effect (Marsh and Parker, 1984). Recent evidence indicates that this mechanism might differ by gender during high school, as adolescence is a period when girls, but not boys, experience a dramatic decline in social confidence (Alan et al., 2019). Social comparisons can also drive gender differences, as female students tend to make more upward and fewer downward social comparisons than male students (Pulford et al., 2018).

Consistent with this evidence, I find that boys are more confident than girls in their social and academic abilities, even after controlling for observable measures of social and academic skills. Peers also affect the self-confidence of boys and girls differently. While socially central neighbors increase low-centrality boys' confidence in their social abilities, they have the opposite impact on girls. The estimates show that girls' self-reported popularity declines when they are exposed to socially central peers. Gender differences in self-confidence in academic abilities are also consistent with the main effects on grades and test scores.

Unlike self-confidence, I show that forming friendships or study groups in and of itself is not enough for peers to influence students' outcomes in this context. Specifically, I test whether the students most affected by their peers are more likely to form friendships or study groups with their neighbors. All students are equally likely to befriend their neighbors regardless of gender, student, or peer characteristics. Hence, this evidence suggests that peers may not influence their friends' outcomes even when social connections are formed.

The paper has three main contributions to the literature. First, the experimental design generates systematic and random variation in peers' characteristics, overcoming weak variation concerns of random allocation to groups. Second, the experimental design manipulates two significant peer characteristics: academic achievement and social centrality. Third, I use rich administrative and survey data to assess peer influences on various measures of cognitive and non-cognitive skills.

The paper contributes to understanding how research designs affect peer-effects estimates (Sacerdote, 2014). Designs in the literature that guarantee substantial variation in peer achievement generally find little evidence of academic peer effects. Most studies find small positive peer effects when schools randomly allocate students to small groups, such as small dorms (Epple and Romano, 2011; Sacerdote, 2001), and sizable effects when schools randomly allocate students to large groups, such as classrooms (Duflo et al., 2011; Carrell et al., 2009; Garlick, 2018). However, larger groups in random assignment designs are usually prone to weak-variation problems by construction. This paper does not rely on random assignments to groups. While I conduct an experiment, my results are aligned with quasi-experimental research studies. Like this paper, quasi-experimental studies do not exploit variation across randomly formed groups and find little academic gains from higher-achieving peers (Abdulkadiroğlu et al., 2014; Duflo et al., 2011).

The paper also contributes to understanding how social skills are formed. Most existing evidence has focused on peer effects on test scores, but the results in this study suggest that peer effects on social skills can be stronger. I show that socially central peers can improve students' social skills in high school, and these findings indicate that social skills are malleable during adolescence. While a substantial body of evidence documents the positive and increasing returns to social skills in the labor market (Deming, 2017), less is known about the formation of those skills. My findings extend the evidence on early childhood (Falk et al., 2018) and primary schools (Rao, 2019; Alan et al., 2021), which has mainly focused on prosociality.

Finally, the paper contributes to understanding possible mechanisms driving peer effects. The findings in this paper are consistent with literature showing that students affect their peers' self-confidence (Marsh and Parker, 1984) and that self-confidence affects performance (Compte and Postlewaite, 2004). This study shows how students' beliefs in their abilities are shaped differently for boys and girls by peer interactions. It adds to the broader evidence, mainly from laboratory studies, on gender differences in belief formation (Mobius et al., 2014; Bordalo et al., 2019; Coffman and Kulkarni., 2020).

The rest of the paper is organized as follows. Section 2 explains the experimental design. Section 3 describes the research context and the implementation of the experiment. Section 4 shows the balance of the experiment and the first stage. Section 5 describes the main outcomes and outlines the empirical strategy. Section 6 documents the results on skill formation. Section 7 discusses the evidence on mechanisms. Section 8 concludes.

2 Experimental Design

In this section, I explain my experimental design and provide a step-by-step guide for its implementation. In the following section (Section 3), I describe the setting and the application of the design at selective boarding schools in Peru.

I use an experimental approach to estimate peer effects. This experimental design accounts for recent concerns in the peer effects literature of weak variation (Angrist, 2014; Booij et al., 2017). In a typical random group assignment (especially to large groups), the composition of all groups will be approximately the same by construction. Given these average similarities across the groups, there will be weak variation in peer characteristics. Therefore, peer effects estimates can be unreliable and exposed to bias.[3] To bypass this obstacle, I introduce an alternative research design. The experimental design aims to generate strong and random variation in peer characteristics in the allocation to dormitories. As dormitories can vary in size and structure across the twenty-five selective boarding schools in the sample, the experimental design must also be adaptable to different dorms of various sizes.

Rather than estimating peer effects directly after building random groups of students and placing them in different dormitories, I randomly assign students to peers categorized by the median in the distribution of the peer attributes of interest. More precisely, I classify peers into two treatments based on where they stand relative to the median of two relevant peer attributes: (i) more socially central versus less socially central peers, and (ii) higher-achieving versus lower-achieving peers. These treatments can be globally defined as peer types. There would only be two peer types in this paper: those with a score above the median are the high types, and those with a score below the median are the low types. I control students' random exposure to different peer types by systematically (but randomly) manipulating the assignment of dorm peers in boarding schools. The student's assignment to each treatment (peer type) serves as an instrumental variable for the average peer characteristics in a peer group.

This experimental design can be fully executed by following four steps that guarantee random and systematic variation in peer attributes and help identify causal peer effects:

1. First, the researcher classifies students into peer types determined by the quantiles in the distribution of the peer attribute of interest. In the simplest case, the classification is determined by the median, with only two peer types[4].

2. In the second step, conditional on a student's type, each student is assigned to a peer type; in my case, either to a high-type-peer treatment (matched with high-type students) or to the control group (matched with low-type students). Treatment arms are equivalent to assigning students to combinations of a student's and a peer type. These combinations guarantee the treatment's predictive power on peer attributes (a strong first stage) as they vary the proportion of each type: 0%, 50%, or 100%.

3. Third, student names are organized on a list that will guide students' allocation to groups. In my case, to dormitories of different sizes. Lists are determined by combinations of student-peer types and are adaptable to dorms of various sizes.

4. Fourth, rather than estimating peer effects directly, the treatment assignment serves as an

[3] Angrist (2014) formulates the weak variation concern, linking it to a weak instrument problem. However, Angrist (2014) does not formally show the direction of the bias. Appendix B illustrates why peer effects estimates relying on random groups can be overestimated. This appendix also reviews the studies using random allocation to groups that show that peer effects estimates increase with group size.

[4] Since there is a trade-off between the number of treatments arms and statistical power, I implement the simplest design with just two types of peers using the median.

instrumental variable for average peer characteristics. A 2SLS approach safeguards this design from imperfect compliance and the exclusion bias.

The identification of peer effects relies on the variation across treatment arms rather than the variation of peer characteristics across the groups (dorms in my setting). This feature eliminates the possibility of measurement error and other factors that Angrist (2014) points out as potential confounders of social influences.

This design is also not subject to the exclusion bias described by Caeyers and Fafchamps (2016) and Guryan et al. (2009). The exclusion bias may arise when the assignment of peers is done without replacement: a student cannot be her own peer. In this paper, however, the identification stems entirely from variation across treatment and control groups. After conditioning on type, all students are equally likely to receive the high-type-peer treatment. Hence, the treatment is uncorrelated to individual characteristics, circumventing exclusion bias concerns. In Appendix Section C, I introduce an example with 12 students to illustrate why my design guarantees strong variation and is not subject to exclusion bias.

2.1 Student-Peer Type Combinations

Following steps (1) and (2) above, the treatment assignment produces student-peer type combinations. In the simplest case, with two types of students, I assign high- and low-type students to high- and low-type peers. If we look at a student and any of her peers in my research design, there would only be three combinations of a student's own type and the peer type.

a) Combination A: composed of high-type students assigned to high-type peers.

b) Combination B: a mixed combination where half of the members are high-type students assigned to low-type peers and the other half are low-type students assigned to high-type peers.

c) Combination C: composed of low-type students assigned to low-type peers.

The following matrix illustrates the composition of these combinations, which are a function of a student's type and her assigned peer type:

Student Type		Peer Type: High	Peer Type: Low
	High	Combination A Proportion High=100% Proportion Low=0%	Combination B Proportion High=50% Proportion Low=50%
	Low	Combination B Proportion High=50% Proportion Low=50%	Combination C Proportion High=0% Proportion Low=100%

Each row in this matrix represents a type of student, and each column is the assigned peer type. The diagonal of the matrix shows all combinations composed of a single student type. Off-diagonal elements of this matrix are symmetrical, as students are matched to peers of the opposite type in Combination B.[5] The size of each of these combinations is determined by the

[5]Notice that for all three combinations to have the same size, two-thirds of students are assigned to peers of their same type, and one-third to the mixed combination.

sample size of randomization strata. For example, if the total sample is thirty students, fifteen students are high-type, fifteen are low-type, and each of these combinations would have ten students. Combination A would have ten high-type students, Combination B, five high- and five low-type students, and Combination C, ten low-type students.

While the entire sample is used to estimate peer effects, the treatment predicts variation in peer characteristics coming from only half of the peers. The difference in the proportion of high-type peers for a high-type student (the matrix's top row, Combination A vs. B) is equal to 0.5.[6] Likewise, the difference in the proportion of high-type peers for a low-type student (the matrix's bottom row, Combination B vs. C) also amounts to 0.5. For both high- and low-type students, half of the peers (the fifty percent point difference) drives changes in average peer characteristics across treatment and control groups.

2.2 Allocation to Groups

Expanding on step (3) above, the experimental design is flexible enough to adapt to groups of various sizes. Participants' names are sorted on a list based on their student-peer type combination, and schools use the list to assign students to specific groups. In my case, dorms or beds in large dormitories. Under this design, the position on the list predicts the final assignment to groups as well as the physical distance between two students in a dorm. For example, students whose names are adjacent on the list are more likely to be roommates or in neighboring beds.

Each student's position on the list is random and determined by the assignment to the treatment as follows. First, student-peer-type combinations are randomly ordered on the list. Second, the students' order in the list is also randomized with one condition; that the list alternates the two student types in the mixed combination (Combination B). This rotation guarantees that the closest neighbors' type on the list coincides with the student's treatment arm. For example, for a student assigned to the high-type peers, the two adjacent names on the list (the one before and the one after) would be of the high type.

Let's consider an example with 12 students (6 low and 6 high types) who get assigned either to the treatment (high-type peers) or the control (low-type peers). First, the student-peer-types combinations are randomly ordered on the list, and one of six potential orders is selected.[7] In this example, I assume that the selected random order is A-C-B. Within each combination, students are randomly ordered while adhering to the condition that students in the mixed group alternate. The following list illustrates this example, with the letters H and L representing the student's type and the blue font identifying students assigned to the treatment (high-type peers).

$$\underbrace{H - H - H - H}_{\text{Group A}} - \underbrace{L - L - L - L}_{\text{Group C}} - \underbrace{H - L - H - L}_{\text{Group B}}$$

This illustrative list represents how the experimental design is adaptable in allocating students to dorms of various sizes. For example, if students were assigned to six dorms of two students each, the assignment would look as follows:

$$\underbrace{H - H}_{\text{Dorm 1}} - \underbrace{H - H}_{\text{Dorm 2}} - \underbrace{L - L}_{\text{Dorm 3}} - \underbrace{L - L}_{\text{Dorm 4}} - \underbrace{H - L}_{\text{Dorm 5}} - \underbrace{H - L}_{\text{Dorm 6}},$$

[6] If we focus on the leave-out proportion, this difference would be higher the smaller the groups.

[7] The six potential orders are: (i) A-B-C, (ii) A-C-B, (iii) B-A-C, (iv) B-C-A, (v) C-A-B, (vi) C-B-A.

Each student ends up with exactly one roommate whose type always corresponds to the assigned treatment arm. The list is also flexible for larger dorms. For example, if dormitories carried four students, the dorms' composition would perfectly align with student-peer-type combinations:

$$\underbrace{H - H - H - H}_{\text{Dorm 1}} - \underbrace{L - L - L - L}_{\text{Dorm 2}} - \underbrace{H - L - H - L}_{\text{Dorm 3}}$$

Noncompliance can nonetheless occur in dormitories of sizes that do not fully conform to student-peer-type combinations. For example, consider dorms with three students. The allocation would be as follows:

$$\underbrace{H - H - H}_{\text{Dorm 1}} - \underbrace{H - L - L}_{\text{Dorm 2}} - \underbrace{L - L - H}_{\text{Dorm 3}} - \underbrace{L - H - L}_{\text{Dorm 4}},$$

There is noncompliance between the treatment and the neighbor's type for some students. For example, the last student in Combination A ends up with low-type roommates despite being assigned to the high-type-peers treatment. While this noncompliance produced by the dorm size would weaken the first stage, the allocation for most students would still guarantee that the assignment to the treatment can be used as an instrument for peer characteristics in the dorm.

2.3 2SLS Framework

Finally, following step (4) of my experimental design, I use the treatment as an instrument for average peer characteristics. To explain my approach, consider a traditional peer effects model describing how peer characteristic x affects students' outcomes:

$$y_{ig} = \alpha + \pi_0 x_{ig} + \pi_1 \overline{x}_{(i)g} + \pi_2 x_{ig} + \varepsilon_{ig}, \tag{1}$$

where y_{ig} is the outcome for individual i when assigned to group g, x_{ig} is a pre-specified exogenous characteristic of i in group g, and $\overline{x}_{(i)g}$ is the leave-out mean of the exogenous characteristic x among students in group g. The parameter π_1 is the causal effect of a change in the leave-out group average of x on students' outcomes.

Consider a researcher interested in estimating parameter π_1 in equation 1. The treatment (high-type peers) can be used as an instrument for average group composition. In particular, being assigned to high-type peers predicts average peer characteristics in a student's group. The following equation shows the first stage of this model:

$$\overline{x}_{(i)g} = \mu_0 + \lambda h_{ig} + \mu_1 x_{ig} + \gamma H_{ig} + \nu_{ig}. \tag{2}$$

Here, h_{ig} takes the value of 1 when student i in group g is assigned to the high-type peer treatment and 0 otherwise. The parameter of interest of this first stage is λ, which captures the impact of the treatment on leave-out peer characteristics $\overline{x}_{(i),g}$. The model includes student-type fixed effect (H_{ig} indicates whether the student is a high-type) as the randomization is performed conditional on student's type. In my design, high-type students are twice as likely as low-type students to receive the treatment. Using student-type fixed effects allows the propensity to receive the treatment to vary across student types. The model also controls for individual attribute x_{ig} at baseline, and ν_{ig} is an error term.

Booij et al. (2017) also use a design that varies the proportion of high-, middle-, and low-GPA peers in tutorial groups for undergraduate students in economics. My design importantly differs from theirs in three main aspects. First, my student allocation to groups goes one step further (step 3). After randomly crafting student-peer-type combinations that vary the proportion of each peer type, I use lists to place these combinations into groups (or dorms) of various sizes and configurations, making my design adaptable to diverse settings. Second, I use an experimental approach with a binary treatment. While Booij et al. (2017) exploit the direct variation in peers' GPA mean and variance across their tutorial groups, my identification strategy relies on the variation across treatment arms. This feature prevents empirical concerns in my design, such as the exclusion bias. Third, as described in Section 3.3, I use my experimental design to analyze peer attributes beyond academic abilities and study peer effects for another substantively interesting peer attribute: social centrality.

3 Setting and Implementation

3.1 Exam Schools in Peru

The Peruvian government runs a network of exam schools—called Colegios de Alto Rendimiento, or the COAR Network −to provide high-quality education to the most talented low-income students during the last three years of secondary school. The first exam school opened near Lima in 2010. As of 2017, there is a COAR school in all 25 regions of Peru. COAR schools are boarding schools, where students stay for the entire academic year and get placed in dormitories with peers. Students can visit their families on weekends, as long as a family member can pick them up from school. Because many students come from a different region than the school's region, they stay at school on weekends, increasing their odds of peer interactions. These added interactions in boarding schools relative to day schools and the high presence of COAR schools across all regions of Peru make this an incredibly convenient context to study peer effects.

The COAR Network comprises twenty-five schools and enrolls approximately three thousand students every year. It is also one of the largest programs in the national budget for education. Every school serves one hundred students per cohort, except for the school in Lima, which serves three hundred students per cohort. Students typically range from ages 14-15 at school entry to 17-18 at graduation. The schools operate Monday through Friday from approximately 7:30 a.m. to 3:45 p.m. and Saturday from 7:30 a.m. to 12:45 p.m. Outside of school hours, students can study, play with their classmates, and do homework in their dormitories. Before the experiment, school directors implemented their own system for allocating students to classrooms and dormitories.

The COAR Network meets the standards of elite Latin American private high schools, where students have access to all the required inputs for high-quality education. COAR schools are deliberately located close to each region's capital city to reduce transportation costs for both families and the government. Upon admission, students receive school materials, uniforms, and a laptop for school use. All schools have a high-quality infrastructure, including a library and excellent scientific laboratories. Students can optionally pursue an International Baccalaureate (IB) degree. Teachers are hired from outside the public school system and receive higher salaries. The government covers all the necessary operating expenses, including laundry service and food.

Applicants are eligible for admission to COAR if they ranked in the top ten of their public school cohort in the previous academic year. The admissions process consists of two rounds. In the first round, applicants take a written test assessing reading comprehension and math skills. The highest-scoring applicants move on to a second round. In this second round, psychologists rate candidates based on two activities: a one-on-one interview, and applicants' observed peer interactions as they complete a set of tasks. I will refer to these as the admissions interview and the social-fitness scores, respectively. Admissions decisions are determined by a composite score of all three tests, the region of origin, and the applicants' school preferences.

The analysis sample includes all students enrolled in the COAR Network in 2017. This sample encompasses three cohorts from the 2015- to the 2017-admission cycles. Figure 1 presents the timeline for the project.

3.2 Data

This study uses both administrative and survey data to implement the experimental design described in Section 2. Administrative data includes the admissions scores listed in the previous subsection: (i) the written test assessing math and reading comprehension, (ii) the admissions interview, and (iii) the social-fitness score determined by a team of psychologists after the student interacts with other applicants (MINEDU, 2017a). I also use administrative data from government databases: socio-demographic data characterizing the population of students (available for 85% of the sample) (MINEDU, 2017d). The latter helps to describe whether a student comes from a household classified as poor or from a rural area. Other administrative data includes scores from a pre-enrollment national-wide standardized test, which is available for the 2016-17 student cohorts (MINEDU, 2016).

Column 1 of Table 1 reports descriptive statistics for students in the COAR Network. While these schools target students in the public school system, who are usually from low-income households, COAR students have diverse socioeconomic backgrounds. For example, 41% of students come from poor households and 59% from non-poor households. Similarly, 26% of students come from rural Peru, and around 50% receive subsidized health insurance. On academic achievement, students enrolled at the COAR Network have higher test scores (1.81 standard deviations, on average) than the average student in the country.[8]

The Ministry of Education also collects administrative data on psychological tests (MINEDU, 2017c). Some of these tests incorporate measures of social skills, including emotional intelligence (Law et al., 2004) and the score in the Reading the Mind in the Eyes test (Declerck and Bogaert, 2008). The latter measure is not self-reported, as the test is a multiple-choice questionnaire with objectively correct answers. It also predicts teamwork abilities at both the group (Woolley et al., 2010) and individual levels (Weidmann and Deming, 2020). Appendix D describes these tests in detail. Using principal component analysis on these tests, I build a baseline social-skills index.

We also ran surveys to characterize students' social interactions (MINEDU and Zárate, 2017). In December 2016, we administered an online survey measuring social interactions and non-cognitive skills. The survey was computer-based and conducted during class hours, ensuring a high participation rate of over 95% in every school. Teams of psychologists in each school proctored the survey. This survey asked students to list the names of their peers in four distinct categories of social interactions: (i) preferred roommates, (ii) friends, (iii) study partners, and (iv) people with whom they play sports or games. These survey questions consisted of a drop-down list of all students in the school cohort, and there were no restrictions on the number of peers

[8]This fact stems from a national-wide test that students take before enrolling at the COAR Network. The Ministry of Education began to collect this data in 2015. Therefore, these scores are not available for the 2015 cohort.

students could list. Column 1 in Table 1 shows that, on average, students have 14.7 connections with a standard deviation of 6.49. As this data was collected in 2016, social centrality and network statistics at baseline are unavailable for the 2017 cohort.

Additionally, the same survey included questions assessing students' perceptions of their peers. Students were asked to rank up to five peers along the dimensions of leadership, friendliness, popularity, and shyness. On average, students were ranked by two to three peers in each social category. Like I did with the social-skills index, I again applied principal component analysis to these four questions to build a peers' perception index.

3.3 Peer Attributes: Social Centrality and Academic Achievement

Using the described data and implementing the first step of this paper's experimental design described in Section 2, I classify students by social centrality and academic achievement at baseline. To classify students as more or less central, I rely on the baseline network survey described in Section 3.2. I use the eigenvector centrality of an aggregate undirected social network that groups the four categories of social interactions listed above.[9] To characterize students as lower or higher-achieving, I use their scores in the admission test to the COAR Network that assess their math and reading comprehension skills.

As the classification is done for two positively correlated attributes (see Appendix Figure A.1), the procedure needs to account for this correlation. To do this, I first perform the classification for one of the two attributes (social centrality or academic achievement) using the school-by-grade-by-gender cell median. Then for the second attribute, the classification is cell and first-attribute types specific.[10] The order of the two attributes, social centrality and academic achievement, is randomized across the cells.

Columns 2-3 of Table 1 present descriptive statistics by social centrality type. More socially central students are less likely to come from poor or rural households. They are also less likely to have subsidized health insurance. As expected, there is also a large gap in measures of social skills between the two groups. More socially central students have more connections, a higher social-skills index score, and are perceived to be more social by their peers than less socially central students are.

Table 1, columns 4-5, report summary statistics by students' academic type. Importantly, in a national standardized test before the application, higher-achieving students scored 0.68σ higher than lower-achieving students did.[11] This shows that even though COAR schools target very talented students, students' achievement still varies widely within them.

[9] Appendix Table A.1 reports standardized coefficients of an OLS regression of social skills measures on the three admissions scores and eigenvector centrality, controlling for school-by-grade-by-gender fixed effects. Eigenvector centrality has a stronger correlation than admissions-test scores do. These results confirm that individuals assessed as very central in the schools' social networks at baseline also have highly developed social skills.

[10] For example, when the first attribute is social centrality, students with an eigenvector centrality above the cell-specific median are classified as more central, and those below the cell-specific median as less socially central. The median for academic achievement, the second attribute, is now cell and social centrality type specific. The reference median for a student is calculated among those that share their gender, school, grade, and social centrality type (either less or more central). Students with a score above this median are classified as higher-achieving and those with a score below as lower-achieving. This procedure guarantees that the proportions of each type in student-peer-types combinations in Figure 2 are 0%, 50%, or 100%.

[11] This data was not available at the time of the experiment. This test is also not available for the 2015 cohort. I defined the academic treatment using the admission test score for these reasons.

3.4 Randomization

To estimate the impact of peers' social centrality and academic achievement on students' outcomes, I follow the second step of the experimental design described in Section 2 with two treatments: (1) more socially central peers and (2) higher-achieving peers. The randomization is analogous to the general case using one peer attribute. However, in this case, I study two peer attributes that yield four types of students: (i) less socially central and lower-achieving, (ii) less socially central and higher-achieving, (iii) more socially central and lower-achieving, and (iv) more socially central and higher-achieving.

This design has two treatments—more socially central peers and higher-achieving peers—and the interaction between them. With four student types, there are ten student-peer type combinations. Figure 2 exhibits these ten possible combinations. Each row represents the student type, each column the peer type by peer attribute, and each cell the student-peer type combination.[12] Each combination takes a different cell color in the symmetrical matrix of Figure 2.

I run the randomization by stratifying at the school-by-grade-by-gender level and by the student's type. The first stratification (school-by-grade-by-gender) is performed since the allocation to dormitories is specific to these strata. The second stratification (student type) is necessary as students were assigned to student-peer type combinations conditional on their type, as described in Section 2.1. The average number of students in each combination depends on the total enrollment by gender at each school grade. On average, 65 students of each gender are in each school peer cohort. Hence, the average size for each combination is 6.5 students.

3.5 Assigning Students to Dormitories

After randomizing students to student-peer-type combinations, I follow the third step of the design and use these combinations to allocate students to COAR school dormitories. The structure of dormitories varies across the schools. For example, while the school in Lima has dormitories of three to five students, the one in Cusco has four dormitories, with approximately seventy-five students per dormitory. Figure 3 shows a picture of school dormitories in Lima, Piura, and Cusco illustrating this variation.

To make the treatment assignment consistent with the various school dorms, I sorted students' names on a list based on the ten student-peer types combinations described in Figure 2. The list was used to allocate students to dorms of small size or to specific beds in large dormitories. Each student's position on the list was determined by randomly ordering student-peer type combinations, and within each combination, randomly ordering students (subject to alternating between the two student types in mixed combinations). This last condition guarantees that adjacent neighbors on the list are always of the type of the assigned treatment, as in Section 2.2.

Most schools (twenty-three out of twenty-five) in the COAR Network used the lists to allocate students to dormitories. There were coordination problems with the other two schools. School administrators generally followed the design protocol, but exceptionally, compliance between the order of students on the list and the actual assignment to dormitories was imperfect.

Finally, first-year students' (the 2017 cohort) and newly enrolled students' (the 2015-16 co-

[12] Combination 1, for example, only includes more socially central and higher-achieving students. Combination 3 comprises (i) less socially central and higher-achieving students and (ii) more socially central and lower-achieving students.

horts) assignments to dorms occurred differently in two aspects. First, as pointed out before, the social centrality measure at baseline is unavailable for the 2017 cohort, as they had not enrolled at the schools when the survey at baseline was collected. However, this did not prevent students from getting assigned to the higher-achieving peers treatment. Second, some schools used the list to allocate first-year students to dorms and classrooms. I, therefore, include a classroom-gender fixed effect in the estimations to maximize statistical power for the higher-achieving peers and avoid confounding roommates and classmates' peer effects.[13] Hence, estimates only consider peer variation originating from neighbors in dormitories.

School administrators generally followed the design protocol, but exceptionally, compliance between students' order on the list and the actual assignment to dormitories was imperfect. Imperfect compliance especially happened when dorm structures and sizes did not conform to the size of the ten student-peer type combinations in Figure 2. But also, occasionally, school administrators decided not to follow the list and changed students' dorm assignments on health- or behavior-related grounds. In gauging the extent of non-compliance issues, I examine whether the distance on the list predicts the likelihood of being actual neighbors in dormitories. I define neighbors in dormitories as roommates for small dormitories (fewer than five students). For larger dormitories (more than five students), neighbors are students in either the same or the adjacent bunk bed. I estimate the following equation to test how the distance on the list affects the likelihood of being neighbors:

$$y_{ij} = \gamma_0 + \sum_{k=1}^{9} \gamma_k \mathbf{1}_{d_{ij}=k} + \nu_{ij}, \tag{3}$$

where y_{ij} is a dummy variable equal to 1 when students i and j are neighbors, and $\mathbf{1}_{d_{ij}=k}$ are dummy variables indicating a distance of k between students i and j on the list. The equation includes nine dummy variables, each representing a distance of 1–9 on the list. A distance of 1 between students i and j implies that the name of student j is either below or above the name of student i on the list.

Panel A of Figure 4 shows that the distance between students on the list predicts whether students are neighbors in the dormitories. The plots show the estimates of γ_k with the respective 95% confidence intervals. A distance of one on the list increases the likelihood of being neighbors by 72 percentage points p.p (p-value 0.000). A distance of two or three is also large and statistically significant, with an increase of 65 p.p. (p-value 0.000) and 48 p.p. (p-value 0.000), respectively. Overall, Panel A of Figure 4 shows a monotonically decreasing effect of the distance on the list and the likelihood of being neighbors. All estimates are weaker from a distance of four upward, with a precise zero at a distance of six.

4 Balance and First Stage

4.1 Balance

First, I show that both treatments are uncorrelated with students' characteristics at baseline by estimating the following equation:

$$y_i = \alpha + \lambda_s s_i + \lambda_a a_i + \sum_{\tau \in \mathcal{T}} \gamma_\tau t_{i\tau} + \nu_{i\tau}, \tag{4}$$

[13] By including this fixed effect, the results stem from the comparison between two students in the same classroom assigned to different types of neighbors in the dormitories.

where s_i and a_i are dummy variables indicating whether individual i is assigned to the more-central peers and the higher-achieving peers' treatment, respectively. I control for student-type fixed effects as the propensity of receiving the treatment varies by student type; $\mathcal{T}$ is the set of students' types by social centrality and academic achievement at baseline, and $t_{i\tau}$ are dummy variables equal to one when student i is of type τ. The parameters of interest are λ_s and λ_a, the correlations of more central and higher-achieving peers with characteristic y_i at baseline. As the randomization is stratified by school×grade×gender×student-type, I also control for these strata.

Tables 2 and 3 report equation 4 estimates on social and academic variables at baseline for all students, boys, and girls. Furthermore, Table 2 also report balance tests by social centrality subgroups, and Table 3 by academic achievement subgroups. As expected from an RCT, I do not reject a zero correlation of the treatments with baseline characteristics. Furthermore, Appendix Tables A.2 and A.3 show balance tests on all other variables available at baseline. These tables include F-statistics for multivariate regressions, displaying balance for both treatments across all student subgroups..

4.2 First Stage

Next, I explore the impact of both treatments on the number of assigned peers of each type and their average characteristics. First, I estimate equation 4 on the number of more central and higher-achieving assigned peers. Table 4, columns 1 and 2 indicate that each treatment changes the number of more central and higher-achieving peers assigned to each group. As a general rule, being assigned to more central peers increases the number of more central peers in a student's group by three, and the same holds for higher-achieving peers.

I also estimate how both treatments impact average peer characteristics; this constitutes the first stage, depicted in equations 5a and 5b:

$$\bar{s}_{p_i} = \theta_s + \delta_s s_i + \phi_s a_i + \sum_{\tau \in \mathcal{T}} \rho_{s,\tau} t_{i\tau} + \xi_i, \tag{5a}$$

$$\bar{a}_{p_i} = \theta_a + \delta_a s_i + \phi_a a_i + \sum_{\tau \in \mathcal{T}} \rho_{a,\tau} t_{i\tau} + \nu_i, \tag{5b}$$

where, δ_s and δ_a are the effects of the more-central-peers treatment on the average social centrality and academic achievement of peers of individual i, respectively. Likewise, ϕ_s and ϕ_a represent the effects of the higher-achieving-peers treatment on the same variables.

Table 4, columns 3 and 4 capture the treatments' effect on the average characteristics of the assigned peers. The more-central-peers treatment increases the average social centrality of the assigned peers by 0.89 standard deviations. Similarly, the higher-achieving-peers treatment raises the average academic achievement of the assigned peers by 0.94 standard deviations. Results also reveal that social centrality and academic achievement are positively correlated at baseline. The higher-achieving-peers treatment positively impacts peers' average social centrality, and the more-central-peers treatment raises peers' average academic achievement. This indirect influence is small compared with the direct impact on peer characteristics emerging from each treatment.

I also estimate equations 5a and 5b on actual neighbors' characteristics rather than on the peers in the student-peer-types combinations. As discussed in Section 3.5, non-compliance between the list and the actual assignment to dormitories could affect the predictive power of the treatments on neighbors' characteristics. The data elucidates that the treatments predict the neighbors' characteristics, confirming that schools followed the list implementation procedures described in the previous section.

Table 4, columns 5 to 8 show the effect of each treatment on students' neighbors in the dormitories. Columns 5 and 6 gather estimates from equation 4 on more central and higher-achieving neighbors. Overall, both treatments increase the number of neighbors of their respective types by about 1.6. Columns 7 and 8 show the effect on the average characteristics of neighbors. Being assigned to more central peers increases the average social centrality of neighbors by 0.54 standard deviations. Similarly, the higher-achieving-peers treatment increases the average academic achievement of neighbors by 0.57 standard deviations. As expected, these effects are smaller for the non-compliance reasons mentioned above than those reported in columns 1 to 4 of Table 3 based on assigned peers. However, they are still very strong and highly significant, supporting the notion of a strong first stage.

5 Outcomes and Empirical Strategy

5.1 Outcomes

Outcomes are grouped into two categories, mapping directly with this paper's main results in Section 6: social and academic outcomes. Social outcomes encompass network degree and centrality, self-reported psychological instruments, and peers' perceptions of students. Academic outcomes include school grades and test scores collected by the Ministry of Education. I also examine data on longer-term outcomes, including dropouts from the COAR Network and college enrollment.

5.1.1 Social Skills Outcomes

Finding reliable measures of social skills is a big challenge. The first outcomes are social networks' statistics after the intervention (MINEDU and Zárate, 2017). These include the network degree (the number of connections) and the social centrality level measured by eigenvector centrality. We collected two waves of network surveys after the intervention (MINEDU and Zárate, 2017) (see the timeline in Figure 1). In each of these, students listed their friends, study partners, and peers with whom they play games and sports. Like the baseline survey, these questions provided a drop-down list of all the students enrolled in the school cohort, and there were no restrictions on the number of peers students could list. I constructed a global network aggregating all questions from both waves. As with other network studies (Breza and Chandrasekhar, 2019; Banerjee et al., 2013, 2014), I consider an undirected network. My results are robust to a network of mutual connections.

I also measure social skills using a battery of psychological tests (MINEDU and Zárate, 2017; MINEDU, 2017c). My main outcome is a psychological social-skills index; built from the first component of a principal component analysis over the entire set of tests. These tests measure openness, extraversion, and agreeableness (among the Big Five personality characteristics) as well as altruism, empathy, leadership, emotional intelligence, and intercultural sensitivity. The index also incorporates the results of the Reading the Mind in the Eyes Test, a test that predicts teamwork abilities at both the group (Woolley et al., 2010) and individual levels (Weidmann and Deming, 2020). Appendix D describes the features of each test.

To account for potential biases in self-reported answers, I consider a third variety of social outcomes: peers' perceptions of social skills (MINEDU and Zárate, 2017). While self-reported psychological tests are frequently used to measure social skills, they are subject to social desirability bias and respondent manipulation. Since social skills surface when interacting with peers, I introduce questions to measure how peers perceive students. Students were asked to rank up to five of their peers along four dimensions: leadership, friendliness, popularity, and shyness (reversed). I construct an index of peers' perceptions using the number of peers that named the student in each category.

Comprehensive of all social outcomes, I use an index that aggregates the four types described above: connections, centrality, psychological tests, and peers' perceptions. I reproduce a similar social skills index with the available measures at baseline. Panel B in Appendix Figure A.1 displays a scatter plot of the two general measures of social skills before and after the intervention. I find a large, positive correlation between the two measures. An OLS regression reveals that a one-standard-deviation increase in the social-skills index at baseline correlates with a 0.41 standard deviation increase in the social-skills index after the intervention.

5.1.2 Academic Outcomes

Students' performance in the 2016 and 2017 cohorts is measured with standardized tests designed by the Ministry of Education (MINEDU, 2017b). For the 2015 cohort, the Ministry relied on students' IB test results. I combine these outcomes to measure test scores, as they are comparable across schools. I also use the student grades assigned by teachers for the 2016 and 2017 cohorts, which are not available for the 2015 cohort, as students' IB test results are taken as their final grades.

5.1.3 Longer-Term Outcomes

I also test the intervention's impact on two types of longer-term outcomes. First, I observe whether students dropped out from the COAR Network. Second, I use administrative data to track students' progression into higher education (MINEDU, 2020). I consider three college outcomes assessing students' enrollment in university and the quality of the university.

There is a vast number of private universities of low quality in Peru. As a result, in 2014, the Peruvian government issued the Universities Law and created the National Superintendence of Higher Education (SUNEDU) to regulate universities. As part of this law, all public and private universities must fulfill a minimum quality requirement to receive government certification. Up to 2019, only seventy-three institutions have received this certification. I use whether a university is certified as the first measure of college quality. The SUNEDU also ranks higher education institutions to inform families about the quality of universities. I use whether a university ranks in the top twenty as a second measure of college quality.

5.2 Empirical Strategy

I begin by estimating the effect of my two treatments—more socially central and higher-achieving peers—on the social skills and academic outcomes described in section 5.1. The following equation estimates the impact of each treatment:

$$y_i = \alpha + \lambda_s s_i + \lambda_a a_i + X_i' \delta + \sum_{\tau \in \mathcal{T}} \gamma_\tau t_{i\tau} + \varepsilon_i. \tag{6}$$

Equation 6 shows how the more-central-peers treatment, s_i, and the higher-achieving-peers treatment, a_i, affect the outcome y_i of individual i. I include student-type fixed effects, γ_τ, as the propensity of receiving the treatments varies by student type (Rosenbaum and Rubin, 1983), and student-type fixed effects account for these differences. I also include a gender-classroom fixed effect for the first-year students as described above.

The parameters of interest in equation 6, λ_s and λ_a, denote the causal impact of the more-central-peers and higher-achieving-peers treatments, respectively. The vector X_i' is a set of pre-defined covariates at baseline that include the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. The results are robust to selecting these covariates using the post-double-selection Lasso method developed by Belloni et al. (2014a,b). The standard errors are clustered at the student-type×group-type-level since all the students within this unit share the same treatment peers (Abadie et al., 2017). I also report the randomization inference p-values for my main results (Athey and Imbens, 2017; Young, 2018).

To estimate heterogeneous effects by gender, I estimate equation 6, including the interaction of the two treatments with the dummy variable *boy*. The following equation describes this model:

$$y_i = \alpha + \lambda_s s_i + \lambda_a a_i + \phi_s s_i \times boy_i + \phi_a a_i \times boy_i + X_i'\delta + \sum_{\tau \in \mathcal{T}} \gamma_\tau t_{i\tau} + \varepsilon_{i\tau}, \tag{7}$$

where ϕ_s and ϕ_a are the differentiated impacts on boys of each treatment.

I also use the 2SLS framework from section 2.3 to exploit experimental variation in a model with two endogenous variables. I use this to jointly estimate the effect of peers' characteristics on students' social and academic outcomes. The following equation introduces this model:

$$y_i = \theta + \beta_s \overline{s}_{n_i} + \beta_a \overline{a}_{n_i} + X_i'\delta + \sum_{\tau \in \mathcal{T}} \varrho_\tau t_{i\tau} + \varepsilon_{i\tau}. \tag{8}$$

Here, $\overline{s}_{n_i}$ and $\overline{a}_{n_i}$ denote the average baseline social centrality and academic achievement of neighbors of student i. The parameters of interest are β_s and β_a: the effect of a one-standard-deviation increase in the average centrality and academic achievement of neighbors on students' outcomes. The first stage of this model is depicted by equations 5a and 5b and columns 7 and 8 of Table 4 presents these estimates.

5.3 RCT Registry

The experiment was registered in the AEA RCT Registry (Zárate, 2017). The original design considered the impact of each of the four peer types, which is equivalent to adding an interaction term in equation 6. The original project had two main hypotheses: (1) socially central peers can improve social outcomes, and (2) the interaction of socially central and higher-achieving peers can generate positive academic peer effects and explain the heterogeneity found in the literature.

The main empirical strategy in the design's current form does not include the interaction term to gain precision. The interaction term does not affect the main results on social outcomes, and the coefficient associated with it is a precise zero. For academic outcomes, I cannot reject the hypothesis that the interaction term is equal to zero, but the estimates are less precise.[14]

[14] Appendix Table A.7 reports the treatment effects with the interaction on the composite index of social skills. Appendix Table A.8 reports the treatment effects with the interaction on math and reading test scores.

The randomization was also stratified by gender, which helps identify heterogeneous effects for boys and girls, even when that was not pre-registered. I also present multiple robustness checks supporting the consistency of my results by gender for various outcomes.

6 Main Results

6.1 Social Outcomes

I start describing the results by reporting the impact of the two treatments on social outcomes. Panel A of Table 5 reports the reduced-form estimates of equations 6 and 7 for all students on social-outcomes indicators.

The results reveal that having more central peers improves social outcomes, but only for boys. Columns 1 and 2 in Panel A report the post-intervention effects on the number of connections. The impact of having more central peers on the number of connections for all students is close to zero (0.006, p-value 0.967). However, column 2 reveals how this average impact masks some heterogeneity by gender. While the impact is negative for girls (-0.334, s.e. 0.181), the effect is large and positive for boys, who end up having 0.50 (p-value 0.029) more connections after the intervention. The results for network centrality (columns 3 and 4) have a similar pattern: boys with more central peers have a better network position after the intervention (0.10σ, p-value 0.009), but for girls, the point estimate is -0.048 (s.e. 0.031).

I also find that having more central neighbors only increases boys' scores in social psychology tests (columns 5 and 6). Estimates in column 5 show an ATE of 0.070σ (s.e. 0.027) for psychological tests. This positive impact is mainly driven by boys, for whom having more central neighbors increases the social-skills index by 0.143σ (p-value 0.001). Although the results on peers' perception (columns 7 and 8) are weaker, the same conclusion applies. While the ATE of having more central neighbors on peers' perception (column 7) is 0.029σ (s.e. 0.020), the impact for boys has a larger magnitude of 0.054σ (p-value 0.091).

By contrast, I do not find that higher-achieving peers affect social outcomes for either boys or girls. Overall, the estimates for all students in Panel A are precise zeros. This is true for the network-centrality measure (column 3, effect of 0.011σ, s.e. 0.019), the social-skills index (column 5, effect of -0.018σ, s.e. 0.021), and peers' perceptions (column 7, effect of 0.015σ, s.e. 0.017). Both the point estimates and standard errors are small for every social outcome. I also find no differences by gender when testing for heterogeneous impacts in the even columns of Table 5.

Next, I explore whether these effects vary according to students' social centrality at baseline by estimating equations 6 and 7 by subgroups: less and more socially central students at baseline (Panels B and C, respectively). I then compare these results to equation 6 estimates for all students, presented in Panel A.

The positive effects of having more central neighbors on boys' social skills mainly originate from the impact on students who were less socially central at baseline (Panel B of Table 5). Having more central peers increases connections for less socially central students by 0.956 (p-value 0.001). Estimates for network centrality, psychological tests, and peers' perceptions are all consistent with this conclusion. All of the point estimates are larger than those reported in Panel A, and the p-values range between 0.000 and 0.012.[15] By contrast, I do not find robust evidence

[15] The difference in the effect of the more-socially-central peers treatment for less and more-central boys is statistically significant for three out of four social outcomes.

that more central neighbors affect the social outcomes for less socially central girls. Likewise, higher-achieving neighbors do not appear to change the social outcomes for less socially central students. While I observe some negative effects on peers' perceptions (column 7), I do not put much weight on this result, as it is inconsistent with the impact on other social outcomes.

The more-central-peers treatment does not affect the formation of social skills for students assessed as more socially central at baseline. Panel C supports this conclusion by showing the reverse side of the story. I cannot reject a zero treatment effect for most outcomes in this table for both boys and girls. Having higher-achieving peers, however, appears to increase the social perceptions of lower-achieving girls. As this effect is not consistent with the effect for other social outcomes, I refrain from drawing general conclusions from these estimates.

The positive impacts on social skills for the less socially central boys translate into longer-term outcomes such as lower dropout rates and higher enrollment rates at better colleges. Appendix Table A.4 shows the correlation between the general social-skills index (grouping all social skills measures), math, and reading scores with dropout rates, college enrollment, and college quality. Column 1 shows how the social skills index has the largest predictive power on the COAR Network dropout rates. A one-standard-deviation increase in the social-skills index is correlated with a decrease in dropout rates of 0.8 p.p. The three types of skills —social skills, math, and reading scores— are also positively correlated with college enrollment and quality.[16] The best predictor is generally math scores. Still, social skills are crucial as they are better than reading scores in predicting college enrollment and enrollment at certified colleges.

Furthermore, Table 6, Panel B, shows that having more central neighbors also influences the longer-term outcomes of less socially central boys. Column 2 shows a negative effect of 2.4 p.p. (p-value 0.004) on the dropout rate.[17] Columns 4 and 6 show an increase of 6.8 p.p. (p-value 0.028) and 5.4 p.p. (p-value 0.049), respectively, in the likelihood of enrolling at certified or top-twenty colleges. Results broadly support that social skills matter for later-life outcomes.[18]

The social skills improvement for less socially central boys remains after multiple robustness checks. Appendix Figure A.3 presents the effect of the more-central-peers treatment on all the individual outcomes related to social skills. I measure these social outcomes at different moments after the experiment's implementation (see timeline in Figure 1). Yet, results are consistently positive regardless of the time of measurement. These outcomes include: (i) the degree and centrality of friendship, study, and other social activities networks, (ii) openness, extraversion, and agreeableness (among the Big Five), as well as other psychological-test measures, and (iii) the number of peers who perceive the student as a leader or as a friendly, popular, or shy person. Panel A displays the point estimates and 90% confidence intervals for the less socially central boys. Point estimates are positive for thirty-eight out of thirty-nine outcomes, and statistically different from zero in twenty-nine cases.

Moreover, Appendix Table A.5 presents p-values following Young (2018), showing that the above results are robust to randomization inference. Likewise, I can also reject a zero effect after

[16] A one-standard-deviation increase in social skills correlates with a 2.4-percentage-point increase in college enrollment. This is one-quarter of the correlation between college enrollment and math scores. These results are more important for enrollment at certified colleges, where the correlation with social skills is about 45% of the correlation with math scores.

[17] As we collected different surveys over time, the effect on the dropout rate does not generate attrition on the main outcomes.

[18] While more central peers negatively impact college enrollment for girls with higher centrality at baseline, they do not affect college quality.

accounting for multiple hypotheses testing. Appendix Table A.6 presents p-values for multiple hypotheses across different student groups by gender and social centrality at baseline.

6.1.1 2SLS Estimates

To account for imperfect compliance between assigned peers and actual neighbors and provide comparable estimates to other peer-effects studies, I estimate equation 8. Table 7 presents the results of the 2SLS model with two endogenous variables described in equation 8 for social and academic outcomes. The table reports the estimates of parameters β_s and β_a—the impact of neighbors' average social centrality and academic achievement on students' outcomes, respectively. There are two endogenous variables: neighbors' social centrality and neighbors' academic achievement (both calculated at baseline). I instrument for these by using whether the student was assigned to the more-central-peers or higher-achieving-peers treatment. Table 7 reports peer effects on the composite social-skills index aggregating the four social outcomes.

The results of this 2SLS model mirror the treatment effects described above. Neighbors' social centrality positively impacts social skills, but only for boys. Table 7 (Panel A) show the results for all students, boys and girls. A one-standard-deviation increase in neighbors' social centrality has an impact of 0.039σ (s.e. 0.046) on the average student's social-skills index score (column 1). This slightly positive impact is driven by boys boys in column 2, with an estimate of 0.246σ (s.e. 0.084). By contrast, column 3 shows that the social peer-effects estimate for girls is small (-0.076) and relatively precise (s.e. 0.053). Social outcomes are also not affected by the academic achievement of students' neighbors.

The positive social peer effects on boys are higher for the less central students. Panel B reports these results for the less central students. Estimates are all-around larger than in the combined sample in Panel A. Less central boys (Panel B, column 2) benefit the most from more central neighbors. A one-standard-deviation in neighbors' centrality increases the social skills index for the less central boys by 0.486σ (p-value 0.000). By contrast, results in Panel C indicate that peers' centrality and achievement do not affect social outcomes for the more central students.

6.2 Academic Outcomes

Next, I estimate treatment effects on academic outcomes. Table 8 reports estimates of equations 6 and 7. Columns 1 and 2 report effects on math and reading grades. These outcomes are only available for the 2016 and 2017 cohorts. Analogously, columns 3 and 4 show the impact of each treatment on math and reading test scores.

Consistent with the peer effects estimates reported by previous quasi-experimental studies (Angrist and Lang, 2004; Duflo et al., 2011; Abdulkadiroğlu et al., 2014) that generate large variation in peers' skills, I find that the impact of higher-achieving peers on students' academic achievement is a precisely estimated zero. The odd columns in Table 8, Panel A, present the ATEs for all students in my sample. These are precise estimates in the context of my study. The 95% confidence interval for math test scores (column 5) ranges between -0.058 and 0.004σ. For reading (column 7), it ranges between -0.072 and 0.006σ. These confidence intervals allow me to rule out positive peer effects on the average student. Likewise, I do not find evidence that having more central peers affects the academic achievement of the average student. And I cannot reject homogeneous treatment effects by gender, except for reading test scores. In column 8, I find a negative effect on girls of 0.072σ (p-value 0.007).

I also examine treatment effects' heterogeneity by academic achievement. I estimate equations 6 and 7 for two academic-achievement subgroups: lower- and higher-achieving students at baseline. Table 8, Panels B and C, report the reduced-form estimates for lower- and higher-achieving students at baseline.

Higher-achieving peers have heterogeneous treatment effects on academic achievement. Columns 1 and 3 in Panel B of Table 8 show that the higher-achieving-peers treatment negatively affects both math and reading grades. Higher-achieving neighbors reduce students' math grades by 0.061σ (p-value 0.092) and reading grades by 0.075σ (p-value 0.043). The treatment effects on test scores are also negative. Table 8, Panel B, columns 5 and 7, show that the effects of higher-achieving peers on lower-achieving students are -0.043σ (p-value 0.090) on math scores and -0.041σ (p-value 0.193) on reading scores. For the more-central-peers treatment, there is no consistent evidence of an effect on academic performance.

The negative academic peer effects on lower-achieving students are starker for girls. The even columns in Table 8, Panel B, report the estimates of equation 7 for lower-achieving students. These results indicate that for lower-achieving girls, the academic treatment academic effect is particularly negative, as reflected in math grades (column 2, -0.114σ, p-value 0.018), math test scores (column 6, -0.071σ, p-value 0.013), and reading-test scores (column 8, -0.078σ, p-value 0.067). The point estimate for reading grades (column 2) is also negative (-0.066σ, p-value 0.172), but it is more negative for boys. This evidence suggests that higher-achieving neighbors can reduce the academic performance of lower-achieving girls. I cannot reject the null hypothesis of a zero impact for lower-achieving boys.

These results are also robust to randomization inference (Table A.5, Panel B). However, the effects are weaker than those on social skills once we account for multiple hypotheses testing (Table A.6, Panel B). Under the traditional multiple-hypotheses tests, the treatment effects on math for lower-achieving girls are significant at the 10% level. In contrast, estimates are not statistically significant for reading scores.

I do not find that having more central or higher-achieving neighbors affects the academic performance of higher-achieving students. Table 8, Panel C, reports these estimates. Estimates are generally small and fairly precise. This is true for both grades and test scores and for both boys and girls (even columns in Table 8). Neighbors' characteristics do not appear to affect the academic achievement of the academically strongest students. If anything, higher-achieving peers reduce reading performance for girls by 0.073σ (column 8).

6.2.1 2SLS Estimates

Table 9 reports the 2SLS estimates of equation 8 for both math (columns 1-3) and reading test scores (columns 4-6). These estimates account for imperfect compliance and are comparable to the findings of other peer effects studies. Panel A presents the results for all students, and Panel B for subgroups by academic achievement at baseline. Columns 1 and 4 show a precise zero estimate for average academic peer effects. The impact of a one-standard-deviation increase in neighbors' academic achievement at baseline is -0.046 (s.e. 0.029) on math scores and -0.059 (s.e. 0.035, p-value 0.094) on reading scores. My estimates rule out even small positive peer effects: the upper limit of the 95% confidence interval is just 0.011 for math scores and 0.010

for reading scores. I also find fairly precise estimates for neighbors' social centrality on academic outcomes. Estimates in columns 2 and 3 show that peer effects on math scores are similar for boys and girls. However, academic peer effects are negative for girls on reading (columns 5 and 6); a one-standard-deviation increase in neighbors' academic achievement reduces girls' reading performance by 0.12σ (s.e. 0.044).

Results in Panel B of Table 9 show that academic peer effects are negative for lower-achieving students but statistically indistinguishable from zero at the 95% confidence level. However, columns 3 and 6 in Panel B indicate that these effects are more negative and statistically significant for girls. For them, a one-standard-deviation increase in peers' achievement at baseline reduces math performance by 0.116σ (p-value 0.015) and reading performance by 0.136σ (p-value 0.060). In contrast, for boys (columns 2 and 5), the estimates for academic peer effects are very close and indistinguishable from zero. For higher-achieving students (Panel C), peer effects are indistinguishable from zero.

In summary, higher-achieving peers have, on average, zero effect on students' academic outcomes. And they appear to be detrimental to the performance of lower-achieving students, especially lower-achieving girls.

7 Mechanisms

I now study what mechanisms may explain the results described in Section 6. While my goal is not to establish the causal impact of any particular mechanism, as this was not the experiment's purpose, I present consistent evidence with the main findings. I show that boys' and girls' beliefs respond differently to peers and that friendships are not enough to cause peer effects.

7.1 Self-Confidence

This section examines whether beliefs about one's abilities (self-confidence) can explain my findings. In Appendix E, I present a simple framework based on previous theoretical results to illustrate how beliefs affect student outcomes and the impact of peer characteristics on beliefs. Three reasons below could help explain the role of self-confidence in peer effects.

First, the literature can help identify two channels for beliefs to affect performance. On the one hand, when effort and abilities are complements, more self-confident individuals exert more effort (Benabou and Tirole, 2002). Second, as argued by Compte and Postlewaite (2004), self-confidence could directly affect performance.

The second reason is that, by interacting with peers, students receive signals about their skills that could change their beliefs. While it is beyond this paper's scope to study these signals, a natural example is the "big fish, little pond" effect: students can lose confidence in their own abilities through social comparisons. Still, students might as well receive positive signals from peer relationships. For example, a student might feel more popular if she befriends the most-central students in her class. Previous evidence shows that students may be discouraged due to these social comparisons (Antecol et al., 2016; Rogers and Feller, 2016).

Third, the interpretation of a signal might depend on gender. Men and women differ in how they form beliefs about themselves and others (Bordalo et al., 2019). Recent evidence in psychology shows that female students tend to make more upward social comparisons and fewer downward comparisons than male students in assessing their math abilities (Pulford et al., 2018).

Similarly, an extensive literature in economics shows that men and women differ in their levels of confidence (Sarsons and Guo, 2016), how they respond to feedback (Mobius et al., 2014), and their preferences for competition (Gneezy et al., 2003; Buser and Yuan, 2019).

In the following two subsections sections, I explore whether self-confidence might be a mechanism driving this paper's results. First, I study to what extent male and female students differ in their beliefs. Second, I estimate treatment effects on self-reported measures of ability.

7.1.1 Gender Differences

To determine whether gender differences affect students' beliefs, I study whether boys and girls report different beliefs in their skills. In the endline survey, we asked students to rank their own academic skills and popularity from 0 (lowest) to 100 (highest). Another measure is whether a student identifies herself as being in her cohort's top five along the dimensions of academic skill, leadership, friendliness, popularity, and shyness (reversed).

Figure 5 presents the cumulative distribution of the self-reported academic and popularity rankings by gender (Panels A and B, respectively). The left column displays quantile regressions in the gender gap of these self-reports after controlling for observable characteristics: test scores, the number of friends, centrality, and peers' perceptions of academic skills and popularity.

In general, boys report higher self-confidence in both academic skills and popularity. The left column in both panels shows that the distribution of boys' self-reported academic and popularity rankings has first-order stochastic dominance over girls' distribution for the same variables. Furthermore, estimates for the quantile regressions in the right column reveal that these differences remain even after controlling for observable characteristics. The estimates suggest that men are more confident than women. The male-female gap is positive across the entire distribution, and in most cases, it is statistically significant at the 95% level. For example, at the median of the distributions, the difference in the ranking is approximately five positions (0.25σ for the academic ranking and 0.20σ for the popularity ranking).

7.1.2 Peers and Beliefs

Social Outcomes: I examine whether having more central peers affects students' perception of their own social skills after the intervention. Table 10 reports these effects. Panel A presents the results for the less socially central students at baseline and Panel B for the more socially central. Columns 1 to 3 show the effect on self-reported popularity rankings (all between 0 and 100) in the dorm[19], the classroom, and the cohort. Columns 4 to 8 report estimates on whether students entered their own names when asked to list up to five top peers in leadership (column 4), popularity (column 5), friendliness (column 6), and shyness (reversed in column 7). Finally, column 8 presents an index combining these measures.

While the less socially central girls negatively updated their beliefs on their own social skills, the less socially central boys positively updated theirs. Results in column 8 show the treatment effect for the less socially central boys on the index is about 0.139σ (p-value 0.027). Less socially central girls, by contrast, have a negative treatment effect of about 0.120σ (p-value 0.013).

The table also exhibits the effects on the index's constituting measures. Results in Table 10, Panel A, column 1, are very telling about differences in belief formation by gender. There is a

[19]For large dormitories, the dorm is defined as students in neighboring bunk beds.

negative mechanical relationship between being assigned to more central peers and popularity ranking within the dorm by construction. However, this negative relationship only holds for girls, who report a 2.91 p.p. lower ranking (p-value 0.049) when assigned to more-central neighbors. By contrast, we cannot reject a zero impact for boys. The interaction term of the treatment with the boy dummy is positive and marginally significant (p-value 0.096), which suggests that boys' perception of their popularity increases when they interact with more central peers.

The results also show that beyond the negative mechanical effect within the dorm, girls also report a lower ranking in their classroom and cohort when assigned to more central peers (columns 2 and 3). The intervention caused them to weaken their beliefs in their own popularity. This result is in line with previous evidence that women tend to make more upward social comparisons than men do. The impact of the treatment on self-reported rankings also varies by gender. The treatment effects on the classroom and dorm rankings are 5.2 and 5.3 positions greater for boys than for girls. Both differences are statistically significant at the 95% level. Furthermore, the estimate in column 3 shows that the treatment effect is positive for boys, with a ranking increase of 2.60 positions (p-value 0.103) in the ranking. This result supports that boys believe they are more popular after interacting with more central neighbors.

The estimates on whether students list their own names in the survey (columns 4 to 7) are consistent with these results. The positive impact on the beliefs of the less socially central boys is driven by their self-perceived levels of leadership, popularity, and, especially, shyness. In general, less socially central boys are 4.0 p.p (p-value 0.009) less likely to report themselves as being among the shyest in the school after the intervention.

Overall, this evidence suggests that more socially central neighbors affect boys' and girls' beliefs in their abilities differently.

Academic Achievement: Changes in beliefs on academic skills are also a valid mechanism to explain the academic peer effects in this paper. Evidence from Table 8 shows that having higher-achieving peers decreases the academic scores of lower-achieving students, especially lower-achieving girls. Here, I explore whether these changes align with changes in self-confidence.

Table 11 displays estimates for equation 7 on self-confidence in academic skills with three measures of self-confidence and an aggregate index. The first measure is self-reported beliefs of academic rankings within the dorm, classroom, and cohort. The second measure is whether a student names herself as one of the five most skilled students in the cohort. The third measure comprises two factors from the Achievement Goals Questionnaire: (i) the performance-approach goal, assessing whether a student wants to do better than her peers, and (ii) the performance-avoidance goal, measuring whether a student avoids doing worse than her peers.

The results suggest that lower-achieving girls lose self-confidence in their academic skills when paired with higher-achieving peers. The main estimates in column 7 convey a negative effect of 0.093 (p-value 0.033). The table also displays treatment effects on individual measures. Results in the first column show that while lower-achieving girls report a ranking within the dorm 1.49 positions lower (p-value 0.116), the effect for boys is around -0.010 positions and statistically indistinguishable from zero. Results for relative-performance goals are yet more striking, with a negative impact on the performance-approach goal of 0.19σ (p-value<0.001) and a negative impact on the performance-avoidance goal of 0.129σ (p-value 0.019). Conversely, the impact on boys is slightly positive but indistinguishable from zero.[20]

Panel B of Table 11 presents the results for higher-achieving students. Although the estimates are, in general, indistinguishable from zero, there is some evidence of gender differences in the formation of beliefs. Columns 1 to 3 show that while higher-achieving peers reduce the self-reported academic rankings within a dorm, classroom, and cohort for girls, this is not the case for boys. This impact is in line with the negative effects on reading scores.

Overall, gender differences in self-confidence are consistent with the results in Section 6. Gender differences in psychological factors appear to be an important mediator of peer effects.

7.2 Social Interactions

I also study whether social connections with neighbors can explain the results in this paper. Intuitively, the effects of friends should be different from those of other peers. For example, Carrell et al. (2013) find that peers who were expected to increase the performance of low-skilled students ended up harming them. When low-skilled students are in groups with high-skilled peers, they segregate by academic achievement, and the performance of the lower-skilled students worsens. Recognizing the evidence on the importance of social interactions for the direction and magnitude of peer effects, I test whether this mechanism is driving my results.

I find that the intervention globally influences friendships and social interactions in my setting. I estimate equation 3 on the likelihood that individuals i and j form a social connection. Figure 4, Panel B, displays the likelihood that two students will form a social connection as a function of their distance on the list. A distance of one on the list (being in the row above or below) increases the likelihood of becoming friends, engaging in joint social activities, or studying together by approximately 23 p.p (p-value 0.000). I also find a decreasing pattern with distance, and distance impacts social interactions regardless of dorm size.

To assess if these social interactions drive the peer effects in this study, I estimate the impact of each treatment (equation 7) on the number of connections students made with their neighbors while looking at heterogeneity by gender and baseline characteristics. In a scenario where social interactions are a major driver of peer effects, we would expect the following heterogeneous treatment effects. Less socially central boys and more socially central neighbors form more connections than other groups, and lower-achieving girls study less with their neighbors when these are higher-achieving. However, I find no evidence of it. Figure 6, Panel A, shows that less socially central boys form connections with their neighbors like other groups do. For all groups of less socially central students, the distance on the list reduces the average number of connections. The number of connections is also relatively similar across groups at each distance value.

I formally test whether the treatment or its interaction with gender predicts connections in Appendix Table A.9, Panel A. The estimates show that I cannot reject the hypothesis that less socially central boys form more social connections with more socially central neighbors than other groups. In particular, column 1 shows that neither the more-central-peers-treatment status nor gender explain social connections with neighbors. Other than a marginally significant effect in column 6, I cannot reject that these parameters are equal to zero. These results suggest that other groups for which there is no evidence of an improvement in social skills also formed similar

[20] The results on rankings and self-confidence are more negative for first-year students, who have less information about their academic abilities relative to their peers.

connections with their neighbors.

Changes in social interactions are also inconsistent with academic peer effects findings. Figure 6, Panel B, reports the average number of connections by distance with neighbors for lower-achieving students.[21] This figure shows a similar pattern to Panel A and Figure 4, where increases in the distance on the list are associated with fewer social interactions for the three groups and a similar average number of connections across groups for each distance value. The estimates in Appendix Table A.9, Panel B, confirm this. Neither the higher-achieving-peers treatment nor its interaction with gender predicts social connections (column 1). Strikingly enough, this result also holds for study partnerships (column 3). Indeed, the results also show that, counter-intuitively, lower-achieving girls receive more support from their neighbors in dealing with academic and personal problems (columns 5 and 6, respectively) when the neighbors are higher achieving. By contrast, the estimates in Table 9 reveal negative academic peer effects for lower-achieving girls.

Taken together, this evidence rules out social connections as the ultimate driver of peer effects. All students are equally likely to befriend their neighbors, and yet, estimates of peer effects vary widely across outcomes, student characteristics, and peer type.

8 Conclusion

This paper presents the results of a field experiment designed to estimate causal peer effects on social and academic outcomes. The study was conducted in twenty-three out of twenty-five exam schools in Peru, with a sample of approximately six thousand students. The experimental design alleviates recent concerns with the traditional approach to estimating peer effects—random allocation to groups. The experiment guarantees strong variation in peer characteristics by randomly manipulating the peer type, and using an identification strategy that relies on the variation in peer characteristics across treatments rather than groups.

Students were classified by baseline social centrality and academic achievement using centrality measures of social networks and admission test scores. I found that more socially central peers positively impact boys' development of social skills. The effects are mainly driven by the impact on boys assessed as less socially central at baseline. This group of boys ended up with more connections and a higher centrality in their networks. These results are consistent with the impact on psychological tests and peers' perceptions of students' social skills. These effects translate into longer-term outcomes. More central neighbors prevent less socially central boys from dropping out of the COAR Network, making them more likely to enroll in good universities.

By contrast, I reject positive academic peer effects on academic achievement. The evidence suggests that higher-achieving peers reduce the performance of lower-achieving students at baseline. This result is stronger for lower-achieving girls. These findings are inconsistent with peer effects estimates from other studies, especially those that use random allocation to groups (Sacerdote, 2011; Epple and Romano, 2011). My conclusions are similar to the evidence on peer effects from quasi-experimental studies (Angrist and Lang, 2004; Abdulkadiroğlu et al., 2014; Duflo et al., 2011) that also ensure substantial variation in peers' skills.

A potential limitation of this paper is that it does not allow for nonlinearities in peer effects. However, while the main estimation is based on a linear-in-means peer-effects model, I do allow

[21] These numbers are higher than for less socially central students because first-year students form, on average, more links.

for heterogeneity by gender and baseline characteristics. Furthermore, the experimental design can be adapted to include nonlinearities, but as in other experimental studies, there is a trade-off between more treatments and greater statistical power.

I rule out social interactions as a driving mechanism of this paper's peer effects. For example, although lower-achieving girls befriend and study with their higher-achieving neighbors, they have lower test scores. This result counters previous literature where students only benefit from higher-achieving peers when interacting with them (Carrell et al., 2013). Further studies are needed to elucidate the differences between peer effects from friends and others.

Overall, the results show that policies that affect peer characteristics need to account for gender differences in psychological factors. Less socially central boys and less socially central girls experience different impacts on their beliefs in their own social skills after interacting with more central neighbors. These results are consistent with a broad literature studying how men and women form beliefs about themselves and others.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When Should You Adjust Standard Errors for Clustering? Working Paper 24003, National Bureau of Economic Research.

Abdulkadiroğlu, A., Angrist, J., and Pathak, P. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1):137–196.

Acemoglu, D. and Angrist, J. (2000). How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws. *NBER Macroeconomics Annual*, 15:9–59.

Akee, R., Copeland, W., Costello, E. J., and Simeonova, E. (2018). How Does Household Income Affect Child Personality Traits and Behaviors? *American Economic Review*, 108(3):775–827.

Alan, S., Baysan, C., Gumren, M., and Kubilay, E. (2021). Building Social Cohesion in Ethnically Mixed Schools: An Intervention on Perspective Taking*. *The Quarterly Journal of Economics*, 136(4):2147–2194.

Alan, S., Ertac, S., Kubilay, E., and Loranth, G. (2019). Understanding Gender Differences in Leadership. *The Economic Journal*, 130(626):263–289.

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Chapter 1 - Personality Psychology and Economics1. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of The Economics of Education*, volume 4 of *Handbook of the Economics of Education*, pages 1 – 181. Elsevier.

Angrist, J. D. (2014). The Perils of Peer Effects . *Labour Economics*, 30:98 – 108. Special Section articles on "What determined the dynamics of labour economics research in the past 25 years? edited by Joop Hartog and and European Association of Labour Economists 25th Annual Conference, Turin, Italy, 19-21 September 2013 Edited by Michele Pellizzari.

Angrist, J. D. and Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5):1613–1634.

Antecol, H., Eren, O., and Ozbeklik, S. (2016). Peer Effects in Disadvantaged Primary Schools: Evidence from a Randomized Experiment. *Journal of Human Resources*, 51(1):95–132.

Arli, S. K. and Bakan, A. B. (2018). An Investigation of the Relationship between Intercultural Sensitivity and Compassion in Nurses. *International Journal of Intercultural Relations*, 63:38 – 42.

Athey, S. and Imbens, G. (2017). Chapter 3 - The Econometrics of Randomized Experimentsa. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Field Experiments*, volume 1 of *Handbook of Economic Field Experiments*, pages 73 – 140. North-Holland.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The Diffusion of Microfinance. *Science*, 341(6144).

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2014). Gossip: Identifying Central Individuals in a Social Network. Working Paper 20422, National Bureau of Economic Research.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014a). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives*, 28(2):29–50.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014b). Inference on Treatment Effects after Selection among High-Dimensional Controls. *The Review of Economic Studies*, 81(2):608–650.

Benabou, R. and Tirole, J. (2002). Self-Confidence and Personal Motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Booij, A. S., Leuven, E., and Oosterbeek, H. (2017). Ability Peer Effects in University: Evidence from a Randomized Experiment. *The Review of Economic Studies*, 84(2):547–578.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2019). Beliefs about Gender. *American Economic Review*, 109(3):739–73.

Breza, E. and Chandrasekhar, A. G. (2019). Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment. *Econometrica*, 87(1):175–216.

Buser, T. and Yuan, H. (2019). Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad. *American Economic Journal: Applied Economics*, 11(3):225–52.

Caeyers, B. and Fafchamps, M. (2016). Exclusion bias in the estimation of peer effects. Working Paper 22565, National Bureau of Economic Research.

Carre, A., Stefaniak, N., D'Ambrosio, F., Bensalah, L., and Besche, C. (2013). The Basic Empathy Scale in Adults (BES-A): Factor structure of a revised form. 25:679–691.

Carrell, S., Fullerton, R. L., and West, J. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, 27(3):439–464.

Carrell, S. E., Sacerdote, B. I., and West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81(3):855–882.

Castro-Solano, A. (2007). *Teoría y Evaluación dell Liderazgo*. Paidos.

Chen, G.-M. and Starosta, W. J. (2000). The Development and Validation of the Intercultural Sensitivity Scale. *Human Communication*.

Cobb-Clark, D. A. and Tan, M. (2011). Noncognitive Skills, Occupational Attainment, and Relative Wages. *Labour Economics*, 18(1):1 – 13.

Coffman, Katherine B., M. C. and Kulkarni., L. (2020). Stereotypes and Belief Updating. Working Paper 19-068, Harvard Business School.

Colman, A. M. (2009). *A Dictionary of Psychology*.

Compte, O. and Postlewaite, A. (2004). Confidence-Enhanced Performance. *American Economic Review*, 94(5):1536–1557.

Declerck, C. H. and Bogaert, S. (2008). Social Value Orientation: Related to Empathy and the Ability to Read the Mind in the Eyes. . *The Journal of social psychology*, 148.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics*, 132(4):1593–1640.

Dobbie, W. and Fryer, Roland G., J. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, 6(3):58–75.

Donato, K., Miller, G., Mohanan, M., Truskinovsky, Y., and Vera-Hernández, M. (2017). Personality Traits and Performance Contracts: Evidence from a Field Experiment among Maternity Care Providers in India. *American Economic Review*, 107(5):506–10.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5):1739–74.

Epple, D. and Romano, R. E. (2011). Chapter 20 - Peer Effects in Education: A Survey of the Theory and Evidence . volume 1 of *Handbook of Social Economics*, pages 1053 – 1163. North-Holland.

Falk, A., Kosse, F., Deckers, T., Pinger, P., and Schildberg-Hörisch, H. (2018). The Formation of Prosociality: Causal Evidence on the Role of Social Environment. Technical report, University of Bonn.

Feld, J. and Zölitz, U. (2017). Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. *Journal of Labor Economics*, 35(2):387–428.

Fletcher, J. M. (2013). The Effects of Personality Traits on Adult Labor Market Outcomes: Evidence from Siblings. *Journal of Economic Behavior & Organization*, 89:122 – 135.

Gambin, M. and Sharp, C. (2018). The Relations between Empathy, Guilt, Shame and Depression in Inpatient Adolescents. *Journal of Affective Disorders*, 241:381 – 387.

Garlick, R. (2018). Academic Peer Effects with Different Group Assignment Policies: Residential Tracking versus Random Assignment. *American Economic Journal: Applied Economics*, 10(3):345–69.

Glaeser, E. L., Laibson, D., and Sacerdote, B. (2002). An Economic Approach to Social Capital. *Economic Journal*, pages 437–458.

Glaeser, E. L., Sacerdote, B. I., and Scheinkman, J. A. (2003). The Social Multiplier. *Journal of the European Economic Association*, 1(2-3):345–353.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Gordon, S. R. and Mwavita, M. (2018). Evaluating the International Dimension in an undergraduate Curriculum by Assessing Students' Intercultural Sensitivity. *Studies in Educational Evaluation*, 59:76 – 83.

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

Heckman, J. J. and Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics*, 6(1):689–733.

J. Elliot, A. and Murayama, K. (2008). On the Measurement of Achievement Goals: Critique, Illustration, and Application. 100.

Javed, B., Khan, A. K., Arjoon, S., Mashkoor, M., and Haque, A. u. (2020). Openness to Experience, Ethical Leadership, and Innovative Work Behavior. *The Journal of Creative Behavior*, 54(1):211–223.

John, O. P. and Srivastava, S. (1999). *Handbook of Personality: Theory and Research*, chapter The Big Five trait taxonomy: History, measurement, and theoretical perspectives, pages 102–138. Guilford Press, New York, second edition.

Jolliffe, D. and Farrington, D. P. (2006). Development and Validation of the Basic Empathy Scale. *Journal of Adolescence*, 29(4):589 – 611.

Jones, D., Molitor, D., and Reif, J. (2019). What do Workplace Wellness Programs do? Evidence from the Illinois Workplace Wellness Study*. *The Quarterly Journal of Economics*, 134(4):1747–1791.

Kranton, R. E. and Sanders, S. G. (2017). Groupy versus Non-Groupy Social Preferences: Personality, Region, and Political Party. *American Economic Review*, 107(5):65–69.

Law, K., Wong, C.-S., and J Song, L. (2004). The Construct and Criterion Related Validity of Emotional Intelligence and Its Potential Utility for Management Studies. 89:483–96.

Lleras-Muney, A., Miller, M., Sheng, S., and Sovero, V. T. (2020). Party On: The Labor Market Returns to Social Networks and Socializing. Working Paper 27337, National Bureau of Economic Research.

Marsh, H. W. and Parker, J. W. (1984). Determinants of Student Self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, 47(1):213–231.

McCrae, R. R. and John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2):175–215.

Mehu, M., Grammer, K., and Dunbar, R. I. (2007). Smiles when Sharing. *Evolution and Human Behavior*, 28(6):415 – 422.

MINEDU (2016). Evaluación Censal de Estudiantes 2015-2016. Technical report, Ministry of Education, Peru.

MINEDU (2017a). Admissions and Enrollment Data COAR Network 2015-2017. Technical report, Ministry of Education, Peru.

MINEDU (2017b). COAR Standardized Tests and Grades. Data, Ministry of Education, Peru.

MINEDU (2017c). Encuesta de Bienestar y Desarrollo COAR 2016-2017. Data, Ministry of Education, Peru.

MINEDU (2017d). Sociodemographic Information of Students in the COAR Network. Technical report, Ministry of Education, Peru.

MINEDU (2020). Higher Education Applications, Admissions and Enrollment 2018-2020. Data, Ministry of Education, Peru.

MINEDU and Zárate, R. A. (2017). Social Interactions and Non-Cognitive Skills Survey 2016-2017. Data, Ministry of Education, Peru.

Mobius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2014). "Managing Self-Confidence: Theory and Experimental Evidence". Technical report.

Nieb, C. and Zacher, H. (2015). Openness to Experience as a Predictor and Outcome of Upward Job Changes into Managerial and Professional Positions. *PLoS ONE*, 10:e0131115.

Otto, P. E. and Bolle, F. (2011). Multiple Facets of Altruism and their Influence on Blood Donation. *The Journal of Socio-Economics*, 40(5):558 – 563.

Pulford, B. D., Woodward, B., and Taylor, E. (2018). Do Social Comparisons in Academic Settings Relate to Gender and Academic Self-confidence? *Social Psychology of Education*, 21(3):677–690.

Rao, G. (2019). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, 109(3):774–809.

Rogers, T. and Feller, A. (2016). Discouraged by Peer Excellence: Exposure to Exemplary Peer Performance Causes Quitting. *Psychological Science*, 27(3):365–374. PMID: 26825105.

Rosenbaum, P. and Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies For Causal Effects. *Biometrika*, 70:41–55.

Rushton, J. P., Chrisjohn, R. D., and Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, 2(4):293 – 302.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, 116(2):681–704.

Sacerdote, B. (2011). Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? volume 3 of *Handbook of the Economics of Education*, pages 249 – 277. Elsevier.

Sacerdote, B. (2014). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, 6(1):253–272.

Sarsons, H. and Guo, X. (2016). Confidence Men? Gender and Confidence: Evidence among Top Economists.

Villadangos, M., Errasti, J., Amigo, I., Jollife, D., and García-Cueto, E. (2016). Characteristics of Empathy in Young People Measured by the Spanish Validation of the Basic Empathy Scale. *Psicothema*.

Wai, M. and Tiliopoulos, N. (2012). The Affective and Cognitive Empathic nature of the Dark Triad of Personality. *Personality and Individual Differences*, 52(7):794 – 799.

Weidmann, B. and Deming, D. J. (2020). Team Players: How Social Skills Improve Group Performance. Working Paper 27071, National Bureau of Economic Research.

Weinberger, C. J. (2014). The Increasing Complementarity between Cognitive and Social Skills. *The Review of Economics and Statistics*, 96(5):849–861.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, 330(6004):686–688.

Young, A. (2018). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*. *The Quarterly Journal of Economics*, 134(2):557–598.

Yukl, G. (2013). *Leadership in Organizations*. Pearson, eighth edition edition.

Zárate, R. A. (2017). Social Connections and Peer Effects: An Experiment at Selective Enrollment High Schools in Peru. *AEA RCT Registry*.

Özbağ, G. K. (2016). The Role of Personality in Leadership: Five Factor Personality Traits and Ethical Leadership. *Procedia - Social and Behavioral Sciences*, 235:235 – 242. 12th International Strategic Management Conference, ISMC 2016, 28-30 October 2016, Antalya, Turkey.

Table 1: Summary Statistics

	All Students	By Social Centrality		By Academic Achievement	
		Less Central	More Central	Lower-Achieving	Higher-Achieving
Variable	(1)	(2)	(3)	(4)	(5)
Demographics					
Female (%)	0.57	0.59	0.60	0.57	0.57
Poor (%)	0.41	0.47	0.39	0.46	0.36
Rural (%)	0.26	0.30	0.22	0.31	0.22
Subsidized health insurance	0.50	0.55	0.46	0.53	0.47
Baseline characteristics					
National standardized score*	1.81	1.37	1.68	1.47	2.15
	(0.95)	(0.98)	(0.95)	(0.84)	(0.92)
Connections	14.69	12.00	15.27	14.49	14.89
	(6.49)	(5.28)	(5.76)	(6.46)	(6.51)
Social skills index	-0.00	-0.10	0.10	-0.03	0.03
	(1.00)	(1.00)	(0.98)	(0.99)	(1.00)
Peers' perception	0.00	-0.28	0.28	-0.11	0.11
	(1.00)	(0.65)	(1.19)	(0.87)	(1.10)
Treatments					
Social centrality	-0.00	-0.75	0.75	-0.05	0.05
	(0.99)	(0.40)	(0.83)	(0.97)	(1.01)
Academic achievement	0.00	-0.08	0.08	-0.78	0.78
	(0.99)	(0.96)	(1.02)	(0.47)	(0.71)
N	6,147	1,832	1,822	3,069	3,078

Notes: This table reports summary statistics for all students and by student type. Standard deviations are in parentheses. Column 1 shows statistics for all students, columns 2-3 by social centrality, and columns 4-5 by academic achievement. Columns 2-3 exclude the 2017 cohort because there is no available measure of centrality. *Scores in the national standardized test before the application to the COAR Network are not available for the 2015 cohort as this was the first year of this test. The table includes a set of students' demographic characteristics from government administrative data.

Table 2: Balance on Social Skills at Baseline

Dependent variable:	Social Skills Index		
	All Students	Boys	Girls
	(1)	(2)	(3)
Panel A: All students			
More central	0.000	-0.040	0.028
	(0.030)	(0.046)	(0.040)
Higher-achieving	-0.019	-0.068	0.015
	(0.030)	(0.046)	(0.040)
Control mean	-0.18	-0.26	-0.12
N	3,654	1,490	2,164
Panel B: Less central students at baseline			
More central	0.031	-0.045	0.084*
	(0.033)	(0.052)	(0.042)
Higher-achieving	0.017	-0.022	0.044
	(0.033)	(0.052)	(0.043)
Control mean	-0.76	-0.81	-0.73
N	1,832	753	1,079
Panel C: More central students at baseline			
More central	-0.031	-0.036	-0.028
	(0.051)	(0.076)	(0.068)
Higher-achieving	-0.055	-0.114	-0.016
	(0.050)	(0.076)	(0.067)
Control mean	0.71	0.59	0.79
N	1,822	737	1,085

Notes: This table reports balance checks of being assigned to more central and higher-achieving peers on social skills for all students and subgroups by social centrality at baseline. All regressions include strata fixed effects. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 3: Balance on Academic Performance at Baseline

Dependent variable:	Math score			Reading score		
	All Students	Boys	Girls	All Students	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All students						
More central	0.027	0.000	0.046	-0.001	0.011	-0.009
	(0.023)	(0.037)	(0.029)	(0.021)	(0.032)	(0.028)
Higher-achieving	0.006	0.013	0.001	-0.013	-0.014	-0.013
	(0.019)	(0.030)	(0.024)	(0.018)	(0.027)	(0.025)
Control mean	-0.10	0.01	-0.20	-0.06	-0.15	0.01
N	6,031	2,614	3,417	6,029	2,613	3,416
Panel B: Lower-achieving students at baseline						
More central	0.048	0.004	0.079	0.038	0.042	0.036
	(0.032)	(0.052)	(0.041)	(0.031)	(0.047)	(0.042)
Higher-achieving	-0.016	0.002	-0.028	-0.017	0.025	-0.046
	(0.032)	(0.052)	(0.041)	(0.031)	(0.046)	(0.042)
Control mean	-0.24	-0.04	-0.39	-0.20	-0.25	-0.16
N	1,830	753	1,077	1,829	752	1,077
Panel C: Higher-achieving students at baseline						
More central	0.007	-0.004	0.014	-0.040	-0.020	-0.053
	(0.031)	(0.051)	(0.039)	(0.028)	(0.042)	(0.037)
Higher-achieving	-0.004	0.002	-0.007	-0.043	-0.076	-0.022
	(0.031)	(0.052)	(0.039)	(0.029)	(0.045)	(0.038)
Control mean	0.10	0.25	0.00	0.17	0.14	0.19
N	1,821	736	1,085	1,820	736	1,084

Notes: This table reports balance checks of being assigned to more central and higher-achieving peers on academic performance for all students and subgroups by academic achievement at baseline. All regressions include strata fixed effects. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 4: First Stage on Assigned Peers and Actual Neighbors in Dormitories

	Assigned peers				Neighbors			
	Number		Average characteristics		Number		Average characteristics	
	More central	Higher-achieving	Social centrality	Academic achievement	More central	Higher-achieving	Social centrality	Academic achievement
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
More central	3.177***	-0.009	0.886***	0.088***	1.575***	0.060	0.544***	0.087***
	(0.107)	(0.105)	(0.016)	(0.017)	(0.049)	(0.048)	(0.019)	(0.020)
Higher-achieving	0.016	2.980***	0.036***	0.942***	-0.031	1.526***	0.022*	0.570***
	(0.067)	(0.081)	(0.010)	(0.013)	(0.032)	(0.038)	(0.013)	(0.015)
Control mean	0.38	0.92	-0.22	-0.53	0.59	1.39	-0.13	-0.36
N	6,079	6,079	6,079	6,079	6,079	6,079	6,079	6,079

Notes: This table reports the effect of being assigned to more central and higher-achieving peers on the number of more central and higher-achieving assigned peers and neighbors, and on the average centrality and academic achievement for each of these groups. Assigned peers are students in the in the student-peer-type combinations to which the student was assigned. Neighbors are students in the same dormitory for small dorms and students in the same or adjacent bunk bed for large dorms. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 5: Reduced-Form Effects on Social Skills

Dependent variable:	Connections		Centrality		Psychological tests		Peers' perception	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Panel A: All students					
More central	0.006	-0.334*	0.012	-0.048	0.070***	0.020	0.029	0.011
	(0.143)	(0.181)	(0.024)	(0.031)	(0.027)	(0.034)	(0.020)	(0.025)
Higher-achieving	-0.021	-0.128	0.011	-0.013	-0.018	-0.010	0.015	0.036
	(0.127)	(0.169)	(0.019)	(0.025)	(0.021)	(0.028)	(0.017)	(0.022)
More central × boy		0.834***		0.148***		0.123**		0.043
		(0.293)		(0.049)		(0.055)		(0.041)
Higher-achieving × boy		0.249		0.056		-0.020		-0.049
		(0.256)		(0.038)		(0.043)		(0.036)
mean control	13.78	13.78	-0.05	-0.05	-0.03	-0.03	-0.05	-0.05
p-val mc boys		0.029		0.009		0.001		0.091
p-val ha boys		0.528		0.138		0.372		0.636
N	6,079	6,079	6,079	6,079	6,079	6,079	6,079	6,079
			Panel B: Less central students at baseline					
More central	0.290	-0.159	0.051	-0.046	0.133***	0.063	0.045**	0.012
	(0.197)	(0.261)	(0.032)	(0.041)	(0.039)	(0.048)	(0.022)	(0.027)
Higher-achieving	-0.133	-0.177	0.007	-0.011	-0.032	-0.085*	-0.064***	-0.052*
	(0.198)	(0.261)	(0.033)	(0.042)	(0.038)	(0.047)	(0.024)	(0.030)
More central × boy		1.093***		0.236***		0.169**		0.080*
		(0.397)		(0.065)		(0.080)		(0.045)
Higher-achieving × boy		0.096		0.040		0.127		-0.029
		(0.390)		(0.064)		(0.078)		(0.047)
mean control	11.24	11.24	-0.25	-0.25	-0.16	-0.16	-0.29	-0.29
p-val mc boys		0.002		0.000		0.000		0.012
p-val ha boys		0.783		0.558		0.499		0.027
N	1,832	1,832	1,832	1,832	1,832	1,832	1,832	1,832
			Panel C: More central students at baseline					
More central	-0.312	-0.572**	-0.031	-0.058	0.009	-0.021	0.017	0.011
	(0.204)	(0.246)	(0.035)	(0.044)	(0.037)	(0.047)	(0.032)	(0.040)
Higher-achieving	0.185	0.159	0.057	0.038	-0.017	0.063	0.079***	0.112***
	(0.208)	(0.258)	(0.036)	(0.046)	(0.037)	(0.047)	(0.031)	(0.039)
More central × boy		0.639		0.068		0.075		0.013
		(0.422)		(0.072)		(0.075)		(0.065)
Higher-achieving × boy		0.059		0.046		-0.199***		-0.082
		(0.431)		(0.072)		(0.075)		(0.066)
mean control	14.28	14.28	0.27	0.27	0.04	0.04	0.21	0.21
p-val mc boys		0.844		0.866		0.353		0.628
p-val ha boys		0.526		0.131		0.020		0.554
N	1,822	1,822	1,822	1,822	1,822	1,822	1,822	1,822

Notes: This table reports the effect of being assigned to more central and higher-achieving peers on social skills outcomes. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. The sample in Panel A includes students from all the cohorts. The sample in Panels B and C includes students from the 2015-16 cohorts as there is no information on centrality at baseline for the 2017 cohort. The table also reports the p-value for the more central peers ("p-val mc boys") and the higher-achieving peers ("p-val ha boys") treatment for boys. These tests correspond to the sum of parameters $\lambda_s + \phi_s$ and parameters $\lambda_a + \phi_a$ in equation 7 being equal to zero, respectively. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 6: Reduced-Form Effects on Longer-term Outcomes

Dependent variable:	Dropout		College enrollment		Certified college		Top 20 college	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All students								
More central	-0.002	0.005	-0.026*	-0.046**	-0.010	-0.030*	-0.013	-0.031*
	(0.004)	(0.006)	(0.014)	(0.019)	(0.014)	(0.017)	(0.013)	(0.017)
Higher-achieving	0.005	0.002	-0.017	-0.024	-0.015	-0.018	-0.017	-0.023
	(0.004)	(0.006)	(0.014)	(0.019)	(0.014)	(0.017)	(0.013)	(0.017)
More central × boy		-0.018**		0.050*		0.050*		0.043*
		(0.008)		(0.028)		(0.028)		(0.026)
Higher-achieving × boy		0.007		0.018		0.006		0.014
		(0.008)		(0.028)		(0.028)		(0.026)
mean control	0.02	0.02	0.62	0.62	0.32	0.32	0.26	0.26
p-val mc boys		0.013		0.847		0.394		0.539
p-val ha boys		0.081		0.742		0.599		0.662
N	3,654	3,654	3,654	3,654	3,654	3,654	3,654	3,654
Panel B: Less central students at baseline								
More central	-0.001	0.015	-0.013	-0.020	0.017	-0.018	-0.001	-0.040*
	(0.007)	(0.010)	(0.021)	(0.028)	(0.019)	(0.024)	(0.018)	(0.023)
Higher-achieving	0.019***	0.019**	-0.030	-0.045	-0.015	-0.018	-0.019	-0.029
	(0.007)	(0.010)	(0.020)	(0.028)	(0.019)	(0.025)	(0.018)	(0.023)
More central × boy		-0.039***		0.016		0.086**		0.094***
		(0.013)		(0.040)		(0.039)		(0.035)
Higher-achieving × boy		0.001		0.036		0.008		0.024
		(0.014)		(0.041)		(0.040)		(0.036)
mean control	0.02	0.02	0.58	0.58	0.26	0.26	0.22	0.22
p-val mc boys		0.004		0.901		0.028		0.049
p-val ha boys		0.041		0.770		0.735		0.854
N	1,832	1,832	1,832	1,832	1,832	1,832	1,832	1,832
Panel C: More central students at baseline								
More central	-0.004	-0.005	-0.043**	-0.080***	-0.035*	-0.040	-0.024	-0.020
	(0.005)	(0.008)	(0.019)	(0.027)	(0.020)	(0.025)	(0.019)	(0.025)
Higher-achieving	-0.010*	-0.016*	-0.008	-0.007	-0.020	-0.022	-0.016	-0.018
	(0.005)	(0.008)	(0.019)	(0.026)	(0.020)	(0.024)	(0.019)	(0.025)
More central × boy		0.003		0.092**		0.012		-0.010
		(0.010)		(0.038)		(0.041)		(0.039)
Higher-achieving × boy		0.014		-0.004		0.004		0.004
		(0.010)		(0.038)		(0.040)		(0.039)
mean control	0.03	0.03	0.68	0.68	0.41	0.41	0.33	0.33
p-val mc boys		0.657		0.666		0.393		0.313
p-val ha boys		0.838		0.698		0.593		0.641
N	1,822	1,822	1,822	1,822	1,822	1,822	1,822	1,822

Notes: This table reports the effect of being assigned to more central and higher-achieving peers on longer-term outcomes. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. The sample in Panel A includes students from all the cohorts. The sample in Panels B and C includes students from the 2015-16 cohorts as there is no information on centrality at baseline for the 2017 cohort. The table also reports the p-value for the more central peers ("p-val mc boys") and the higher-achieving peers ("p-val ha boys") treatment for boys. These tests correspond to the sum of parameters $\lambda_s + \phi_s$ and parameters $\lambda_a + \phi_a$ in equation 7 being equal to zero, respectively. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 7: 2SLS Effects on Social Skills Index

Group:	All Students (1)	Boys (2)	Girls (3)
Panel A: All students			
Neighbors' centrality	0.039	0.246***	-0.076
	(0.046)	(0.084)	(0.053)
Neighbors' achievement	0.014	-0.000	0.002
	(0.043)	(0.073)	(0.053)
F centrality	805.84	224.70	646.08
F achievement	817.96	323.53	508.51
N	3,654	1,490	2,164
Panel B: Less central students at baseline			
Neighbors' centrality	0.137**	0.486***	-0.047
	(0.069)	(0.115)	(0.083)
Neighbors' achievement	-0.093	-0.093	-0.111
	(0.058)	(0.096)	(0.071)
F centrality	326.91	81.87	285.72
F achievement	374.30	198.76	195.68
N	1,832	753	1,079
Panel C: More central students at baseline			
Neighbors' centrality	-0.053	0.029	-0.091
	(0.061)	(0.119)	(0.069)
Neighbors' achievement	0.130**	0.095	0.131*
	(0.062)	(0.111)	(0.076)
F centrality	476.10	137.67	381.54
F achievement	463.25	139.94	338.17
N	1,822	737	1,085

Notes: This table reports 2SLS estimates of neighbors' average social centrality and academic achievement on an index of social skills, using treatment assignments as instruments. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. The sample only includes students from the 2015-16 cohorts as there is no information on centrality at baseline for the 2017 cohort. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 8: Reduced-Form Effects on Test Scores

Dependent variable:	Grades (2016-17 cohorts)				Test scores			
	Math		Reading		Math		Reading	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Students								
More central	0.022	0.001	0.041	0.044	-0.022	-0.019	0.025	-0.007
	(0.035)	(0.045)	(0.036)	(0.050)	(0.021)	(0.026)	(0.026)	(0.035)
Higher-achieving	0.009	-0.012	-0.018	-0.005	-0.027*	-0.030	-0.033	-0.072***
	(0.022)	(0.029)	(0.024)	(0.032)	(0.016)	(0.020)	(0.020)	(0.027)
More central × boy		0.051		-0.006		-0.006		0.074
		(0.071)		(0.073)		(0.043)		(0.053)
Higher-achieving × boy		0.046		-0.031		0.006		0.091**
		(0.045)		(0.048)		(0.034)		(0.040)
mean control	-0.06	-0.06	-0.05	-0.05	-0.04	-0.04	-0.04	-0.04
p-val mc boys		0.349		0.481		0.468		0.089
p-val ha boys		0.315		0.329		0.387		0.536
N	4,419	4,419	4,418	4,418	5,681	5,681	5,796	5,796
Panel B: Lower-achieving students at baseline								
More central	0.013	-0.010	0.067	0.061	-0.015	-0.062*	0.041	0.020
	(0.050)	(0.066)	(0.052)	(0.070)	(0.030)	(0.035)	(0.039)	(0.053)
Higher-achieving	-0.061*	-0.114**	-0.075**	-0.066	-0.043*	-0.071**	-0.041	-0.078*
	(0.036)	(0.048)	(0.037)	(0.048)	(0.025)	(0.028)	(0.032)	(0.042)
More central × boy		0.052		0.016		0.110*		0.050
		(0.100)		(0.104)		(0.061)		(0.078)
Higher-achieving × boy		0.119		-0.022		0.066		0.085
		(0.072)		(0.076)		(0.052)		(0.064)
mean control	-0.27	-0.27	-0.21	-0.21	-0.29	-0.29	-0.11	-0.11
p-val mc boys		0.576		0.319		0.338		0.226
p-val ha boys		0.925		0.133		0.900		0.886
N	2,195	2,195	2,195	2,195	2,778	2,778	2,860	2,860
Panel C: Higher-achieving students at baseline								
More central	0.044	0.025	0.011	0.025	-0.028	0.020	0.010	-0.030
	(0.051)	(0.063)	(0.052)	(0.071)	(0.031)	(0.038)	(0.037)	(0.048)
Higher-achieving	0.048	0.067	0.037	0.045	-0.026	-0.012	-0.033	-0.073*
	(0.035)	(0.044)	(0.037)	(0.049)	(0.025)	(0.031)	(0.030)	(0.039)
More central × boy		0.044		-0.032		-0.116*		0.096
		(0.103)		(0.104)		(0.062)		(0.074)
Higher-achieving × boy		-0.046		-0.017		-0.032		0.096
		(0.070)		(0.074)		(0.051)		(0.060)
mean control	0.25	0.25	0.18	0.18	0.31	0.31	0.07	0.07
p-val mc boys		0.402		0.925		0.054		0.246
p-val ha boys		0.701		0.631		0.271		0.620
N	2,211	2,211	2,210	2,210	2,890	2,890	2,923	2,923

Notes: This table reports the effect of being assigned to more central and higher-achieving peers identified at baseline on academic outcomes. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. The table also reports the p-value for the more central peers ("p-val mc boys") and the higher-achieving peers ("p-val ha boys") treatment for boys. These tests correspond to the sum of parameters $\lambda_s + \phi_s$ and parameters $\lambda_a + \phi_a$ in equation 7 being equal to zero, respectively. Grades are standardized at the school-by-grade level and test scores at the grade level. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 9: 2SLS Effects on Academic Achievement

Dependent Variable::	Math Test Scores			Reading Test Scores		
	All Students	Boys	Girls	All Students	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Students						
Neighbors' centrality	-0.032	-0.043	-0.027	0.054	0.132	0.007
	(0.038)	(0.074)	(0.043)	(0.049)	(0.086)	(0.058)
Neighbors' achievement	-0.046	-0.039	-0.044	-0.059*	0.023	-0.120***
	(0.029)	(0.053)	(0.033)	(0.035)	(0.059)	(0.044)
F centrality	62.04	18.95	49.90	62.04	18.95	49.90
F achievement	100.56	42.05	62.95	100.56	42.05	62.95
N	5,681	2,505	3,176	5,796	2,540	3,256
Panel B: Lower-achieving students at baseline						
Neighbors' centrality	-0.013	0.093	-0.083	0.086	0.149	0.053
	(0.053)	(0.102)	(0.056)	(0.071)	(0.119)	(0.088)
Neighbors' achievement	-0.076*	-0.020	-0.116**	-0.076	-0.004	-0.136*
	(0.046)	(0.087)	(0.048)	(0.057)	(0.094)	(0.072)
F centrality	39.61	19.05	31.45	39.61	19.05	31.45
F achievement	38.40	16.29	26.64	38.40	16.29	26.64
N	2,778	1,236	1,542	2,860	1,260	1,600
Panel C: Higher-achieving students at baseline						
Neighbors' centrality	-0.046	-0.205*	0.034	0.027	0.138	-0.033
	(0.057)	(0.119)	(0.064)	(0.070)	(0.133)	(0.081)
Neighbors' achievement	-0.040	-0.049	-0.013	-0.055	0.028	-0.117*
	(0.040)	(0.074)	(0.048)	(0.049)	(0.084)	(0.061)
F centrality	26.20	8.40	23.71	26.20	8.40	23.71
F achievement	53.48	25.18	33.76	53.48	25.18	33.76
N	2,890	1,259	1,631	2,923	1,270	1,653

Notes: This table reports 2SLS estimates of neighbors' average social centrality and academic achievement on students' academic outcomes, using the treatment assignment as instruments. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 10: Self-confidence in Social Skills

Dependent variable:	Popularity Ranking			Self-nomination (in the top-5)				Index
	Dorm	Classroom	Cohort	Leader	Popular	Friendly	No shy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Less central students at baseline								
More central	-2.912**	-3.356**	-2.723*	-0.070***	-0.022	0.015	-0.006	-0.120**
	(1.476)	(1.392)	(1.403)	(0.023)	(0.025)	(0.017)	(0.016)	(0.048)
Higher-achieving	0.143	0.787	-0.589	0.023	0.025	-0.009	0.023	0.040
	(1.480)	(1.434)	(1.440)	(0.023)	(0.025)	(0.016)	(0.016)	(0.049)
More central × boy	3.873*	5.226**	5.324**	0.110***	0.081**	0.007	0.046**	0.259***
	(2.323)	(2.182)	(2.113)	(0.038)	(0.038)	(0.032)	(0.022)	(0.078)
Higher-achieving × boy	2.115	2.188	1.383	-0.080**	-0.066*	0.009	-0.050**	-0.055
	(2.273)	(2.150)	(2.117)	(0.038)	(0.038)	(0.031)	(0.023)	(0.078)
mean control	67.29	64.85	58.77	0.28	0.22	0.14	0.91	-0.12
p-val mc boys	0.594	0.270	0.103	0.173	0.043	0.403	0.009	0.027
p-val ha boys	0.193	0.064	0.612	0.061	0.156	0.992	0.088	0.806
N	1,662	1,666	1,665	1,682	1,682	1,682	1,682	1,832
Panel B: More central students at baseline								
More central	0.830	1.347	0.852	-0.003	-0.014	0.005	0.012	0.022
	(1.362)	(1.157)	(1.257)	(0.023)	(0.025)	(0.019)	(0.014)	(0.046)
Higher-achieving	-1.209	-0.148	-0.914	0.034	0.009	-0.015	0.003	-0.010
	(1.341)	(1.150)	(1.220)	(0.024)	(0.024)	(0.020)	(0.013)	(0.046)
More central × boy	0.124	-0.100	0.065	-0.033	-0.021	-0.005	-0.017	-0.038
	(2.067)	(1.880)	(1.954)	(0.038)	(0.038)	(0.032)	(0.021)	(0.071)
Higher-achieving × boy	1.598	-0.516	-0.573	0.000	-0.013	0.009	0.032	0.016
	(2.016)	(1.858)	(1.891)	(0.040)	(0.038)	(0.032)	(0.021)	(0.072)
mean control	72.22	70.75	65.26	0.30	0.25	0.16	0.92	0.09
p-val mc boys	0.539	0.400	0.541	0.232	0.224	0.981	0.764	0.768
p-val ha boys	0.797	0.654	0.308	0.274	0.882	0.812	0.028	0.911
N	1,700	1,699	1,701	1,710	1,710	1,710	1,710	1,822

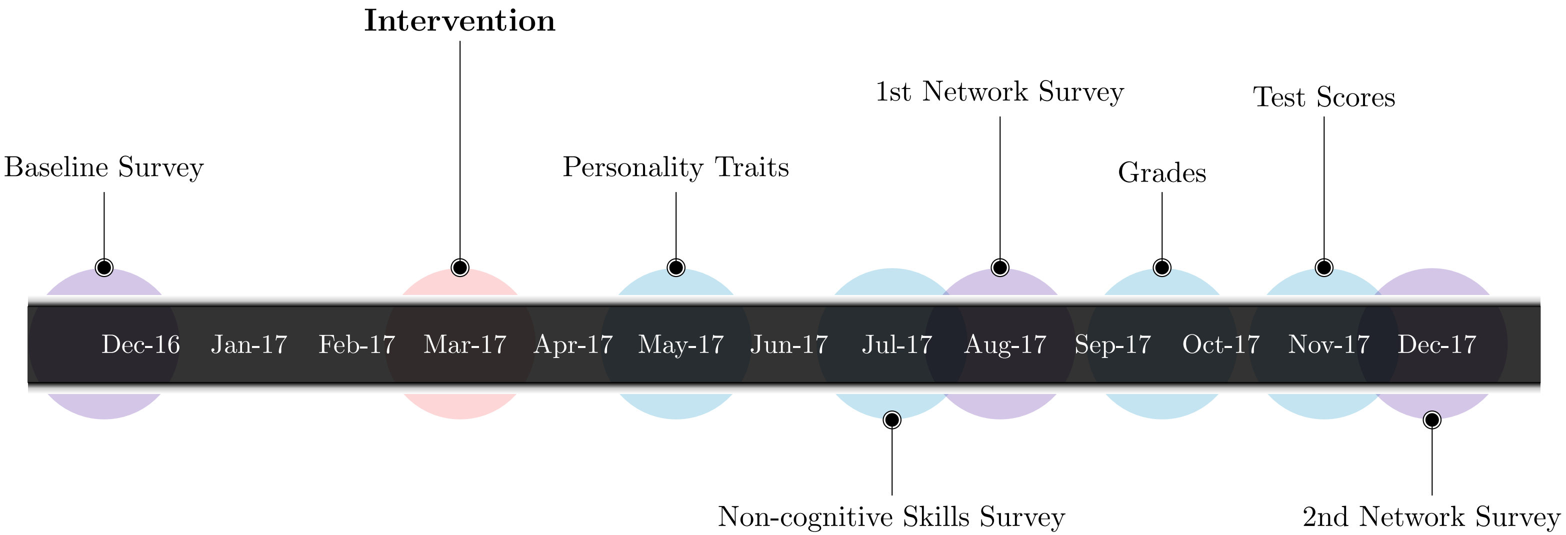
Notes: This table reports the effect of being assigned to more central and higher-achieving peers identified at baseline on self-confidence in social skills. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. The control group is defined as being assigned to less central and lower-achieving peers. The sample only includes students from the 2015-16 cohorts as there is no information on centrality at baseline for the 2017 cohort. The table also reports the p-value for the more central peers ("p-val mc boys") and the higher-achieving peers ("p-val ha boys") treatment for boys. These tests correspond to the sum of parameters $\lambda_s + \phi_s$ and parameters $\lambda_a + \phi_a$ in equation 7 being equal to zero, respectively. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 11: Self-confidence in Academic Skills

Dependent variable:	Academic Ranking			Competition		Self-nominate	Index
	Dorm	Classroom	Cohort	Want to do better than peers	Avoid doing worse than peers	top-5 skilled	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Lower-achieving students at baseline							
More central	-0.560	0.091	-0.491	-0.023	0.011	-0.010	-0.016
	(1.038)	(1.041)	(0.986)	(0.062)	(0.063)	(0.016)	(0.048)
Higher-achieving	-1.485	0.053	-0.770	-0.190***	-0.129**	0.015	-0.093**
	(0.945)	(0.873)	(0.878)	(0.053)	(0.055)	(0.015)	(0.044)
More central × boy	-1.650	-1.374	-0.883	0.020	0.003	-0.043	-0.100
	(1.644)	(1.604)	(1.551)	(0.093)	(0.097)	(0.028)	(0.077)
Higher-achieving × boy	1.475	-0.400	0.995	0.260***	0.224***	-0.057**	0.090
	(1.417)	(1.346)	(1.291)	(0.080)	(0.085)	(0.025)	(0.068)
mean control	72.46	69.62	65.75	-0.01	-0.05	0.15	-0.09
p-val mc boys	0.082	0.293	0.251	0.965	0.848	0.018	0.055
p-val ha boys	0.992	0.737	0.812	0.241	0.139	0.036	0.945
N	2,801	2,805	2,805	2,672	2,672	2,831	3,025
Panel B: Higher-achieving students at baseline							
More central	-0.264	-0.719	-0.632	-0.053	0.009	-0.000	-0.033
	(0.963)	(0.865)	(0.859)	(0.060)	(0.055)	(0.019)	(0.046)
Higher-achieving	-0.694	-1.617**	-1.294*	-0.001	-0.008	0.028*	-0.022
	(0.829)	(0.755)	(0.722)	(0.048)	(0.045)	(0.017)	(0.037)
More central × boy	1.844	3.104**	1.667	-0.053	-0.136	-0.034	0.029
	(1.609)	(1.565)	(1.436)	(0.091)	(0.092)	(0.035)	(0.078)
Higher-achieving × boy	1.663	2.359*	2.422**	0.019	0.022	-0.029	0.087
	(1.354)	(1.283)	(1.187)	(0.072)	(0.072)	(0.029)	(0.064)
mean control	74.77	73.47	69.50	-0.04	0.05	0.19	0.09
p-val mc boys	0.226	0.070	0.372	0.123	0.083	0.232	0.949
p-val ha boys	0.367	0.475	0.231	0.750	0.799	0.998	0.207
N	2,848	2,851	2,850	2,765	2,765	2,868	3,041

Notes: This table reports the effect of being assigned to more central and higher-achieving peers identified at baseline on self-confidence in academic skills. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. The table also reports the p-value for the more central peers ("p-val mc boys") and the higher-achieving peers ("p-val ha boys") treatment for boys. These tests correspond to the sum of parameters $\lambda_s + \phi_s$ and parameters $\lambda_a + \phi_a$ in equation 7 being equal to zero, respectively. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Figure 1: Timeline of the Project



Notes: This figure presents the timeline of the project. The purple circles represent data collection with surveys, the blue circles the collection of administrative data through the Ministry of Education, and the red circle the implementation of the intervention.

Figure 2: Student-Peer Type Combinations in the Experimental Design

Student Type	Peer Types: higher-achieving more central	higher-achieving less central	lower-achieving more central	lower-achieving less central
higher-achieving more central	Combination 1	Combination 2	Combination 3	Combination 4
higher-achieving less central	Combination 2	Combination 5	Combination 6	Combination 7
lower-achieving more central	Combination 3	Combination 6	Combination 8	Combination 9
lower-achieving less central	Combination 4	Combination 7	Combination 9	Combination 10

Notes: this figure shows the ten student-peer type combinations in my experimental design. It represents all possible combinations between student type and peer types. Rows represent student types, and columns show the types of peers to which they were randomly assigned. The diagonal of the matrix is composed of groups of a single type. The matrix is symmetric since students are matched with peers of the assigned type.

Figure 3: Dorm Structure

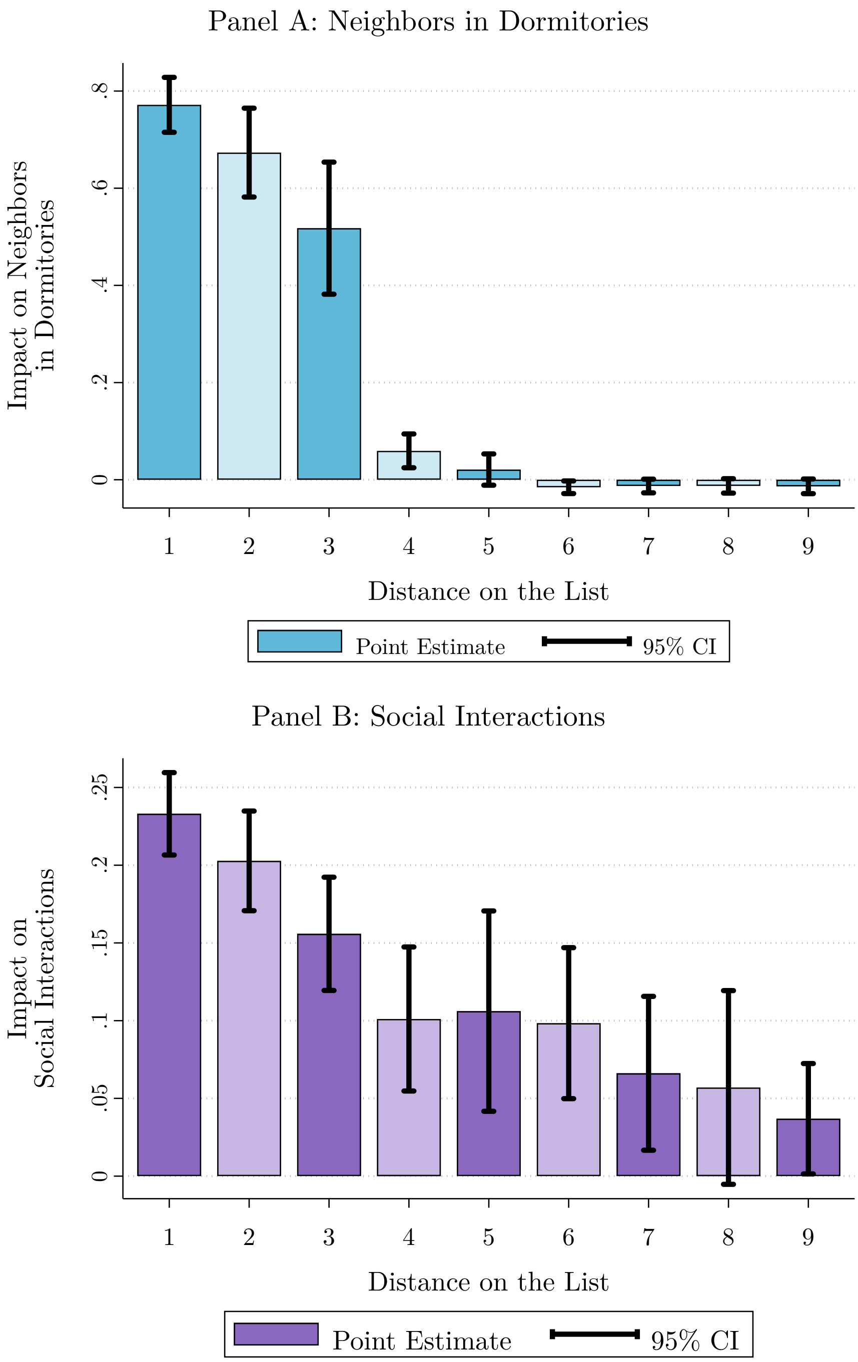
School in Lima

School in Piura

School in Cusco

Notes: This figure displays pictures of the dorms for Lima, Piura, and Cusco schools. It illustrates the vast heterogeneity in the type of dorms across the schools.

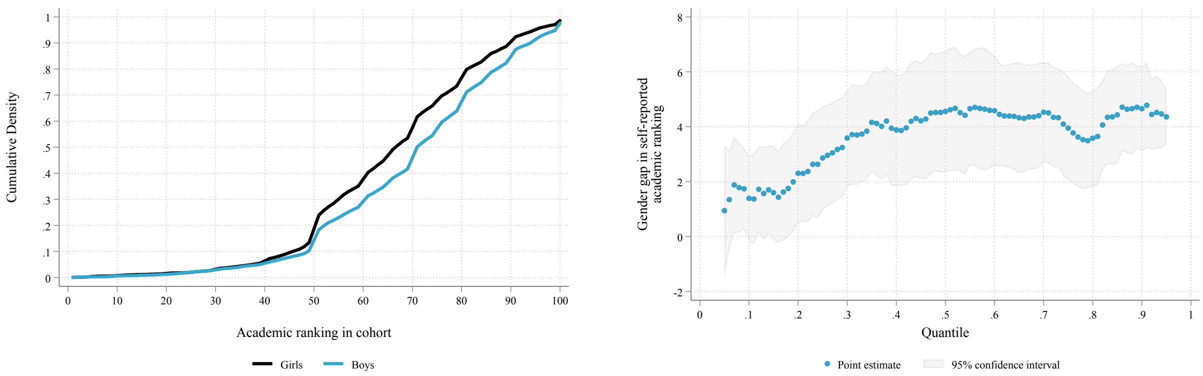
Figure 4: Effects of Proximity on the List on Neighbors and Social Interactions



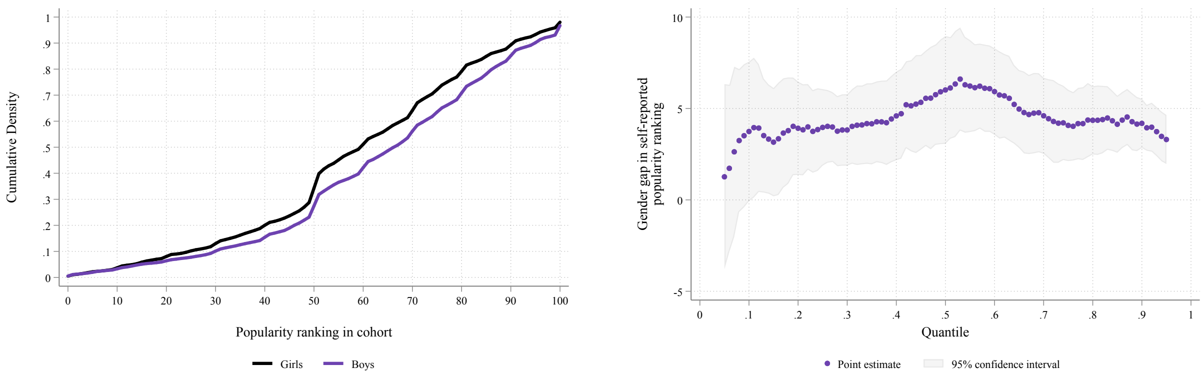
Notes: This figure shows the impact of distance between a pair of students on the likelihood of being neighbors and social interaction (friends, study, and playing games or sports). Nine distance dummies capture the effect of distance on the list. Students are at an odd distance from peers that provide the treatment and at an even distance from peers of their same type. The figure also displays 95% confidence intervals for each of these dummy variables. All estimations control for strata fixed effects. Standard errors are clustered at the school-by-cohort level.

Figure 5: Gender Differences in Self-reported Rankings

Panel A: Academic ranking in cohort



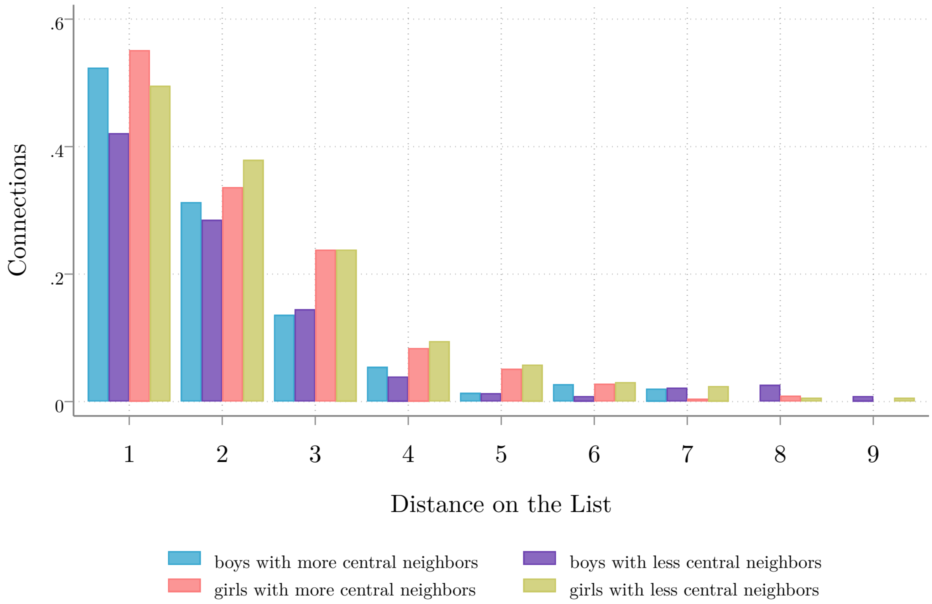
Panel B: Popularity ranking in cohort



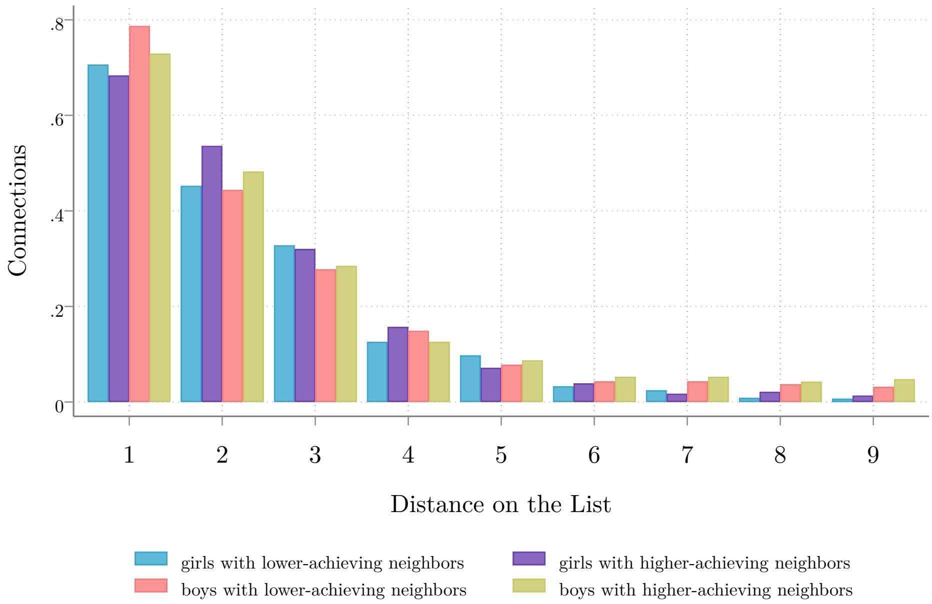
Notes: This figure plots differences by gender in self-reported academic and popularity ranking within the cohort. The left column presents the cumulative distribution function and the right column the estimates from quantile regressions of the gender gap after controlling for observable characteristics. These covariates include scores in mathematics and reading tests, network degree and centrality, and peers' perception of social and academic skills. Standard errors are clustered at the school-by-cohort level.

Figure 6: Social Interactions of Most Affected vs. Comparable Groups

Panel A: Less central students



Panel B: Lower-achieving students



Notes: This figure shows the average number of connections with neighbors using nine dummies of the distance on the list and by student's type, treatment, and gender. Students are at an odd distance from peers that provide the treatment and at an even distance from peers of their same type.

Uncovering Peer Effects in Social and Academic Skills

Román Andrés Zárate

Online Appendix

A Supplementary Material

Table A.1: Correlation of Sociability and Social Skills Outcomes

	Big Five Personality Traits					Peers' Perception				Other measures of social skills	Social Skills Index at baseline	Social Skills Index at endline
	Openness	Conscientiousness	Emotional Stability	Extraversion	Agreeableness	Leadership	Friendliness	Popularity	Shyness			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Academic achievement	0.090***	-0.001	0.065***	-0.015	-0.015	0.217***	0.044***	0.105***	-0.066***	0.040**	0.056***	0.040**
	(0.021)	(0.022)	(0.022)	(0.019)	(0.020)	(0.020)	(0.016)	(0.022)	(0.019)	(0.016)	(0.017)	(0.017)
Sociability	0.103***	0.062**	0.032*	0.142***	0.121***	0.230***	0.360***	0.215***	-0.103***	0.092***	0.125***	0.117***
	(0.024)	(0.025)	(0.017)	(0.021)	(0.017)	(0.021)	(0.020)	(0.029)	(0.021)	(0.019)	(0.019)	(0.020)
Social-fit score	0.072***	0.022	0.001	0.063***	0.027	0.136***	0.075***	0.103***	-0.112***	0.027	0.057***	0.042**
	(0.018)	(0.020)	(0.020)	(0.019)	(0.020)	(0.015)	(0.016)	(0.017)	(0.020)	(0.018)	(0.018)	(0.019)
Interview score	0.092***	0.072***	0.072***	0.090***	0.048***	0.069***	0.058***	0.050***	-0.036**	0.066***	0.118***	0.080***
	(0.019)	(0.016)	(0.016)	(0.018)	(0.018)	(0.015)	(0.015)	(0.017)	(0.017)	(0.016)	(0.016)	(0.015)
N	3,106	3,106	3,106	3,106	3,106	3,637	3,637	3,637	3,637	3,654	3,654	3,654

Notes: This table reports standardized estimates of an OLS regression on social skills outcomes of social centrality at baseline and the score in the three tests of the admission process to the COAR Network. All regressions include school-by-grade-by-gender fixed effects. Academic achievement and social centrality are measured at baseline. Centrality at baseline is measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality is a measure of the influence of a student in the network. Academic achievement at baseline is the score on the admission test to the COAR Network, which evaluates the applicants in math and reading comprehension. In columns 1 to 5, the dependent variables are personality traits from the Big Five. In columns 6 to 9, the dependent variables are the number of peers who perceive the student as part of the top 5 of leadership, friendliness, popularity, and shyness. In column 10, the dependent variable is an index excluding social network outcomes, personality traits, and peers' perceptions. In columns 11 and 12, the dependent variable is a social skills index that excludes social network outcomes. Column 11 presents the correlations on this index at baseline and column 12 at endline. All indexes are constructed using Principal Component Analysis (PCA) on all the variables that measure social skills (see Appendix D for details). *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.2: Balance Tests for the More Central Peers Treatment

Variable	All Students		Less Central at Baseline		More Central at Baseline	
	Control mean	Difference	Control mean	Difference	Control mean	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Admission test	-0.016	0.002	-0.064	-0.024	0.058	0.028
		(0.019)		(0.027)		(0.026)
Interview score	-0.006	0.017	-0.006	0.041	-0.006	-0.006
		(0.030)		(0.042)		(0.042)
Social-fit score	-0.024	0.045	-0.028	0.027	-0.019	0.062
		(0.031)		(0.042)		(0.044)
Female (%)	0.591	0.000	0.589	0.000	0.594	0.000
		(0.000)		(0.000)		(0.000)
Poor (%)	0.431	0.014	0.460	0.026	0.389	0.003
		(0.014)		(0.021)		(0.019)
Rural household (%)	0.284	-0.018	0.314	-0.024	0.241	-0.013
		(0.014)		(0.020)		(0.019)
Subsidized health insurance	0.508	0.008	0.552	-0.011	0.443	0.027
		(0.016)		(0.022)		(0.022)
Math scores	-0.043	0.027	-0.141	0.048	0.103	0.007
		(0.023)		(0.032)		(0.031)
Reading scores	-0.029	-0.001	-0.128	0.038	0.120	-0.040
		(0.021)		(0.031)		(0.028)
Social skills	-0.108	-0.000	-0.522	0.023	0.515	-0.024
		(0.022)		(0.024)		(0.037)
Degree friends	7.331	0.230	5.609	0.163	9.918	0.295
		(0.130)		(0.113)		(0.234)
Centrality friends	-0.110	0.021	-0.521	0.032	0.505	0.011
		(0.025)		(0.023)		(0.044)
Degree study	4.560	-0.007	3.741	-0.060	5.789	0.046
		(0.069)		(0.081)		(0.111)
Centrality study	-0.071	-0.006	-0.350	-0.056	0.340	0.043
		(0.027)		(0.028)		(0.046)
Degree all	10.475	0.067	8.261	-0.060	13.802	0.193
		(0.135)		(0.127)		(0.238)
Centrality all	-0.152	0.023	-0.705	-0.010	0.679	0.056
		(0.018)		(0.016)		(0.031)
Reading the mind in the eyes	20.521	-0.060	20.224	0.184	20.960	-0.304
		(0.130)		(0.186)		(0.180)
Peers' perception leadership	2.499	-0.185	1.586	-0.111	3.874	-0.259
		(0.150)		(0.132)		(0.271)
Peers' perception friendliness	2.550	-0.008	1.836	0.195	3.625	-0.214
		(0.080)		(0.083)		(0.135)
Peers' perception popularity	2.201	0.108	1.465	0.127	3.310	0.089
		(0.159)		(0.142)		(0.285)
Peers' perception shyness	2.083	0.025	2.561	-0.137	1.364	0.189
		(0.144)		(0.226)		(0.178)
Total score Rosenberg Scale	32.991	0.101	32.777	0.119	33.306	0.081
		(0.154)		(0.224)		(0.212)
Total score Grit Scale	43.707	-0.248	43.340	-0.171	44.251	-0.325
		(0.198)		(0.285)		(0.274)
Multivariate F p-value		0.756		0.479		0.594

Notes: This table reports balance checks of being assigned to more central peers on baseline characteristics. All regressions include strata fixed effects and include the higher-achieving peers treatment. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. The "F p-value" corresponds to the F-statistic of the more central peers treatment of multivariate regressions that include all the variables at baseline. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.3: Balance Tests for the Higher-Achieving Peers Treatment

Variable	All Students		Lower-achieving at Baseline		Higher-achieving at Baseline	
	Control mean	Difference	Control mean	Difference	Control mean	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Admission test	-0.163	0.022	-0.787	0.004	0.764	0.036
		(0.015)		(0.018)		(0.026)
Interview score	0.049	-0.016	0.212	-0.031	-0.192	-0.015
		(0.024)		(0.031)		(0.041)
Social-fit score	0.045	-0.030	0.164	-0.014	-0.130	-0.050
		(0.022)		(0.029)		(0.039)
Female (%)	0.567	0.000	0.566	-0.000	0.567	0.000
		(0.000)		(0.000)		(0.000)
Poor (%)	0.418	0.010	0.456	0.015	0.362	-0.002
		(0.012)		(0.018)		(0.017)
Rural household (%)	0.267	0.004	0.308	-0.019	0.206	0.017
		(0.011)		(0.018)		(0.015)
Subsidized health insurance	0.504	0.013	0.517	0.045	0.485	-0.019
		(0.015)		(0.021)		(0.022)
Math scores	-0.081	0.007	-0.345	-0.023	0.310	0.028
		(0.019)		(0.026)		(0.030)
Reading scores	-0.047	-0.014	-0.234	-0.033	0.229	0.006
		(0.018)		(0.027)		(0.027)
Social skills	-0.015	-0.017	-0.090	-0.029	0.097	-0.006
		(0.022)		(0.030)		(0.032)
Degree friends	8.023	-0.438	7.787	-0.418	8.374	-0.457
		(0.134)		(0.183)		(0.195)
Centrality friends	0.029	-0.081	-0.025	-0.082	0.107	-0.080
		(0.025)		(0.033)		(0.037)
Degree study	4.719	0.012	4.609	-0.145	4.882	0.167
		(0.069)		(0.097)		(0.096)
Centrality study	-0.011	-0.004	-0.058	-0.028	0.059	0.020
		(0.027)		(0.034)		(0.042)
Degree all	11.125	-0.244	10.946	-0.394	11.392	-0.094
		(0.137)		(0.187)		(0.201)
Centrality all	-0.002	-0.012	-0.031	-0.036	0.041	0.011
		(0.018)		(0.025)		(0.025)
Reading the mind in the eyes	20.602	-0.264	20.248	-0.231	21.127	-0.295
		(0.128)		(0.180)		(0.183)
Peers' perception leadership	2.487	0.011	1.923	-0.026	3.328	0.047
		(0.148)		(0.178)		(0.236)
Peers' perception friendliness	2.688	0.011	2.612	-0.002	2.801	0.024
		(0.078)		(0.112)		(0.109)
Peers' perception popularity	2.361	-0.018	2.006	-0.025	2.890	-0.014
		(0.158)		(0.185)		(0.255)
Peers' perception shyness	2.017	0.038	2.108	0.162	1.881	-0.086
		(0.146)		(0.211)		(0.203)
Total score Rosenberg Scale	33.013	0.138	32.854	0.127	33.249	0.153
		(0.154)		(0.221)		(0.215)
Total score Grit Scale	43.568	0.160	43.330	0.431	43.921	-0.108
		(0.196)		(0.271)		(0.282)
Multivariate F p-value		0.256		0.889		0.232

Notes: This table reports balance checks of being assigned to higher-achieving peers on baseline characteristics. All regressions include strata fixed effects, control for the baseline value of the dependent variable, and include the more central peers treatment. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. The "F p-value" corresponds to the F-statistic of the higher-achieving peers treatment of multivariate regressions that include all the variables at baseline. Standard errors are clustered at the peer-group-type-by-student-type level level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.4: Correlations between Types of Skills and Longer-term Outcomes

Dependent variable:	Dropout	College	Certified	Top 20
	(1)	(2)	(3)	(4)
Social skills	-0.008***	0.024***	0.036***	0.026***
	(0.001)	(0.008)	(0.008)	(0.008)
Math scores	-0.002	0.069***	0.079***	0.090***
	(0.001)	(0.010)	(0.010)	(0.010)
Reading scores	0.002*	0.019**	0.025***	0.036***
	(0.001)	(0.009)	(0.009)	(0.009)
mean control	0.02	0.62	0.33	0.27
N	6,147	3,654	3,654	3,654

Notes: This table reports the correlation of social skills with longer-term outcomes. The three variables of interest are standardized. All models include school-by-cohort-by-gender fixed effects. Column 1 presents the results on the dropout rate with data available for all cohorts. Columns 3 to 4 present the estimates on college outcomes only available for the 2015-16 cohorts. Robust standard errors in parentheses; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.5: Randomization Inference

Group (1)	Treatment (2)	Dependent variables (3)	(4)	(5)	(6)
		Panel A: Results on Social Skills			
		Connections	Centrality	Psychological Tests	Peers' perception
	More central peers	0.003	0.012	0.070	0.030
		[0.978]	[0.623]	[0.006]	[0.142]
All students	Higher-achieving peers	-0.117	-0.001	-0.017	0.012
		[0.378]	[0.960]	[0.409]	[0.479]
	Joint test	[0.668]	[0.878]	[0.023]	[0.232]
	More central peers	0.498	0.099	0.144	0.057
		[0.038]	[0.011]	[0.002]	[0.092]
Boys	Higher-achieving peers	0.042	0.032	-0.028	-0.017
		[0.834]	[0.262]	[0.399]	[0.546]
	Joint test	[0.115]	[0.027]	[0.003]	[0.222]
	More central peers	0.946	0.195	0.237	0.099
		[0.004]	[0.001]	[0.001]	[0.010]
Less central boys	Higher-achieving peers	-0.232	0.009	0.052	-0.090
		[0.466]	[0.865]	[0.383]	[0.021]
	Joint test	[0.012]	[0.001]	[0.001]	[0.005]
		Panel B: Results on Academic Skills			
		Grades		Test Scores	
		Math	Reading	Math	Reading
	More central peers	0.022	0.041	-0.022	0.025
		[0.517]	[0.246]	[0.297]	[0.365]
All students	Higher-achieving peers	0.009	-0.018	-0.027	-0.033
		[0.691]	[0.442]	[0.125]	[0.102]
	Joint test	[0.742]	[0.398]	[0.184]	[0.201]
	More central peers	0.000	0.044	-0.020	-0.007
		[0.997]	[0.416]	[0.449]	[0.865]
Girls	Higher-achieving peers	-0.012	-0.006	-0.027	-0.073
		[0.677]	[0.834]	[0.165]	[0.005]
	Joint test	[0.906]	[0.686]	[0.317]	[0.026]
	More central peers	0.013	0.067	-0.015	0.041
		[0.790]	[0.183]	[0.631]	[0.295]
Lower-achieving	Higher-achieving peers	-0.061	-0.075	-0.043	-0.041
		[0.075]	[0.039]	[0.098]	[0.170]
	Joint test	[0.220]	[0.060]	[0.223]	[0.266]
	More central peers	-0.010	0.063	-0.061	0.020
		[0.858]	[0.369]	[0.081]	[0.709]
Lower-achieving girls	Higher-achieving peers	-0.115	-0.068	-0.069	-0.080
		[0.010]	[0.137]	[0.015]	[0.048]
	Joint test	[0.044]	[0.277]	[0.016]	[0.161]

Notes: This table reports randomization inference p-values for social and academic outcomes by groups of students and treatments. All the estimates come from a separate regression for each subgroup. The first column presents the group for which the test is performed, and the second column the respective treatment. Columns 3 to 6 show the set of outcomes: social skills outcomes in Panel A and academic outcomes in Panel B. All p-values are in square brackets. The "Joint test" corresponds to the p-value of the joint test of at least one treatment being statistically significant. The p-values were calculated using the procedure developed by Young (2018) with 1,000 randomization iterations.

Table A.6: Multiple-Hypotheses Testing

Group	Test	Dependent variables			
(1)	(2)	(3)	(4)	(5)	(6)
		Panel A: Results on Social Skills			
		Connections	Centrality	Psychological Tests	Peers' perception
Boys	Point estimate	0.510	0.100	0.144	0.061
	Sidak and Holm	0.060	0.031	0.002	0.118
	Bonferroni and Holm	0.061	0.031	0.002	0.122
	Westfall and Young	0.107	0.058	0.004	0.188
Less sociable boys	Point estimate	0.946	0.200	0.237	0.099
	Sidak and Holm	0.012	0.002	0.001	0.032
	Westfall and Young	0.012	0.002	0.001	0.032
		0.029	0.004	0.002	0.079
		Panel B: Results on Academic Skills			
		Grades		Test Scores	
		Math	Reading	Math	Reading
Lower-achieving	Point estimate	-0.061	-0.075	-0.043	-0.041
	Sidak and Holm	0.176	0.083	0.172	0.348
	Bonferroni and Holm	0.184	0.085	0.180	0.385
	Westfall and Young	0.259	0.135	0.284	0.414
Girls	Point estimate	-0.012	-0.006	-0.027	-0.073
	Sidak and Holm	0.686	0.842	0.316	0.011
	Bonferroni and Holm	0.686	0.842	0.346	0.012
	Westfall and Young	0.749	0.880	0.400	0.018
Lower-achieving girls	Point estimate	-0.115	-0.068	-0.069	-0.080
	Sidak and Holm	0.063	0.422	0.055	0.193
	Bonferroni and Holm	0.065	0.512	0.056	0.209
	Westfall and Young	0.113	0.567	0.103	0.262

Notes: This table reports multiple-hypotheses testing p-values. The first column presents the group for which the test is performed among all the possible classifications of students. For instance, if the groups are boys, the reported test is on the treatment effect for boys of multiple hypotheses that considers the impact on both boys and girls. The second column corresponds to the respective test of multiple hypotheses. Columns 3 to 6 show the set of outcomes: social skills outcomes in Panel A and academic outcomes in Panel B. Calculations were performed using the *wyoung* command developed by Jones et al. (2019).

Table A.7: Treatment Effects on Social Skills (with an Interaction Term)

Group:	All students			Less central			More central		
	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
More central	0.052	0.121**	0.002	0.065	0.236***	-0.042	0.060	0.023	0.081
	(0.034)	(0.056)	(0.044)	(0.046)	(0.070)	(0.060)	(0.053)	(0.091)	(0.065)
Higher-achieving	0.039	0.021	0.048	-0.046	-0.006	-0.079	0.161***	0.045	0.232***
	(0.034)	(0.056)	(0.043)	(0.042)	(0.067)	(0.052)	(0.058)	(0.101)	(0.069)
Interaction	-0.059	-0.009	-0.095	-0.006	-0.048	0.022	-0.146**	0.015	-0.253***
	(0.049)	(0.077)	(0.064)	(0.067)	(0.105)	(0.084)	(0.074)	(0.121)	(0.094)
N	3,654	1,490	2,164	1,832	753	1,079	1,822	737	1,085

Notes: This table reports the effect of being assigned to more central peers, higher-achieving peers, and the interaction of both treatments on social skills. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. The sample includes students from the 2015-16 cohorts as there is no information on centrality at baseline for the 2017 cohort. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.8: Treatment Effects on Academic Achievement (with an Interaction Term)

Group:	All students			Lower-achieving			Higher-achieving		
	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Dependent variable math scores									
More central	-0.028	-0.061	-0.004	-0.011	0.033	-0.047	-0.061	-0.190**	0.030
	(0.027)	(0.047)	(0.033)	(0.037)	(0.066)	(0.041)	(0.044)	(0.073)	(0.054)
Higher-achieving	-0.031	-0.045	-0.017	-0.039	-0.017	-0.057*	-0.046	-0.096*	-0.002
	(0.020)	(0.034)	(0.023)	(0.031)	(0.057)	(0.032)	(0.031)	(0.052)	(0.038)
Interaction	0.012	0.073	-0.031	-0.011	0.032	-0.033	0.056	0.160*	-0.018
	(0.036)	(0.061)	(0.044)	(0.056)	(0.095)	(0.065)	(0.052)	(0.089)	(0.064)
N	5,681	2,505	3,176	2,778	1,236	1,542	2,890	1,259	1,631
Panel B: Dependent variable reading scores									
More central	0.041	0.113**	-0.014	0.015	0.089	-0.030	0.074	0.161*	0.004
	(0.035)	(0.052)	(0.048)	(0.048)	(0.070)	(0.066)	(0.056)	(0.084)	(0.075)
Higher-achieving	-0.023	0.047	-0.078**	-0.065	0.014	-0.125**	0.005	0.080	-0.055
	(0.024)	(0.037)	(0.031)	(0.040)	(0.061)	(0.053)	(0.036)	(0.056)	(0.045)
Interaction	-0.031	-0.093	0.015	0.063	-0.032	0.119	-0.110*	-0.164	-0.059
	(0.043)	(0.066)	(0.058)	(0.065)	(0.101)	(0.087)	(0.066)	(0.101)	(0.087)
N	5,796	2,540	3,256	2,860	1,260	1,600	2,923	1,270	1,653

Notes: This table reports the effect of being assigned to more central peers, higher-achieving peers, and the interaction of both treatments on academic achievement. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include gender-by-classroom fixed effects. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

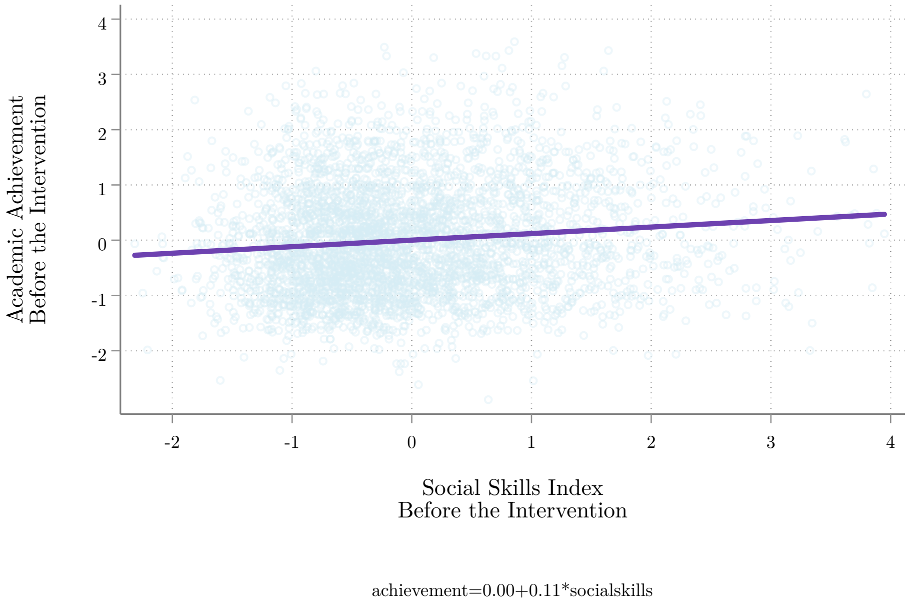
Table A.9: Social Connections with Neighbors

Dependent variable: (network)	Friend	Study	Social	Any	Help academic	Help personal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Less central students at baseline						
More central	0.002	-0.010	-0.032	0.018	-0.000	-0.036
	(0.046)	(0.033)	(0.042)	(0.053)	(0.032)	(0.034)
Higher-achieving	-0.051	-0.021	-0.064	-0.027	0.059*	0.027
	(0.043)	(0.033)	(0.041)	(0.051)	(0.033)	(0.035)
More central × boy	0.035	0.007	0.076	0.051	0.033	0.075*
	(0.065)	(0.050)	(0.062)	(0.072)	(0.046)	(0.044)
Higher-achieving × boy	0.054	0.014	0.067	0.021	-0.017	-0.017
	(0.065)	(0.050)	(0.060)	(0.072)	(0.045)	(0.044)
mean control	0.57	0.39	0.54	0.74	0.20	0.28
p-val mc boys	0.426	0.936	0.328	0.159	0.317	0.156
p-val ha boys	0.954	0.850	0.947	0.909	0.180	0.721
N	1,829	1,829	1,829	1,829	1,829	1,829
Panel B: Lower-achieving students at baseline						
More central	-0.126*	-0.052	-0.105	-0.061	0.009	-0.022
	(0.065)	(0.050)	(0.068)	(0.079)	(0.048)	(0.054)
Higher-achieving	0.041	0.055	-0.014	0.034	0.092**	0.106**
	(0.053)	(0.046)	(0.054)	(0.058)	(0.038)	(0.046)
More central × boy	0.199**	0.037	0.141	0.122	0.008	0.041
	(0.090)	(0.073)	(0.089)	(0.104)	(0.068)	(0.064)
Higher-achieving × boy	0.014	-0.051	-0.116	-0.022	-0.027	-0.151***
	(0.082)	(0.067)	(0.076)	(0.085)	(0.055)	(0.055)
mean control	1.15	0.84	1.04	1.45	0.47	0.51
p-val mc boys	0.255	0.781	0.560	0.377	0.743	0.613
p-val ha boys	0.381	0.946	0.016	0.857	0.103	0.147
N	2,269	2,269	2,269	2,269	2,269	2,269

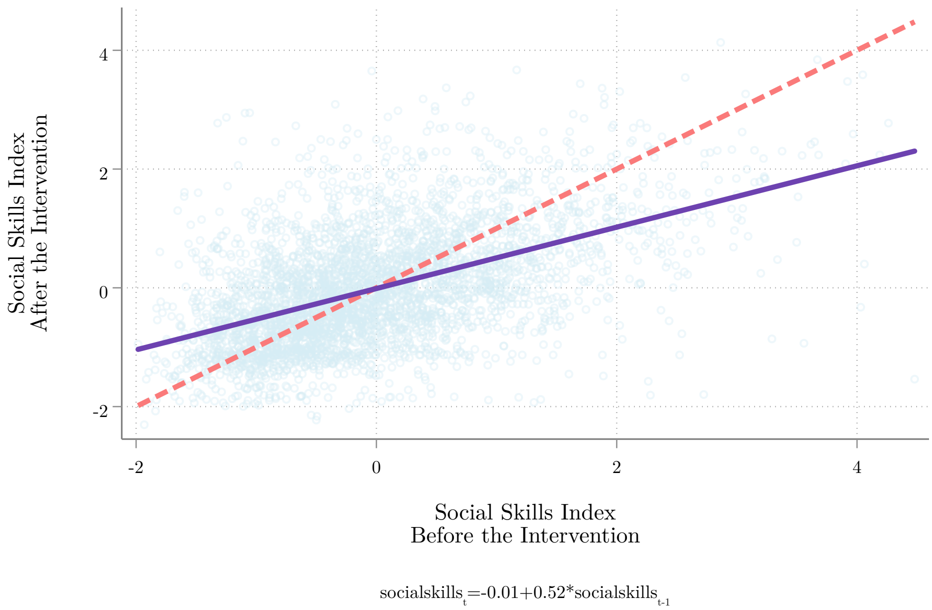
Notes: This table reports the effect of being assigned to more central and higher-achieving peers on the number of social connections with neighbors in dormitories. All regressions control for strata fixed effects and selected covariates at baseline, including the score on the admission test, math and reading scores, centrality and degree in the social network, peers' perception of social skills, and own perception of academic and social skills. For the 2017 cohort, all regressions include strata-by-classroom fixed effects. The control group is defined as being assigned to less central and lower-achieving peers. Standard errors are clustered at the peer-group-type-by-student-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Figure A.1: Social Skills Index

Panel A: Correlation with Academic Achievement



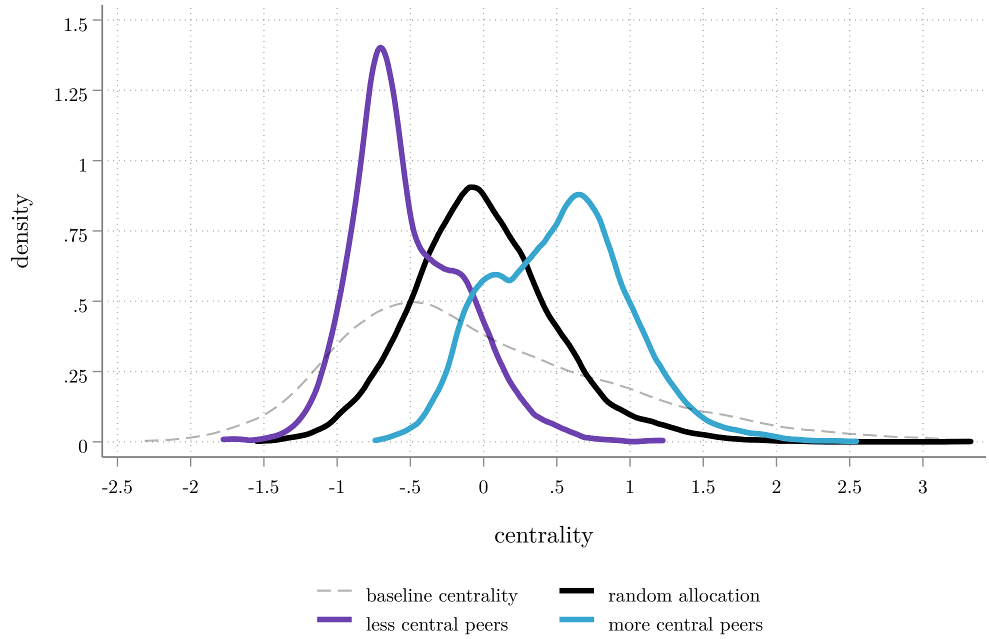
Panel B: Correlation over Time



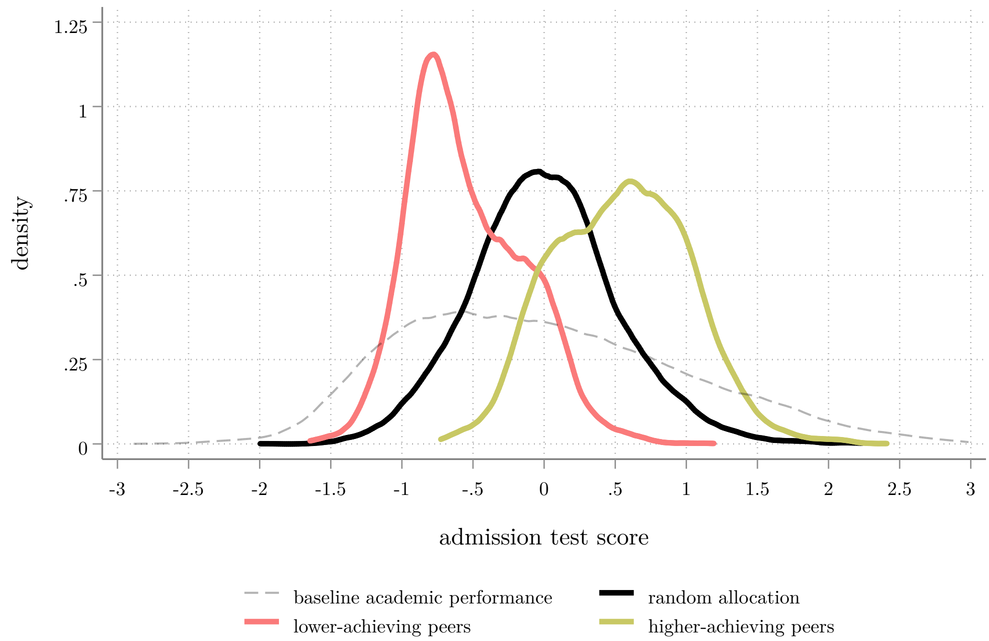
Notes: Panel A shows a scatter plot of academic achievement and the social skills index at baseline for the 2015-16 cohorts by student type. A one-standard-deviation of the social skills index predicts an increase in 0.11 standard deviations of academic achievement at baseline. Panel B shows a scatter plot and the linear prediction of the social skills index before and after the intervention. A one-standard-deviation of the social skills index before the intervention predicts an increase of 0.42 in the social skills index after the intervention.

Figure A.2: Distribution of Baseline and Peer Characteristics

Panel A: Social Centrality

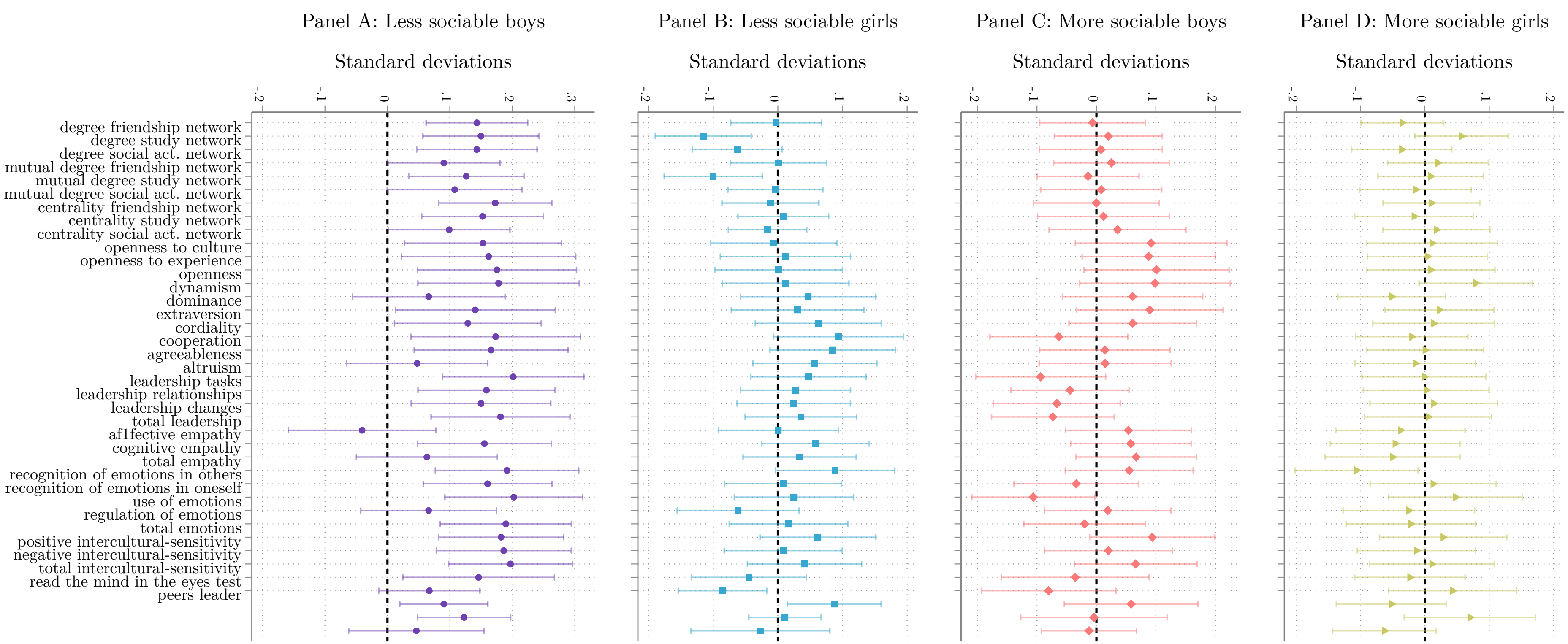


Panel B: Academic Achievement



Notes: This figure plots the distribution of baseline and peer characteristics in the allocation to the student-peer type combinations. It also shows the distribution of peer characteristics using random assignment to groups for comparison.

Figure A.3: Effects of More Central Peers on Social Skills



Notes: This figure reports treatment effects and 90% confidence intervals of being assigned to more central peers on social skills outcomes. All regressions include strata fixed effects and control for the baseline value of the dependent variable. The control group is defined as being assigned to less central peers. Standard errors are clustered at the peer-group-type-by-student-type level.

B Estimation of Peer Effects

This appendix describes the methodological concerns about exploiting random allocation to groups to identify peer effects and how my experimental design addresses them.

B.1 Random Allocation to Groups

A widely used research method to estimate peer effects - is to exploit random allocation to groups. In settings where schools and colleges apply this method, there is no self-selection of peers, and ex-ante individual and peer characteristics are unrelated. Hence, random group allocation allows researchers to estimate the causal impact of predetermined peer characteristics on individual outcomes.

However, there are some problems with using random allocation to groups to estimate peer effects, which can either be under- or over-estimated. By construction, the variation in peer characteristics from random groups is small. As groups get larger, this problem is aggravated. As Manski in Epple and Romano (2011) comments, "Random assignment will not work well in a large group setting, because all groups will have essentially the same distribution of types." Similarly, Angrist (2014) argues that "the interpretation of results from models that rely solely on chance variation in peer groups is therefore complicated by bias from weak instruments."

I now build on Angrist (2014) to describe methodological concerns about random allocation to groups. I then relate this to the existing literature and explain the advantages of my experimental design.

To introduce the problem, consider the following peer-effects model:

$$y_{ig} = \alpha + \pi_0 x_{ig} + \pi_1 \overline{x}_g + \varepsilon_{ig} \tag{B.1}$$

where y_{ig} is the outcome of individual i when assigned to group g, x_{ig} is a pre-specified exogenous characteristic of individual i in group g, and $\overline{x}_g$ is the mean of the exogenous characteristic x among those in group g.

Parameter π_1 is the causal effect of a change in the group average of x over students' outcomes. Acemoglu and Angrist (2000) show that equation B.1 relates to whether a 2SLS estimator using group dummies to instrument individual characteristics differs from OLS estimates of the effect of these characteristics. Specifically:

$$\pi_1 = \frac{\psi_1 - \psi_0}{1 - \tau^2}, \tag{B.2}$$

where ψ_0 is the OLS estimator of the parameter ψ in the following model:

$$y_{ig} = \alpha + \psi x_{ig} + \varepsilon_i, \tag{B.3}$$

and ψ_1 is the 2SLS estimator of this model, using the vector of group dummies g as instruments for x_i. The parameter τ^2 is the first-stage R-squared associated with this 2SLS estimate; the variation in x_i explained by the group dummies.

Angrist (2014) argues that because of the relationship in equation B.2, the estimation of peer effects using random allocation to groups can suffer from weak instruments. Furthermore, even with systematic variation in group composition, the 2SLS estimates can exceed the OLS estimates for other reasons unrelated to social effects, such as measurement error. The use of variation across groups to estimate peer effects can confound peer effects with factors unrelated to social influences.

Nevertheless, as pointed out by Feld and Zölitz (2017), Angrist (2014) does not explicitly show under what conditions an upward bias exists and how it depends on the underlying parameters of the model. In fact, under regular conditions, 2SLS estimates with weak instruments are biased towards OLS, which implies that π_1 tends towards zero. Feld and Zölitz (2017) also show that with classical measurement error in the exogenous characteristics x, and with a random group assignment, peer-effects estimates are biased towards zero. In this vein, exploiting random allocation to groups seems to underestimate rather than overestimate peer effects.

Still, the evidence in the literature suggests that estimates of peer effects increase with group size when the variation in peer characteristics weakens. For instance, with an average classroom size of forty-four students (ranging from nineteen to ninety-one), Duflo et al. (2011) find that a one–standard deviation increase in average peer test scores would increase the test score of a student by 0.445 standard deviations, an effect they claim is comparable to previous findings. Similarly, in Carrell et al. (2009), a one-hundred-point increase in peer SAT verbal scores has negligible peer effects on grades when roommates are the relevant peer group (0.003 (s.e. 0.019)) but sizable and significant effects when the peers are other freshmen in the squadron (0.338 (s.e. 0.107)), where group size is larger.[22] Carrell et al. (2013) use the last set of estimates in a posterior experiment that estimates the effect of optimal groups. Contrary to the prediction, they find a negative treatment effect. While the authors attribute this disappointing result to the endogenous patterns of social interactions, Angrist (2014) argues that it might be driven in part by the imprecision of a 2SLS design without a real first stage.

Epple and Romano (2011) reach a similar conclusion with respect to group size. Their handbook chapter concludes that a one-unit increase in peer average ability increases a student's achievement by 0.2 to 0.6 points. Epple and Romano (2011) also consider it surprising that studies that exploit randomization tend to find larger peer effects than those typically found with other identification strategies. For instance, studies using quasi-experimental variation such as Dobbie and Fryer (2014) and Abdulkadiroğlu et al. (2014) find little evidence of peer effects on test scores and college outcomes. Likewise, in the same context of the large peer effects mentioned above, Duflo et al. (2011) find little evidence of peer effects when exploiting an RD on the median student of a tracking system.

This pattern is not limited to classrooms or groups of very large size. Garlick (2018) estimates an impact of 0.216 s.d. when students are assigned to dorms with an average size of 128 students in a South African university. Glaeser et al. (2003) find that the impact on fraternity participation of the average fraction of peers that drink in high school increases with the size of the reference group, even when the differences are small (see Table 1 in Angrist (2014)). While the estimated effect is 0.098 at the dorm level (average size of 2.3 students), it increases by 50% to 0.145 at the floor level (average size of 8.0 students). The impact is even larger at the building level (0.232), where the average group size is 28 students.

Two explanations for this phenomenon can be extracted from equation B.2. The first one is that as groups get larger and the variation in peer characteristics gets weaker, estimates of peer effects become more imprecise. To see this, notice that the variance of the estimator of π_1 in

[22] A squadron comprises approximately 120 students (freshmen through seniors).

equation 1 is given by the following:

$$var(\hat{\pi}_1) = \frac{1}{N_s}\frac{N_g}{N_g - 1}\frac{\sigma^2_{\varepsilon}}{var(\overline{x}_g)} \tag{B.4}$$

where N_s is the sample size, $var(\overline{x}_g) = \frac{\sigma^2_{x}}{N_g}$, and N_g is the group size. The variance of π_1 (equation B.4) is an increasing function in $N_g \geq 2$. Intuitively, as groups get larger, the variation in peer characteristics is lower and hence the precision of the estimate decreases. This argument has been previously explored by Angrist (2014) as an explanation for the estimates in Glaeser et al. (2003) and the differences between Carrell et al. (2009) and Carrell et al. (2013).

The second and less explored explanation of positive peer effects that increase with group size is the amplification of bias when the variation in peer characteristics is weak. This is a similar situation to the one encountered when instruments explain little of the variation in the endogenous variables. A very small violation of the exclusion restriction can lead to a large (asymptotic) bias. Following equation B.2, this would imply that estimates of peer effects grow with group size, as the difference between 2SLS and OLS estimates is increasing.

The probability limit of ψ_1 is:

$$plim\ \psi_1 = \pi_0 + \frac{cov(\varepsilon_{ig}, \overline{x}_g)}{var(\overline{x}_g)}.$$

As groups get larger and the variance of $\overline{x}_g$ decreases, any correlation between the error term and the average peer characteristics is amplified. Notice that this is the case even if the covariance between the error term and $\overline{x}_g$ also decreases with group size but at a lower rate than the decrease in the variance of $\overline{x}_g$. Any model with this feature will amplify the bias with group size much like weak instruments do.

Figure B.1 introduces simulations of the linear-in-means peer effects model (equation 1), illustrating both problems. In particular, the left column in Panel A presents the distribution of the estimates of π_1 in equation 1, assuming that $\pi_1 = 0$. In general, and as expected from equation B.4, estimates become imprecise as the group size increases. However, these losses in precision imply that we should observe both large positive and large negative estimates across studies, which is inconsistent with the empirical evidence.

A second explanation for the increase in peer effects estimates when group size is larger is the amplification of bias. The right plot in Panel A of Figure B.1 illustrates this concern. For this plot, I consider a small nonlinear peer effect in the error term. In particular, all groups with an average score above the median receive a positive shock of 0.1 –a relatively small nonlinearity. As the plot shows, the misspecification of the functional form amplifies the bias of the linear-in-means estimate when the group size increases.[23] While the average bias is only about 0.056 when the groups are pairs, it rapidly grows to 0.082 when the group size is four. The increase in the magnitude of the bias is explosive. With a group size of seven students, the bias is twice as large as the one with two students. A larger group size of twenty-five students quadruples the bias, with an average and median estimate of 0.20.

My experimental design does not lose precision or increase bias with group size. This is because there is substantial variation in peer characteristics by virtue of the treatment arms and because

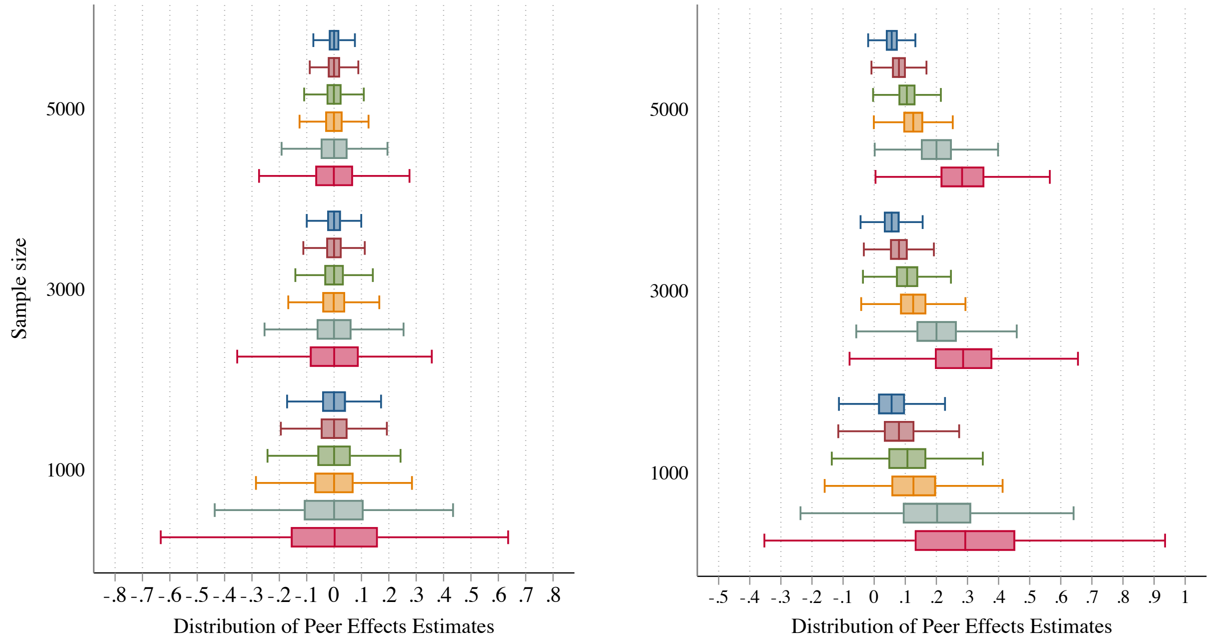
[23] In an individual model, this correlation with the error term would generate a bias of $\hat{\pi}_1$ of 0.05. However, as illustrated by the plot, the bias amplifies with the group size.

the identification strategy does not rely on variation across groups. Parameter π_1 is estimated via a 2SLS model with a single instrument and a strong first stage. Simulations in Panel B of Figure B.1 numerically illustrate how my experimental design preserves precision and limits the bias of estimates with different group sizes. In the left panel, I assume that $\pi_1 = 0$ and plot the corresponding peer-effects estimates. The precision of the estimates remains constant with group size and only varies with the sample size. Similarly, in the right panel, I consider a nonlinear correlation with the error term, but the positive bias remains constant regardless of group size. Both plots sharply contrast with the findings from a random allocation to groups.

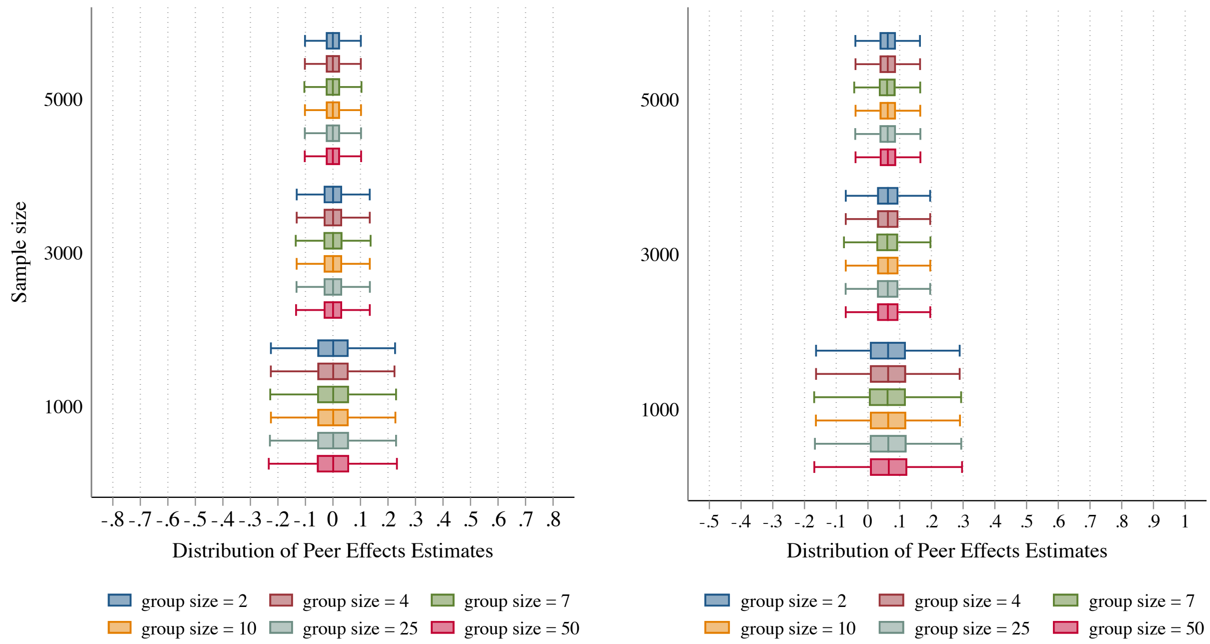
FIGURE B.1: Simulations of Peer Effects

No peer effects | Correlation with the error term

Panel A: Random Allocation to Groups



Panel B: Experimental Design



Notes: Monte Carlo simulations based on 10,000 repetitions of the estimate of parameter π_1 in equation B.3. The simulations assume that $x_i \sim N(0,1)$, and $\nu_{ig} \sim N(0,1)$. In the left column, $\pi_0 = 1$, $\pi_1 = 0$, and $\varepsilon_{ig} = \nu_{ig}$. In the right column, $\pi_0 = 1$, $\pi_1 = 0$, and $\varepsilon_{ig} = 0.1 \times I(\overline{x}_g \geq 0) + \nu_{ig}$.

C Experimental Design: An Application

To understand how the design works in practice, consider an example similar to the one described by Guryan et al. (2009). I use this example to explain why my design guarantees strong variation and is not subject to exclusion bias.

In this example, twelve individuals are randomly assigned to groups of four. To easily navigate through this example, let the individuals have pre-determined levels of an attribute x_i, where $x_i = i$ and i has a range from 1 (the lowest skill level) to 12 (the highest skill level). Following the first step of my research design, individuals 1-6 are classified as low type and individuals 7-12 as high type. Second, these individuals are randomly assigned to the high-type peer treatment:

- Four of the individuals 7 to 12 are randomly assigned to high-type peers and put together in the same group. This is Combination A composed of high-type students assigned to high-type peers.
- The other two high-type individuals are allocated to two low-type individuals (selected individuals in the 1-6 range) who were randomly assigned to the high-peers treatment. This is Combination B, a mixed combination of high- and low-type peers.
- The remaining four low-type individuals from 1 to 6 were assigned to the low-type peer treatment and belong to the same group. This is Combination C, composed of low-type students assigned to low-type peers.

From our simple example above, we have four students in A, four in B, and four in C. As groups are of size four, all students will have three peers in their group.

Let's move now to step 3 of the research design, focusing on the strength of the first stage and the exclusion bias. I will do this for low-type students. There will only be low types in Combinations B and C. For those assigned to the treatment (Combination B), two out of three peers are randomly chosen from individuals 7-12. For a low-type in Combination B, only one out of 3 peers will come from individuals 1 to 6 (low-types too). The expected value of the leave-out mean for low-types who are assigned to the treatment (Combination B) can then be described by:

$$\mathbb{E}\left[\overline{x}_{g,-i}|i, H_i = 1\right] = \underbrace{\frac{2}{3}\left(\frac{\sum_{j=7}^{12} j}{6}\right)}_{\substack{\text{peer variation coming}\\ \text{from being treated}}} + \underbrace{\frac{1}{3}\left(\frac{\sum_{j=1}^{6} j}{5} - \frac{i}{5}\right)}_{\substack{\text{same-type peer}\\ \text{assigned to the same group}}}, \tag{C.1}$$

In Combination C, low-type students are assigned to the control group; all three peers come from individuals 1 to 6. In this case, the expected value of the leave-out mean for low-types who are assigned to the control (Combination C) can be described by:

$$\mathbb{E}\left[\overline{x}_{g,-i}|i, H_i = 0\right] = \underbrace{\frac{2}{3}\left(\frac{\sum_{j=1}^{6} j}{5} - \frac{i}{5}\right)}_{\substack{\text{peer variation coming}\\ \text{from being in the control group}}} + \underbrace{\frac{1}{3}\left(\frac{\sum_{j=1}^{6} j}{5} - \frac{i}{5}\right)}_{\substack{\text{same-type peer}\\ \text{assigned to the same group}}} \tag{C.2}$$

Then, for individual i, the expected value of the difference in peer characteristics, conditional on being treated (equation C.1) and in the control group C.2 is:

$$\lambda_i = \mathbb{E}\left[\overline{x}_{g,-i}|i, H_i = 1\right] - \mathbb{E}\left[\overline{x}_{g-i}|i, H_i = 0\right] \quad = \tfrac{53-2i}{15} \tag{C.3}$$

Note how the second term from expressions C.1 and C.2 cancel out together. Hence, any expected difference between treatment and control groups will arise from the first term (the peers that provide the treatment or the control). We can now apply the law of iterated expectations to calculate the strength of the first stage for low-type individuals (the parameter λ in equation 2:

$$\lambda = \mathbb{E}_i\left[\lambda_i\right] = \frac{53 + 2\mathbb{E}\left[i\right]}{15} = \frac{60}{15} = 4, \tag{C.4}$$

where the expected value of the attribute is $\mathbb{E}\left[i\right] = \frac{7}{2}$, as attributes from low-type individuals range from 1 to 6.

There are two things to notice from equation C.4. First, not only is the difference of 4 driven by the treatment assignment, as previously mentioned, but it is different from the near-zero result we would have found if the variation in peer characteristics were weak. A first stage of 4 is also strong and larger than a one-standard deviation in the distribution of attribute x_i. Second, while expressions C.1 and C.2 are subject to the exclusion bias (term $\frac{i}{5}$), the main source of identification comes from variation across treatment and control groups. As the treatment is uncorrelated with individual attributes, the first stage is in expectation always equal to 4. Under this research design, it is perfectly possible to study peer effects like in any standard 2SLS model described by equations 1 and 2.

If we were to run the same example under a typical design of random assignment to groups, we would still have twelve individuals who are randomly assigned to three groups of four. In this case, however, all groups would have the same expected value of peer attributes, $\overline{x}_g = \frac{1+12}{2} = 6.5$. This illustrates what Angrist (2014) describes as the weak variation problem. The only variation in the expectation across individuals comes from the exclusion bias. In particular, as individual 1 cannot be her own peer, the leave-out mean of the three students she can be paired with ranges from 2 to 12, with an average $\overline{x}_{g,-1} = \frac{2+12}{2} = 7$. On the other extreme, we have individual 12, with a leave-out mean of the three students she can be paired with that ranges from 1 to 11, with an average $\overline{x}_{g,-12} = \frac{1+11}{2} = 6$. The higher the level of attributes, the lower the peer leave-out mean. This negative correlation between individual attributes and the peer leave-out mean signals there is exclusion bias in random assignments to groups.

D Psychological Tests

This section describes in detail the psychological tests that were used to construct the sociability index.

In addition to the Big Five personality traits and the peers' perceptions measures described in section 5.1, the tests used to construct the sociability index are:

D.1 The Big Five

The most widely accepted taxonomy of psychological traits, both in the literature and in my data, is the Big Five (McCrae and John, 1992; John and Srivastava, 1999).[24] The *American Psychology Association Dictionary* defines the Big Five personality traits as follows (Table 1.1 in Almlund et al. (2011)):

1. Conscientiousness: the tendency to be organized, responsible, and hardworking.

2. Openness to Experience: the tendency to be open to new aesthetic, cultural, or intellectual experiences.

3. Extraversion: an orientation of one's interests and energies toward the outer world of people and things rather than the inner world of subjective experience; characterized by positive affect and sociability.

4. Agreeableness: the tendency to act in a cooperative, unselfish manner.

5. Neuroticism or Emotional Stability: Emotional Stability is "predictability and consistency in emotional reactions, with absence of rapid mood changes." Neuroticism is a chronic level of emotional instability and proneness to psychological distress.

Two traits from the Big Five are linked to social skills: extraversion[25] agreeableness[26]. Empirical evidence shows that extraversion is associated with good labor market outcomes (Fletcher, 2013), and that agreeableness influences occupational decisions (Almlund et al., 2011; Cobb-Clark and Tan, 2011). These results are consistent with a study by Deming (2017) that concludes that the labor market increasingly rewards social skills. I also include openness to experience[27] in the index as previous research shows that it is associated with leadership (Nieb and Zacher, 2015; Özbağ, 2016; Javed et al., 2020). In the COAR Network, it is also the trait with the largest predictive power on the number of peers that identify a student as a leader. The results are robust to excluding openness to experience from the index.

D.2 Altruism

The altruism self-reported scale was developed by Rushton et al. (1981). The test used in the COAR network is composed of 17 items. The score on the test is found to predict criteria such as peer ratings of altruism, completing an organ donor card, and paper-and-pencil measures of prosocial orientation (Rushton et al., 1981). More recent evidence shows that the score on the test is related to spontaneous smiles—which is an important signal in the formation and maintenance of cooperative relationships (Mehu et al., 2007). Likewise, there is evidence that the score on the test is related to charity giving but not to blood donor donation behavior (Otto and Bolle, 2011).

[24] Almlund et al. (2011) summarizes the Big Five personality traits and their application to economics. Likewise, Akee et al. (2018); Donato et al. (2017); Kranton and Sanders (2017) provide recent evidence of the Big Five in economics research.

[25] The facets of extraversion are: warmth (friendly), gregariousness (sociable), assertiveness (self-confident), activity (energetic), excitement seeking (adventurous), and positive emotions (enthusiastic).

[26] The facets of agreeableness are: trust (forgiving), straight-forwardness (not demanding), altruism (warm), compliance (not stubborn), modesty (not show-off), tender-mindedness (sympathetic).

[27] Openness involves six facets or dimensions, including active imagination (fantasy), aesthetic sensitivity, attentiveness to inner feelings, preference for variety, and intellectual curiosity.

D.3 Leadership

The leadership scale corresponds to the leader behavior questionnaire developed in Spanish by Castro-Solano (2007). It is based on the theory of Yukl (2013). The scale measures three components of leadership: (1) behaviors guided towards tasks, (2) behaviors guided towards others, and (3) behaviors guided towards changes. In my data, there is a positive correlation between the score on the scale and the number of peers who perceived the student as a leader.

D.4 Empathy

The empathy scale corresponds to the Basic Empathy Scale developed by Jolliffe and Farrington (2006). The scale is composed of two factors: cognitive and emotional empathy. The scale has been validated in other contexts: when applied to adults (Carre et al., 2013) and the Spanish version (Villadangos et al., 2016). It has also been affirmed that students who report higher scores in socially aversive personalities (psychopathy, narcissism, and Machiavellianism) have a low score on the scale (Wai and Tiliopoulos, 2012). Likewise, Gambin and Sharp (2018) show that a low score on the test is associated with guilt and depressive symptoms.

D.5 Intercultural Sensitivity

This 24-item scale of intercultural sensitivity was developed by Chen and Starosta (2000). The authors define intercultural sensitivity as: "*a person's ability to develop a positive emotion towards understanding and appreciating cultural differences that promotes appropriate and effective behavior in intercultural communication.*" The scale comprises two factors: positive and negative reactions to intercultural interactions. Evidence shows that there is a positive correlation between intercultural sensitivity and compassion in nurses (Arli and Bakan, 2018), that American student scores depend on religious affiliation and the number of times they have traveled outside the US (Gordon and Mwavita, 2018), and that Iranian university students have demonstrated a strong relationship between intercultural sensitivity and ethnic background.

D.6 Emotional Intelligence

Emotional intelligence is defined as individuals' ability to recognize their own emotions and those of others, discern between different feelings and label them appropriately, use emotional information to guide thinking and behavior, and manage and/or adjust emotions to adapt to environments or achieve one's goal(s) (Colman, 2009). The emotional intelligence test corresponds to the scale developed by Law et al. (2004). The test comprises 16 items and has four factors: self-emotional appraisal, uses of emotion, regulation of emotion, and others' emotional appraisal.

D.7 The Reading the Mind in the Eyes Test

This test aims to assess how well people can read others' emotions just by looking at pictures of their eyes. It is a multiple-choice test with 36 items. For each item, the respondent has to identify the corresponding emotion expressed in a pair of eyes; four choices are given for each question. According to Deming (2017), this test is a reliable measure of social skills since it relates to social value orientation (Declerck and Bogaert, 2008), a social intelligence factor, and performance in groups (Woolley et al., 2010), and individual teamwork abilities (Weidmann and Deming, 2020).

D.8 Achievement Goals

While not part of the construction of the social skills index, students completed the *Achievement Goal Questionnaire* (J. Elliot and Murayama, 2008). Achievement goals are conceptualized as cognitive–dynamic aims that focus on competence. The test comprises 12 items and has four factors: mastery approach goal items, mastery avoidance goal items, performance-approach goal items, and performance-avoidance goal items. The last two items are related to goals in comparison with peers and are the ones I use as part of self-confidence in academic skills.

E Theoretical Framework for the Role of Beliefs

In this section, I present a simple theoretical framework to understand how the formation of beliefs about abilities can drive peer effects. Overall, there are two mechanisms for self-confidence to improve students' outcomes. First, if ability and effort are complements in the education production function, students with higher confidence will exert more effort (Benabou and Tirole, 2002).

To illustrate this, let's consider the following education production function that depends on effort e_i and ability a_i.

$$y_i = a_i + \theta e_i + \gamma a_i e_i, \tag{E.1}$$

with $\theta > 0$ (effort improves the output) and $\gamma > 0$ (effort and ability are complements). The utility of student i is $u_i = y_i - \frac{c}{2}e_i^2$, where c parametrizes the marginal cost of effort. The optimal effort level of the student would be given by: $e_i^* = \frac{\theta + \gamma a_i}{c}$. When students have imperfect information then students take expectation over the ability distribution, such that:

$$e_i^* = \frac{\theta + \gamma \mathbb{E}\left[a_i\right]}{c}.$$

Hence, two students with the same level of ability (a_i) but different beliefs ($\mathbb{E}\left[a_i\right]$) would have different outcomes. By having higher self-confidence, students are incentivized to exert more effort, and this can improve their performance.

The second mechanism for self-confidence to affect performance is a direct one. Compte and Postlewaite (2004) introduce a model that explains how a person's psychological state can affect performance. In their model, the probability of success depends on a person's level of confidence, captured by her perception of success in previous cases. For example, a student who is more confident about her chances of making friends is more likely to make these friendships, and a student who is more confident in her math skills would have a higher score on a test.

A simple way of introducing the direct effect into the education production function is by including a parameter of self-confidence, $\kappa(\cdot)$ in equation E.1:

$$y_i = \kappa\left(\mathbb{E}\left[a_i\right]\right)\left(a_i + \theta e_i\right), \tag{E.2}$$

with $0 \leq \kappa(\cdot) \leq 1$, and $\kappa'(\cdot) > 0$. Notice that in equation *E.2*, I set $\gamma = 0$. The idea behind this production function is that even without complementarity between effort and ability, higher self-confidence can increase output.