

Online Appendices for “Student Performance, Peer Effects, and Friend Networks: Evidence from a Randomized Peer Intervention”

Jia Wu, Junsen Zhang, and Chunchao Wang*

September 3, 2021

* Jia Wu, Department of Economics, Shandong University; Weihai Institute for Interdisciplinary Research, Weihai, China; Junsen Zhang, Corresponding Author, School of Economics, Zhejiang University, Hangzhou, China; Department of Economics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (email: jszhang@cuhk.edu.hk). Chunchao Wang, Department of Economics, Jinan University; Jinan University–University of Birmingham Joint Institute, Guangzhou, China.

I. Introduction of the Longhui county

Longhui is a relatively poor rural area. According to the National Bureau of Statistics of China, the county's per capita GDP was 12,466 CNY (1,945 USD) in 2016, which was less than a quarter of the national level (53,980 CNY or 8,487 USD). However, the educational resource of the county is close to the national average. For example, according to the Longhui Statistical Yearbook 2015, the enrollment rate was universal (100%) for primary school and almost universal (99.2%) for middle school due to the nine-year compulsory education in China. About three-quarters of students (74.7%) proceeded to study in high school in Longhui. The student–fulltime teacher ratio is 17.2, indicating that an average of 17 students are instructed by one teacher, which is similar to the national level (17.05). In 2016, the average public expenditure per incumbent student in Longhui was 5,770 CNY (907 USD), which was higher than the average level in China (4,408 CNY, or 693 USD). For the college admission rate, approximately 12.5% of high school graduates enrolled in tier 1 universities, which is similar to the national average (12.9%). In summary, the education resource and teaching conditions of primary schools in our experimental county are similar to the national averages.

II. Tables and Figures

TABLE A1—Comparison Between Non-randomly Assigned Students and Compliers

	Non-randomly assigned students				Compliers
	All	Non-compliers	Non-compliers' original assigned deskmates	Extended non-compliers	
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Achievement and characteristics</i>					
Chinese test score	66.61 (19.23)	69.28 (16.75)	64.99 (21.56)	66.09 (19.27)	66.77 (18.98)
Mathematic test score	56.36 (25.40)	60.92 (25.09)	57.54 (23.51)	55.50 (26.38)	57.09 (24.03)
Male	0.57 (0.50)	0.39 (0.50)	0.63 (0.49)	0.62 (0.49)	0.57 (0.50)
Age	8.55 (0.91)	8.66 (0.97)	8.53 (0.89)	8.50 (0.90)	8.65 (0.89)
Height (inch)	51.52 (4.56)	51.09 (4.71)	51.42 (4.05)	51.79 (4.75)	52.09 (4.12)
Weight (kg)	25.14 (5.04)	25.86 (5.53)	24.38 (3.88)	25.16 (5.30)	25.75 (6.17)
Non-minority (yes=1)	0.06 (0.24)	0.11 (0.31)	0.05 (0.23)	0.04 (0.20)	0.04 (0.21)
<i>Hukou</i> registration status (urban <i>hukou</i> =1)	0.99 (0.08)	1.00 (0.00)	1.00 (0.00)	0.99 (0.11)	0.99 (0.09)
Interest in Chinese (yes=1)	0.57 (0.50)	0.55 (0.50)	0.66 (0.48)	0.53 (0.50)	0.60 (0.49)
Interest in mathematics (yes=1)	0.52 (0.50)	0.55 (0.50)	0.61 (0.50)	0.46 (0.50)	0.56 (0.50)
Father's education (# of years)	9.61 (2.40)	9.24 (2.45)	9.63 (2.62)	9.78 (2.27)	9.74 (2.55)
Mother's education (# of years)	8.87 (2.38)	8.37 (2.43)	8.61 (2.43)	9.25 (2.29)	9.16 (2.51)
Student's household has a computer (yes=1)	0.28 (0.45)	0.16 (0.37)	0.34 (0.48)	0.30 (0.46)	0.29 (0.45)
Student's household has a car (yes=1)	0.17 (0.38)	0.13 (0.34)	0.24 (0.43)	0.16 (0.37)	0.17 (0.38)
<i>Panel B. Big Five personality traits</i>					
Neuroticism	39.13 (5.20)	40.34 (5.38)	38.03 (4.37)	39.08 (5.42)	38.60 (5.54)
Extraversion	38.43 (5.97)	38.47 (5.98)	37.79 (5.82)	38.74 (6.10)	38.84 (5.62)
Openness	40.62 (4.60)	40.39 (4.99)	40.24 (3.91)	40.92 (4.75)	40.66 (4.30)
Agreeableness	34.08 (7.91)	34.37 (7.79)	32.37 (7.40)	34.79 (8.19)	34.95 (7.23)
Conscientiousness	38.84 (5.72)	38.61 (5.67)	37.53 (5.92)	39.62 (5.60)	38.65 (5.64)
Number of individuals	152	38	38	76	1,650

Notes: This table reports the mean and standard deviation of the variables. The sample in column 1 includes students who were not randomly allocated. Samples in columns 2 and 3 include students who rejected the original seating arrangement and their original deskmates. The sample in column 4 includes non-compliers' new deskmates and students who eventually sat with non-compliers' original deskmates (i.e., 74 students whose seats were reassigned due to the 37 non-compliers). The sample in column 5 includes students who accepted the original seating arrangement.

TABLE A2—Effects of the Interventions on Sample Attrition

	Transferring to other schools or having grade retention			Missing values of the key variables			Overall attrition		
	Overall students (1)	Lower-track students (2)	Upper-track students (3)	Overall students (4)	Lower-track students (5)	Upper-track students (6)	Overall students (7)	Lower-track students (8)	Upper-track students (9)
MS	-0.019 [0.028]	-0.038 [0.026]	0.001 [0.033]	-0.007 [0.026]	-0.006 [0.027]	-0.009 [0.032]	-0.023 [0.045]	-0.039 [0.041]	-0.007 [0.054]
MSR	-0.005 [0.025]	-0.017 [0.025]	0.007 [0.028]	0.004 [0.024]	0.010 [0.029]	-0.002 [0.026]	-0.001 [0.041]	-0.007 [0.040]	0.005 [0.046]
Mean of the dependent variable	0.057	0.061	0.053	0.064	0.066	0.061	0.114	0.120	0.108
Observations	1,911	960	951	1,925	965	960	2,034	1,024	1,010

Notes: This table reports the treatment effects of the MS and MSR interventions on the sample attrition. The dependent variable is a dummy that takes a value of 1 if the missing observation is due to students who transferred to other schools or had grade retention (columns 1–3), missing values of key variables (columns 4–6), and missing observation due to either case (columns 7–9). Robust standard errors clustered at the class level are reported in brackets.

TABLE A3—Prediction of Student Characteristics on Deskmate Pairing

	Control group		MS group		MSR group	
	Estimated coefficient (1)	<i>p</i> -value (2)	Estimated coefficient (3)	<i>p</i> -value (4)	Estimated coefficient (5)	<i>p</i> -value (6)
Percentile rank	0.000	(0.93)	0.055	(0.00)	0.061	(0.00)
Male	−0.003	(0.12)	0.005	(0.10)	0.004	(0.19)
Age	−0.001	(0.53)	0.000	(0.91)	0.001	(0.10)
Height (inch)	−0.025	(0.19)	−0.052	(0.01)	−0.025	(0.07)
Weight (kg)	0.000	(0.22)	0.000	(0.33)	0.000	(0.49)
Non-minority (yes=1)	0.003	(0.05)	−0.001	(0.46)	−0.002	(0.31)
<i>Hukou</i> register status (urban <i>hukou</i> =1)	0.000	(0.04)	0.005	(0.00)	0.000	(0.94)
Interest in Chinese (yes=1)	0.001	(0.76)	0.001	(0.44)	−0.005	(0.16)
Interest in Mathematics (yes=1)	−0.001	(0.68)	−0.002	(0.35)	0.001	(0.55)
Father's education (# of years)	0.000	(0.03)	0.000	(0.65)	0.000	(0.19)
Mother's education (# of years)	0.000	(0.76)	0.000	(0.28)	0.000	(0.34)
Student's household has a computer (yes=1)	−0.001	(0.59)	−0.001	(0.10)	−0.001	(0.71)
Student's household has a car (yes=1)	0.000	(0.97)	0.000	(0.90)	0.001	(0.55)
Class fixed effect	Yes		Yes		Yes	
Number of observations	14,403		17,093		14,541	

Notes: This table reports the estimates of the effects of the similarity of characteristics of two students on the probability of students who can be deskmates. We run OLS regression on the dyadic data, which are obtained by pairing each individual with his/her classmates to finalize a dataset. Each observation of the data represents a pair of students. The dependent variable is a dummy that has a value of 1 if the paired students are deskmates. The independent variables are the absolute difference of characteristics between a student and his/her potential deskmate. *P*-values are obtained by permuting student pairs in the control, MS, and MSR classes 500 times and reported in the parentheses.

We employ a nonlinear model to estimate a complex set of the MS and MSR effects.¹ We separate 1,802 students or the entire sample into quartiles according to the baseline score in each class (bottom, second-to-bottom, second-to-top, and top quarters). Four dummy variables, which indicate whether students belong to each quarter are generated. We then interact four dummies with MS and MSR. Therefore, eight interaction terms are created. Individual controls and school-by-grade fixed effects are added. Standard errors are clustered at the class level.

Table A3 reports the results. Column 1 shows that students in the lower and upper middle reap significant score gain in their average scores from the MSR intervention. The MSR effects are 0.17σ (s.e.=0.104) and 0.1σ (s.e.=0.059), respectively. However, students in the upper middle and top quarters gain little from the MSR intervention. Column 2 shows that for Chinese language scores, the MSR intervention cannot alter scores of students in each of the quartiles. Column 3 shows that for mathematics, the magnitudes of the MSR effects among low- and middle-achieving students are positive and statistically significant. For example, for the bottom three quarter students, the estimated MSR effects are 0.22σ (s.e.=0.129), 0.15σ (s.e.=0.08), and 0.21 (s.e.=0.092), respectively. For the top-quarter students, their score gain in mathematics is the smallest and insignificant. To sum up, our results indicate that the mixed-seating strategy accompanied by incentivized high-achieving deskmates is effective for mathematics and the low- and middle-achieving students. Despite the heterogeneity observed in the MSR effects between students in the bottom and top quartiles, the F -test of whether the MSR effect in the bottom-quarter students is the same as the effect in each of the other three quarters show that they are not significantly different. Besides, the effects of interaction between the MS and student previous test scores on all subjects are small and insignificant.

¹ The specification is as follows: $zscore_{ijg}^{endline} = \beta_0 + \beta_1 zscore_{ijg}^{baseline} + \gamma_1 MS_{ijg} \cdot Q_{ijg} + \gamma_2 MSR_{ijg} \cdot Q_{ijg} + X_{ijg}\Gamma + \eta_g + \varepsilon_{ijg}$, where Q_{ijg} is a column vector that indicate the initial quartile of the student's baseline score distribution. γ_1 and γ_2 are two row vectors to be estimated. We consider this equation in which the two treatment variables are interacted with the initial quartile of the baseline score to determine whether student test scores can vary with student's initial academic performance.

TABLE A4—Non-linear Effects of the Interventions on Test Scores in the Entire Sample

	Average score (1)	Chinese score (2)	Mathematics score (3)
(1) Bottom quarter × MS	0.047 [0.106]	-0.152 [0.101]	-0.067 [0.128]
(2) Second-to-bottom quarter × MS	-0.017 [0.081]	-0.065 [0.066]	-0.028 [0.102]
(3) Second-to-top quarter × MS	-0.056 [0.062]	0.007 [0.050]	-0.043 [0.083]
(4) Top quarter × MS	-0.052 [0.069]	0.006 [0.052]	0.091 [0.102]
(5) Bottom quarter × MSR	0.173 [0.104]	-0.132 [0.085]	0.216 [0.129]
(6) Second-to-bottom quarter × MSR	0.102 [0.059]	-0.051 [0.065]	0.152 [0.080]
(7) Second-to-top quarter × MSR	0.103 [0.068]	0.038 [0.052]	0.210 [0.092]
(8) Top quarter × MSR	-0.020 [0.077]	0.059 [0.071]	0.098 [0.084]
Controls	Yes	Yes	Yes
<i>p</i> -value of total MS effect=0	0.824	0.585	0.671
<i>p</i> -value of total MSR effect=0	0.255	0.588	0.172
<i>p</i> -value of MSR effect (bottom quarter=upper middle quarter)	0.393	0.358	0.592
<i>p</i> -value of MSR effect (bottom quarter=lower middle quarter)	0.540	0.104	0.964
<i>p</i> -value of MSR effect (bottom quarter=top quarter)	0.210	0.129	0.503
Observations	1,802	1,802	1,802

Notes: This table reports the nonlinear treatment effects of the MS and MSR interventions on students' endline average scores and Chinese and mathematics test scores. A group of dummies indicating the quartile of the student's baseline score distribution is interacted with two treatment variables. We estimate the nonlinear treatment effect using the entire sample (1,802 students). Controls include the corresponding own baseline scores, gender, age, height, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A5—Effects of the Interventions on Test Scores of Lower-track Students

	By Deskmates' Baseline Score Ranking			By Baseline Score Gaps between Deskmates		
	Average score (1)	Chinese score (2)	Mathematics score (3)	Average score (4)	Chinese score (5)	Mathematics score (6)
MS	0.008 [0.101]	-0.009 [0.078]	-0.013 [0.124]	-0.093 [0.110]	0.182 [0.120]	0.043 [0.124]
MSR	0.103 [0.094]	-0.024 [0.069]	0.204 [0.112]	0.049 [0.116]	0.163 [0.102]	0.283 [0.137]
MS × Top 10 students	0.022 [0.202]	-0.115 [0.180]	0.060 [0.221]			
MSR × Top 10 students	0.285 [0.198]	0.119 [0.196]	0.288 [0.237]			
MS × Large difference between deskmates				0.005 [0.003]	-0.008 [0.003]	-0.002 [0.003]
MSR × Large difference between deskmates				0.005 [0.003]	-0.006 [0.003]	-0.001 [0.003]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901

Notes: This table reports the treatment effects of the MS and MSR interventions on 901 lower-track students' endline average scores and Chinese and mathematics test scores. "Top 10 students" is a dummy that takes a value of 1 if students are assigned with deskmates who rank in the top 10 in the baseline examination in their classes and 0 otherwise. "Large difference between deskmates" is a dummy that takes a value of 1 if the difference in baseline average score between deskmates is above the deskmate-level median difference score in the class and 0 otherwise. Controls include the corresponding own baseline scores, gender, age, height, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A6—Effects of the Interventions on “Big Five” Personality Traits (By Baseline Personality Traits)

	Extraversion	Agreeableness	Openness	Neuroticism	Conscientiousness
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Lower-track students</i>					
MS	0.585 [0.816]	1.586 [1.686]	0.804 [0.660]	1.021 [0.759]	0.116 [1.023]
MSR	3.003 [0.734]	3.581 [1.284]	2.013 [0.904]	0.944 [0.671]	1.199 [0.807]
MS × Below the median	-0.110 [0.947]	-0.065 [1.434]	-1.685 [0.852]	-0.032 [1.134]	1.248 [1.106]
MSR × Below the median	-1.316 [0.980]	-1.884 [1.508]	-1.745 [0.847]	-0.747 [1.077]	0.238 [1.041]
Controls	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901
<i>Panel B. Upper-track students</i>					
MS	-1.309 [1.140]	-0.975 [1.488]	-0.162 [0.567]	0.004 [0.956]	-1.210 [1.061]
MSR	2.107 [0.912]	2.607 [1.422]	0.860 [0.632]	-0.099 [0.852]	0.480 [0.883]
MS × Below the median	0.951 [1.369]	0.615 [1.636]	-0.084 [0.988]	1.079 [1.047]	2.864 [1.394]
MSR × Below the median	-0.102 [1.313]	-0.452 [1.894]	-0.586 [0.964]	0.339 [0.982]	1.147 [1.195]
Controls	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901

Notes: This table reports estimated results of checking heterogeneity in treatment effects on students’ endline personality traits with respect to the baseline personality traits. The dependent variable is each of the five domains following the “Big Five” model surveyed in the endline questionnaire. “Below the median” is a dummy that takes a value of 1 if the variable of interest measured in the baseline is below the sample median and 0 otherwise. Panel A reports the estimated results for lower-track students and Panel B reports estimated results for upper-track students. Controls include “Below the median,” the corresponding own baseline personality trait, gender, age, height, health status, *hukou* registration status, minority, father’s education, mother’s education, and whether the student’s household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A7—Effects of the Interventions on Test Scores (By Baseline Personality Traits)

	Lower-track students			Upper-track students		
	Average score	Chinese score	Mathematics score	Average score	Chinese score	Mathematics score
	(1)	(2)	(3)	(4)	(5)	(6)
MSR	0.184 [0.113]	0.036 [0.102]	0.317 [0.132]	0.046 [0.073]	0.042 [0.059]	0.062 [0.095]
MSR × High extraversion	-0.099 [0.116]	-0.299 [0.125]	0.026 [0.126]	0.044 [0.062]	-0.069 [0.066]	0.134 [0.069]
MSR × High agreeableness	-0.026 [0.092]	0.063 [0.104]	-0.096 [0.112]	0.056 [0.069]	0.093 [0.051]	0.029 [0.098]
MSR × High openness	0.118 [0.093]	0.090 [0.122]	0.164 [0.091]	-0.113 [0.080]	-0.108 [0.086]	-0.097 [0.092]
MSR × High neuroticism	0.006 [0.094]	0.010 [0.096]	0.019 [0.108]	0.092 [0.090]	0.016 [0.059]	0.139 [0.134]
MSR × High conscientiousness	-0.080 [0.093]	0.066 [0.125]	-0.215 [0.089]	-0.034 [0.069]	0.022 [0.074]	-0.088 [0.107]
High extraversion	0.066 [0.061]	0.090 [0.047]	0.057 [0.075]	-0.047 [0.036]	0.018 [0.045]	-0.067 [0.051]
High agreeableness	-0.109 [0.055]	-0.077 [0.069]	-0.107 [0.059]	0.018 [0.038]	-0.001 [0.030]	0.031 [0.061]
High openness	-0.055 [0.062]	-0.020 [0.057]	-0.090 [0.064]	0.023 [0.045]	0.023 [0.041]	0.039 [0.060]
High neuroticism	0.069 [0.057]	0.051 [0.059]	0.073 [0.070]	-0.068 [0.028]	-0.064 [0.033]	-0.077 [0.044]
High conscientiousness	0.092 [0.051]	0.037 [0.057]	0.152 [0.066]	0.070 [0.054]	-0.004 [0.046]	0.116 [0.082]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901

Notes: This table reports estimated results of checking heterogeneity in treatment effects on student test scores with respect to the baseline personality traits. For each domain in “Big Five” model, we generate a dummy that takes a value of 1 if a student scores higher than the sample median and 0 otherwise. These dummies are interacted with MS and MSR treatment variables. Columns 1–3 report the estimated results for lower-track students, and columns 4–6 report estimated results for upper-track students. Controls include the corresponding own baseline test scores, gender, age, height, health status, *hukou* registration status, minority, father’s education, mother’s education, and whether the student’s household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A8—Effects of the Interventions on Test Scores and Personality Traits (By Parental Educational Level)

	Average score	Chinese score	Mathematics score	Extraversion	Agreeableness	Openness	Neuroticism	Conscientiousness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Lower-track students</i>								
MS	0.073 [0.174]	0.080 [0.125]	-0.017 [0.206]	-0.500 [1.505]	0.900 [2.067]	-0.146 [1.062]	0.677 [1.444]	-0.369 [1.715]
MSR	0.176 [0.118]	0.020 [0.100]	0.262 [0.141]	2.107 [0.905]	2.939 [1.654]	0.678 [0.767]	0.497 [0.704]	1.523 [0.901]
MS × Low-educated parents	-0.083 [0.145]	-0.136 [0.115]	0.018 [0.170]	1.360 [1.283]	0.678 [1.569]	0.251 [0.985]	0.337 [1.284]	1.559 [1.497]
MSR × Low-educated parents	-0.056 [0.099]	-0.043 [0.121]	-0.032 [0.120]	0.333 [0.945]	-0.397 [1.431]	0.726 [0.702]	0.069 [0.771]	-0.342 [0.879]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901
<i>Panel B. Upper-track students</i>								
MS	-0.065 [0.068]	-0.077 [0.064]	-0.043 [0.092]	-1.673 [1.045]	0.070 [1.367]	-0.407 [0.580]	0.760 [0.771]	-0.421 [1.068]
MSR	0.048 [0.076]	-0.100 [0.071]	0.161 [0.098]	1.926 [0.932]	1.756 [1.368]	0.326 [0.640]	-0.152 [0.769]	0.663 [0.962]
MS × Low-educated parents	0.004 [0.077]	0.049 [0.075]	0.014 [0.096]	1.144 [0.953]	-1.104 [1.098]	0.309 [0.734]	-0.399 [0.772]	0.542 [1.044]
MSR × Low-educated parents	-0.006 [0.076]	0.171 [0.078]	-0.088 [0.112]	0.186 [1.045]	1.168 [1.142]	0.480 [0.767]	0.480 [0.956]	0.495 [1.062]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901

Notes: This table reports estimated results of checking heterogeneity in treatment effects with respect to parental education. Students' personality traits are measured by the "Big Five" model. "Low-educated parents" is a dummy that takes a value of 1 if both parents have not received a high school education and 0 otherwise. Panel A reports the estimated results for lower-track students and Panel B reports estimated results for upper-track students. Controls include "Low-educated parents", the corresponding own baseline performances, gender, age, height, health status, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A9—Effects of the Interventions on Test Scores and Personality Traits (By Family Income)

	Average score	Chinese score	Mathematics score	Extraversion	Agreeableness	Openness	Neuroticism	Conscientiousness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Lower-track students</i>								
MS	−0.014	−0.023	−0.043	0.331	1.339	−0.133	0.770	0.525
	[0.067]	[0.058]	[0.093]	[0.818]	[1.340]	[0.601]	[0.721]	[0.901]
MSR	0.123	0.002	0.212	2.458	2.927	1.359	0.622	1.485
	[0.066]	[0.051]	[0.086]	[0.704]	[1.120]	[0.757]	[0.589]	[0.757]
MS × High income	0.262	0.069	0.364	2.470	2.366	1.822	2.223	3.365
	[0.237]	[0.234]	[0.246]	[2.454]	[2.981]	[1.837]	[2.272]	[2.383]
MSR × High income	0.135	−0.076	0.241	−1.252	−2.191	−1.980	−0.920	−1.665
	[0.245]	[0.213]	[0.283]	[1.569]	[2.445]	[0.592]	[1.285]	[1.322]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901
<i>Panel B. Upper-track students</i>								
MS	−0.054	−0.040	−0.027	−0.651	−0.449	−0.178	0.685	−0.057
	[0.059]	[0.042]	[0.093]	[0.912]	[1.258]	[0.437]	[0.728]	[0.916]
MSR	0.047	0.013	0.107	2.568	2.900	0.867	0.350	1.395
	[0.058]	[0.046]	[0.083]	[0.714]	[1.207]	[0.498]	[0.637]	[0.774]
MS × High income	−0.045	−0.059	−0.019	−1.636	−1.285	−0.100	−1.056	0.093
	[0.094]	[0.061]	[0.129]	[1.483]	[2.340]	[1.185]	[1.403]	[1.358]
MSR × High income	0.013	−0.034	0.037	−3.546	−3.477	−2.064	−1.750	−3.162
	[0.080]	[0.066]	[0.105]	[1.436]	[2.591]	[1.295]	[1.512]	[1.645]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901

Notes: This table reports estimated results of checking heterogeneity in treatment effects with respect to family income. Students' personality traits are measured by the "Big Five" model. "High income" is a dummy that takes a value of 1 if students' family possess a car and a computer and 0 otherwise. Panel A reports the estimated results for lower-track students and Panel B reports estimated results for upper-track students. Controls include the corresponding own baseline performances, gender, age, height, health status, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A10—Cross-skill Peer Effects

	Estimation for MS classes						Estimation for MSR classes					
	Average score (1)	Extraversion (2)	Agreeableness (3)	Openness (4)	Neuroticism (5)	Conscientiousness (6)	Average score (7)	Extraversion (8)	Agreeableness (9)	Openness (10)	Neuroticism (11)	Conscientiousness (12)
<i>Panel A. Lower-track students</i>												
Deskmate's baseline average z-scores		-1.107 [0.748]	-0.773 [0.893]	-0.871 [0.775]	-1.070 [0.723]	-0.666 [0.796]		0.163 [0.822]	-0.684 [0.757]	-0.529 [0.440]	-0.759 [0.510]	-0.809 [0.679]
Deskmate's baseline extraversion	0.003 [0.008]						0.009 [0.008]					
Deskmate's baseline agreeableness	-0.011 [0.007]						0.003 [0.009]					
Deskmate's baseline openness	-0.011 [0.012]						0.010 [0.013]					
Deskmate's baseline neuroticism	0.001 [0.008]						-0.007 [0.009]					
Deskmate's baseline conscientiousness	0.006 [0.011]						0.009 [0.011]					
Class-by-height group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	317	317	317	317	317	317	317	297	297	297	297	297

Panel B. Upper-track students

Deskmate's baseline average z-scores		0.144	0.446	-0.326	0.899	0.367		-0.142	-0.605	-0.400	-0.387	-0.391
		[0.230]	[0.362]	[0.267]	[0.262]	[0.285]		[0.424]	[0.705]	[0.399]	[0.518]	[0.373]
Deskmate's baseline extraversion	0.004						-0.002					
	[0.010]						[0.017]					
Deskmate's baseline agreeableness	0.01						0.003					
	[0.009]						[0.007]					
Deskmate's baseline openness	0.007						0.008					
	[0.006]						[0.015]					
Deskmate's baseline neuroticism	-0.011						0.016					
	[0.011]						[0.011]					
Deskmate's baseline conscientiousness	-0.012						-0.015					
	[0.010]						[0.012]					
Class-by-height group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	317	317	317	317	317	317	317	297	297	297	297	297

Notes: This table reports the regression estimates of the effect of peer quality by examining the effect of peer baseline test scores and personality traits on the corresponding endline performance. A peer is defined as one's deskmate. Regressions are run for upper- and lower-track students in the MS and MSR classes separately. The estimated equation is Equation (3). Controls include gender, age, height, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A11—Effects of Interventions on the Fraction of Upper-Track Students in one’s Friend

Network

	Relatively low-performing students (1)	Relatively high-performing students (3)
MS	0.034 [0.020]	0.023 [0.024]
MSR	0.062 [0.024]	0.039 [0.022]
Controls	Yes	Yes
<i>p</i> -value of the effect (MS=MSR)	0.212	0.430
Mean of the dependent variable for students in control classes	-0.058	-0.035
Observations	900	902

Notes: This table reports the regression estimates of treatment effects of MS and MSR interventions on the fraction of upper-track students in one’s friend network. The dependent variable is the first-difference of the fraction of upper-track students in one’s friend network. The sample is divided into subsamples according to the relative baseline score of two students sharing the same desk [The relatively high-performing (low-performing) students of the deskmates are the upper-track (lower-track) students in the treatment classes, but this may not be the case in the control classes]. All specifications include school-by-grade fixed effects. Controls include gender, age, height, *hukou* registration status, minority, father’s education, mother’s education, and whether the student’s household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A12—Effects of Interventions on Friendship Ties

	Estimated coefficient	<i>p</i> -value
	(1)	(2)
Lower track	0.004	(0.00)
MS × Lower track	−0.015	(0.00)
MSR × Lower track	−0.008	(0.01)
MS	0.002	(0.50)
MSR	0.008	(0.00)
Male	−0.003	(0.10)
Age	0.006	(0.00)
Height (inch)	−0.014	(0.10)
Non-minority (yes=1)	0.002	(0.81)
<i>Hukou</i> registration status (urban <i>hukou</i> =1)	0.000	(0.90)
Health status (good=1)	−0.001	(0.68)
Number of siblings	0.004	(0.14)
Interest in Chinese (yes=1)	0.007	(0.00)
Interest in mathematics (yes=1)	−0.003	(0.12)
Father's education (# of years)	−0.001	(0.60)
Mother's education (# of years)	0.006	(0.01)
Student's household has a computer (yes=1)	−0.007	(0.00)
Student's household has a car (yes=1)	−0.001	(0.57)
Dyad	1,622,701	

Notes: This table reports the estimates of the effects of the similarity of characteristics of two students on the probability of students who can be friends. We run OLS regression on the dyadic data, which are obtained by pairing each individual with other students to finalize a dataset. Each observation of the data represents a pair of students. The dependent variable is a dummy, which takes a value of 1 if a pair of students are friends (including A nominates B as friends or vice versa). The independent variables are a set of dummies indicating whether pairs of students have the same characteristics (i.e. both belong to the lower track, its interactions with MS and MSR, both receive MS or MSR treatment, have the same gender, height, weight, race, *hukou* registration status, parental educational level, both have a computer or a car, and both are interested in Chinese or mathematics). *P*-values are obtained by permuting network 100 times and reported in the parentheses.

TABLE A13—Effects of the Interventions on Test Scores and Personality Traits (By Female Share in Neighboring Five Students)

	Average score	Chinese score	Mathematics score	Extraversion	Agreeableness	Openness	Neuroticism	Conscientiousness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. By Female share in neighboring five students (Overall lower-track students)</i>								
MS	-0.025	-0.133	0.023	-1.695	0.285	-1.123	-0.372	-0.758
	[0.127]	[0.151]	[0.160]	[1.429]	[1.946]	[1.109]	[1.197]	[1.513]
MSR	0.307	0.190	0.355	2.669	3.427	1.178	1.444	1.130
	[0.144]	[0.204]	[0.146]	[1.081]	[1.712]	[1.175]	[0.975]	[1.308]
MS × Female share in neighboring five	0.073	0.214	-0.048	4.265	2.327	2.165	2.572	3.047
	[0.192]	[0.266]	[0.195]	[2.332]	[2.330]	[1.589]	[1.744]	[2.235]
MSR × Female share in neighboring five	-0.310	-0.351	-0.221	-0.494	-1.196	0.005	-1.560	0.419
	[0.204]	[0.341]	[0.208]	[1.887]	[2.938]	[1.413]	[1.557]	[2.075]
Female share in neighboring five	0.159	-0.019	0.298	-2.140	-2.209	-1.287	-1.000	-1.583
	[0.124]	[0.144]	[0.141]	[1.468]	[2.000]	[1.349]	[1.385]	[1.711]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901

Panel B. By Female share in neighboring five students (Excluding the first and last rows)

MS	-0.058	-0.096	-0.050	-1.940	0.131	-1.241	-0.439	-1.090
	[0.101]	[0.153]	[0.146]	[1.298]	[2.048]	[1.012]	[1.076]	[1.379]
MSR	0.293	0.194	0.328	2.783	3.478	1.526	1.690	1.146
	[0.143]	[0.201]	[0.151]	[0.928]	[1.652]	[1.152]	[0.846]	[1.113]
MS × Female share in neighboring five	0.124	0.150	0.066	4.780	2.697	2.475	2.756	4.032
	[0.185]	[0.279]	[0.200]	[2.373]	[2.561]	[1.623]	[1.665]	[2.100]
MSR × Female share in neighboring five	-0.285	-0.367	-0.175	-0.548	-1.291	-0.627	-1.803	0.404
	[0.230]	[0.354]	[0.236]	[1.865]	[2.981]	[1.409]	[1.487]	[1.982]
Female share in neighboring five	0.144	0.016	0.250	-2.109	-2.089	-0.531	-0.847	-1.706
	[0.132]	[0.157]	[0.146]	[1.377]	[2.045]	[1.270]	[1.209]	[1.495]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	764	764	764	764	764	764	764	764

Notes: This table reports estimated results of checking heterogeneity in treatment effects with respect to the share of female students in the neighboring five students. Panel A reports estimated results for overall lower-track students and Panel B reports estimated results after excluding students sitting in the first and last rows. “Female share in neighboring five” measures the share of female students in neighboring five students. Controls include the corresponding own baseline scores, gender, age, height, *hukou* registration status, minority, father’s education, mother’s education, and whether the student’s household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A14—Items from Student Evaluation of the Class Head Teacher

Item domain (1)	Item statement (2)	Answer scale (3)
Learning instruction	He looks impatient when I ask him questions. (R)	1 (rarely)–4 (often)
	He criticizes me when I cannot solve math problems well. (R)	1 (rarely)–4 (often)
	He praises me when I make progress in my study.	1 (rarely)–4 (often)
	He patiently gives answers when I ask him questions.	1 (rarely)–4 (often)
	I feel he does not concern about my learning. (R)	1 (rarely)–4 (often)
Daily care	I feel he does not care too much about my existence. (R)	1 (rarely)–4 (often)
	He praises me when I do well in other things except in studying.	1 (rarely)–4 (often)
	He criticizes me when I do not do well. (R)	1 (rarely)–4 (often)
	He is very concerned about my feelings.	1 (rarely)–4 (often)
	He communicates with me.	1 (rarely)–4 (often)
	I feel that he does not pay attention to me. (R)	1 (rarely)–4 (often)
Fairness	He is very friendly to me.	1 (rarely)–4 (often)
	He shows partiality to some students. (R)	1 (rarely)–4 (often)
	He despises me because of my dress. (R)	1 (rarely)–4 (often)
	He despises me because of my appearance. (R)	1 (rarely)–4 (often)
	He despises me because of my family background. (R)	1 (rarely)–4 (often)
	He scores fairly.	1 (rarely)–4 (often)
	He does not distinguish between boys and girls.	1 (rarely)–4 (often)

Notes: The class head teacher of the respondent is referred to as “he” in the questionnaire. “(R)” indicates the reverse-scored item, which negatively assesses teachers.

TABLE A15—Effects of Interventions on Student Evaluations on Teacher Behaviors

	Learning instruction		Daily care		Fairness	
	Lower-track (1)	Upper-track (2)	Lower-track (3)	Upper-track (4)	Lower-track (5)	Upper-track (6)
MS	-0.069 [0.159]	-0.078 [0.137]	-0.127 [0.138]	-0.023 [0.146]	0.053 [0.174]	0.180 [0.164]
MSR	-0.013 [0.188]	0.012 [0.161]	-0.060 [0.150]	0.112 [0.160]	-0.019 [0.140]	-0.119 [0.152]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>p</i> -value of intervention (MS=MSR)	0.754	0.573	0.643	0.407	0.668	0.091
Mean of the dependent variable for students in control classes	0.076	0.013	0.080	-0.031	0.046	-0.077
Observations	901	901	901	901	901	901
<i>p</i> -value of MS intervention (lower-track=upper-track)		0.824		0.479		0.321
<i>p</i> -value of MSR intervention (lower-track=upper-track)		0.896		0.164		0.371

Notes: This table reports the regression estimates of treatment effects of MS and MSR interventions on student evaluation on teacher behavior regarding the three domains, including learning instruction, daily care, and fairness. The dependent variable is the first-difference of each of the evaluation on teacher behavior before and after the experiment. All specifications include school-by-grade fixed effects. Controls include the corresponding outcome variable measured in the baseline survey, gender, age, height, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A16—Peer Effects and Teacher Quality

	Estimation for MS classes						Estimation for MSR classes					
	Average z-scores	Extra-version	Agree-ability	Openness	Neuro-ticism	Conscientio-usness	Average z-scores	Extra-version	Agree-ability	Openness	Neuro-ticism	Conscientio-usness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Panel A. Lower-track students</i>												
Deskmate's baseline performance	0.012	0.059	-0.037	-0.057	0.052	0.085	0.116	0.175	0.129	0.044	-0.118	-0.060
	[0.131]	[0.052]	[0.022]	[0.046]	[0.065]	[0.058]	[0.073]	[0.154]	[0.087]	[0.103]	[0.133]	[0.134]
Deskmate's baseline performance × Teacher with a bachelor degree	-0.159	-0.025	-0.120	0.045	-0.114	-0.247	-0.052	0.092	-0.040	0.127	0.109	0.079
	[0.240]	[0.093]	[0.109]	[0.157]	[0.146]	[0.140]	[0.135]	[0.221]	[0.152]	[0.113]	[0.180]	[0.156]
Class-by-height group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	317	317	317	317	317	317	297	297	297	297	297	297
<i>Panel B. Upper-track students</i>												
Deskmate's baseline performance	-0.040	0.046	0.051	0.059	0.075	0.073	0.025	0.205	0.300	-0.029	0.105	0.026
	[0.052]	[0.053]	[0.077]	[0.101]	[0.069]	[0.107]	[0.063]	[0.160]	[0.075]	[0.067]	[0.073]	[0.099]
Deskmate's baseline performance × Teacher with a bachelor degree	0.023	0.038	0.154	-0.019	0.101	0.006	0.017	-0.112	-0.264	-0.167	-0.119	-0.116
	[0.066]	[0.178]	[0.138]	[0.196]	[0.111]	[0.133]	[0.077]	[0.201]	[0.152]	[0.171]	[0.105]	[0.150]
Class-by-height group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	317	317	317	317	317	317	297	297	297	297	297	297

Notes: This table reports estimated results to test whether the deskmate-level peer effects vary by teacher quality. A peer is defined as one's deskmate. The outcome variables are the student's endline average scores (columns 1 and 7) or Big Five personality traits (columns 2–6 and 8–12). The estimated equation is similar to Equation (3) except that we include the interaction between deskmate's test scores/personality traits and teacher quality. Teacher quality is measured by a dummy indicating whether the class head teacher has a bachelor's degree. Controls include teacher quality, gender, age, height, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car. Robust standard errors clustered at the class level are reported in brackets.

TABLE A17—Effects of Interventions on Test Scores and Personality Traits (Corrected Standard Error)

	Average score	Chinese score	Mathematics score	Extraversion	Agreeableness	Openness	Neuroticism	Conscientiousness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Lower-track student</i>								
MS	0.013	-0.020	-0.004	0.553	1.525	0.019	0.985	0.847
	(-0.130, 0.156)	(-0.150, 0.109)	(-0.182, 0.174)	(-1.108, 2.214)	(-1.182, 4.231)	(-1.431, 1.469)	(-0.457, 2.426)	(-0.920, 2.613)
MSR	0.139	-0.009	0.242	2.325	2.677	1.140	0.535	1.317
	(-0.008, 0.286)	(-0.142, 0.125)	(0.059, 0.425)	(0.622, 4.028)	(-0.105, 5.459)	(-0.346, 2.627)	(-0.948, 2.018)	(-0.495, 3.129)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901
<i>Panel B. Upper-track student</i>								
MS	-0.063	-0.051	-0.031	-0.918	-0.676	-0.203	0.507	-0.067
	(-0.183, 0.058)	(-0.135, 0.034)	(-0.207, 0.145)	(-2.393, 0.556)	(-2.915, 1.563)	(-1.078, 0.672)	(-0.787, 1.801)	(-1.711, 1.577)
MSR	0.045	0.004	0.109	2.066	2.386	0.589	0.072	0.953
	(-0.079, 0.170)	(-0.083, 0.092)	(-0.073, 0.291)	(0.574, 3.559)	(0.103, 4.670)	(-0.300, 1.478)	(-1.250, 1.394)	(-0.725, 2.632)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	901	901	901	901	901	901	901	901

Notes: This reports the treatment effects of the MS and MSR interventions on students' test scores and personality traits using the corrected standard errors. We obtain Moulton standard error, which is then magnified by $\sqrt{K(K-1)}$, where K is the number of classes. Panel A reports the estimated results for lower-track students and Panel B reports the estimated results for upper-track students. The estimated equation is specified by Equation (2). The 95% confidence intervals are reported in parentheses. Controls include the baseline score of the corresponding outcome variable, gender, age, height, *hukou* registration status, minority, father's education, mother's education, and whether the student's household has a computer or a car.

TABLE A18—Comparison with the Literature

	Estimated magnitude (1)	Standard deviation (2)	Finding (3)
<i>Panel A. Treatment effects</i>			
This study	0.24	23.59	The MSR intervention increases lower-track students' mathematics scores by 0.24 standard deviations.
Li, Han, Zhang, and Rozelle (2014)	0.265	Not reported	Pairing high and low achieving classmates as deskmates and offering them rewards improves low achieving students' average test scores by 0.265 standard deviations.
Hoffer (1992)	0.18–0.26	11.05–12.27	High-ability students who are placed in the high-ability group gain 0.18–0.26 standard deviations in scores in tracking schools.
Argys, Rees, and Brewer (1996)	0.25	14.9–17.3	Tracking boosts scores of students in high-ability classes by 0.25 standard deviations.
Duflo, Dupas, and Kremer (2011)	0.16–0.19	23.4–23.6	Tracking promotes students in the top half of the pre-assignment score distribution by 0.19 standard deviations and those in the bottom half by 0.16 standard deviations.
Card and Giuliano (2016)	0.5	Not reported	Placement in a fourth-grade gifted/high-achieving class increases the scores of high-achieving minority students by 0.5 standard deviations.
<i>Panel B. Peer effects on test scores/GPA</i>			
This study	0.09 σ	19.17	Being assigned with upper-track incentivized peers promotes low-ability students' test scores by 0.09 standard deviations.
Burke and Sass (2013)	0.23–0.28 σ	Not reported	Low-ability students can benefit by 0.23–0.28 standard deviations from middle-ability classroom peers.
Feld and Zölitz (2017)	0.126 σ	Peers' GPA: 0.471 Grade: 1.687	Being assigned to peers with a one standard deviation higher GPA increases the student's grade by 0.126 standard deviations.
Booij, Leuven, and Oosterbeek (2017)	0.095–0.148 σ	22.1–24.8	A one standard deviation increase of the mean performance of peers raises the number of credit points that a student collects by 0.095–0.148 standard deviations.
Carrell, Sacerdote, and West (2013)	0.053 GPA grades	0.661	Students in the bottom third of the academic distribution would rises 0.053 grades of GPA.
<i>Panel C. Peer effects on personality traits</i>			
This study	Extraversion: 0.035 Agreeableness: 0.017	Extraversion: 6.43 Agreeableness: 8.52	An extraverted or agreeable peer can improve the extraversion and agreeableness of lower-track students by 0.035 and 0.017 standard deviations, respectively.
Zárate (2019)	Extraversion: 0.067 Agreeableness: 0.066	Not reported	Being assigned with sociable roommate peers increases students' extraversion and agreeableness scale by 0.067 and 0.066 standard deviations, respectively.



Figure A1. Typical Desk in Chinese Elementary Schools

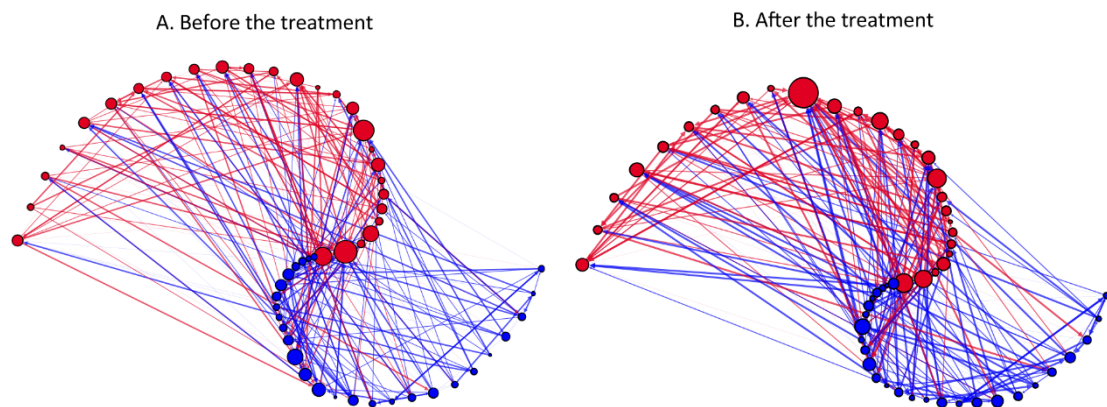


Figure A2. Structure of Student Networks in a Selected Class with the MSR Intervention

Notes: This figure presents a picture of the friend networks between low- and high-achieving students in a selected MSR class. The figure on the left shows the pattern before the treatment, while the figure on the right shows the pattern after the treatment. A red dot indicates an upper-track student, whereas a blue dot represents a lower-track student. The size of the dot means the number of other students who nominate this student as a friend. Dots closer to the center indicate that the student's baseline achievement is closer to the class median baseline score. Clear evidence that can be found in the left figure is that students with similar abilities are likely to group together: the red lines dominate the top semi-cycle, indicating that upper-track students mostly select upper-track students as friends; similarly, the blue lines dominate the bottom semi-cycle, indicating that lower-track students mostly select lower-track students. In comparison, the right figure shows that blue lines in the top semi-cycle (red area) increase considerably, implying that more low-achieving students select high-achieving students as friends after the treatment.