Peer skill identification and social class: Evidence from a referral field experiment*

Jhon Díaz[†], Manuel Munoz[‡], Ernesto Reuben[§‡], Reha Tuncer[¶]

March 20, 2025

Abstract

Cognitive and social skills are both increasingly valued in the labor market, but social skills are difficult to observe. In the absence of observable signals, peer assessments can be valuable screening tools. We study how well individuals identify productive peers across cognitive and social skills in a lab-in-the-field experiment with 849 university students. After students interact for an entire term, we collect incentivized skill measures from all classmates. We then ask for referrals of the highest scoring peers in each skill, incentivizing referrals based on the nominee's score. To examine potential social class barriers in referrals, we randomly assign half of the participants to receive additional incentives for identifying high-skilled peers from low-socioeconomic status. We find that peers can successfully identify cognitive skills but not social skills of their classmates. There is only evidence of a bias against low-SES peers in unique cognitive skill referrals,

*We obtained Institutional Review Board approvals from NYU Abu Dhabi (HRPP 2024-50) and the University of Luxembourg (ERP 24–028). The study design was preregistered in the OSF Registries prior to data collection (see https://doi.org/10.17605/OSF.IO/V9T3W).

[†]Universidad Autónoma de Bucaramanga

[‡]Luxembourg Institute of Socio-Economic Research

[§]Division of Social Science, New York University Abu Dhabi

[¶]University of Luxembourg

and the treatment incentives helps mitigate it. Our findings suggest that the accuracy of peer assessments varies substantially across skill dimensions and appropriate changes in the incentivization structure can make peer assessments robust to existing biases.

JEL Classification: C93, D03, D83, J24

Keywords: network homophily, labor market, performance evaluation, hiring screening, human capital, incentive mechanisms, workplace diversity, academic performance, socioeconomic barriers, information asymmetry

1 Introduction

Evaluating the productivity of others is a standard feature of the labor market. Employers assess job candidates, managers evaluate workers for promotion, and team leaders select collaborators based on beliefs about others' capacity to perform well in different tasks. Whenever observable productivity signals such as test scores or past experience are available, decision-makers rely on those to make accurate evaluations. But such signals are scarce for tasks that are interpersonal in their nature and difficult to quantify. In these settings, peer assessments akin to referrals can be a particularly strong screening tool which combines cost-efficiency and accuracy, as sustained interactions among people who work together provide opportunities to directly observe each other's productive qualities in various domains.

However, identifying productive peers across a multitude of productivity dimensions is not straightforward. First, peers could accurately assess productivity in one dimension, but they may struggle to evaluate it in another because of its harder to observe nature. Cognitive and social (interpersonal) skills are two such dimensions of human capital that are increasingly rewarded in the labor market (Deming, 2017, 2023). Second, biases in productivity beliefs can lead to systematic deviations in assessment accuracy. The case for low-socioeconomic status (low-SES) individuals is particularly concerning, as peers may systematically underestimate their abilities due to stereotypes or lack of information. Such biased assessments could contribute to their worse labor market outcomes despite having the necessary skills (Stansbury & Rodriguez, 2024).

The overall purpose of this paper is twofold: To evaluate how accurately peers identify productive others in cognitive and social skills, and whether disadvantaged low-SES individuals face barriers in selection when peers assess productivity across these skills.

We conducted a lab-in-the-field experiment in a Colombian university to answer these questions. After interacting for an entire term (about 4 months) in small classrooms (average 26 students per class), we collected incentivized cognitive and social skill measures from all participants to obtain objective productivity distributions. Participants then assessed classmates' productivity across these dimensions by making referrals, allowing us to compare referred peers to those who were not. We incentivized referrals by bonuses contingent on the nominee's score in the skill measures. Nominees did not receive any benefit from being referred. Both features allowed us to rule out concerns of potential social transfers (i.e., nepotism or favoritism) and reputational costs typical in the referral literature (see for example Bandiera, Barankay, and Rasul (2009); Witte (2021)). Once we abstracted away from these elements, the referral decision became one of measuring productivity beliefs through nominated candidates.

Even in an incentivized setting like ours, biases about low-SES individuals could be at play because of the underlying beliefs classmates hold about their productivity. To address this we designed two treatments. In the **Baseline** treatment, we gave pure performance incentives to referrals regardless of social class. Participants in the **Quota** treatment received additional incentives to identify high-skilled low-SES peers. To be able to make comparisons within the same referral choice sets, we assigned half of the participants within each classroom to either treatment. This setup allows us to assess how well incentives mitigate the said biases in peer productivity beliefs across the different referral behaviors that we observe.

Our first goal is understanding how well peers identify cognitive and social skills of their classmates under pure performance incentives at **Baseline**. We find that peers have distinct screening abilities for skills, and use different types of referral strategies because of it. Specifically, peers successfully identify cognitive skill but not social skill of their classmates. They also frequently refer the same peers for both skills, at rates much higher than the actual overlap between those who are productive at both cognitive and social skill. For this reason we separately analyzed the three referral types: Those made in common for both skills, and those made uniquely for cognitive or social skill. Common referrals for both skills identified classmates with higher grades but not higher skills. This suggests an observable proxy such as academic performance influences peer productivity assessments in the absence of credible skill information. For unique cognitive skill referrals, both grades and measured cognitive skill are equally good predictors. Unique social skill referrals are not predicted by either academic performance or social skill, suggesting that social skills might be less observable in classroom settings or require different measures to evaluate accurately. These findings reveal a nuanced picture of how peer assessments of productivity may depend on how discernible the skill in question is, and how they can be influenced by the availability of other observable proxies for productivity.

We find limited support for a bias affecting low-SES individuals. Of the three referral types, we find bias only in unique cognitive skill referrals when accounting for peer skills. This characterizes the decisions of about 75% of participants who made at least one unique cognitive skill referral, and about half of all cognitive skill referrals overall. The **Quota** treatment mitigates the bias for this subset of referrals, while not changing the referral rates of low-SES individuals for the rest of the referral strategies that were not biased in the first place. There is also no meaningful efficiency-equity tradeoff affecting productivity of peers referred in the **Quota** treatment. Our findings show peer productivity assessments are robust to salient differences between social classes, and provide evidence that existing biases can be remedied with changes in the incentivization structure without compirimising productivity.

Our paper contributes to various strands of the literature. First, we contribute the literature on referral experiments that strives to understand how referrals help screeening for productive workers. Past work provides causal evidence that peer productivity assessments using referrals bring in productive workers (Pallais & Sands, 2016), and that performance-contingent incentives lead to improvements in the productivity of referred candidates (Beaman, Keleher, & Magruder, 2018; Beaman & Magruder, 2012). These studies allow referrals to be made from different candidate pools where referrers are free to nominate any candidate, and as a result confound screening ability with advantages arising from access to different candidate pools (Montgomery, 1991). We implement common choice sets for referrals which allow us to isolate peers' true screening ability and enable straightforward comparison between experimental treatments in terms of referral choice sets. Our paper complements the literature on referral experiments by providing causal evidence that peers have skill-dependent screening abilities that go beyond the differences in candidate pools under performance-contingent incentives.

Second, we contribute to the growing body of work on the relevance of noncognitive skills in the labor market. This literature examines dimensions of human capital such as patience, self-control, conscientiousness, teamwork, and critical thinking that contribute positively to labor market returns (Heckman & Kautz, 2012; Heckman, Stixrud, & Urzua, 2006; Lindqvist & Vestman, 2011; Weinberger, 2014). Among these, interpersonal skills are exceptionally relevant for labor market gains in the last two decades as a complement to cognitive skill (Deming, 2017, 2023). Yet, hiring firms report difficulties in assessing social skills in candidates, and applicants are willing to pay substantial sums to convey social skill feedback to employers (Bassi & Nansamba, 2022). We contribute to this literature with our peer productivity assessments across two dimensions of skills, and show that peers can identify cognitive skill but struggle to assess social skills. Our results suggest that referrals may be ineffective for screening attributes that are less visible or harder to proxy through standard productivity measures in the assessment environment.

Finally, we contribute to the literature on diversity considerations in referrals. Homophily[1] in referrals drives correlations among social groups' employment and wages (Calvo-Armengol & Jackson, 2004; Calvó-Armengol & Jackson, 2007), as individuals are

[1] A well-documented empirical consistency in sociology where individuals form ties more often with others who are similar to themselves across observable characteristics (McPherson, Smith-Lovin, & Brashears, 2006; McPherson, Smith-Lovin, & Cook, 2001).

more often tied to others with comparable socioeconomic status (Chetty et al., 2022b). Limited interaction across social classes due to spatial segregation is shown to drive at least some of the differences (Chetty et al., 2022a). In this context, efficiency of diversity treatments in endogenous networks may be constrained by availability. To counter this, we consider a socially diverse university setting where we use exogenously imposed networks, and required participants to refer among classmates. Anticipating differences in referral outcomes for low-SES individuals even when networks across social classes overlap by design, we introduced quota-like incentives as a treatment arm to increase referrals to low-SES peers.[2] Our findings complement the literature on biases in referrals (Beugnot & Peterlé, 2020; Hederos, Sandberg, Kvissberg, & Polano, 2024) by first showing the existence of a social class bias and then providing the causal evidence for targeted incentives that effectively reduce the bias in our setting without compirimising productivity.

The remainder of the paper is organized as follows. Section 2 begins with the background and setting in Colombia. In Section 3 we present the design of the experiment, including the skill assessment, referral and guessing tasks. In Section 4 we describe the data and procedures. Section 5 discusses the results of the experiment. Section 6 concludes. The Appendix presents additional tables and figures as well as the experiment instructions.

2 Background and Setting

Our study takes place at UNAB, a medium-sized private university in Bucaramanga, Colombia with approximately 6,000 enrolled students. The university's student body is remarkably diverse with slightly more than half of the students classified as low-SES. This diversity provides a unique research setting, as Colombian society is highly unequal and generally characterized by limited interaction between social classes, with different

[2]We design the treatment incentives in inspiration from the success of gender quotas in the affirmative action literature (e.g., Balafoutas and Sutter (2012); Bertrand, Black, Jensen, and Lleras-Muney (2019); Niederle, Segal, and Vesterlund (2013)).

socioeconomic groups separated by education and geographic residence.[3] Despite significant financial barriers, many lower middle-class families prioritize university education for their children (Hudson & Library of Congress, 2010, p. 103), with UNAB representing one of the few environments where sustained inter-class contact occurs naturally.

In 1994, Colombia introduced a nationwide classification system dividing the population into 6 strata based on housing characteristics and neighborhood amenities.[4] We use this exogenous cutoff as the measure of social class in our experiment: Students in strata 1 to 3 are categorized as low-SES, and those in strata 4 to 6 as high-SES (see Appendix Figure A.1 for a detailed stratum distribution of our sample).

We invite all students enrolled in two compulsory courses to participate in our experiment. Throughout the term, students meet weekly for three-hour sessions where attendance is mandatory. Both courses are university-wide graduation requirements which result in large variations in academic programs (see Appendix Table A.3) and socioeconomic backgrounds across the classrooms. This setup provides a unique opportunity for collaborative inter-class contact on equal status, whose positive effects on reducing discrimination are casually documented (Lowe, 2021; Mousa, 2020; Rao, 2019).

3 Design

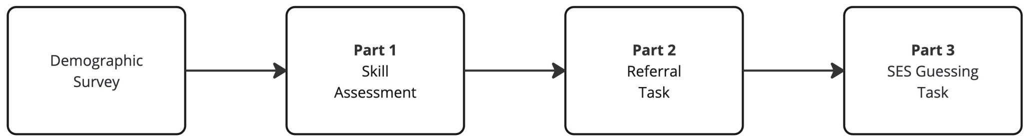
We designed an experiment to assess the peer screening ability for different skills and to measure biases related to social class. The study design consists of a single experiment with sessions organized at the classroom level (see Figure 1). The instructions are

[3]Colombia has consistently ranked as one of the most unequal countries in Latin America (World Bank, 2024), with the richest decile earning 50 times more than the poorest decile (United Nations, 2023). This economic disparity is reflected by a highly stratified society with significant class inequalities and limited class mobility (Angulo, Gaviria, Páez, & Azevedo, 2012; García, Rodríguez, Sánchez, & Bedoya, 2015).

[4]Initially designed for utility subsidies from higher strata (5 and 6) to support lower strata (1 to 3), it now extends to university fees and social program eligibility. Stratum 4 neither receives subsidies nor pays extra taxes. This stratification system largely aligns with and potentially reinforces existing social class divisions (Guevara S & Shields, 2019; Uribe-Mallarino, 2008).

provided in Appendix B.

Figure 1: Experiment Timeline



Note: Participants first complete incentivized skill tests, then refer classmates for skills. In the final part, they guess the social class of their peers. This order is implemented in all sessions.

3.1 Skill Assessment

To understand the basis for referral decisions, we collect objective measures of cognitive and social skills. These two distinct skills are crucial for the labor market and suitable to assess given classmates interact through the term. By measuring skills before the referral stage, we eliminated the need for referred students to take additional action. Participants perform two incentivized skill tests. They have 5 minutes to complete each test. We provide test-specific instructions and an example item before participants begin. Correctly solved items increase chances to earn a fixed bonus.[5]

We use Raven's Progressive Matrices to measure cognitive skills (Raven, 1936; Raven, Raven, & Court, 1976). Raven's test is a well-established measure of fluid intelligence, i.e., an individual's capacity to reason and solve problems in novel situations independent of past knowledge (Schilbach, Schofield, & Mullainathan, 2016). In this test, participants see series of images where there is a pattern with a piece that has been intentionally removed. They are tasked with choosing the piece that completes the pattern among available options. For each image, there is only one correct answer. We implement an

[5] The tests are presented in a randomized order. No performance feedback is provided. Participants see one item at a time and cannot return to previous screens once they start a test. They are not required to answer items and can skip them if they choose to do so. We elicit beliefs about performance after each test.

18-item version featuring increasingly difficult questions, with 6 response options for the first 9 items and 8 thereafter.

We measure social skills with the Multiracial Reading the Mind in the Eyes Test (MRMET) from Kim et al. (2022).[6] The test is an established measure for the ability to recognize emotions in others, and it has been previously used in economic experiments (van Leeuwen et al., 2018; Weidmann & Deming, 2021; Zárate, 2023). MRMET tends to correlate with fluid intelligence as measured by Raven's (Alan & Kubilay, 2025). It consists of photos of human faces portraying different emotions, cropped so that only the eye region is visible. Participants must choose the emotion that best describes the photo from the available answers. For each photo, there is only one correct answer and 4 response options. We administer the first 36 items in MRMET.

3.2 Referral Task

After the skill assessment, we create the referral task to screen for high skilled peers. For each skill, participants make incentivized referrals by nominating classmates. We first explain the measured skill accompanied by an example test item. We them provide an alphabetically ordered list of all classmates. Participants make three referral choices per skill. They are instructed to exclude themselves from referrals. A classmate may be nominated once per triad. The order in which participants refer for a skill test is randomized. We incentivize referrals with classroom-level performance rankings. The three highest-scoring classmates are designated as the top 3 for a skill. Referrers are eligible for a fixed bonus for referrals among the top 3.[7]

We have two between subject treatments that varies the top 3 selection. In the **Baseline** treatment, the top 3 selection is based solely on performance ranking, regardless of other participant characteristics. The **Quota** treatment modifies the top 3 selection

[6] We choose MRMET because it is a race- and gender-inclusive test suitable for application in non-WEIRD (Western, Educated, Industrial, Rich, Democratic) populations like the one we sample from. The test is based on the original RMET (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001).

[7] We solve ties among the top 3 randomly. We describe only the top 3 selection mechanism and provide no feedback about the top 3 composition to participants.

to prioritize low-SES individuals. We reserve the first spot in the top 3 for the highest-scoring low-SES peer, and assign the remaining two places based on performance (see Table 1). This guarantees at least one low-SES participant in the top 3 per skill. Participants are informed about the top 3 selection mechanism before making referral choices (Appendix Figure B.1 provides illustrations explaining the treatments). Assignment to the treatment is at the individual level within each classroom. This allows comparing the effect of the treatment while keeping the referral choice set constant.

Table 1: Places in the Top 3 according to composition rule

	Baseline	**Quota**
Merit-only	3	2
Reserved for low-SES	0	1

3.3 Socioeconomic Status Guessing Task

Participants make guesses about the anticipated SES of their classmates. We inform participants that a computer algorithm randomly selects three students belonging to strata 1, 2, or 3. They are tasked with nominating the people they believe the computer could choose at random (Appendix Figure B.2 provides the illustration explaining the task). Participants select three classmates from an alphabetically ordered list containing all their classmates. This task measures the ability to distinguish SES independent of test performance, as SES identification is relevant to our study.

4 Sample, Incentives, and Procedure

We invited 849 UNAB undergraduate students to participate in the experiment. Our final sample consists of 702 individuals who completed the study, resulting in an 83% participation rate.[8] We block randomized participants into treatments balancing gender

[8] The missing students did not come to class on the day of the experiment.

and social class. Table 2 presents key demographic characteristics and academic performance indicators across treatments (Appendix Table A.1 illustrates the selection into the experiment). The sample is well-balanced between the **Baseline** and **Quota** conditions and we observe no statistically significant differences in any of the reported variables (all p values > 0.1). Our sample is characterized by a majority of low-SES students with about one-third of the sample being first-generation college students. The gender distribution is balanced. The mean GPA of 3.95 is consistent across both treatments.

Table 2: Balance between treatment conditions

	Baseline	**Quota**	***p***
Low-SES	59%	55%	0.297
Female	52%	47%	0.195
Cognitive score (Raven's)	10.04	10.27	0.322
Social score (MRMET)	18.45	18.50	0.886
GPA	3.95	3.95	0.828
Entry exam score	61.85	62.17	0.638
Age	19.33	19.02	0.228
First generation	34%	37%	0.386
Ethnic minority	1%	3%	0.133
Rural community	30%	27%	0.308
Scholarship	1%	1%	0.916
# semesters at UNAB	3.18	3.17	0.916
N	368	334	702

Note: Low-SES indicates strata 1, 2, or 3. Cognitive score measures Raven's performance out of 18 questions. Social score reflects MRMET performance out of 36 questions. GPA indicates average grades out of 5. Entry exam represents the average score across reading, math, social sciences, and science components of Colombia's standardized university entrance exam ICFES. First generation indicates neither parent attended university. Rural community denotes residence in a non-urban area. p-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample t-tests with equal variances. All reported p-values are two-tailed.

Participants could earn bonuses worth 100,000 Pesos (about 26 US Dollars) in each part of the experiment. In the first part, we incentivized performance in the skill tests. 20% of participants were eligible for the bonus. We randomly picked one skill test for each eligible participant and drew a number between 1 and 100. The participant received the bonus if the percentage of correct answers in the selected test exceeded the drawn number. Chances of earning the bonus increased with each correctly solved question by 5.5% (=1/18) for the Cognitive Skill test and by 2.78% (=1/36) for the Social Skill test.

In the second part, we incentivized referrals among the top 3 performers. 40% of participants were eligible for the bonus. We randomly selected one skill test and one referral for each eligible participant. The participant received the bonus if their referral was among the top 3. In the third part, we incentivized the correct identification of low-SES peers. 20% of participants in each classroom were eligible for the bonus. We randomly selected one guess for each eligible participant. The participant received the bonus if their guess correctly identified a low-SES peer. Draws for the bonuses were independent meaning participants could earn multiple bonuses.

Data collection occurred during the last two weeks of April 2024. Our local partner at UNAB coordinated scheduled classroom visits and recruited research assistants to administer the experiment. Students present in class on the scheduled visit dates participated. Each classroom visit constituted a separate session. There were in total 35 sessions.[9] Participants accessed the Qualtrics-based experiment using their smartphones during these visits. The median time to complete the survey was 20 minutes, with a compensation of $26 for 117 lottery winners.

5 Results

5.1 Can peers screen cognitive and social skills?

Our first goal is understanding whether higher skilled individuals get more referrals. Because every referrer nominates 3 classmates per skill, analyzing only the extensive margin, i.e., whether an individual gets a referral, is not very informative.[10] We consider the percentage share of referrals from individuals in **Baseline** condition as our dependent variable. This approach combines the intensive and extensive margins and also makes

[9] See Appendix Figures A.2a, A.2b and A.2c for the distribution of skills and GPA across classrooms and Appendix Table A.3 for diversity in program choices.

[10] Only 86 of the 849 students (10%) never get a referral for either skill.

comparisons across classrooms with different sizes easier.[11]

Formally, we define the percentage share of referrals received by individual i from participants j in classroom c and in **Baseline** condition ($\forall j \in B_c$) for skill s as:

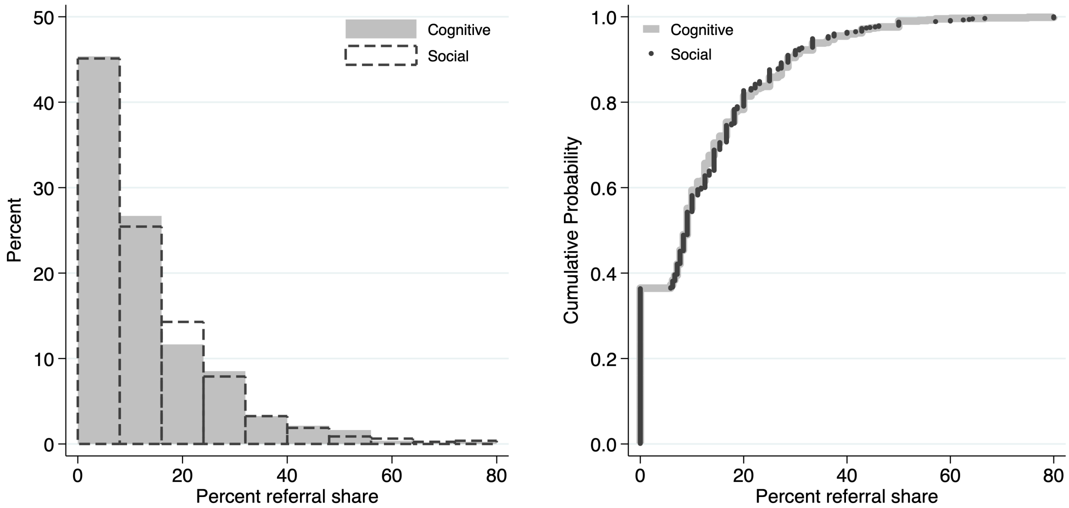
$$y_{ic}^{s} = \frac{\sum_{j \neq i} r_{ijc}^{s}}{n_c - \mathbb{1}(i \in B_c)} \times 100 \tag{1}$$

where n_c represents the number of participants in the **Baseline** condition in classroom c. The indicator r_{ijc}^{s} takes value 1 if participant j in the **Baseline** condition refers individual i for skill s, and 0 otherwise, and require both i and j to be in the same classroom c. The denominator $n_c - \mathbb{1}(i \in B_c)$ accounts for the maximum possible referrals that individual i could receive. If i is in the **Baseline** condition ($\mathbb{1}(i \in B_c) = 1$), we subtract one from n_c to account for the self-referral restriction.[12] This normalized measure represents the percentage of potential referrals actually received by each individual, adjusting for classroom size and treatment status. By construction, $y_i^s \in [0, 100]$ for all c, and we can compare referrals across classrooms of different sizes. Figures 2a and 2b present the distribution of our dependent variable.

[11]The number of participants in a classroom mechanically drives the number of total referrals that could be received by an individual. By normalizing referrals we focus on differences within classrooms.

[12]33.8 percent of participants in the sample for cognitive and social skills self-referred, while explicitly instructed not to do so. In Appendix Table A.4 we compare the outcomes of those who self-refer. Self-referrers are more likely to be low-SES, and have significantly lower cognitive skill (0.2 SD) and GPA (0.25 SD). We rule out the hypothesis that self-referrers nominate themselves strategically. As self-referrerals are not informative and add noise to our estimates, we drop these instances from our paired referral-referrer sample in subsequent analyses. Self-referrering participants' remaining referral choices are kept in the dataset.

Figure 2: Distribution of referrals by skill in Baseline



(a) Frequency Histogram

(b) ECDF

Note: Figures show the percantage of referrals recieved from participants in the **Baseline** condition for cognitive and social skills. The left panel shows the frequency histogram and the right panel shows the empirical cumulative distribution function (ECDF). A two-sample Kolmogorov-Smirnov test shows no statistically significant difference between the share of referrals received across the skill distributions ($D = 0.0363$, $p = 0.668$).

Under performance pay in the **Baseline** condition, classmates with higher scores in the skill tests should collect more referrals if classmates can screen skills. Our independent variables are the standardized skill test scores. We estimate referral percentage shares y_i^s:

$$y_i^s = \alpha^s + \beta_1^s Score_i^s + \epsilon_i^s \tag{2}$$

Table 3 illustrates our first findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the point estimates for different test scores. In column (2), a one standard deviation increase in cognitive skill score causes a 1.5 percentage point increase in the share of referrals received. On a base rate of 13%, this is a modest increase of 11.5 percent. In column (4), 95% confidence intervals rule out that a one standard deviation increase in the social skill score results in more than a 0.1 percentage point difference in the share of referrals received.

Result 1 *Participants have difficulties screening skills in the* ***Baseline*** *condition, with modest screening ability for cognitive and no screening ability at all of social skill test scores.*

Table 3: Share of referrals received conditional on skill test score

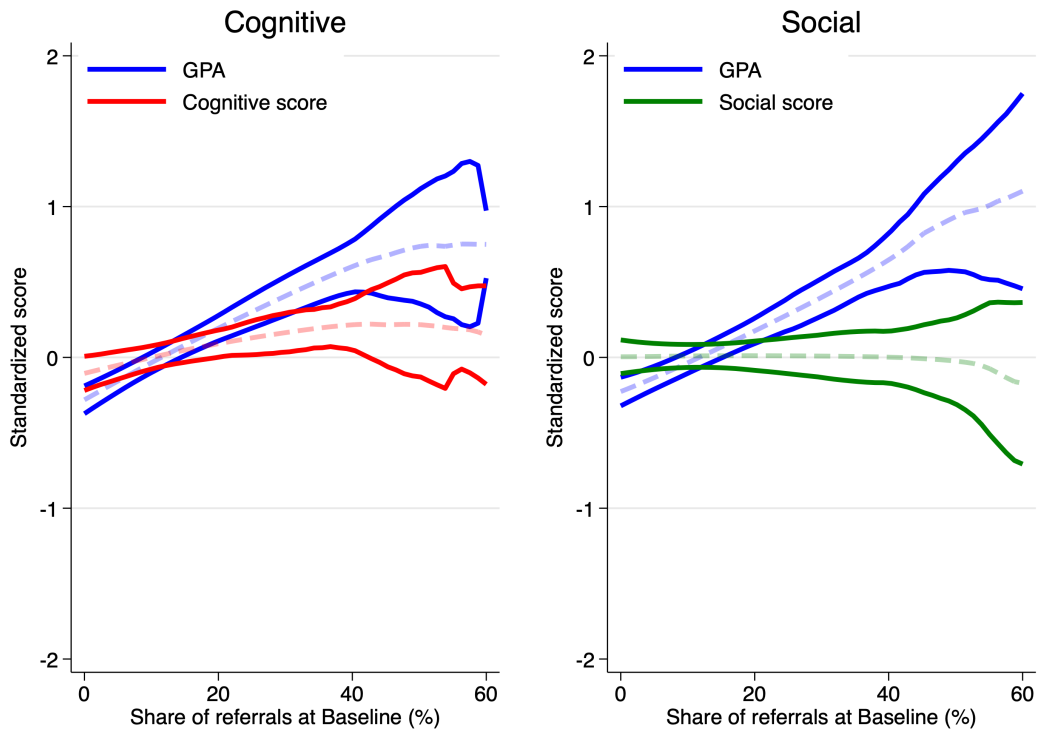
	Cognitive		Social	
	(1)	(2)	(3)	(4)
Score	1.197**	1.497***	0.037	-0.080
	(0.479)	(0.464)	(0.474)	(0.461)
Dep. var. mean	12.986	12.981	13.049	13.050
Classroom FE	No	Yes	No	Yes
R^2	0.008	0.116	0.000	0.100
Observations	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables are the percentage of referrals received relative to all referrals. "Score" refers to standardized test scores for cognitive and social skills. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

5.2 Grades as a proxy for skills

Absence of a clean skill-signal or the lacking the screening ability for skills may have pushed partipants to refer classmates using proxies of skills. Proxies are peer beliefs about strong correlates for skills. A potential proxy for cognitive skill (i.e., "smart students") would be the "students with good grades" in the classroom, as measured by GPA. Figure 3 illustrates the relationship between grades, skill score, and the share of referrals received.

Figure 3: Referral shares by GPA and skill test scores



Note: The left panel shows how GPA and cognitive skill scores vary with the share of cognitive skill referrals received, while the right panel shows the same for GPA and social skill score for the share of social skill referrals received. Solid lines indicate 95% confidence intervals and dashed lines indicate the means. Output is truncated at 60 percent of referral share for the sake of having meaningful confidence intervals.

The idea that grades signal cognitive skill is a common belief among researchers and practitioners alike. Yet, cognitive skill and grades are far from perfectly correlated (Heckman & Kautz, 2012; Heckman et al., 2006), and screening with such beliefs may not lead to good referrals. Indeed, GPA correlates very weakly with skill test scores in our sample (see Appendix Table A.2). We capture the screening behavior using proxies by including the standardized GPA of referrals as an independent variable. We reestimate referral percentage shares for the **Baseline** condition:

$$y_i^s = \alpha^s + \beta_1^s Skill_i^s + \beta_2^s GPA_i^s + \epsilon_i^s \tag{3}$$

Table 4 illustrates our findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the difference between point estimates for skill test scores and GPA. In column (2), a one standard deviation increase in cognitive skill score causes a 1.1 percentage point increase in the share of referrals received when controlling for GPA. On a base rate of 12.8%, this is a comparable increase in magnitude of about 8.6 percent to our previous estimate in Table 3, and suggests cognitive skills have an independent effect on referrals. However, a one standard deviation increase in GPA causes a substantial 4.4 percentage point increase in the share of referrals received when controlling for cognitive skill score. This is an increase of four times in terms of magnitude (34 percent) when compared to cognitive skill, and suggestive of the extent to which academic performance is easier to screen among peers in our setting.

In column (4), 95% confidence intervals rule out that a one standard deviation increase in the social skill score results in more than a 0.5 percentage point difference in the share of referrals received. This is consistent with our previous estimate confirming participants cannot screen social skill scores. On the other hand, a one standard deviation increase in GPA causes a substantial 3.8 percentage point increase in the share of referrals received when controlling for social skill. This is a 30 percent increase in the share of referrals when including controls for social skill.

Result 2 *For both skills, we find strong evidence that grades act as a proxy for referral decisions.*

Table 4: Share of referrals received conditional on skill test score and academic performance

	Cognitive		Social	
	(1)	(2)	(3)	(4)
Score	0.873*	1.080**	-0.278	-0.527
	(0.467)	(0.455)	(0.460)	(0.409)
GPA	3.949***	4.364***	3.429***	3.789***
	(0.664)	(0.684)	(0.581)	(0.651)
Dep. var. mean	12.806	12.783	12.891	12.876
Classroom FE	No	Yes	No	Yes
R^2	0.095	0.204	0.064	0.165
Observations	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables are the percentage of referrals received relative to all referrals. "Score" refers to standardized test scores for cognitive and social skills. GPA is standardized to mean zero and unit variance. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

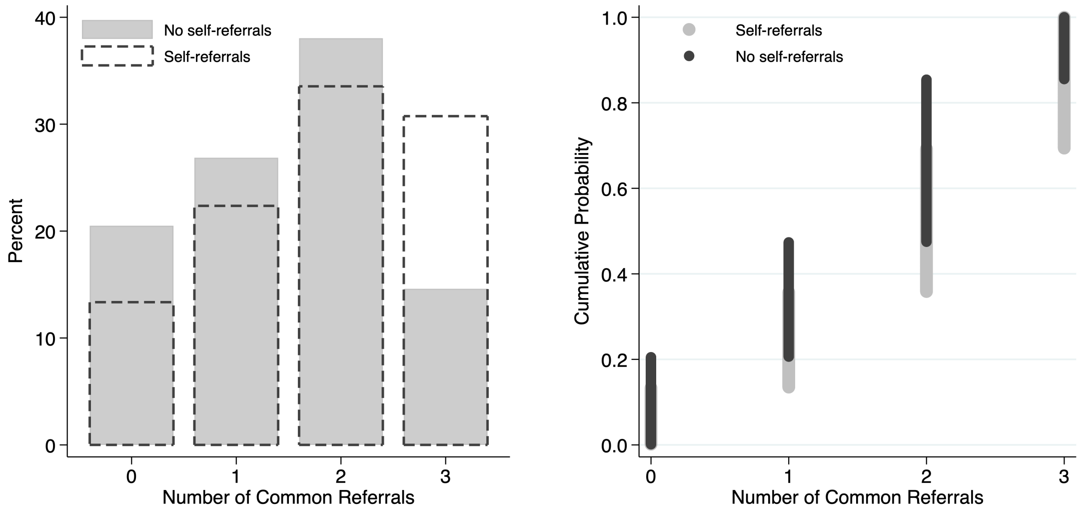
5.3 Types of Referrals

In this section, we expand on the diversity in referral choices to differentiate between referrers using GPA proxy and others. Despite having the opportunity to nominate up to six different classmates across two skills, referrals choices were highly concentrated. The median participant nominated two classmates in common, effectively using four of their six referral slots for the same individuals. Considering self-referrals which illustrate participants' original choices,[13] the majority of participants nominated two classmates in common for both skills, and picked themselves or someone else with almost equal probability. We visualize referral concentration by plotting the number of common

[13]Self-referrals were not valid and are excluded from the main analyses.

referrals made across skills in Figures 4a and 4b.[14]

Figure 4: Common referrals between skills at Baseline



(a) Frequency Histogram

(b) ECDF

Note: Figures show the distribution of common referrals with and without self-referrals. The first bar (value of 0) indicates the share of participants with 6 unique referrals. The last bar (value of 3) indicates the share of participants with 3 identical referral choices across both skills.

With such a large share common referrals across skills, it is possible that participants believed classmates with a higher score in one skill would also have a higher score in the other. Would such beliefs be accurate? There is modest ($\rho = 0.267$) correlation between the two skill test scores (see Appendix Table A.2). To understand whether making common referrals is strategic, we turn to the incentives. Participants were incentivized to pick the top 3 performers for each skill to earn a fixed bonus. Looking at the characteristics of top skilled participants in Appendix Table A.6, we find that conditional on

[14] In Appendix Table A.5 we compare the characteristics of referrers who make unique referrals to those who made at least one common referral. Results suggest minimal differences in GPA, skills, and social class.

being among the top 3 for one of the skills, only 1 in 3 participants were in the top 3 for the other skill too. This suggests *ex-post* making more than 1 common referral across skills would decrease the chances to win the bonus.

A competing explanation for the amount of common referral choices between skills coupled with the notable difficulties in screening for skills would be that individuals who refer classmates twice for both skills are worse at screening. This imples the underlying heterogeneity in skill identification results in differential referral strategies where participants with a good signal for a skill choose to refer classmates only once for that skill, and those without a good signal use the grades proxy and refer classmates for both skills. We can test both hypotheses in our data: If "common" referrers -defined as those who refer an individual for both skills- are better at screening at least one of the skills, point estimates for skills in common referrals would be larger than those made uniquely for a skill. This would give credence to beliefs about correlated skills. On the other hand, if common referrers are worse in skill identification compared to unique-referrers and use GPA proxy for referrals, we can infer that they have no additional information about skills.

We compare the outcomes of participants who recieve common referrals from their classmates to those who recieve unique referrals per skill. Formally, let indicator r_{ijc}^{common} take value 1 if individual j referred individual i for both skills. The percentage share of referrals received by individual i from participants in classroom c and in **Baseline** condition ($\forall j \in B_c$) is:

$$y_{ic}^{common} = \frac{\sum_{j \neq i} r_{ijc}^{common}}{n_c - \mathbb{1}(i \in B_c)} \times 100 \tag{4}$$

where n_c represents the number of participants in the **Baseline** condition in classroom c. The indicator r_{ijc}^{common} takes value 0 if participant j in the **Baseline** condition does not refer individual i for both skills. The denominator $n_c - \mathbb{1}(i \in B_c)$ accounts for the maximum possible "common" referrals that individual i could receive as before. Similarly, let $r_{ijc}^{s,unique}$ take value 1 if individual j referred individual i only for skill s. The percentage share of "unique" referrals received by individual i from participants in classroom c and in **Baseline** condition ($\forall j \in B_c$) for skill s is:

$$y_{ic}^{s,unique} = \frac{\sum_{j \neq i} r_{ijc}^{s,unique}}{n_c - \mathbb{1}(i \in B_c)} \times 100 \tag{5}$$

and it follows that for any s, percentage share of "unique" and "common" referrals received by individual i from participants in classroom c and in **Baseline** condition ($\forall j \in B_c$) must add up to the total share of referrals received:

$$y_{ic}^{s} = y_{ic}^{s,unique} + y_{ic}^{common} \tag{6}$$

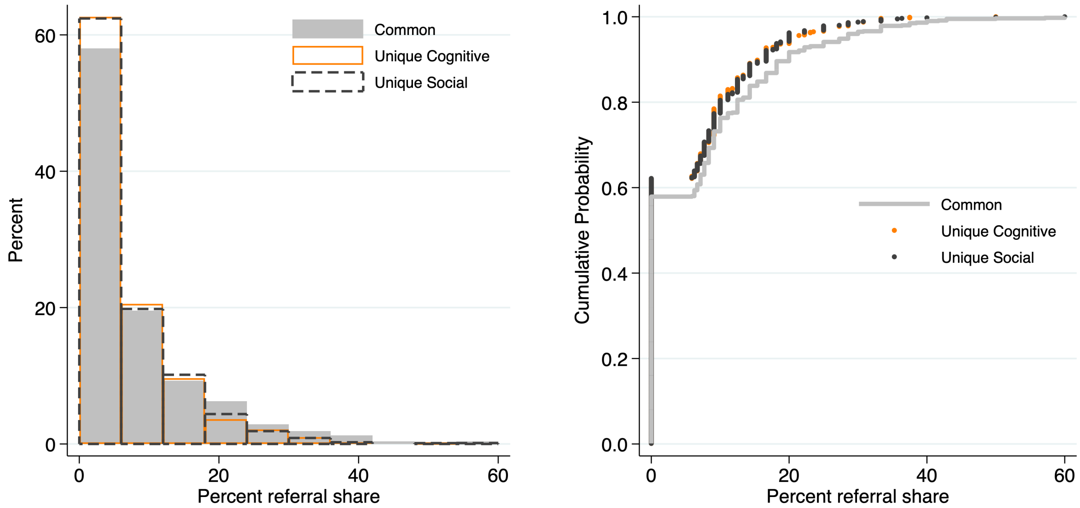
Table 5 shows the distribution of the referrals types in our sample. 37% of referrals fall under the type y_{ic}^{common} as pairs. This is equivalent to saying 54% of cognitive and social skill referrals were made in common. Figures 5a and 5b present the distributions of the three referral types.

Table 5: Distribution of Referral Types

	Frequency	Share (%)
Common	945	37.06%
Unique Cognitive	794	31.14%
Unique Social	811	31.80%
Total	2,550	100.00%

Note: Common referrals indicate the pair when the same classmate was referred for both cognitive and social skills. Unique referrals indicate when a classmate was referred for only one of the skills.

Figure 5: Distribution of common and unique referrals in Baseline



(a) Frequency Histogram

(b) ECDF

Note: Figures show the percantage of referrals recieved from participants in the **Baseline** condition depending on the referral type. The left panel shows the frequency histogram and the right panel shows the empirical cumulative distribution function (ECDF). Two-sample Kolmogorov-Smirnov tests show no statistically significant differences between the share of referrals received between "unique" cognitive and "unique" social referral distributions ($D = 0.0125$, $p = 1.000$) as well as "common" referrals ($D = 0.0602$, $p = 0.111$ for cognitive and $D = 0.0551$, $p = 0.177$ for social).

We regress Equation 3 for our three new dependent variables and report our findings in Table 6. Our preferred specification includes classroom fixed effects. The comparison of interest is the skill test scores and GPA estimates across columns. In column (2), we find that a one standard deviation increase in GPA causes a 3.7 percentage point increase in the share of "common" referrals received when controlling for skill test scores. This is a substantial 50 percent increase on a base rate of 7.4%. Cognitive skills remain statistically insignificant and social skills show a marginally significant negative coefficient, suggesting that participants who nominate the same individuals for both skills primarily make referrals based on academic performance.

For participants who receive unique cognitive skill referrals, in column (4), we find that a one standard deviation increase in GPA causes a 0.75 percentage point increase in the share of referrals when controlling for cognitive skill test score. A one standard deviation increase in cognitive skill test score causes a larger 1.1 percentage point increase in referrals when controlling for GPA. These are respectively 14 and 20 percent increases in the share of referrals received, and suggest participants are able to screen higher skilled peers when uniquely referring for cognitive skill. The lower base rate of 5.4% compared to 7.4% in column (2) suggests less than half of referrals came from "unique" referrals. The GPA estimate is five times smaller in magnitude compared to column (2), and suggests a smaller weight put on the grades proxy. Nevertheless, the comparable magnitudes of GPA and cognitive skill point estimates still suggest participants refer peers with higher grades much more often than the correlation between the two supported by the data ($\rho = 0.085$). There is heterogeneity in skill identification ability when uniquely referring for cognitive skill.

For participants who receive unique social skill referrals, in column (6), 95% confidence intervals rule out that a one standard deviation increase in social skill test score or GPA result in more than a 0.1 percentage point difference in the share of referrals received. These results further support our previous finding that peers cannot screen social skills in our sample, and do not attempt to screen social skills with the GPA proxy.

Result 3 *The majority of participants nominate the same individuals in common for both skills, cannot screen for skills and refer instead using the GPA proxy.*

Result 4 *Those who refer uniqely for cognitive skill can identify the skill test score, and drive the entierity of the results in terms of peer skill identification. Still, they confound cognitive skill with academic performance, and put comparable weights on the two. Those who refer uniqely for social skill can neither screen social skill or use the GPA proxy.*

Table 6: Share of "common" versus "unique" referrals received conditional on skill test score and academic performance

	Common		Unique Cognitive		Unique Social	
	(1)	(2)	(3)	(4)	(5)	(6)
GPA	3.172***	3.670***	0.801**	0.752*	0.260	0.108
	(0.464)	(0.501)	(0.391)	(0.401)	(0.334)	(0.360)
Cognitive score	-0.042	0.139	1.006***	1.084***		
	(0.416)	(0.388)	(0.270)	(0.281)		
Social score	-0.353	-0.553*			0.086	-0.011
	(0.304)	(0.272)			(0.381)	(0.357)
Dep. var. mean	7.407	7.382	5.400	5.401	5.485	5.493
Classroom FE	No	Yes	No	Yes	No	Yes
R^2	0.093	0.194	0.028	0.130	0.001	0.090
Observations	665	665	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable in columns (1)-(2) is the percentage share of "common" referrals received from the same referrer, in columns (3)-(4) "unique" referral share for cognitive skill, and in columns (5)-(6) for social skill. Independent variables are the respective standardized test scores for skills and GPA. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

5.4 Social class bias across common and unique referral types

In this section, we analyze referrals from the perspective of social class while accounting for the referral types described in this part. Based on the referral types from the previous section, we document the existence of a social class bias in referrals when controlling for skill test scores and academic performance at **Baseline**. Our dependent variables are the percentage shares of referrals received at **Baseline** as defined in Equation 6, and we include a social class dummy for the participant receiving the referrals. We estimate for our three dependent variables:

$$y_i^s = \alpha^s + \beta_1^s GPA_i + \beta_2^s Score_i^s + \beta_3^s SES_i + \epsilon_i^s \tag{7}$$

Table 7 summarizes our findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the SES estimates for the three referral strategies. In column (2), controlling for skill test scores and GPA, the point estimate for low-SES is not statistically significant. Skill score and GPA estimates are robust to the inclusion of this variable and remain close to those in Table 6.

For participants who receive unique cognitive skill referrals in column (4), we find that being low-SES causes a 1.8 percentage point decrease in the share of referrals when controlling for cognitive skill and GPA. This is a substantial 28 percent difference in the share of referrals received, confirming participants are biased against low-SES peers when uniquely referring for cognitive skill. Skill test scores and GPA estimates are robust to the inclusion of this variable. GPA and low-SES are not confounders as there are no significant differences across social classes in terms of GPA (see Appendix Figure A.3). The low-SES bias is consistent with the data where low-SES students underperform in the cognitive skill test (see Appendix Figure A.4a).

For participants who receive unique social skill referrals, in column (6), the point estimate for low-SES is not statistically significant. GPA and social skill estimates remain similar to those in Table 6. The finding that low-SES students underperform across skill dimensions is also consistent with earlier research (Falk, Kosse, Pinger, Schildberg-Hörisch, & Deckers, 2021), though we find that low-SES bias manifests only in unique cognitive skill referrals.

Result 5 *We document a sizeable low-SES bias for unique cognitive skill referrals when controlling for cognitive skill test score and academic performance of peers.*

Table 7: Share of "common" versus "unique" referrals received conditional on skill test score, academic performance, and social class

	Common		Unique Cognitive		Unique Social	
	(1)	(2)	(3)	(4)	(5)	(6)
GPA	3.170***	3.663***	0.797**	0.766*	0.260	0.111
	(0.462)	(0.499)	(0.386)	(0.388)	(0.334)	(0.360)
Cognitive score	0.000	0.167	0.869***	0.973***		
	(0.411)	(0.382)	(0.261)	(0.274)		
Social score	-0.306	-0.524*			0.047	-0.027
	(0.315)	(0.286)			(0.372)	(0.354)
Low-SES	0.799	0.568	-2.017***	-1.814**	-0.549	-0.260
	(0.939)	(0.934)	(0.711)	(0.713)	(0.610)	(0.593)
Dep. var. mean	6.948	7.056	6.558	6.442	5.800	5.642
Classroom FE	No	Yes	No	Yes	No	Yes
R^2	0.094	0.194	0.044	0.142	0.002	0.090
Observations	665	665	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable in columns (1)-(2) is the percentage share of "common" referrals received from the same referrer, in columns (3)-(4) "unique" referral share for cognitive skill, and in columns (5)-(6) for social skill. Independent variables are the respective standardized test scores for skills, GPA, and a dummy for low socioeconomic status. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

5.5 Social class bias and the Quota treatment

In the following empirical specification, we document whether there is a social class bias in aggregate, and whether the **Quota** treatment causes referral shares of low-SES participants to change when controlling for skills and academic performance. We hypothesized that the **Quota** treatment should increase referrals to low-SES peers because of the additional incentive to refer low-SES. The dependent variable is the percentage share of referrals received as defined for the **Baseline** treatment in Equation 1, now extended to the referrals from the **Quota** treatment. It is trivial to see y_{ic}^{s} can also be calculated for the **Quota** treatment as participants in every classroom are randomized into either treatment. Now, every participant is observed twice in the data for the share of referrals they received from participants in either treatment. We add a treatment dummy to indicate whether the referrals came from participants in the **Baseline** or the **Quota** treatment. We also add a social class dummy for the participant receiving the referrals to our specification and estimate:

$$y_i^s = \alpha^s + \beta_1^s Quota_i + \beta_2^s SES_i + \beta_3^s (Quota_i \times SES_i) + \beta_4^s Score_i^s + \beta_5^s GPA_i + \epsilon_i^s \quad (8)$$

Table 8 illustrates our findings. Our preferred specification includes classroom fixed effects. Our comparison of interest is the effect of the **Quota** treatment on low-SES peers. In column (2) for cognitive skill, we find that being low-SES decreases the share of referrals received by about 1.3 percentage points when controlling for the skill test score and academic performance. This difference is not statistically significant, but its direction and magnitude suggests a relatively large bias against low-SES classmates: A one standard deviation increase in cognitive skill test score has a similar magnitude (0.8 percentage points). This finding suggests the low-SES bias is driven by those who made unique cognitive referrals but it is not large enough to carry over to all cognitive skill referrals considered together. In column (4) for social skill, we find that being low-SES has no statistically significant effect on the share of referrals received when controlling for the skill test score and academic performance.

Result 6 *The low-SES bias is not large enough to carry over to all cognitive skill referrals when referrals are aggregated.*

Table 8: Share of referrals received by treatment, controlling for skill test score, academic performance, and social class

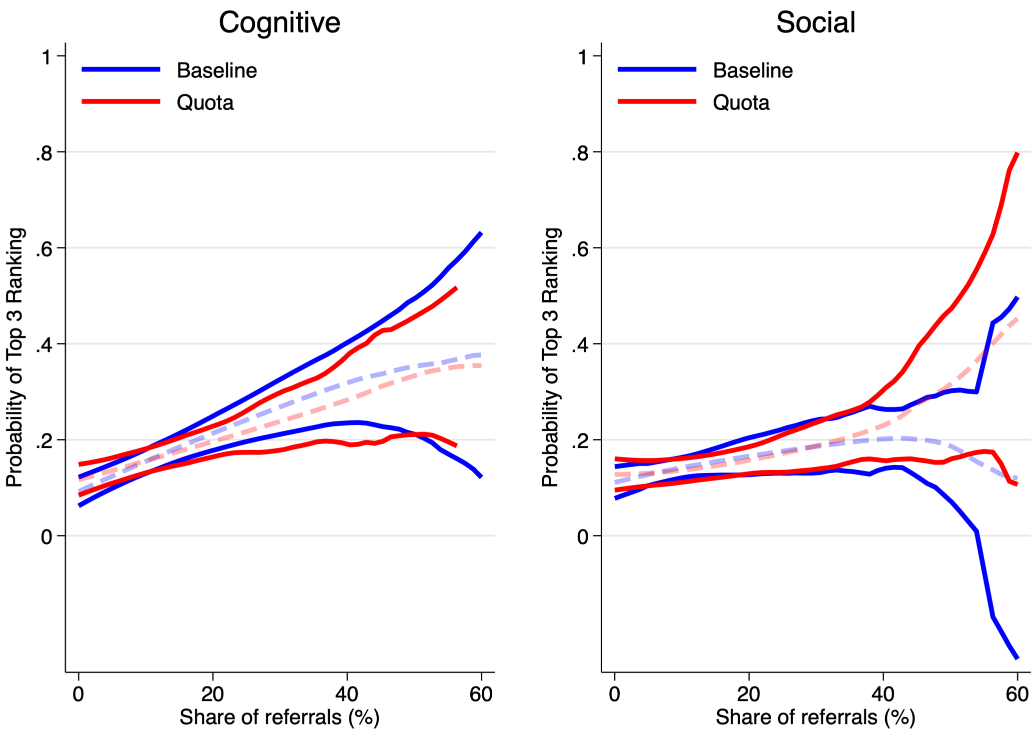
	Cognitive		Social	
	(1)	(2)	(3)	(4)
Quota	-0.073	-0.073	0.299	0.299
	(0.755)	(0.755)	(0.716)	(0.716)
Low-SES	-1.230	-1.276	0.364	0.324
	(1.079)	(1.014)	(1.282)	(1.361)
Quota × Low-SES	-0.167	-0.167	-0.835	-0.835
	(1.117)	(1.117)	(1.181)	(1.181)
Score	0.594	0.811*	0.201	-0.006
	(0.448)	(0.424)	(0.426)	(0.458)
GPA	3.184***	3.522***	2.819***	3.174***
	(0.517)	(0.552)	(0.493)	(0.621)
Dep. var. mean	13.551	13.558	12.706	12.714
Classroom FE	No	Yes	No	Yes
R^2	0.060	0.158	0.044	0.134
Observations	1,330	1,330	1,330	1,330

Note: Standard errors in parentheses are clustered at both classroom and individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are the percentage share of referrals received for cognitive and social skills. Quota is a dummy for the referrals received from classmates in the **Quota** treatment. Low-SES is a dummy for participant's socioeconomic status. Remaining independent variables are the respective standardized test scores for skills and GPA. Sample includes 1,330 observations with complete administrative and experimental data.

5.6 Quota treatment and referral productivity

As any intervention that changes the nomination decisions in terms of SES composition should not reduce the productivity of referrals, the equity-efficiency traedeoff is a valid concern for the **Quota** treatment. To address it, in Figure 6, we plot the share of referrals received across the two conditions and the probability of being among the Top 3 in the classroom for either skill. We find first that the slopes of the distributions are always positive for the **Quota** treatment, indicating a positive relationship with the share of referrals received. Second, a two-sample Kolmogorov-Smirnov test reveals no statistically significant differences in the distribution of referrals between the two conditions for both cognitive and social referrals. These suggest that the **Quota** treatment does not impact the positive relationship between the share referrals received and productivity in skills.

Figure 6: Referral shares and the probability of being in the Top 3



Note: The left panel shows how Baseline and Quota referral shares vary with the probability of being in the Top 3 of the classroom for cognitive skill scores, while the right panel shows the same figure for social skill scores. Solid lines indicate the 95% confidence intervals, with dashed lines representing means. The output is truncated at 60 percent of referral share to ensure meaningful confidence intervals. Two-sample Kolmogorov-Smirnov tests reveal no statistically significant differences in the distribution of referrals between Baseline and Quota conditions for both cognitive referrals ($D = 0.0351$, $p = 0.710$) and social referrals ($D = 0.0439$, $p = 0.427$).

5.7 Effects of the Quota treatment across referral types

Effects of the social class bias gets diluted across common and unique referral types. A large proportion of participants -“common” referrers- who struggle with skill identification and screen for skills using the academic performance proxy. But there are no SES differences for GPA in our sample. When referrals are made with academic performance in mind, it seems reasonable not to observe a negative bias against low-SES. Then what about skills, knowing that high-SES score higher in both measures?

We observe a bias in undersampling from equally well performing low-SES only for "unique" cognitive skill referrals, where referrers screen better compared to "unique" social skill referrals. We expect the **Quota** treatment be effective in increasing referrals for low-SES only in a scenario where the skill can be screened, and turn toward our classification of different referral types to test this hypothesis. To get clearer estimates for the effects of the **Quota** treatment on low-SES referrals, we re-estimate the shares of "common" and "unique" referrals. Following the same logic in the section before, we observe every participant twice in each specification, and add a treatment dummy to indicate whether the referrals came from referrers in the **Baseline** or the **Quota** treatment. We keep the social class dummy and regress Equation 8 for the three dependent variables.

Table 9 illustrates our findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the SES of the participant receiving the referrals and the effect of the **Quota** treatment across "common" and "unique" referral types. In column (2), for participants who refer the same peers in common using the academic performance proxy, we find no statistically significant effect of participant SES or the **Quota** on the referrals share when controlling for skill test scores and academic performance.

For unique cognitive skill referrals, in column (4), we find that being low-SES in the **Baseline** treatment reduces the percentage share of referrals received by 1.9 percentage points when controlling for the skill test score and academic performance. This is a very large effect size which translates to a decrease in referral share by 29 percent on a base rate of 6.5%, and is similar to the one found in Table 7. In turn, the **Quota** treatment increases referrals to low-SES by 1.42 percentage points when controlling for the skill test score and academic performance. This is also a large effect size that results in an increase in low-SES referral share by 22 percent.

For participants who make unique social skill referrals, in column (6), we find no statistically significant effect of participant SES or the **Quota** on the referrals share when controlling for the skill test score and academic performance. These are in accordance with our previous findings that social skills cannot be identified in our setting and it is possible that we do not observe the low-SES bias in this skill domain for this reason.

Result 7 *There is a bias against low-SES peers only for the skill that is well-identified by peers, and in which low-SES underperform. We find no evidence of a bias when referrals are made based on academic performance where both social classes perform equally well.*

Result 8 *The bias in unique cognitive skill referrals is partially alleviated by the **Quota** treatment. Because there is remarkable heterogeneity in the ability to detect SES for both social classes (see Appendix Figure A.5), this significant increase in low-SES referrals is satisfying in our setting.*

Table 9: Share of "common" and "unique" referrals received by treatment, controlling for skill test score, academic performance, and social class

	Common		Unique Cognitive		Unique Social	
	(1)	(2)	(3)	(4)	(5)	(6)
Quota	0.436	0.436	-0.509	-0.509	-0.136	-0.136
	(0.817)	(0.817)	(0.598)	(0.598)	(0.523)	(0.523)
Low-SES	0.857	0.598	-2.074***	-1.891**	-0.510	-0.256
	(0.920)	(0.897)	(0.722)	(0.710)	(0.613)	(0.594)
Quota × Low-SES	-1.584	-1.584	1.417**	1.417**	0.750	0.750
	(1.159)	(1.159)	(0.656)	(0.656)	(0.717)	(0.717)
Cognitive score	-0.079	0.095	0.658***	0.739***		
	(0.374)	(0.346)	(0.201)	(0.210)		
Social score	0.062	-0.091			0.158	0.061
	(0.283)	(0.236)			(0.312)	(0.269)
GPA	2.322***	2.727***	0.858***	0.804**	0.502*	0.439
	(0.330)	(0.366)	(0.312)	(0.340)	(0.278)	(0.292)
Dep. var. mean	6.952	7.080	6.591	6.488	5.765	5.623
Classroom FE	No	Yes	No	Yes	No	Yes
R^2	0.052	0.139	0.028	0.099	0.005	0.071
Observations	1,330	1,330	1,330	1,330	1,330	1,330

Note: Standard errors in parentheses are clustered at both classroom and individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable in columns (1)-(2) is the percentage share of "common" referrals received from the same referrer, in columns (3)-(4) "unique" referral share for cognitive skill, and in columns (5)-(6) for social skill. Quota is a dummy for the referrals received from classmates in the **Quota** treatment. Low-SES is a dummy for participant's socioeconomic status. Remaining independent variables are the respective standardized test scores for skills and GPA. Sample includes 1,330 observations with complete administrative and experimental data.

6 Conclusion

In this paper, we study how accurately individuals assess productivity of their peers across different skill dimensions and whether these assessments systematically disadvantage low-SES individuals in a diverse university setting. Through a lab-in-the-field experiment that isolates screening ability, we find that the accuracy of peer productivity assessments varies significantly across skill types, with implications for referral-based screening.

Our findings reveal that peers can effectively identify cognitive skills but struggle to assess social skills in their classmates. This differential screening ability appears to stem from the inherent challenges in evaluating interpersonal capabilities compared to cognitive abilities. When faced with uncertainty in skill assessment, peers often rely on observable proxies like academic performance which may be misleading. This suggests that the effectiveness of peer assessments depends crucially on how discernible the target skill is, rather than indicating a fundamental limitation of referrals as a screening mechanism.

These results complement the broader literature showing referrals' effectiveness in worker screening by highlighting how skill visibility affects assessment accuracy. While previous work demonstrates that referrals successfully identify productive workers overall (Pallais & Sands, 2016), our findings suggest their effectiveness may vary across different dimensions of human capital. This variation is particularly relevant given the growing importance of social skills in the labor market as found in other research (Deming, 2017). Our evidence also supports earlier evidence that accurate assessment of social skills remains challenging (Bassi & Nansamba, 2022), suggesting the need for either longer periods of interaction to discern these skills or development of alternative assessment methods that can better capture interpersonal capabilities in referral settings.

Looking forward, our findings suggest several implications for improving screening mechanisms in similar settings. First, institutions that implement referral programs may need to develop complementary tools for evaluating less visible skills like interpersonal capabilities, perhaps in the likes of the social skill certificates in Bassi and Nansamba (2022). Second, our results on social class bias - finding it only in unique cognitive skill referrals and its mitigation through quota incentives without compirimising productivity- indicate that targeted interventions can effectively address specific biases without compromising the overall screening process. Future research could investigate how to optimize referral programs to leverage their strengths in identifying easier to discern skills while developing better methods for assessing harder to observe skills.

References

Alan, S., & Kubilay, E. (2025). Empowering adolescents to transform schools: Lessons from a behavioral targeting. *American Economic Review*, *115*(2), forthcoming. doi: 10.1257/aer.20240374

Angulo, R., Gaviria, A., Páez, G. N., & Azevedo, J. P. (2012). Movilidad social en colombia. *Documentos CEDE*.

Balafoutas, L., & Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, *335*(6068), 579–582.

Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, *77*(4), 1047–1094.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251. (Publisher: Cambridge University Press) doi: 10.1017/S0021963001006643

Bassi, V., & Nansamba, A. (2022). Screening and signalling non-cognitive skills: experimental evidence from uganda. *The Economic Journal*, *132*(642), 471–511.

Beaman, L., Keleher, N., & Magruder, J. (2018). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics*, *36*(1), 121–157. doi: 10.1086/693869

Beaman, L., & Magruder, J. (2012). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review*, *102*(7), 3574–3593. doi: 10.1257/aer.102.7.3574

Bertrand, M., Black, S. E., Jensen, S., & Lleras-Muney, A. (2019). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, *86*(1), 191–239.

Beugnot, J., & Peterlé, E. (2020). Gender bias in job referrals: An experimental test. *Journal of Economic Psychology*, *76*, 102209. doi: 10.1016/j.joep.2019.102209

Calvo-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American economic review*, *94*(3), 426–454.

Calvó-Armengol, A., & Jackson, M. O. (2007). Networks in labor markets: Wage and employment dynamics and inequality. *Journal of economic theory*, *132*(1), 27–46.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., … others (2022a). Social capital II: determinants of economic connectedness. *Nature*, *608*(7921), 122–134.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., … others (2022b). Social capital I: measurement and associations with economic mobility. *Nature*, *608*(7921), 108–121.

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

Deming, D. J. (2023). Multidimensional human capital and the wage structure. *National Bureau of Economic Research*.

Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., & Deckers, T. (2021). Socioeconomic status and inequalities in children's iq and economic preferences. *Journal of Political Economy*, *129*(9), 2504–2545.

García, S., Rodríguez, C., Sánchez, F., & Bedoya, J. G. (2015). La lotería de la cuna: La movilidad social a través de la educación en los municipios de colombia. *Documentos CEDE*.

Guevara S, J. D., & Shields, R. (2019). Spatializing stratification: Bogotá. *Ardeth. A Magazine on the Power of the Project*(4), 223–236.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, *19*(4), 451–464.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, *24*(3), 411–482.

Hederos, K., Sandberg, A., Kvissberg, L., & Polano, E. (2024). Gender homophily in job referrals: Evidence from a field study among university students. *Labour*

Economics, *87*, 102461. doi: 10.1016/j.labeco.2024.102461

Hudson, R. A., & Library of Congress (Eds.). (2010). *Colombia: a country study* (5th ed.). Washington, D.C: Federal Research Division, Library of Congress: For sale by the Supt. of Docs., U.S. G.P.O. Retrieved from the Library of Congress, https://www.loc.gov/item/2010009203/.

Kim, H., Kaduthodil, J., Strong, R. W., Germine, L., Cohan, S., & Wilmer, J. B. (2022). *Multiracial Reading the Mind in the Eyes Test (MRMET): an inclusive version of an influential measure* (preprint). doi: 10.31219/osf.io/y8djm

Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101–128.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, *111*(6), 1807–1844.

McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in america: Changes in core discussion networks over two decades. *American sociological review*, *71*(3), 353–375.

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

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *American Economic Review*.

Mousa, S. (2020). Building social cohesion between christians and muslims through soccer in post-isis iraq. *Science*, *369*(6505), 866–870.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1–16.

Pallais, A., & Sands, E. G. (2016). Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, *124*(6), 1793–1828. doi: 10.1086/688850

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review*, *109*(3), 774–809.

Raven, J. C. (1936). *The Performances of Related Individuals in Tests Mainly Educative and Mainly Reproductive Mental Tests Used in Genetic Studies* (PhD Thesis). University of London (King's College).

Raven, J. C., Raven, J., & Court, J. H. (1976). *Standard progressive matrices: sets A, B, C, D & E.* Oxford, UK: Oxford Psychologists Press.

Schilbach, F., Schofield, H., & Mullainathan, S. (2016). The Psychological Lives of the Poor. *American Economic Review*, *106*(5), 435–440. doi: 10.1257/aer.p20161101

Stansbury, A., & Rodriguez, K. (2024). The class gap in career progression: Evidence from US academia. *Working Paper*.

United Nations. (2023). *Social panorama of latin america and the caribbean 2023: labour inclusion as a key axis of inclusive social development.* ECLAC and United Nations. Retrieved from https://www.cepal.org/es/publicaciones/68702-panorama-social-america-latina-caribe-2023-la-inclusion-laboral-como-eje-central

Uribe-Mallarino, C. (2008). Estratificación social en bogotá: de la política pública a la dinámica de la segregación social. *Universitas humanistica*(65), 139–172.

van Leeuwen, B., Noussair, C. N., Offerman, T., Suetens, S., van Veelen, M., & van de Ven, J. (2018). Predictably Angry—Facial Cues Provide a Credible Signal of Destructive Behavior. *Management Science*, *64*(7), 3352–3364. doi: 10.1287/mnsc.2017.2727

Weidmann, B., & Deming, D. J. (2021). Team Players: How Social Skills Improve Team Performance. *Econometrica*, *89*(6), 2637–2657. doi: 10.3982/ECTA18461

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, *96*(5), 849–861.

Witte, M. (2021). Why do workers make job referrals? experimental evidence from ethiopia. *Working Paper*.

World Bank. (2024). *Regional poverty and inequality update spring 2024* (Poverty and Equity Global Practice Brief). Washington, D.C.: World

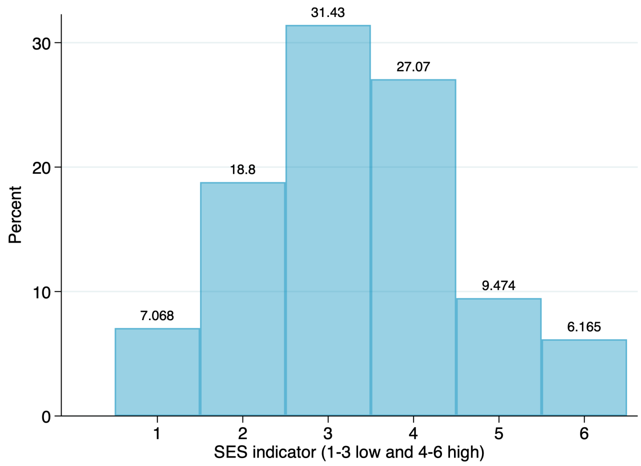
Bank Group. Retrieved from http://documents.worldbank.org/curated/en/099070124163525013/P17951815642cf06e1aec4155e4d8868269

Zárate, R. A. (2023, July). Uncovering peer effects in social and academic skills. *American Economic Journal: Applied Economics*, *15*(3), 35–79. doi: 10.1257/app.20210583

A Additional Figures and Tables

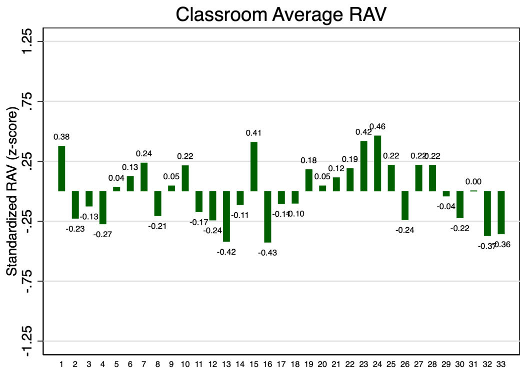
A.1 Additional Figures

Figure A.1: Stratum distribution of the sample

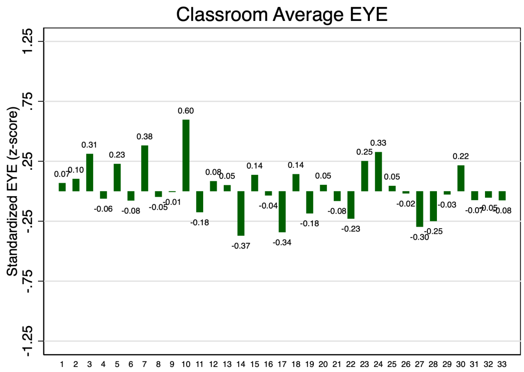


Note: This figure shows the distribution of strata in the sample of students that participated in the study.

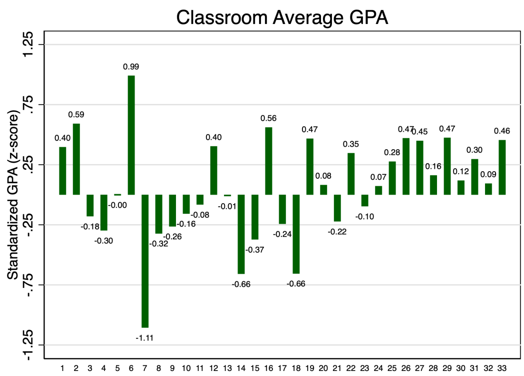
(a) Cognitive score across classrooms



(b) Social score across classrooms

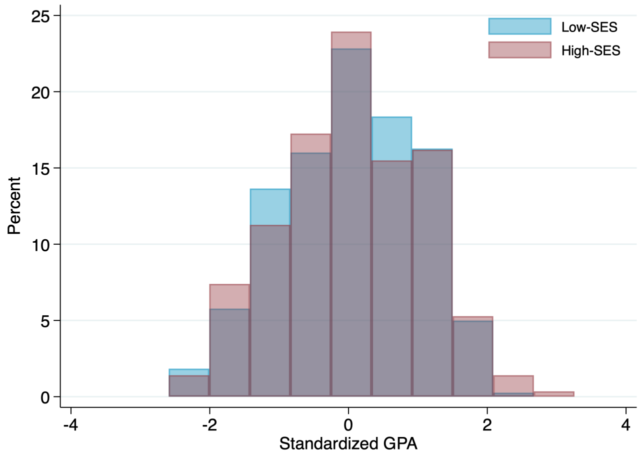


(c) GPA across classrooms



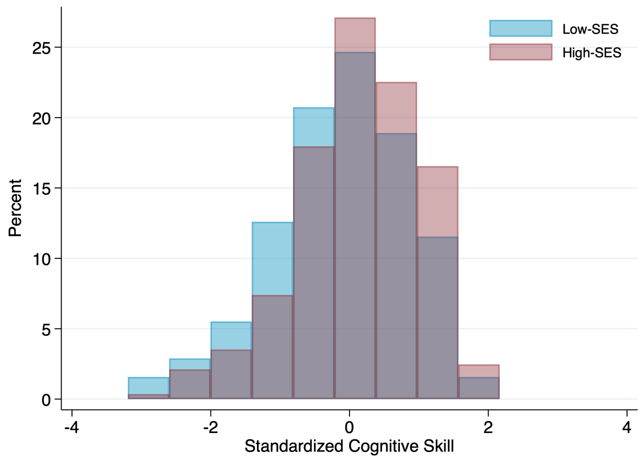
Note: These figures show the respective distribution of standardized scores for cognitive skill, social skill, and GPA across sampled classrooms.

Figure A.3: GPA by SES

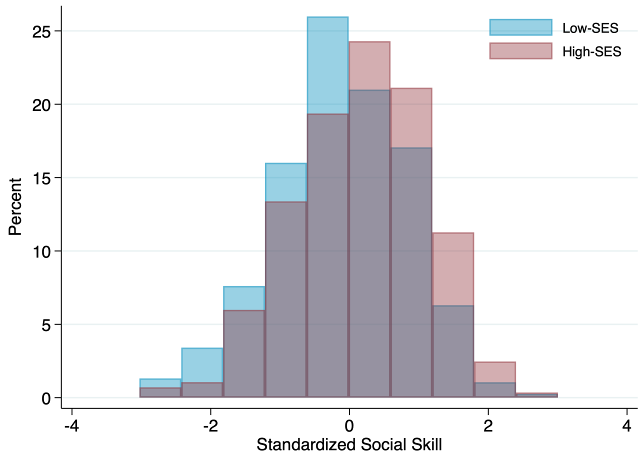


Note: This figure shows the distribution of GPA across SES. There are no significant differences in the mean standardized GPA scores between high-SES and low-SES participants (t test $p = 0.695$).

(a) Cognitive score by SES

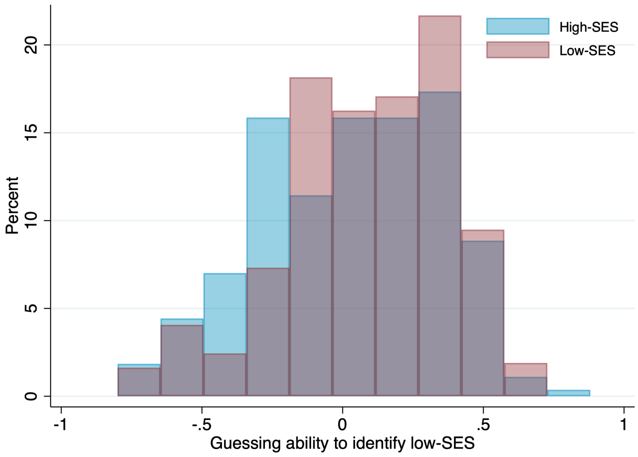


(b) Social score by SES



Note: These figures show the respective distribution of cognitive and social skills across SES. High social class outperform Low-SES in both skills (t tests have p values < 0.001). We can visually verify that larger share of high-SES in quantiles above median for both skills.

Figure A.5: Distribution of guessing ability across SES



Note: This figure shows the distribution of the guessing ability across SES. We calculate the guessing ability as the share of succesful low-SES guesses minus the expected probability of randomly drawing low-SES in class c. A score of 0 indicates an accuracy as good as random draws, below 0 drawing worse than chance, and above 0 better than chance. There are significant differences in the mean guessing ability between high-SES (M = 0.022, SD = 0.325, n = 271) and low-SES participants (M = 0.093, SD = 0.302, n = 369), $t(638) = -2.85$, $p = 0.005$, $d = 0.226$. Low-SES participants have higher guessing ability compared to their high-SES counterparts, with a mean difference of 7 percentage points.

A.2 Additional Tables

Table A.1: Selection into the experiment

	Sample	**Missing**	***p***
Referral share (both skills)	0.127	0.043	0.000
GPA (standardized)	0.044	-0.273	0.001
Entry Exam (standardized)	0.028	-0.168	0.046
# Semesters at UNAB	3.171	3.188	0.884
Age	19.182	20.287	0.001
Female	49.8%	48.5%	0.788
Ethnic Minority	2.1%	4.4%	0.114
Rural Community	28.8%	31.6%	0.501
Has Scholarship	0.8%	0.7%	0.899

Note: Values for female, ethnic minority, rural community, and scholarship represent percentage proportions. All other variables represent means. p-values for gender, ethnic, rural, and scholarship are from two-sample tests of proportions. For all other variables, p-values are from two-sample t-tests with equal variances. All tests compare the sample and missing students. All reported p-values are two-tailed.

Table A.2: Correlation between GPA, entry exam, and skill test scores

	GPA	Cognitive score	Social score	Entry Exam
GPA	1.000			
Cognitive score	0.083	1.000		
Social score	0.091	0.266	1.000	
Entry Exam	0.229	0.403	0.267	1.000

Note: Pairwise correlation between GPA, entry exam, and skill test scores. Sample is restricted to 655 participants with complete administrative and experimental data.

Table A.3: Between-Classroom Variation in Academic Programs

Statistic	Most common program share
Mean	0.424
Standard Deviation	0.216
10th percentile	0.174
25th percentile	0.292
Median	0.345
75th percentile	0.533
90th percentile	0.696
# classrooms with share 1	3
Most diverse classroom	0.154
# classrooms	35

Note: Table shows the distribution of academic programs across classrooms, measured by the share of students from the most common program in each classroom. Three classrooms are completely homogeneous (share = 1). In the median classroom, the most common program accounts for 34.5% of students. The most diverse classroom has only 15.4% of students in the same program. Data based on 849 students across 35 classrooms.

Table A.4: Characteristics of self-referrers

	No self-referral	Any self-referral	Δ	p
GPA	0.132	-0.120	0.252	0.002
	(1.003)	(0.966)		
Cognitive score	0.087	-0.118	0.205	0.013
	(0.988)	(1.023)		
Social score	0.034	-0.038	0.072	0.374
	(1.003)	(0.959)		
Low-SES	0.605	0.511	0.094	0.021
	(0.490)	(0.501)		
N	440	225	665	
Share (%)	66.2	33.8	100	

Note: Table compares standardized scores between participants who self-referred at least once (N = 225) and those who did not (N = 440). Positive differences indicate higher scores for those who never self-referrered. p-values from two-sided t-tests (GPA, Cognitive Skill, Social Skill) and proportion test (Low-SES). The results suggest self-referrers have significantly lower cognitive skills and GPA, and are more likely to be low-SES. Standard deviations in parentheses, samples restricted to participants with complete administrative and experimental data.

Table A.5: Characteristics of participants who make overlapping referrals

	Unique referrals	Common referrals	Δ	p
GPA	0.057	0.045	0.012	0.903
	(0.983)	(1.009)		
Cognitive score	0.110	0.024	0.086	0.371
	(1.005)	(0.979)		
Social score	-0.014	0.033	-0.047	0.621
	(0.938)	(0.981)		
Low-SES	0.530	0.597	-0.067	0.164
	(0.501)	(0.491)		
N	132	512	644	
Share (%)	20.5	79.5	100	

Note: Table compares characteristics between participants who made at least one overlapping referral (N = 512) to those who did not (N = 132, 20.5%). Overlapping referrals indicate cases where a participant referred the same classmate once for cognitive or social skills. Positive differences indicate higher scores for those who made no overlapping referrals. The results suggest minimal differences across all variables. p-values from two-sided t-tests (GPA, Cognitive Skill, Social Skill) and proportion test (Low-SES). Standard deviations in parentheses, sample restricted to participants with complete administrative and experimental data.

Table A.6: Characteristics of Top Performers and Referrals

	Cognitive		Social		Both
	Top 3	Referrals	Top 3	Referrals	Top 3
Cognitive score	1.223	0.112	0.383	0.058	1.201
	(0.419)	(1.009)	(0.922)	(1.015)	(0.458)
Social score	0.357	0.086	1.340	0.042	1.391
	(0.923)	(0.996)	(0.395)	(1.009)	(0.453)
GPA	0.277	0.251	0.264	0.212	0.551
	(0.990)	(1.021)	(1.046)	(1.004)	(0.897)
Low-SES	0.457	0.532	0.456	0.555	0.500
	(0.500)	(0.499)	(0.500)	(0.497)	(0.507)
N	129	1,759	114	1,775	36
Share (%)	20.0	100	17.7	100	5.6

Note: Table shows characteristics of students ranked in the top 3 of their classroom and average characteristics of referred students, by skill. Standard deviations in parentheses. Sample restricted to participants with complete administrative and experimental data. All continous variables are standardized.

B Experiment

We include the English version of the instructions used in Qualtrics. Participants saw the Spanish version. Horizontal lines indicate page breaks, and clarifying comments are inside brackets.

Please enter the password:
[classroom-specific password sent to each participant the day before data collection]

Welcome

Welcome to this study organized by the Social Bee Lab. You have been invited to participate in a survey where you can make a series of decisions. The study takes approximately 20 minutes to complete. During the study, you should not communicate with any other students. If you have any questions at any time, please raise your hand. One of the assistants will help you privately.

In this study, you can win bonus money depending on your choices. In total, we will draw [classroom-specific number equal to 40% of class size] bonuses of 100.000 pesos among the participants of this classroom. It is also possible for the same person to win more than one voucher. The following screens will detail how the bonus draw will be conducted. The UNAB finance office will make the payment of the vouchers through Nequi.

All your decisions in this survey will be anonymized. Therefore, the answers you provide will not affect your grades in this class or your records at the university. We will use your personal information to determine the bonus allocation, but after that, we will remove any data that identifies you.

This survey has several parts. Each of these parts has specific instructions. Please read the instructions for each part carefully because they describe how you can earn bonuses. This study has been approved by the [omitted for anonymous review] on the condition that all the information we provide is true and all the bonuses we offer are real. On the next screen, we present you with an informed consent form that you must accept to participate in this study.

Informed Consent

You have been invited to participate in a study to learn more about how people make decisions in common scenarios.

This study is conducted by [omitted for anonymous review] and the Social Bee Lab at UNAB. The purpose of this study is to broaden our understanding of how people make decisions.

Participation in this study is voluntary. You may opt-out at any time. No known risks are associated with your participation in this project beyond those of everyday life. Apart from the monetary bonuses that will be drawn, participation has no direct benefits.

The Social Bee Lab is in charge of data collection. Your answers in this study are anonymous and will not be shared with anyone. In addition to your answers, UNAB will provide the Social Bee Lab with administrative records of your courses and your university entrance exam score. Your records, decisions, and your identity will be kept strictly confidential. Data about you collected within the scope of the study are used for scientific purposes only and are treated as strictly confidential. The Social Bee Lab will anonymize your data, and the researcher will analyze it without knowing your identity. All data generated will be stored on the researcher's computer. You have the right to access your personal data and request its deletion. You can exercise this right by contacting the researcher.

If something is unclear or you have any questions, you can contact [omitted for anonymous review].

If you have questions about your rights as a participant, you can contact [omitted for anonymous review].

By continuing to the next screen, you agree to participate in this study.

Before you start, please answer these four questions.

What is your gender?

[Male, Female]

What is the socio-economic stratum to which your family belongs?

[Stratum 1 to Stratum 6]

What is your father's highest acquired level of education?

[Primary school, High school, Technical school, Undergraduate, Graduate, Postgraduate, Not applicable.]

What is your mother's highest acquired level of education?

[Primary school, High school, Technical school, Undergraduate, Graduate, Postgraduate, Not applicable.]

Part 1

You will now participate in two quizzes, each lasting five minutes. Please try to answer them to the best of your ability.

We will allocate up to [classroom-specific number equal to 20% of class size] bonuses of 100.000 pesos in this first part. The steps to allocate the bonuses for Part 1 are explained below.

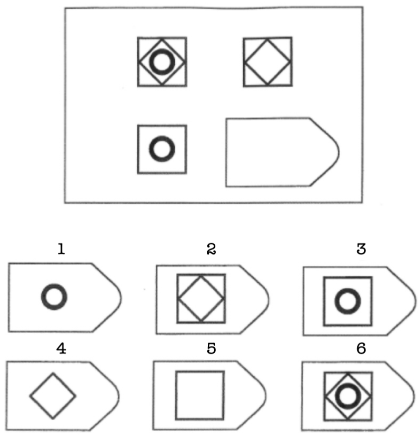
[classroom-specific illustrations explaining the incentive structure]

[random assignment to either cognitive or social skills test]

Test - Cognitive Skill

In this test, you will see a series of images. Below is an example of the images you will solve. At the top of each image, there is a pattern with a piece that has been removed.

Your task is to choose which of the six pieces completes the pattern correctly. For each image, there is only one correct piece. Look at the following example:



First, notice a square in the upper left, the upper right, and the lower left. Also, notice that the circle is eliminated when one moves from the upper left to the upper right. Finally, the rhombus is eliminated when moving from the upper left to the lower left. Therefore, the correct piece should eliminate the circle and the rhombus, leaving only a square. So, the correct answer is piece 5.

To give your answer to each image, you must choose the correct option and then continue to the next screen. After giving your answer you cannot go back.

You will have 5 minutes to complete the test, which consists of 18 images to solve. The percentage of correct answers will determine your chances of winning one of the 100.000 pesos bonuses if you are chosen for the drawing.

Are you ready?

Your 5 minutes will start as soon as you move to the next screen.

Problem 1

[screenshot of Raven's matrix]

[After participants submit an answer, a new matrix appears on the screen. The sequence of matrices is the same for all participants. Participants cannot return to a previous screen. Participants do not have to provide answers for all 18 matrices.]

You have finished the test. You can proceed to the next screen.

How did you do on the test?

If we randomly choose 10 participants from this classroom, how many people do you think solved fewer correct problems than you?

[Slider from 0 to 10]

Test - Emotions

In this test, you will see a series of photographs. Below is an example of the pictures you will see. In each picture, you will see the eyes of a person. Below the picture, you will see four possible emotions that this person is feeling. Your task is to choose which of the four emotions correctly describes what the person is feeling. For each picture, there is only one emotion. Look at the following example:

[Happy, Disappointed, Shocked, Worried]

In this case, the correct answer is: Shocked.

To give your answer to each picture, you must choose the correct option and then continue to the next screen. After giving your answer you will not be able to go back.

You will have 5 minutes to complete the test, which consists of 36 photographs to solve. The percentage of correct answers you get will determine your chances of winning one of the 100.000 pesos bonuses if you are chosen for the drawing.

Are you ready?

Your 5 minutes will start as soon as you move to the next screen.

Photograph 1: Choose the word that best describes the photograph

[photo from Multiracial Reading the Mind in the Eyes Test]

[After participants submit an answer, a new photo appears on the screen. The sequence of photos is the same for all participants. Participants cannot return to a previous screen. Participants do not have to provide answers for all 36 photos.]

You have finished the test. You can proceed to the next screen.

How did you do on the test?

If we randomly choose 10 participants from this classroom, how many people do you think solved fewer correct photographs than you?

[Slider from 0 to 10]

Part 2

At the beginning of this study, all participants took two tests, one on cognitive ability and one on emotions. In this part, we will ask you to recommend the people who in your opinion will score the best on each test.

You may recommend 3 people per test, but you may not recommend yourself.

We will allocate up to [classroom-specific number equal to 40% of class size] bonuses of 100.000 pesos for Part 2. The steps for allocating bonuses are explained below.

[random assignment to either quota or baseline condition]

[classroom-specific illustrations explaining the incentive structure depending on assignment to either baseline or quota conditions]

Figure B.1: Illustrations for the two conditions

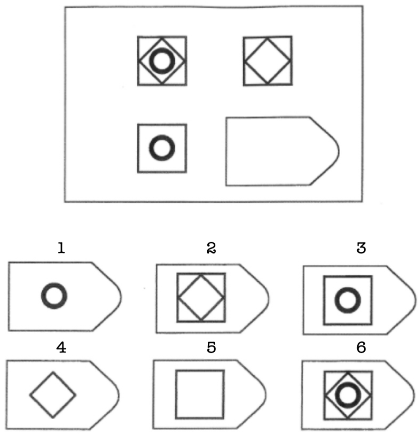
(a) Baseline

(b) Quota

[random assignment to either cognitive or social skills referral task]

Recommendation - Cognitive Skill

All participants took a test to identify the missing pattern in each image, as in the example below. This test is used to measure general intelligence.



Next, we will present you with a list of the names of all the students in this room. We will ask you to recommend the three people you think will score the highest on the general intelligence test.

If you are chosen by the computer, each of your recommendations in the top 3 increases your chances of winning one of the 100.000 pesos bonuses.

Select the students in this classroom who you consider to have the highest scores on the general intelligence test. (Select 3 students)

[Classroom-specific list of all classmate names visible on one screen. Participants have to pick 3 classmates to continue. Picking their own name invalidates their choices.]

Recommendation - Emotions

All participants took a test where they had to identify the emotion that best described the expression of each image as in the example below. This test is used to measure social skills.

Next, we will present you with a list of the names of all the students in this room. We will ask you to recommend 3 people you think will score the highest on the social skills test.

If you are chosen by the computer, each of your recommendations in the top 3 increases your chances of winning one of the 100.000 pesos bonuses.

Select the students in this classroom who you consider to have the highest scores on the social skills test. (Select 3 students)

[Classroom-specific list of all the names visible on one screen. Participants have to pick 3 classmates to continue. Picking their own name invalidates their choices.]

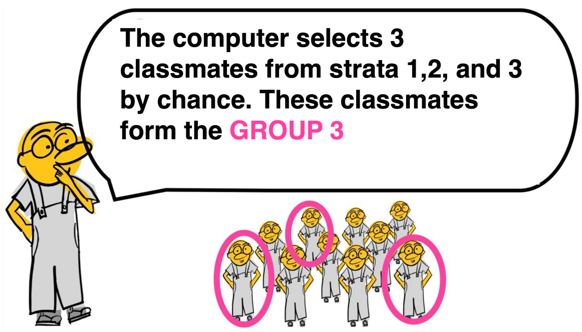
Part 3: Recommendation - Random draw

In this part, the computer will randomly choose three students who belong to strata 1, 2, or 3. We will ask you to nominate three people you think the computer will choose.

We will allocate up to [classroom-specific number equal to 20% of class size] bonuses of 100.000 pesos for Part 3. The steps for allocating the bonuses are explained below.

[classroom-specific illustrations explaining the incentive structure]

Figure B.2: Illustration for the Guessing Task



Select the students in this classroom who belong to strata 1, 2, or 3, who you think will be randomly selected by the computer (Select 3 students).

[Classroom-specific list of all the names visible on one screen. Participants have to pick 3 classmates to continue. Picking their own name invalidates their choices.]

Part 4

Do you want to know your scores on the general intelligence test and the social skills test? We can analyze the data and give you a report that explains your strengths in these two areas. Also, what do these strengths mean, and how can you leverage them for your personal and professional development?

If you want to receive your skills report, we need to contact you again. We also want to be able to invite you to new studies where you can participate for more bonus money. Please indicate if you agree to be contacted again.

[I can be contacted for new studies and to send me my report. I can be contacted to send my report, but not for new studies. No, I do not want to be contacted again.]

[if participant gives consent to be contacted again]

Please enter your contact email:

[student email]