The Journal of Development Studies

From Seating to Success: How Physical Distance in Social Networks Shape Individual Achievement

Submission ID	249026531
Article Type	Original Manuscripts
Keywords	peer effect, peer network, instrumental variable, rural education

Authors	Yujuan Gao, Yu Bai, Yue Ma, Andrew Rule, Sco tt Rozelle

From Seating to Success:

How Physical Distance in Social Networks Shape Individual Achievement

Yujuan Gao[1, 2], Yu Bai [1, *], Yue Ma [3], Andrew Rule[3], Scott Rozelle[3]

1. *China Institute for Vitalizing Border Areas and Enriching the People, Minzu University of China, 27 Zhongguancun South Avenue, Beijing 10081, China. yujuan.gao@ufl.edu*

2. *Food and Resource Economics Department, University of Florida, 1103 McCarty Hall B, Gainesville, FL 32611, USA.*

3. *Stanford Center on China's Economy and Institution, Stanford University, Stanford, CA 94305, USA. yma3@stanford.edu, arule@stanford.edu, rozelle@stanford.edu*

**Corresponding author: Yu Bai, Email: yubai@muc.edu.cn, Phone: 86-186-0040-5923, https://orcid.org/0000-0002-6052-7766*

From Seating to Success:

How Physical Distance in Social Networks Shape Individual Achievement

This study investigates the causal relationship between peer effects and academic performance in classroom microenvironments, addressing the challenge of endogeneity in peer group selection. Using data from 2,956 primary school students in rural China, we employ network theory to model study group structures and an instrumental variable approach to control for selection bias. Our findings reveal that study groups significantly enhance student achievement by 0.11 standard deviations. The marginal benefits are heterogeneous across subgroups, with stronger peer effects observed among male participants, lower-performing individuals and in groups with stronger group cohesion. Through a mediation analysis, we identify intrinsic motivation as the primary mechanism driving these peer effects, suggesting that peer effects primarily work by enhancing autonomous learning behaviors. Our findings inform cost-efficient policy interventions in both educational institutions and corporate environments. The evidence indicates that optimizing spatial proximity in peer networks represents an efficient policy instrument for human capital accumulation, particularly valuable in resource-constrained settings as it leverages existing human capital without substantial additional inputs.

Keywords: peer effect; peer network; instrumental variable; rural education

JEL codes: A20; D85; I25

Introduction

The role of education in driving economic growth extends beyond individual human capital formation to shape broader societal outcomes. While education directly enhances labor productivity and the innovative capacity, leading to increased output and growth (Mankiw, Romer, & N. Weil, 1992), it also facilitates knowledge diffusion and technologies adoption, further fuelling economic development (Aghion, Howitt, Brant-Collett, & García-Peñalosa, 1998; Benhabib & Spiegel, 1994; Romer, 1990). Within education settings, peer interactions-particularly through physical proximity in seating and study arrangements-have emerged as a potentially powerful but understudied channel for academic achievement (Duflo et al., 2011; Marotta, 2017). However, identifying the causality in peer effects research presents significant empirical challenges due to the complex nature of peer interactions, such as endogenous, exogenous, and correlated effects (Manski, 1993). These effects give rise to two fundamental challenges: reflection problem and self-selection bias (Brock & Durlauf, 2001; Manski, 1993). The reflection problem leads to simultaneous behavioral changes among peers, complicating the differentiation between endogenous, correlated, and exogenous effects (Bramoullé, Djebbari, & Fortin, 2020).

This study leverages unique institutional features of Chinese primary schools to address the fundamental identification challenges in peer effects research. In Chinese primary schools, two key institutional features create an ideal setting for studying peer interactions. First, students spend the majority of their school day in fixed seats within a single classroom, while teachers rotate between rooms. Specifically, desks are arranged in a specific way and students are assigned to their seats by their teacher. Second, and crucially, school policy prohibits the disclosure and consideration of students' academic performance in seating assignments, making seat placement effectively exogenous to academic ability. These institutional characteristics help address the self-selection bias that typically plagues peer effects studies. Moreover, the fixed seating arrangement, combined with prolonged daily exposure to the same peers, creates sustained physical proximity that naturally facilitates social interaction and friendship (Back, Schmukle, & Egloff, 2008; Hare & Bales, 1963; McAndrew, 1993). As a result, distance between students serves as an exogenous index for studying peer relationships and their effects on academic performance.

Using data from baseline and follow-up surveys in rural Chinese primary schools, this study examines how physical distance between students shapes their study group formation and how these naturally formed study groups affect academic performance. We employ an Instrumental Variable (IV) analysis where peer groups are constructed based on students' self-reported study partners at baseline. Physical distance between students, calculated from classroom seating distributions at baseline, serves as an instrumental variable to predict actual peer relationships. We measure academic performance using standardized mathematics test scores collected during both baseline and follow-up surveys. Our analysis also examines the mechanisms driving these peer effects and their heterogeneity across different group compositions.

Our analysis reveals three key findings about peer effects and academic performance. First, we find that physical distance between students serves as a reliable instrument for predicting peer relationships. Second, our instrumental variable (IV) analysis shows that peer groups significantly enhance academic achievement, with larger benefits for initially lower-performing students. Third, we identify intrinsic motivation-defined as engaging in activities for their inherent satisfaction rather than external rewards Deci (1973)-as the primary mechanism through which peer effects operate. Students with higher intrinsic motivation exhibit greater peer cooperation and increased study time investment throughout the semester. Together, these findings suggest that peer proximity enhances academic achievement primarily by fostering intrinsic motivation, which in turn facilitates collaborative learning and improved study habits.

This study advances the peer effects literature through its novel identification strategy. While existing research often relies on random assignment to classrooms or dormitories (Carrell, Sacerdote, & West, 2013; Foster, 2006; Lyle, 2007), such approaches overlook the quality and nature of peer relationships. Alternative studies using instrumental variables (IV) typically employ peers' parental education or background characteristics to establish causality (Hinke, Leckie, & Nicoletti, 2019). However, these approaches rest on the questionable assumption that peer backgrounds affect academic outcomes only through peer ability. Our study overcomes these limitations by exploiting exogenous variation in seating assignments as an instrumental variable. This approach offers two key advantages: it captures naturally occurring peer interactions, and it avoids relying on strong assumptions about the exclusion of peer characteristics. By using classroom geography as an identification strategy, we can isolate the causal effects of peer group formation while accounting for the organic development of student relationships.

In addition, this study enriches the peer effects literature by capture both the formation and strength of peer relationships at a more granular level. The existing peer effects literature has largely overlooked the critical role of subgroup formation within classrooms. While many studies examine peer effects at the classroom or school level (Bramoullé et al., 2009; Burke & Sass, 2013; Hinke et al., 2019; Min et al., 2019), this aggregate approach misses crucial within-classroom dynamics. Students in larger groups naturally form subgroups based on shared characteristics such as gender or social identity (McKeown, Stringer, & Cairns, 2016), making classroom-level analyses potentially misleading (Anderson & Lu, 2017). Recent attempts to study smaller subgroups (Berthelon et al., 2019) mark an improvement but still fail to capture the intensity of peer relationships—a factor shown to strengthen academic peer effects (Mora & Oreopoulos, 2011). Our study addresses these limitations by examining naturally forming study groups within classrooms, using physical proximity as a predictor of interaction intensity.

The remainder of this paper is organized as follows. Section 2 develops a conceptual framework for peer effect identification and proposes underlying mechanisms based on existing literature. Section 3 describes study setting and the measurement of key variables. Section 4 presents instrumental variable methodology and identification strategy. Section 5 reports the estimation results. Section 6 examines potential mechanisms, while Section 7 explores heterogeneous effects. Section 8 concludes the paper.

Conceptual Framework

Peer effect in social interactions

Manski defines three ways in which an individual might be affected by their social interactions, all of which are potentially relevant to the study of peer effects in education (Manski, 1993). These three types of effects are: a.) endogenous effects; b.) exogenous effects; and c.) correlated effects. Under endogenous effects, the propensity of an individual to behave in a certain way varies directly with the prevalence of that behavior in the individual's group. For example, if a student belongs to a high-performing peer group, the student is more likely to have high academic performance as well. Under exogenous effects, the propensity of an individual to behave in a certain way varies with the characteristics of the individual's group. One example of an exogenous effect identified by Manksi is the education level of the parents of peers, which may have an indirect effect on student academic performance but must be carefully distinguished from the direct effect of peer academic performance on student performance. Finally, under correlated effects, individuals in a group behave similarly to one another because they have similar individual characteristics or face similar institutional environments.

Stemming from these three types of effects are three fundamental challenges to identifying the causal effects of the academic performance of peers on student scores (Brock & Durlauf, 2001; Manski, 1993). First, there is a reflection problem, as student academic outcomes may affect the outcomes of their peers and vice versa. Second, similar students tend to join or be assigned to the same group, which is referred to as self-selection bias. Correlation in outcomes and endogenous selection of peers are the primary issues with correlated effects. Third, as a consequence of the reflection problem, individual behaviors may change simultaneously among all interacting agents, making it difficult to separate endogenous effects from correlated effects and exogenous effects (Bramoullé et al., 2020).

Given these challenges, identifying the causality of peer effects using a linear-in-means model can be biased. To address this, we will use peers' academic performance from a previous period relative to the student's current performance to avoid reflection issues. Additionally, we will use relative distance in our predictions to minimize endogenous peer selection.

Investigating the possible mechanisms

Many existing studies explore the mechanisms by which peer effects can operate within a study group. The first of these mechanisms is academic anxiety among peer members when they are studying on math, which can be considered a cost of study group membership. Academic anxiety is defined as the feelings of fear, tension, and apprehension that students may experience when engaging with study materials (Ashcraft, 2002), and specifically math in the context of this study. Academic anxiety is negatively related to math achievement because it disrupts the working memory resources with which students use advanced problem-solving strategies to solve difficult math problems (Ashcraft, 2002; Ramirez et al., 2016). One source of anxiety is social penalties within the group. When observed by peers, students within a group may try to avoid social penalties by conforming to the social norms in the group (Santor, Messervey, & Kusumakar, 2000). For example, under the pressure of social penalties, students with high abilities in mathematics may decide to underachieve in order to avoid social exclusion (e.g. being called names associated with being too good in math) in their schools (Boehnke, 2008). At the same time, however, students with low math performance may also experience academic anxiety if they observe the high performance of other group members and feel excluded because of their low grades.

Students can also experience peer effects within their peer groups via intrinsic or instrumental motivation. Intrinsic motivation refers to engagement in an activity for one's own sake, such as to assimilate with one's social and physical surroundings (Ryan & Deci, 2017). Intrinsic motivation has been widely shown to have strong positive effects on academic achievement (Froiland & Oros, 2014). In contrast, instrumental motivation refers to motivation from future goals and activities that have utility value, such as motivations related to academic grades, career opportunities, financial gains, job promotion, etc. Simons et al. (2000). In China, the highly competitive learning environment and the importance of examinations may increase the influence of instrumental motivation, and students may experience high utility from ranking higher than their classmates (Li & Liu, 2020). Because of this phenomenon, it is particularly important to consider the role of instrumental motivation when studying academic achievement in the context of China.

Finally, peer effects can occur because peer groups often serve as a frame of reference or standard of comparison which can help students to form an academic self-concept (ASC). ASC is defined as the mental representation of one's own abilities (Brunner et al., 2010) and refers to how individuals view themselves in specific academic domains (e.g. Byrne, 1984). When studying within a group, students may compare self-beliefs of their own skills with the perceived skills of other students (Marsh et al., 2015). For example, in a situation in which students are working together on a math assignment, they may observe how many problems their peers can solve and compare such observations with their own achievements. When students perform better than their peers, they may form a higher ACS, which has been found to positively predict academic performance (Altermatt, Pomerantz, Ruble, Frey, & Greulich, 2002; Gest, Rulison, Davidson, & Welsh, 2008; Marsh & Hau, 2003).

In our paper, to explore the mechanisms through which peer effects operate, we will use mediators such as motivation and anxiety, as mentioned above. We will also develop hypotheses based on the current study to guide our investigation.

Data

Sampling

The data for the present study were collected from Shaanxi Province. Among the 31 mainland provinces, Shaanxi ranked 14th in terms of GDP per capita in 2019. The per capita disposable income of rural households in the province was 11,213 RMB in 2018, or approximately 1,495 USD (National Bureau of Statistics, 2019). There are 2.7 million students enrolled in primary school in the province, representing 2.6 percent of all primary school students in China (Ministry of Education, 2019).

The sampling strategy for our survey followed a three-step protocol. First, we restricted our sampling frame to two rural prefectures within Shaanxi province. Second, nine counties in these two prefectures were randomly chosen to be included in our sampling frame, all nine of which are nationally designated poverty counties (People's Daily Online, 2014). Third, we obtained a list of all primary schools from the local bureaus of education in each sample county and selected all schools with more than 10 boarding students per classroom as participants in the survey. We set this condition to ensure that all schools in our sample would have a sufficient number of students studying together in groups after class. Out of the 90 schools that met this condition, the principals of four schools declined to participate, resulting in a final sample of 86 schools. Finally, we randomly selected one fourth, fifth, and sixth grade class in each school, enrolling all students in the sample classes into our study. In total, 3,025 students in 86 schools participated.

Data collection

In China, the academic year is typically divided into two semesters. The first semester starts in September and ends in January of the following year. The second semester begins in February and concludes in June. According to this academic calendar, the longitudinal data used for this study were aggregated from two survey waves conducted during the 2017-2018 school year by the enumeration team. The baseline survey was administered in October 2017, one month after September when no students have left or joined the class during that time and the academic term would be well underway with the same group of students who initially enrolled at the beginning of the semester. Follow-up survey was administered in June 2018. Enumerators received two days of training before visiting sample schools. At each sample school, enumerators followed a strict protocol when administering each part of the survey.

The baseline survey consisted of a 35-minute standardized math exam with strict time limits and a standardized questionnaire consisting of four sections. The first section of the questionnaire measured the math learning attitudes of students. This section included questions about math-related anxiety, self-concept, and intrinsic and instrumental motivation, as well as questions about the amount of time they spent studying, reading, relaxing, and sleeping on a typical weekday.[1] Enumerators also asked teachers how often students studied mathematics with classmates each week and how often students were distracted during class each week, on average. Teachers could respond with: "never," "less than once a week," "twice or three times a week," or "three or four times a week." In the second section, students were asked to list the names of their classroom study partners. The third section collected information about student seating arrangements. The final section collected socioeconomic and demographic data from students and their teachers.

The follow-up survey included a 35-minute standardized math exam and the same questions as the first and second sections of the baseline survey. In addition, teachers were asked to describe the rules and process of seating assignments in their classrooms. The description of data collection protocols for each section can be found in more detail below.

1 Specific questions can be found in Appendix III.

Definitions and measurement of variables

Academic performance

In this study, we take math score as the key measure of a student's academic performance. The math exam administered in this study was grade-appropriate and tailored to the national and provincial-level mathematics curriculum and was constructed by trained psychometricians using a multi-stage process. Exam items were first selected from the standardized mathematics curriculum for each grade. The content validity of these items was checked by multiple experts, including local teachers and professors at Shaanxi Normal University. The psychometric properties of the exam were then validated through extensive pilot testing and data analysis. We standardized math score into z-scores using the mean and standard deviation of each grade for both baseline and follow-up surveys before excluding students who had no study partners in their class.

Math learning attitude

Student attitudes toward math learning were measured in the first section of both the baseline and follow-up surveys. Firstly, math anxiety, self-concept, intrinsic motivation, and instrumental motivation were measured using specially designed and validated items from the 2012 Programme for International Student Assessment (OECD, 2014), which have been widely used as measures of math learning attitudes in many countries (Lee, 2009; Pitsia, Biggart, & Karakolidis, 2017; Thien, Darmawan, & Ong, 2015). In each of the four scales, students responded to items with either "strongly agree," "agree," "disagree," or "strongly disagree." We condensed student responses to the items into four single measures using the GLS weighting procedure described in Anderson (2008). Positive values on the math self-concept index and motivation index corresponded to higher student-reported math self-concept or intrinsic and instrumental motivation than the average student. Meanwhile, positive math anxiety scores indicated that a student's level of math anxiety was higher than that of the average student.

The survey also included three continuous variables about math learning attitude. The first variable was studying time ratio per day, or the percentage of time spent studying compared to the total time spent relaxing, reading, sleeping, and studying. The second and third variables were the average cooperative studying occurrences and the average distraction occurrences, which refer to how often students studied mathematics with classmates or were distracted during class on average each week. A summary of the above student and study group outcomes from the baseline and follow-up surveys is reported in Table 1.

<Place Table 1 about here>

Study groups and their structure

In recent years, network theory has provided a conceptual and empirical framework to analyze peer effects among microenvironments. In subgroups within a classroom, network theory can be utilized to define the characteristics of the group and the study relationships among group members. In the literature, a network is defined as a group or system of interconnected people or things, or more technically, a set of nodes and relations or interconnections between nodes, called links (M. O. Jackson, 2008). In the context of academic peer effects, students are the nodes; the relationship in the course of studying between two students is the link; and the complete set of students in a class and their relationships form a study network within the class. In this way, network theory can be used to describe variations in the make-up of each student's study group beyond those captured in non-network-based peer effect studies. Specifically, network theory makes it possible to investigate the variation and cohesiveness of study groups.

In the second section of the baseline and follow-up surveys, we collected a list of each student's study partners by asking each student to list up to 10 classmates with whom they most frequently studied or discussed math together. For each student surveyed, we generated a study group using all students included in their self-reported study partner list. Utilizing these study partner lists and the classroom maps (see below for a description of the maps), we were able to identify links between pairs of students, as well as networks among individual students and their study groups. Links in this study were undirected according to the definition of network theory (M. O. Jackson, 2008) since the students were not asked to provide a clear direction of the study relationship (e.g., who was asking for help and who was providing it). In this mutual academic support partnership, each study partner could be expected to play the same role, which means each student could either ask for help or provide help. We excluded 69 sample students who did not report any study partners, leaving 2,956 students in our final sample.

After creating the networks, the research team then estimated study group academic diversity and cohesiveness. Diversity is defined as the extent of the variation of academic performance among group members (Berthelon et al., 2019; M. O. Jackson, 2008). To measure variation, we used the standard deviation of math scores of the group members as the indicator of diversity in each study group, taking this as zero when a student had only one study partner. For each student, cohesiveness refers to the extent to which each student's study partners also studied with each other (Berthelon et al., 2019; M. O. Jackson, 2008). Additionally, we created a cohesiveness indicator utilizing the ratio of peers from the student study group who, aside from the student themselves, nominated their study companions as belonging to this group.

Distance between students

In this study, we used relative distance as an instrumental variable to the study relationship between classmates, as relative distance is assumed to be exogenous to academic performance (given that seating charts are assigned exogenously with respect to academic performance, as explained above). In the third section of the baseline survey, enumerators drew a chart that reproduced the distribution of the classroom seating assignments.

Based on the distribution data (Figure A1) collected in this section, we calculated the relative distance between pairs of students using the following steps. First, we created a coordinate plane in the seat distribution table by using the first row of the classroom as the horizontal axis and the first column of the classroom as the vertical axis, taking the distance between two adjacent desks on either axis as one desk. Second, we assigned coordinates (a_i,b_j) to the students according to their seat order relative to the origin, defined as the first seat of the first row. Finally, we calculated two relative distances (Figure 1): the direct distance and the step distance. Direct distance refers to the distance calculated by the Pythagorean theorem, whereas the step distance combines horizontal and vertical distances between two students. When there are two students with coordinates (a_{i1},b_{j1}) and (a_{i2},b_{j2}), their direct distance is calculated by model 1, and their step distance is calculated by model 2.

<Place Figure 1 about here>

$$Direct\ distance = \sqrt{(a_{i1} - a_{i2})^2 + (b_{j1} - b_{j2})^{\ 2}} \tag{1}$$

$$Step\ distance = abs(a_{i1} - a_{i2}) + abs\left(b_{j1} - b_{j2}\right) \tag{2}$$

Control variables

Identifying causality in peer effect studies requires overcoming both the reflection problem and the risk of self-selection bias. Most prior studies have sought to resolve the reflection problem by controlling for observable covariables in their analytical models. One common way to achieve this is to measure the academic ability of peers by prior test scores (Lavy, Paserman, & Schlosser, 2012). At the same time, to reduce selection bias, previous studies have taken several approaches. Some studies control for individual student and school characteristics, such as age, ethnicity, and grade level (Burke & Sass, 2013; Hanushek, Kain, & Rivkin, 2004; Lavy et al., 2012). Therefore, in this study, besides assessing students' initial academic performance, we also gather data on their socioeconomic backgrounds during the data collection phase.

The socioeconomic and demographic data that were collected from students and teachers in the final section of the baseline survey were used as control variables. For students, these control variables included age, gender, grade, boarding status, parental education levels (i.e., whether their father or mother graduated from junior high school), number of siblings, and household asset value. Teacher control variables included age, gender, whether they graduated from at least a two-year college, and teaching experience (in years). Individual and household characteristics are summarized in Table 2.

<Place Table 2 about here>

Econometric Approach

One primary goal of this study was to estimate the effect of study groups on student academic performance. To do so, we used a reduced form model:

$$Score_{i,endline} = \beta_0 + \gamma\overline{G_{-i}} + g_{i,baseline} + \beta_1 X_i + \varepsilon_i \quad (3)$$

where $Score_{i,endline}$ is the standardized mathematics score of student i at the time of the follow-up survey; $\overline{G_{-i}}$ represents the average score of the study group, excluding the student's own math score; X_i is a vector of controls (individual, household, or teacher characteristics); and $g_{i,baseline}$ is the standardized mathematics score of student *i* at baseline. ε_i is a random error term.

To explore the mechanisms and analyze the effect of study groups on student math learning attitude, we substituted $Score_{i,endline}$ in model (3) with $Attitude_{i,baseline}$ (math learning attitude variables at baseline survey) and $Attitude_{i,endline}$ (math learning attitude variables at follow-up survey), respectively. When $Attitude_{i,endline}$was included in the outcome variables, $Attitude_{i,baseline}$ was controlled in the model.

Building on the idea that the intensity of the interaction among peers, the relative distance between students may be able to serve as an instrumental variable for student peer relationships at the subgroup level. Research has shown that classmates that sit closer together have a closer relationship. When individuals are repeatedly exposed to a stimulus, such as the frequent communication that may arise between students when they are seated near one another, the literature shows that such students tend to develop feelings of familiarity and positivity toward their nearby peers (Rhodes, Halberstadt, & Brajkovich, 2001; Van Den Berg & Cillessen, 2015). In terms of peer effects, there is ample evidence that when students are in close physical proximity to one another and often meet each other in an academic context, it is possible that they can influence each other's academic performance. In contrast, a lack of such close contact may prevent this influence (Marmaros & Sacerdote, 2006; Rivera, Soderstrom, & Uzzi, 2010).

To control for endogeneity of network formation, we employed an IV strategy where we instrumented for observed study group average scores $\overline{G_{-i}}$ using predicted study group average scores $\widehat{G_{-i}}$. Referencing a social network model from Berthelon et al. (2019), observed group scores were constructed using the following equation:

$$\overline{G_{-i}} = \frac{\sum_{j \neq i} I_{ij,baseline}\, g_{j,baseline}}{\sum_{j \neq i} I_{ij}} \quad (4)$$

where $g_{j,baseline}$ is the mathematics score of student *j* at baseline, and $I_{ij,baseline}$ indicates whether student *i* and student *j* were study partners at baseline. To generate an instrument for $\overline{G_{-i}}$, we needed exogenous instruments or predetermined measures of $I_{ij,baseline}$. Distance between students served as an instrument of study relationships in this paper. Two-Stage least squares (2SLS) regression analysis (2SLS) with linear first and second stages have been used in this paper.

Relationship (network) formation

To study the probability of relationship formation, we first estimated a model of the determinants of forming a relationship (link) between any two given students, based on the relative distance between them and their predetermined characteristics. As an instrument, relative distance is exogenous to academic performance, as seat assignment generates exogenous variation in the probability of link formation between students that is not correlated with academic performance.

The link formation model in its general form can be represented as follows:

$$I_{ij,baseline} = f(\delta distance_{ij} + \beta M_{ij} + w_{ij}) \quad (5)$$

where $I_{ij,baseline}$ is the indicator variable of a relationship between students *i* and *j*, and is equal to one if i and j study together. $Distance_{ij}$ is the relative distance between two students. M_{ij} are variables measuring shared personal characteristics between students, including whether both students are the same gender (yes/no); whether both are boarding students (yes/no); whether both students have a father or mother who completed junior high school (yes/no); absolute difference of sibling numbers between the two students; absolute difference of family asset value between the two students; and absolute difference of baseline math scores. These variables cannot be affected by student decisions made at school. We took w_{ij} as an independently distributed random disturbance.

We estimated Equation (5)—the relationship model—with a Logit regression and estimated probabilities of relationship formation by predicting $\widehat{I_{ij,\ baseline}}$. Following this estimation strategy, Column 1 of Table 3 shows a significant and negative correlation between direct distance and study relationship: students with a smaller direct distance at baseline were more likely to form a study relationship (row 1). Similarly, Column 2 of Table 3 demonstrates that there is a significant and negative correlation between step distance and study relationship (row 2). Therefore, we conclude that distance is a realizable instrumental variable capable of predicting student relationship formation.

<Place Table 3 about here>

Instruments for study group average score

We constructed instruments for study group average score by using the estimated probability of relationship formation ($\widehat{I_{ij,\ baseline}}$) and the baseline math scores of group members. The model is as follows:

$$\widehat{G_{-i}} = \frac{\sum_{j\neq i} \widehat{I_{ij,\ baseline}}\, g_{j,baseline}}{\sum_{j\neq i} I_{ij}} \tag{6}$$

where $\widehat{G_{-i}}$ is the predicted study group average score for $student_i$, and $g_{j,baseline}$ is the standardized mathematics score of student j at baseline. Overall, this procedure allowed us to predict study group average score based on exogenous probabilities of relationships and predetermined math scores of peers.

During follow-up survey, student responses on how and when teachers determined seating assignments, indicate that some students were seated according to math test scores, with teachers often pairing lower-performing students with higher-performing peers to foster mutual support. Additionally, some classes changed their seating arrangements more frequently than others, which could potentially alter the relative distances between students. Regarding those responses, to valid instruments, we will also conduct robustness check after excluding those samples. Please find details in section 5.2.

Main results

Peer effect on student academic performance

To analyze the effect of study group average score on student academic achievement, we conducted both a simple linear regression (OLS) and an IV analysis. We first estimated Eq. (3) with a simple linear regression (OLS) without accounting for the endogeneity of group formation. Results of this regression are reported in column (1) of Table 4. We found that a 1 SD increase in study group's average score result in a 0.084 SDs increase in the student's score, which suggests that the academic ability level of a student's study group members was positively correlated with the student's own grades (row 1).

We then examined the effect of study groups on student academic performance while controlling for the endogeneity of group formation (columns 2 and 3). When using relative distance to predict the formation of study relationships, the instrument was significant in the first stage, as revealed by F-statistics of 350 and 363, and by the fact that the instrument explained about 59% of the variation in the first stage (rows 8 and 9). The second-stage results indicated that study groups had a significantly positive effect on student academic performance. Specifically, when the average score of a student's study group, as predicted by direct distance, increased by one standard deviation, the final score of the student increased by 0.114 standard deviations (column 2, row 2). Similarly, when the study group average score, as predicted by step distance, increased by one standard deviation, the final score of the student increased by 0.113 SDs (column 3, row 3).

These findings reveal substantial peer effects in academic performance, with a one standard deviation increase in study group average scores leading to a 0.113-0.114 standard deviation improvement in individual student performance. Firstly, the peer effect is quite large, considering it is estimated to be one-fifth of their own effect from baseline academic performance, which corresponds to 0.058 standard deviation. In addition, this effect size is also noteworthy when compared to other educational interventions. It's about half the impact of class size reduction (Krueger, 2003) and one-third the impact of attending a high-performing charter school (Angrist, Pathak, & Walters, 2013). Importantly, this effect is comparable to the impact of replacing a bottom 5% teacher with an average one (Chetty, Friedman, & Rockoff, 2014), similar in magnitude to the effects observed from high-dosage tutoring (Fryer & Howard-Noveck, 2020), and on par with some lightweight interventions, such as text message reminders to parents (York, Loeb, & Doss, 2019).These substantial peer effects highlight the potential of structured peer interactions as a cost-effective, scalable approach to improving student outcomes. Unlike interventions requiring significant resources, leveraging peer effects through study groups could be a powerful tool for enhancing educational performance.

<Place Table 4 about here>

Robustness analysis

Two potential sources of bias exist when we use relative distance as an instrumental variable to estimate the effects of study groups on academic performance, necessitating robustness checks. One potential concern is that relative distance between students may be endogenous to their math test score. This concern arises when teachers assign seats to students based on their academic performance. The second is that relative distance between students may not be constant throughout the academic year. While we collected seat distribution tables in the baseline survey, in some cases, student seating arrangement may be re-assigned between the baseline and follow-up surveys. To address these two concerns, we included questions about seating arrangements in the follow-up survey addressed to both students and teachers. Based on these questions, we test the robust of our results in Table 4 by excluding undesirable samples.

Student seating choice

To address the first concern, in the follow-up survey, we asked students two questions about seating arrangements: 1) Were you allowed to choose your own seat last semester? 2) If not, on what basis did the teacher assign the seating arrangement last semester? The second question was presented in multiple-choice format, with answer choices including height, vision, test score, and student personality. Answers from students are summarized in Table A2 and Table A3. We also asked teachers during the follow-up survey to explain the basis on which they assigned seating.

When students reported being allowed to choose their own seats, we asked their teachers the reason for this decision.

According to student responses to the first question (Table A2), 82% of students were not allowed to choose where they sat. For the 18% who were allowed to choose their seats, teachers gave two reasons as to why students were allowed to choose where to sit. The first reason was vision problems, as teachers let students sit closer to the blackboard if they could not see clearly. Secondly, some teachers also reported that students would ask to change their seats when they had conflicts with desk mates. This may lead to overestimates of study group effects on student academic performance in Table 4, as we did not include possibilities of potentially worse relationships between students in this data.

To test the robustness of the results in Table 4, we excluded students who reported being allowed to choose their own seat and re-estimated the regression using model (3). The results in panel A of Table 5 show that the effects of study groups were qualitatively unchanged, suggesting that our earlier results remained robust.

<Place Table 5 about here>

Indicators of seating arrangements

Student responses to the second question, which asked how their teachers determined seating assignments, are reported Table A3. According to these results, more than half of students (54%) were arranged by their height or vision, which does not bias results directly. As height is mostly influenced by genetic or environmental factors, there is no direct relationship between height and student academic performance (Silventoinen et al., 2007; Dubois et al., 2012). In addition, although there are negative associations between grade and vision because poor vision affects sensory perception, cognition, and school connectedness, teachers arrange for students with poor vision to sit in the front of the classroom, decreasing the negative association (Basch, 2011).

In the rest of the sample classrooms, 11% of students were arranged based on their personality, and 37% of students were arranged based on their math test score. Teachers gave two reasons as to why personality was used as a basis by which to arrange seats, both of which may decrease self-selection bias to some extent. Firstly, teachers often seated outgoing students next to introverted students, as they believed that this arrangement could help students learn from each other and avoid chatting during class. Teachers also often separated students who had a close relationship, as they believed a close relationship could decrease both of their academic performances. Teachers who used test scores as a basis for seating arrangements often did so to encourage students to support one another, usually assigning students with poor performance to sit with better performing desk mates.

To prove the robustness of the main results, we excluded classes that fell into the above two categories from our sample and re-estimated the impact of study groups on student academic performance. According to the results presented in panel B of Table 5, the study group coefficients remained positive and significant, although the results were weaker, at the 5% level. This finding suggests that our previous findings were robust.

Consistency of seating arrangements

To address the potential bias from inconsistent seating arrangements, in the follow-up survey, we asked students another two questions: 1) How often did the class change seats throughout the semester? 2) On what basis did a class change seats throughout the semester? Based on the responses to these two questions, which are summarized in Table A4 and Table A5, we find that there was a low probability for students to change their relative distances to each other. Regarding the first question, 21 classes (19%) did not change seating arrangements during the semester. Among the 93 classes that did change seating arrangements, 31 classes (27%) changed seating arrangements more than once a month throughout the semester (Table A4). The rest of the classes changed seating arrangements less than or equal to once a month (54%). Regarding how seats were changed, there was one class in which seating was changed by groups, and therefore the relative distance between students who sat next to each other did not change (Table A5). In addition, more than half of classes rearranged student seating by row (1%) or column (59%) only. 8% of classes changed seats by row and column together.

For classes that changed seats during the first semester (93 classes), we used student responses to both questions to estimate how long the students maintained the same relative distance to classmates throughout the academic year (Table A6).

According to the estimated results, in 66% of the 93 classes, students did not change in relative distance from classmates until the 18th academic week, since for the most part they still sat near the same classmates even when their location in the classroom changed. However, as there are only 16 academic weeks per semester in rural primary schools, and given that most classes use the same seating arrangement strategy for the whole academic year, we could assume that the relative distance between these students did not change throughout the academic year.

In the remaining 34% of classes, relative distance between any two given students changed after 9 or 10 weeks. To prove the robustness of the previous findings, we excluded the classes in which relative distance changed more frequently and re-evaluated the impact of study groups on student academic performance in the remaining samples. The results of this check (Table 5, panel C) were qualitatively similar to previous findings. Taken together, these checks confirmed that our finding of positive and significant effects of study groups on student academic performance was robust.

Peer effect on student academic performance across terciles

Our results show that being part of a higher-achieving study group boosted the performance of all students. However, existing theoretical studies and empirical research suggest that peer effects may operate nonlinearly depending on the ability of individual students (Burke & Sass, 2013; Han & Li, 2009, 2009; Hoxby & Weingarth, 2005). One example of these nonlinear effects is the “Shining Light model,” which posits that a single exemplary student with exceptional academic outcomes can inspire others to improve their academic achievement (Burke & Sass, 2013; Han & Li, 2009; Hoxby & Weingarth, 2005). According to this model, outstanding students may cause other classmates to improve their achievement.

To explore whether this model was supported by our data, we tested the heterogenous effects of study groups by dividing students into terciles based on their initial ranking within their study group. For a given study group, student performance was designated as "low" if the math score of that student fell within the bottom tercile of the study group, "middle" if their math score lay between the 33rd and 67th percentiles, and "high" if their math score fell within the top tercile. Peer group average score was estimated by weighting step distance.

When dividing students into terciles according to math scores within a study group, we found different results as compared to the overall sample of students. As shown in Table 6, increasing peer group average score had a significant effect on low-performing and middle- performing students, but the effect was not significant for high-performing students. [2,3,4] These results provide a strong argument in favour of

[2] The results remain consistent when predicting math scores with direct distance, though these tables are omitted for brevity. They are available upon request.

[3] A robustness check using quintile cut-offs (bottom 20%, middle 20%-80%, top 20%) shows no substantive changes in results. These tables are omitted but available upon request.

[4] Grouping students by classroom ranking yields similar results, shown in Table A1.

distributing top students relatively evenly across classrooms at the elementary school level, rather than isolating them from other students.

<Place Table 6 about here>

Mechanism analysis

As shown in the existing literature reviewed in Section 2, peer effects from study groups may affect student academic performance by influencing the math learning attitudes of students (Gest et al., 2008; Ramirez et al., 2013; Ryan and Deci, 2000) in different ways. Based on the main results and previous studies, we hypothesized that the positive peer effect is due to improvements in self-concept, motivation, and cooperation, as well as reduced anxiety. Additionally, motivated by high-achieving peers, students may pay more attention in class and spend more time studying. This section will test these hypotheses.

We first ran the same regressions as in equation (3) but substituted math test score with math learning attitudes, as measured in the follow-up surveys. We then conducted a subgroup analysis across different quantiles within group, based on student ranking at baseline, as identification of peer effects on math learning attitudes may depend on the distribution of student math score. As in the previous section, we estimated study group average score using step distance as an instrument.

The results of this analysis are reported in Table 7, which partially support the initial hypothesis, revealing significant effects of peer groups on student performance and learning attuited. For middle-tercile students placed in higher-performing study groups (row 3), significant improvements were observed in reduced anxiety (0.137 standard deviations), self-concept (0.234 standard deviations), and intrinsic motivation (0.224 standard deviations by the end of the academic year. These psychological benefits translated into increased study time, with middle-ranked students in higher-performing groups increasing their daily studying time ratio by 2.6%. Interestingly, intrinsic motivation improved across all terciles (0.108 standard deviations on average) when students were in higher-performing groups. However, contrary to the hypothesis, no improvements were observed in cooperation among students or attention in class. Additionally, top and bottom tercile students did not show significant positive effects from belonging to higher-performing study groups.

These findings suggest a "Goldilocks effect," where middle-tercile students may benefit most from higher-performing peer groups because the challenge level is "just right" - not too easy (as it might be for top-tercile students) and not too overwhelming (as it might be for bottom-tercile students). Also, the observed changes were measured at the end of the academic year, but peer effects among top and bottom students may take time to manifest and persist over a more extended period. The results highlight that peer effects primarily operate through enhanced motivation and self-concept for middle-performing students, leading to increased study time and reduced anxiety. While not all aspects of the hypothesis were confirmed, these insights provide valuable information about the mechanisms of peer effects, which could inform educational policies and practices.

<Place Table 7 about here>

Heterogeneous effects

As shown by the past research summarized in Section 2, several study group characteristics can act as sources of potential heterogeneity in the effects of study groups on student academic performance. These characteristics may include study group structure (including diversity and cohesiveness), gender composition, or student seat distribution in the classroom (Hamilton et al., 2012; Van den Berg et al., 2012; Whitmore, 2005; Johannisson, 2000). Analysis along this line has important policy implications that can help teachers to maximize academic outcomes by taking advantage of peer influence. In this section, we explore whether the effects of study groups on student academic performance differ based on these characteristics, again using step distance to estimate study group average score.

Study group structure

We first examined the heterogeneity of peer effects on student academic performance as differentiated by study group structure (i.e., diversity and cohesiveness). We divided students into two subgroups based on study group diversity at baseline: students with less varied study groups (standard deviation <=0.655)[5] and students with more varied study groups (standard deviation >0.655).

[5] 0.655 is the median of student group diversity for all students.

Similarly, we also divide students into two subgroups based on their cohesiveness: those with less cohesive study groups (cohesiveness <=0.5),[6] and those with more cohesive study groups (cohesiveness >0.5).

Panel A of Table 8 demonstrates that the average study group score has significant positive effects on the academic performance of students in study groups with low or high diversity, but the effect is greater for those in low-diversity study groups (0.131 standard deviation vs 0.119 standard deviation) (row 1, columns 1-2). However, in terms of cohesion, only students in more cohesive study groups experience a positive and significant effect on academic performance, and the effect on students in less cohesive study groups is insignificant (row 1, columns 3-4). One possible explanation for this result is that, within a group, high diversity may raise communication costs about problem-solving and can also diminish productivity if cohesiveness is low (Hamilton et al., 2012). Also, as noted in Amason and Schweiger（1994), although a certain amount of diversity is necessary for improving the quality of strategic decision making, it can also increase the likelihood of group conflict that may impede cooperation among team members. However, combining this finding with the non-linear effects between different terciles of students, these findings argue against strictly tracking student performance, as lower-ranked students can feel more motivated when they are in a high-achieving study group.

[6] 0.5 is the median of student group cohesiveness for all students.

<Place Table 8 about here>

Gender

To explore the heterogeneous effects of gender composition of study groups, we first classified students by gender, and then further divided them into two subgroups based on gender composition: one composed of students in study groups with a low proportion of female members (<=50%),[7] and the other composed of students in study groups with a high proportion of female members (>50%). The results displayed in Panel B of Table 8 show that belonging to a higher performing study group significantly improved academic performance of male students by 0.132standard deviations but had no significant impact on female students (column 1 and 2, row 3). The average study group score also has a positive and more significant impact on students in study groups with a lower proportion of female members, a result that holds for both male and female students (column 3 and 4, row 3and 5).

One possible reason for this is that when studying with partners in a group, males may be more cooperative (Cárdenas, Dreber, von Essen, & Ranehill, 2014). Therefore, gender is an important index for seating arrangements. Rather than assigning gender-homogenous study groups, this finding suggests that teachers should consider assigning higher-performing males to work with students with poor grades,

---

[7] 50% is the median study group female ratio for all students.

as the lower-performing students may benefit more from microenvironments with a higher proportion of males.

Conclusion

Peer relationships fundamentally shape educational outcomes, however, the presence of endogenous effects, exogenous effects, and correlated effects in the interactions between peers has historically made identifying the causality of peer effects difficult (Manski, 1993). To address these identification challenges, we construct instrumental variables using weighted average scores of study groups based on physical distances between students. Our estimations revealed that significantly enhance academic performance, with particularly strong effects for lower-ranked students. These results align with the "Shining Light" theory and recent empirical evidence suggesting that low-performing students benefit substantially from interaction with high-performing peers (Berthelon et al., 2019; Card & Giuliano, 2016; Chin & Kwon, 2022; Hoxby & Weingarth, 2005). Moreover, we find that peer effects are more significant in highly cohesive study groups and among male students.

Our study provides valuable insights into the mechanisms of peer effects in educational settings. Contrary to previous research that highlighted anxiety as a negative consequence of classroom competition (Posselt & Lipson, 2016; Sommet, Pulfrey, & Butera, 2013; Wilkinson & Pickett, 2009), we found no evidence of increased academic anxiety resulting from study groups. In fact, middle-tercile students in higher-achieving study groups experienced reduced anxiety over the academic year. This contrast may stem from the distinct nature of study groups versus classroom-wide competition, where teachers and parents in rural China typically evaluate students based on their relative class ranking (Posselt & Lipson, 2016; Tian, Yu, & Huebner, 2017).

Our analysis identifies two key psychological mechanisms through which study groups influence academic performance: academic self-concept (ASC) and intrinsic motivation. For middle-tercile students, exposure to higher-performing peers improves ASC through the "big-fish-little-pond effect," where students begin to view themselves as similar to their more accomplished peers (Koivuhovi et al., 2022). Additionally, participation in higher-performing study groups enhances intrinsic motivation, a finding that aligns with self-determination theory and suggests lasting benefits of peer interactions. These mechanisms offer important implications for both educational policy and workplace management. In educational settings, carefully structured peer groups can serve as a cost-effective tool for enhancing student motivation, reducing anxiety, and sustainably improving academic performance. Similarly, in organizational contexts, these findings suggest that strategic team composition and office layout design could enhance employee performance through peer effects. Managers could leverage these insights by creating mixed-ability work teams, designing collaborative spaces that facilitate peer interaction, and implementing mentorship programs that pair mid-level employees with high performers. By creating supportive environments that strengthen individual self-concept and internal drive, such interventions may produce more durable improvements in both academic and workplace performance.

We acknowledge several limitations in this study. First, our seat distribution data lacks temporal continuity throughout the academic year. While we collected initial seating information at baseline, we did not gather updated seating data during the academic year, making it difficult to verify if relative distances remained constant. To address the first concerns, we conducted robustness checks by excluding samples where students' relative distances potentially changed due to row or column reassignments during the semester and our main results remain consistent in these restricted samples. Second, our data does not capture the intensity of study partnerships—specifically, the amount of time students spent with each study partner. This limitation is important as study time could affect the strength of peer influence, and should be considered in future research.

Data availability statement

Data available on request due to privacy/ethical restrictions.

Reference

Aghion, P., Howitt, P., Brant-Collett, M., & García-Peñalosa, C. (1998). *Endogenous growth theory*. MIT press.

Altermatt, E. R., Pomerantz, E. M., Ruble, D. N., Frey, K. S., & Greulich, F. K. (2002). Predicting changes in children's self-perceptions of academic competence: A naturalistic examination of evaluative discourse among classmates. *Developmental Psychology*, *38*(6), 903. https://doi.org/10.1037/0012-1649.38.6.903

Amason, A. C., & Schweiger, D. M. (1994). Resolving the paradox of conflict, strategic decision making, and organizational performance. *International Journal of Conflict Management*. https://doi.org/10.1108/eb022745

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Anderson, M. L., & Lu, F. (2017). Learning to Manage and Managing to Learn: The Effects of Student Leadership Service. *Management Science*, *63*(10), 3246–3261. https://doi.org/10.1287/mnsc.2016.2483

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2013). Explaining Charter School Effectiveness. *American Economic Journal: Applied Economics*, *5*(4), 1–27. https://doi.org/10.1257/app.5.4.1

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Back, M. D., Schmukle, S. C., & Egloff, B. (2008). Becoming friends by chance. *Psychological Science*, *19*(5), 439. https://doi.org/10.1111/j.1467-9280.2008.02106.x

Basch, C. E. (2011). Healthier students are better learners: A missing link in school reforms to close the achievement gap. *Journal of School Health*, *81*(10), 593–598.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics*, *34*(2), 143–173. https://doi.org/10.1016/0304-3932(94)90047-7

Berthelon, M., Bettinger, E., Kruger, D. I., & Montecinos-Pearce, A. (2019). The Structure of Peers: The Impact of Peer Networks on Academic Achievement. *Research in Higher Education*, *60*(7), 931–959. https://doi.org/10.1007/s11162-018-09543-7

Boehnke, K. (2008). Peer pressure: A cause of scholastic underachievement? A cross-cultural study of mathematical achievement among German, Canadian, and Israeli middle school students. *Social Psychology of Education*, *11*(2), 149–160. https://doi.org/10.1007/s11218-007-9041-z

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, *150*(1), 41–55. https://doi.org/10.1016/j.jeconom.2008.12.021

Bramoullé, Y., Djebbari, H., & Fortin, B. (2020). Peer Effects in Networks: A Survey. *Annual Review of Economics*, *12*, 603–629.

Brock, W. A., & Durlauf, S. N. (2001). Chapter 54—Interactions-Based Models. In J. J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics* (Vol. 5, pp. 3297–3380). Elsevier. https://doi.org/10.1016/S1573-4412(01)05007-3

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology*, *102*, 964–981. https://doi.org/10.1037/a0019644

Burke, M. A., & Sass, T. R. (2013). Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, *31*(1), 51–82.

Byrne, B. M. (1984). The General/Academic Self-Concept Nomological Network: A Review of Construct Validation Research. *Review of Educational Research*, *54*(3), 427–456.

Card, D., & Giuliano, L. (2016). Can Tracking Raise the Test Scores of High-Ability Minority Students? *American Economic Review*, *106*(10), 2783–2816. https://doi.org/10.1257/aer.20150484

Cárdenas, J.-C., Dreber, A., von Essen, E., & Ranehill, E. (2014). Gender and cooperation in children: Experiments in Colombia and Sweden. *PloS One*, *9*(3), e90923. https://doi.org/10.1371/journal.pone.0090923

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, *81*(3), 855–882. https://doi.org/10.3982/ECTA10168

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, *104*(9), 2593–2632. https://doi.org/10.1257/aer.104.9.2593

Chin, S., & Kwon, E. (2022). Learning with Differing-Ability Peers: Evidence from a Natural Experiment in South Korea. *The BE Journal of Economic Analysis & Policy*. https://doi.org/10.1515/bejeap-2021-0306

Deci, E. L. (1973). Paying people doesn’t always work the way you expect it to. *Human Resource Management*, *12*(2), 28–32. https://doi.org/10.1002/hrm.3930120205

Epple, D., & Romano, R. E. (2011). Chapter 20 - Peer Effects in Education: A Survey of the Theory and Evidence. In J. Benhabib, A. Bisin, & M. O. Jackson (Eds.), *Handbook of Social Economics* (Vol. 1, pp. 1053–1163). North-Holland. https://doi.org/10.1016/B978-0-444-53707-2.00003-7

Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics*, *90*(8–9), 1455–1475. https://doi.org/10.1016/j.jpubeco.2005.12.001

Froiland, J. M., & Oros, E. (2014). Intrinsic motivation, perceived competence and classroom engagement as longitudinal predictors of adolescent reading achievement. *Educational Psychology*, *34*(2), 119–132. https://doi.org/10.1080/01443410.2013.822964

Fryer, R. G., & Howard-Noveck, M. (2020). High-Dosage Tutoring and Reading Achievement: Evidence from New York City. *Journal of Labor Economics*, *38*(2), 421–452.

Gest, S. D., Rulison, K. L., Davidson, A. J., & Welsh, J. A. (2008). A reputation for success (or failure): The association of peer academic reputations with academic self-concept, effort, and performance across the upper elementary grades. *Developmental Psychology*, *44*, 625–636. https://doi.org/10.1037/0012-1649.44.3.625

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2012). Diversity and productivity in production teams. In *Advances in the Economic Analysis of participatory and Labor-managed Firms*. Emerald Group Publishing Limited.

Han, L., & Li, T. (2009). The gender difference of peer influence in higher education. *Economics of Education Review*, *28*(1), 129–134. https://doi.org/10.1016/j.econedurev.2007.12.002

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, *88*(9–10), 1721–1746. https://doi.org/10.1016/S0047-2727(03)00063-X

Hare, A. P., & Bales, R. F. (1963). Seating position and small group interaction. *Sociometry*, 480–486. https://doi.org/10.2307/2786150

Hinke, S., Leckie, G., & Nicoletti, C. (2019). The Use of Instrumental Variables in Peer Effects Models. *Oxford Bulletin of Economics and Statistics*, *81*(5), 1179–1191. https://doi.org/10.1111/obes.12299

Hoxby, C. M., & Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects*. Citeseer.

Jackson, C. K., & Bruegmann, E. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, *1*(4), 85–108. https://doi.org/10.1257/app.1.4.85

Jackson, M. O. (2008). Average distance, diameter, and clustering in social networks with homophily. *International Workshop on Internet and Network Economics*, 4–11. Springer. https://doi.org/10.1007/978-3-540-92185-1_3

Koivuhovi, S., Marsh, H. W., Dicke, T., Sahdra, B., Guo, J., Parker, P. D., & Vainikainen, M.-P. (2022). Academic self-concept formation and peer-group contagion: Development of the big-fish-little-pond effect in primary-school classrooms and peer groups. *Journal of Educational Psychology*, *114*(1), 198. https://doi.org/10.1037/edu0000554

Krueger, A. B. (2003). ECONOMIC CONSIDERATIONS AND CLASS SIZE. *THE ECONOMIC JOURNAL*.

Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers In the Classroom. *The Economic Journal*, *122*(559), 208–237. https://doi.org/10.1111/j.1468-0297.2011.02463.x

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, *19*(3), 355–365. https://doi.org/10.1016/j.lindif.2008.10.009

Li, L., & Liu, Y. (2020). An integrated model of principal transformational leadership and teacher leadership that is related to teacher self-efficacy and student academic performance. *Asia Pacific Journal of Education*, 1–18.

Lyle, D. S. (2007). Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point. *Review of Economics and Statistics*, *89*(2), 289–299. https://doi.org/10.1162/rest.89.2.289

Mankiw, N. G., Romer, D., & N. Weil, D. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), 407–437.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531–542. https://doi.org/10.2307/2298123

Marmaros, D., & Sacerdote, B. (2006). How do friendships form? *The Quarterly Journal of Economics*, *121*(1), 79–119.

Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., Nagengast, B., … Abu-Hilal, M. M. (2015). The Internal/External Frame of Reference Model of Self-Concept and Achievement Relations: Age-Cohort and Cross-Cultural Differences. *American Educational Research Journal*, *52*(1), 168–202. https://doi.org/10.3102/0002831214549453

Marsh, H. W., & Hau, K.-T. (2003). Big-Fish–Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, *58*(5), 364. https://doi.org/10.1037/0003-066X.58.5.364

McAndrew, F. T. (1993). *Environmental psychology.* Thomson Brooks/Cole Publishing Co.

McKeown, S., Stringer, M., & Cairns, E. (2016). Classroom segregation: Where do students sit and how is this related to group relations? *British Educational Research Journal*, *42*(1), 40–55. https://doi.org/10.1002/berj.3200

Min, S., Yuan, Z., Wang, X., & Hou, L. (2019). Do peer effects influence the academic performance of rural students at private migrant schools in China? *China Economic Review*, *54*, 418–433. https://doi.org/10.1016/j.chieco.2019.02.004

Mora, T., & Oreopoulos, P. (2011). Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, *30*(4), 575–581. https://doi.org/10.1016/j.econedurev.2011.01.004

National Bureau of Statistics. (2019). *SHAANXI STATISTICAL YEARBOOK 2019*. SHANNXI: China Statistical Press.

OECD. (2014). *PISA 2012 technical report*. Retrieved from www.oecd.org/pisa/pisaproducts/PISA%2020 12%20Technical%20Report_Chapter%203.pdf

People's Daily Online. (2014, October 17). China's first poverty alleviation day, list of 592 national-level poverty-stricken countiesPeople's Daily Online. Retrieved June 8, 2023, from Http://politics.people.com.cn/n/2014/1017/c1026-25854065.html website: http://politics.people.com.cn/n/2014/1017/c1026-25854065.html

Pitsia, V., Biggart, A., & Karakolidis, A. (2017). The role of students' self-beliefs, motivation and attitudes in predicting mathematics achievement: A multilevel analysis of the Programme for International Student Assessment data. *Learning and Individual Differences*, *55*, 163–173. https://doi.org/10.1016/j.lindif.2017.03.014

Posselt, J. R., & Lipson, S. K. (2016). Competition, anxiety, and depression in the college classroom: Variations by student identity and field of study. *Journal of College Student Development*, *57*(8), 973–989. https://doi.org/10.1353/csd.2016.0094

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology*, *141*, 83–100. https://doi.org/10.1016/j.jecp.2015.07.014

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, *14*(2), 187–202. https://doi.org/10.1080/15248372.2012.664593

Rhodes, G., Halberstadt, J., & Brajkovich, G. (2001). Generalization of Mere Exposure Effects to Averaged Composite Faces. *Social Cognition*, *19*(1), 57–70. https://doi.org/10.1521/soco.19.1.57.18961

Rivera, M. T., Soderstrom, S. B., & Uzzi, B. (2010). Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. *Annual Review of Sociology*, *36*(1), 91–115. https://doi.org/10.1146/annurev.soc.34.040507.134743

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5, Part 2), S71–S102.

Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. 370 Seventh Avenue, Suite 1200, New York, NY 10001: The Guilford Press. Retrieved from https://doi.org/10.1521/978.14625/28806

Santor, D. A., Messervey, D., & Kusumakar, V. (2000). Measuring peer pressure, popularity, and conformity in adolescent boys and girls: Predicting school performance, sexual attitudes, and substance abuse. *Journal of Youth and Adolescence*, *29*(2), 163–182. https://doi.org/10.1023/A:1005152515264

Silventoinen, K., Bartels, M., Posthuma, D., Estourgie-van Burk, G. F., Willemsen, G., van Beijsterveldt, T. C., & Boomsma, D. I. (2007). Genetic regulation of growth in height and weight from 3 to 12 years of age: A longitudinal study of

Dutch twin children. *Twin Research and Human Genetics*, *10*(2), 354–363. https://doi.org/10.1375/twin.10.2.354

Simons, J., Dewitte, S., & Lens, W. (2000). Wanting to have vs. wanting to be: The effect of perceived instrumentality on goal orientation. *British Journal of Psychology*, *91*(3), 335–351. https://doi.org/10.1348/000712600161862

Sommet, N., Pulfrey, C., & Butera, F. (2013). Did my MD really go to university to learn? Detrimental effects of numerus clausus on self-efficacy, mastery goals and learning. *PloS One*, *8*(12), e84178. https://doi.org/10.1371/journal.pone.0084178

Thien, L. M., Darmawan, I., & Ong, M. Y. (2015). Affective characteristics and mathematics performance in Indonesia, Malaysia, and Thailand: What can PISA 2012 data tell us? *Large-Scale Assessments in Education*, *3*(1), 1–16. https://doi.org/10.1186/s40536-015-0013-z

Tian, L., Yu, T., & Huebner, E. S. (2017). Achievement goal orientations and adolescents' subjective well-being in school: The mediating roles of academic social comparison directions. *Frontiers in Psychology*, *8*, 37. https://doi.org/10.3389/fpsyg.2017.00037

Van Den Berg, Y. H. M., & Cillessen, A. H. N. (2015). Peer status and classroom seating arrangements: A social relations analysis. *Journal of Experimental Child Psychology*, *130*, 19–34. https://doi.org/10.1016/j.jecp.2014.09.007

Van den Berg, Y. H., Segers, E., & Cillessen, A. H. (2012). Changing peer perceptions and victimization through classroom arrangements: A field experiment. *Journal of Abnormal Child Psychology*, *40*(3), 403–412. https://doi.org/10.1007/s10802-011-9567-6

Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. *American Economic Review*, *95*(2), 199–203. https://doi.org/10.1257/000282805774670158

Wilkinson, R. G., & Pickett, K. E. (2009). Income inequality and social dysfunction. *Annual Review of Sociology*, 493–511. https://doi.org/10.1146/annurev-soc-070308-115926

York, B. N., Loeb, S., & Doss, C. (2019). One Step at a Time: The Effects of an Early Literacy Text-Messaging Program for Parents of Preschoolers. *Journal of Human Resources*, *54*(3), 537–566. https://doi.org/10.3368/jhr.54.3.0517-8756R

Table 1. Student and study group math-related outcomes at baseline and follow-up survey.

		Mean	Std. Dev.	Max.	Min.
Study group measures					
[1]	Average math score of study group at baseline (excluding self)	0.171	0.651	1.907	-3.198
[2]	Average math score of study group at baseline (excluding self, weighted by direct distance)	0.256	0.876	3.855	-5.201
[3]	Average math score of study group at baseline (excluding self, weighted by step distance)	0.254	0.874	3.818	-5.189
Student outcomes at baseline survey					
[4]	Standardized math test score	0.035	0.987	2.023	-4.463
[5]	Anxiety	0.012	0.687	1.920	-1.468
[6]	Self-concept	-0.013	0.675	1.628	-2.011
[7]	Intrinsic motivation	-0.011	0.739	1.394	-2.396
[8]	Instrument motivation	-0.004	0.758	1.034	-3.269
[9]	Frequency of cooperation with classmates	3.057	1.219	5.000	1.000
[10]	Frequency of distraction in class	2.728	1.337	5.000	1.000
[11]	Studying time ratio per day	0.315	0.165	1.000	0.000
Student outcomes at follow-up survey					
[12]	Standardized math test score	0.042	0.983	2.000	-3.611
[13]	Anxiety	0.029	0.739	1.906	-1.433
[14]	Self-concept	-0.029	0.737	1.643	-1.926
[15]	Intrinsic motivation	-0.024	0.785	1.450	-2.311

[16]	Instrument motivation	-0.010	0.800	0.979	-3.470
[17]	Frequency of cooperation with classmates	3.004	1.210	5.000	1.000
[18]	Frequency of distraction in class	2.786	1.282	5.000	1.000
[19]	Studying time ratio per day	0.275	0.154	1.000	0.000
[20]	Observations	2,956			
[21]	Number of classes	114			

Note: Standardized math test scores, anxiety, self-concept, intrinsic motivation, and instrument motivation are reported as z-scores (mean = 0, standard deviation = 1) for the full sample. Frequency measures are on a 1-5 scale. Studying time ratio represents the proportion of a day spent studying.

Table 2. Descriptive statistics.

		Mean	Std. Dev.	Max.	Min.
Demographic and family characteristics at baseline					
[1]	Age (years)	10.949	1.044	17.250	5.917
[2]	Gender (1=male; 0=female)	0.527	0.499	1.000	0.000
[3]	Boarding (1=yes; 0=no)	0.380	0.485	1.000	0.000
[4]	Number of family members	4.949	1.460	15.000	2.000
[5]	Mother graduated from junior high school (1=yes; 0=no)	0.436	0.496	1.000	0.000
[6]	Father graduated from junior high school (1=yes; 0=no)	0.506	0.500	1.000	0.000
[7]	Standardized family asset value	-0.007	1.625	4.890	-2.156
[8]	Fourth grade	0.247	0.431	1.000	0.000
[9]	Fifth grade	0.369	0.483	1.000	0.000
[10]	Sixth grade	0.384	0.486	1.000	0.000
Teacher characteristics at baseline					
[11]	Age (years)	36.905	9.712	59.083	22.667
[12]	Gender (1=female; 0=male)	0.492	0.500	1.000	0.000
[13]	Teacher graduated from junior college (1=yes, 0=no)	0.575	0.494	1.000	0.000
[14]	Teaching experience (years)	15.555	11.606	42.000	0.000
Study partner numbers and average distance at baseline					
[15]	Number of students in the class	31.679	11.721	52.000	5.000
[16]	Study partner numbers	4.683	2.621	10.000	1.000
[17]	Average distance between student and study partner (direct distance)	2.714	1.063	8.062	1.000

[18]	Average distance between student and study partner (step distance)	3.352	1.379	11.000	1.000
[19]	Observations	2,956			
[20]	Number of classes	114			

Note: Age is reported in years. Binary variables (gender, boarding, parental education, and grade levels) are coded as 1 for the stated condition and 0 otherwise. Family asset value is standardized across the sample. All means, standard deviations, maximums, and minimums are calculated based on individual-level data, except for teacher characteristics and class size, which are at the class level.

Table 3. Estimates of the probability of a study relationship between two students (Logit).

VARIABLES		Dependent variable: study relationship=1	
		(1)	(2)
[1]	Direct distance	-0.035***	
		(0.001)	
[2]	Step distance		-0.026***
			(0.001)
[4]	Pseudo R^2	0.182	0.182
[5]	Observations	97,732	
[6]	Number of students	2,956	
[7]	Number of classes	114	

Note: This table presents logit estimates of the probability of a study relationship between two students. The model controls for shared personal characteristics between students, including whether both students were the same gender (yes/no); whether both were boarding students (yes/no); whether both students had a father or mother who completed junior high school (yes/no); absolute difference of sibling numbers between the two students; absolute difference of family asset value between the two students; and absolute difference of baseline math scores. County fixed effects are included in all estimates. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Effect of study group average score on student academic performance.

	VARIABLES	Dependent variable: Standardized math test score at follow-up survey		
		(1)	(2)	(3)
		Ordinary least squares (OLS)	Instrumental Variable (IV)	
[1]	Study group baseline average score (excluding self)	0.084***		
		(0.025)		
[2]	Study group baseline average score (excluding self, direct distance)		0.114***	
			(0.035)	
[3]	Study group baseline average score (excluding self, step distance)			0.113***
				(0.035)
[4]	Student baseline math score	0.564***	0.558***	0.558***
		(0.020)	(0.021)	(0.021)
[5]	Observations	2,956		
[6]	Adjusted R-squared	0.408	0.408	0.408
[7]	First-stage statistics			
[8]	R-squared		0.586	0.587
[9]	Partial R^2 excluded instrument		0.459	0.460
[10]	*F*-test		349.754	363.034

Note: This table presents estimates of the effect of study group average score on individual student math performance at follow-up. Column (1) shows OLS estimates, while columns (2) and (3) present IV estimates using direct distance and step distance as instruments, respectively. Each regression controls for baseline math scores, student and teacher characteristics as detailed in Table 2, as well as county fixed effects. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. The first-stage statistics for IV regressions are reported at the bottom of the table. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5. Robustness analysis

VARIABLES	Dependent variable: Standardized math test score at follow-up survey	
	(1)	(2)
	Instrumental Variable (direct distance)	Instrumental Variable (step distance)
Panel A: Excluding students who chose their own seats		
[1] Study group baseline average score	0.085**	0.085**
(excluding self, step distance)	(0.041)	(0.042)
[2] Observations	2362	2362
[3] Adjusted R-squared	0.380	0.380
Panel B: Excluding students seated based on math test score or personality		
[4] Study group baseline average score	0.082**	0.082**
(excluding self, step distance)	(0.045)	(0.045)
[5] Observations	1853	1853
[6] Adjusted R-squared	0.433	0.433
Panel C: Excluding classes in which relative distance between students was not constant throughout the academic year		
[7] Study group baseline average score	0.091**	0.089**
(excluding self, step distance)	(0.041)	(0.041)
[8] Observations	2085	2085
[9] Adjusted R-squared	0.402	0.402

Note: This table presents robustness checks for the effect of study group average score on individual student math performance at follow-up. All models use Instrumental Variable (IV) estimation with direct distance (column 1) and step distance (column 2) as instruments. Panel A excludes students who chose their own seats. Panel B excludes students seated based on math test scores or personality. Panel C excludes classes where relative distance between students was not constant throughout the academic year. Each regression controls for baseline math scores, student and teacher characteristics as detailed in Table 2, as well as county fixed effects. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Effect of study group average score on student academic performance, as differentiated by student baseline score ranking within study group (IV estimation).

	VARIABLES	Dependent variable: Standardized math test score at follow-up survey		
		(1)	(2)	(3)
		Student is in bottom tercile (0/1)	Student is in middle tercile (0/1)	Student is in top tercile (0/1)
[1]	Study group baseline average score (excluding self, step distance)	0.201*	0.141*	-0.041
		(0.105)	(0.081)	(0.081)
[2]	Student baseline math score	0.413***	0.603***	0.782***
		(0.061)	(0.050)	(0.063)
[3]	Observations	916	1,050	990
[4]	Adjusted R-squared	0.279	0.399	0.369

Note: This table presents IV estimates of the effect of study group average score on individual student math performance at follow-up, differentiated by the student's baseline score ranking within their study group. Each column represents a different tercile of baseline performance: bottom (1), middle (2), and top (3). The dependent variable is the standardized math test score at the follow-up survey. All models control for baseline math scores, student and teacher characteristics as detailed in Table 2, as well as county fixed effects. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Impact of study group average score on students' math learning attitude at follow-up (IV estimation across tercile subgroups).

VARIABLES	Dependent variable: math learning attitude						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Academic anxiety	Self-concept	Intrinsic motivation	Instrumental motivation	Cooperation frequency	Distraction frequency	Studying time ratio
[1] Baseline study group average score (all samples)	-0.016	0.034	0.108**	0.051	-0.059	0.072	0.008
	(0.037)	(0.035)	(0.049)	(0.043)	(0.099)	(0.105)	(0.008)
[2] Baseline study group average score (bottom tercile)	0.067	-0.068	0.004	0.025	-0.271	0.206	-0.018
	(0.080)	(0.074)	(0.100)	(0.126)	(0.208)	(0.211)	(0.019)
[3] Baseline study group average score (middle tercile)	-0.137*	0.234***	0.224**	0.037	-0.075	0.174	0.026*
	(0.083)	(0.085)	(0.094)	(0.091)	(0.148)	(0.182)	(0.016)
[4] Baseline study group average score (top tercile)	0.077	-0.084	0.023	-0.004	-0.028	0.142	0.005
	(0.087)	(0.080)	(0.084)	(0.078)	(0.174)	(0.153)	(0.017)

Note: This table presents IV estimates of the effect of study group average score on various math learning attitudes at follow-up. Each column represents a different attitude measure: (1) Academic anxiety, (2) Self-concept, (3) Intrinsic motivation, (4) Instrumental motivation, (5) Cooperation frequency, (6) Distraction frequency, and (7) Studying time ratio. Results are shown for the full sample [1] and by baseline performance terciles [2-4]. Each regression controls for student and teacher characteristics and baseline math scores as detailed in Table 2, as well as county fixed effects. Study group baseline average score is calculated using step distance, excluding the student's own score. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. The sample includes 2,956 students (916 bottom tercile, 1,050 middle tercile, 990 top tercile). *** p<0.01, ** p<0.05, * p<0.1.

Table 8. Heterogeneity analysis

VARIABLES	Dependent variable: Standardized math test score at follow-up survey			
Panel A: Study group structure	(1)	(2)	(3)	(4)
	Low diversity	High diversity	Low cohesive	High cohesive
[1] Study group baseline average score	0.131***	0.119**	0.033	0.179***
	(0.048)	(0.055)	(0.048)	(0.051)
[2] Observations	1483	1473	1480	1476
Panel B: Student gender and study group female ratio	(1)	(2)	(3)	(4)
	Male	Female	Low female ratio	High female ratio
[3] Study group baseline average score	0.132***	0.079	0.136***	0.078
	(0.044)	(0.060)	(0.042)	(0.064)
[4] Observations	1558	1398	1578	1378
Panel C: Study group female ratio by student gender	(1)	(2)	(3)	(4)
	Male		Female	
	Low female ratio	High female ratio	Low female ratio	High female ratio
[5] Study group baseline average score	0.134***	-0.658	0.363*	0.079
	(0.043)	(0.583)	(0.195)	(0.064)
[6] Observations	1519	39	59	1339

Note: This table presents IV estimates of the effect of study group average score on standardized math test scores at follow-up, examining heterogeneity across different subgroups. Panel A analyzes study group structure (diversity and cohesiveness), Panel B examines student gender and study group female ratio, and Panel C explores the interaction between student gender and study group female ratio. Each regression controls for student and teacher characteristics and baseline math scores as detailed in Table 2, as well as county fixed effects. Study group baseline average score is calculated using step distance, excluding the student's own score. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix to:

From Seating to Success:

How Physical Distance in Social Networks Shape Individual Achievement

Appendix I: Appendix figures

Figure 1: Seat distribution table

Class Id: **Teachers Name:**

Podium

Door

1		2		3		4		5		6		7		8	
Name	ID	Name	ID	Name	ID	Name	ID	Name	ID	Name	ID	Name	ID	Name	ID

Note: The first column starts from door, and the first row starts from podium

Appendix II: Appendix tables

Table A1: Effect of study group average score on student academic performance, as differentiated by student baseline score ranking within class (IV estimation).

VARIABLES	Dependent variable: Standardized math test score at follow-up survey		
	(1)	(2)	(3)
	Student is in bottom tercile (0/1)	Student is in middle tercile (0/1)	Student is in top tercile (0/1)
[1] Study group baseline average math score (excluding self, step distance)	0.179***	0.142*	0.083
	(0.067)	(0.079)	(0.051)
[2] Standardized baseline math score	0.410***	0.503***	0.599***
	(0.047)	(0.094)	(0.063)
[3] Observations	1,087	948	921
[4] Adjusted R-squared	0.249	0.246	0.228

Note: This table presents IV estimates of the effect of study group average score on standardized math test scores at follow-up, differentiated by the student's baseline score ranking within their class. Each column represents a different tercile of baseline performance within the class: bottom (1), middle (2), and top (3). The main independent variable is the study group baseline average math score, calculated using step distance and excluding the student's own score. Each regression controls for student and teacher characteristics and baseline math scores as detailed in Table 2, as well as county fixed effects. Robust standard errors, adjusted for clustering at the class level, are shown in parentheses. The sample includes 2,956 students (1,087 bottom tercile, 948 middle tercile, 921 top tercile). *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Frequency of students choosing their own seats.

	Frequency	Percentage
No	2,362	81.14
Yes	549	18.86
Total	2,911	100.00

Sources: *Based on student responses. 45 students did not respond.*

Table A3: Student-reported seating arrangement criteria in the first semester (multiple-choice question).

	Frequency	Percentage
Height	1,587	53.69
Vision	1,448	48.99
Score	1,103	37.31
Student personality	320	10.83
Total	2,956	100.00

Sources: Based on student responses.

Table A4: How often did the seating arrangement change in the first semester? (Student responses).

	Frequency	Percentage
Never	21	18.42
> 4 weeks	31	27.19
4 weeks	33	28.95
3 weeks	4	3.51
2 weeks	11	9.65
1 week	14	12.28
Total	114	100

Sources: Based on the mode of student responses.

Table A5: How did seating arrangements change in the first semester? (Student responses).

	Frequency	Percentage
By row	1	1.09
By column	54	58.70
By row and column	35	38.04
By group	1	1.09
No rules	1	1.09
Total	92	100

Sources: Based on the mode of student responses. 21 classes of students did not change seats during the last semester, and 1 class had no answer.

Table A6: Projection of consecutive weeks student would spend with the same group of nearby classmates in the first semester.

Week	Frequency	Percent
9	30	33.33
10	1	1.11
18	33	36.67
27	4	4.44
36	22	24.44
Total	90	100.00

Sources: Based on the mode of student responses. 21 classes of students did not change seats during the last semester, and 3 classes had no answer.

Appendix III: Items of math learning attitude scales

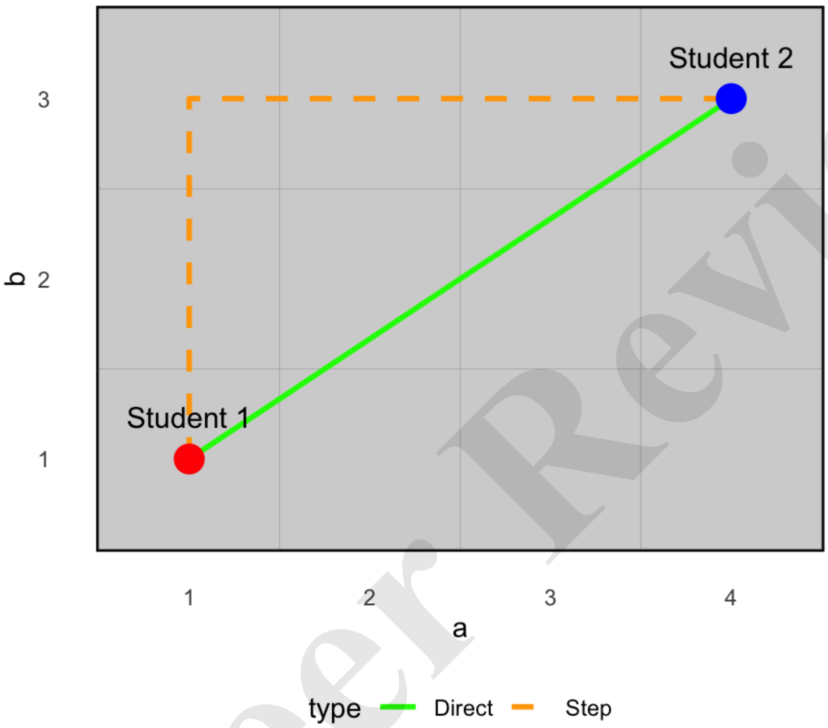
Anxiety included 5 items: a) I often worry that it will be difficult for me in mathematics classes; b) I get very tense when I have to do mathematics homework; c) I get very nervous doing mathematics problems; d) I feel helpless when doing a mathematics problem; e) I worry that I will get poor grades in mathematics.

Self-concept included 5 items: a) I am just not good at mathematics; b) I get good grades in mathematics; c) I learn mathematics quickly; d) I have always believed that mathematics is one of my best subjects; e) In my mathematics class, I understand even the most difficult work.

Intrinsic motivation included 4 items: a) I enjoy reading about mathematics; b) I look forward to my mathematics lessons; c) I do mathematics because I enjoy it; d) I am interested in the things I learn in mathematics.

Instrumental motivation included 4 items: a) Making an effort in mathematics is worth it because it will help me in the work that I want to do later on; b) Learning mathematics is worthwhile for me because it will improve my career prospects and chance; c) Mathematics is an important subject for me because I need it for what I want to study later on; d) I will learn many things in mathematics that will help me get a job.

Figure 1. Calculation of direct distance and step distance



Note: Figure 1 illustrates the calculation of direct distance and step distance between students in a classroom setting. The direct distance (solid green line) is calculated using the Euclidean distance formula: $Direct\ distance = \sqrt{(a_{i1} - a_{i2})^2 + (b_{j1} - b_{j2})^2}$, where (a_{i1}, b_{j1}) and (a_{i2}, b_{j2}) are the coordinates of two students. This represents the straight-line distance between two students, regardless of seating arrangement. The step distance (dashed orange line) is calculated using the Manhattan distance formula: $Step\ distance = abs(a_{i1} - a_{i2}) + abs(b_{j1} - b_{j2})$. This represents the path a student would take to reach another student by moving along rows and columns of desks.