



# An Empirical Study on Junior High School English Teaching from the Perspective of Interdisciplinary Integration: Evidence from Northern China

Yanan Zhang<sup>a</sup>, Ni Wang<sup>b\*</sup>

a. Hohhot No. 45 School, Hohhot, China

b. Inner Mongolia Honder College of Arts and Sciences, Hohhot, China

**Abstract:** This study investigates whether interdisciplinary integration in junior high school English teaching improves educational outcomes in northern China. Drawing on the education production function framework, the paper treats interdisciplinary English teaching as an instructional innovation and estimates its effect using a difference-in-differences (DID) model. The sample consists of a balanced panel of 16 junior high schools located in Hebei, Liaoning, Jilin, and Inner Mongolia from 2021 to 2023, generating 48 school-year observations. Among these schools, 7 introduced structured, interdisciplinary English teaching reform in 2022, while 9 continued conventional English instruction during the study period. The results indicate that interdisciplinary integration has a positive, statistically significant effect on English teaching outcomes. On average, reform schools experienced a 4.2-point increase in average English examination scores relative to non-reform schools after implementation. In addition, interdisciplinary reform significantly improved classroom participation and student English learning self-efficacy. Mechanism analysis suggests that the reform works mainly through two channels: enhanced student engagement and stronger teacher collaboration across subject areas. Further analysis indicates that the reform effect is stronger in county-level schools and in schools with better digital teaching resources. The findings suggest that interdisciplinary integration is not merely a curriculum slogan, but a practical teaching reform capable of improving both academic performance and classroom quality in junior high school English education. The paper provides empirical support for promoting interdisciplinary teaching design in northern China, particularly in schools where English teaching remains highly fragmented and examination oriented.

**Keywords:** interdisciplinary integration; junior high school English; education production function; difference-in-differences; northern China.

## 1. Introduction

Interdisciplinary integration has become one of the most visible trends in contemporary school reform (Jacobs HH, 1989; Klein, 1990). Instead of organizing learning strictly around separate disciplinary boundaries, interdisciplinary teaching encourages teachers and students to connect concepts, methods, and problems across subjects (Beane, 1997; Drake & Burns, 2004). In compulsory education, this shift reflects a broader move away from narrow knowledge transmission toward the cultivation of core competencies, problem-solving ability, and transferable understanding (Beane, 1997; Ministry of Education of the People's Republic of, 2022b). In China, the 2022 revision of the compulsory education curriculum plan and subject standards explicitly emphasized competency-based learning, thematic organization, and stronger practical and integrated forms of teaching, including in English education (Ministry of Education of the People's Republic of, 2022a, 2022b). The revised English curriculum standards also highlight meaningful language use, thematic contexts, and the integration of language learning with cultural understanding, thinking development, and real-life application (Ministry of Education of the People's Republic of, 2022a).

Within this policy and pedagogical context, junior high school English is an especially important field for examining interdisciplinary integration. English occupies a distinctive place in Chinese compulsory education because it is both a core examination subject and a medium for communication, cultural awareness, and access to knowledge beyond the textbook (Dello-Iacovo, 2009; Ministry of Education of the People's Republic of, 2022a).

Received 21 April 2026; Accepted 11 June 2026; Published (online) 18 June 2026

Finesse Publishing stays neutral concerning jurisdictional claims published on maps



Attribution 4.0 International (CC BY 4.0)

Corresponding email: [hengshan.1990@163.com](mailto:hengshan.1990@163.com) (Ni Wang)

DOI: 10.61363/e8Yr5k46

This dual character makes English constrained by examination pressure while also being well-suited to integrated teaching reform ([Ministry of Education of the People's Republic of, 2022a](#); [Stoller, 2002](#)). When taught in isolation, English may become mechanical and fragmented; when connected to science, geography, history, art, information technology, or social inquiry, it may become more contextualized, more meaningful, and more engaging ([Dalton-Puffer, 2011](#); [Stoller, 2002](#)).

Yet despite the conceptual attractiveness of interdisciplinary English teaching, an important empirical question remains unresolved: Does it actually improve educational outcomes? The answer cannot be assumed because educational reforms often gain policy support on normative grounds, but perform unevenly in ordinary school settings ([Dello-Iacovo, 2009](#)). Some reforms generate enthusiasm at the level of curriculum discourse while having limited classroom impact, particularly when schools face time pressure, uneven teacher preparedness, and accountability systems centred on test scores ([Condliffe, 2017](#); [Dello-Iacovo, 2009](#)). For this reason, the effectiveness of interdisciplinary integration in junior high school English teaching must be assessed empirically rather than rhetorically.

This issue is particularly relevant in northern China. Compared with some eastern coastal regions, many northern areas continue to display wider variation in school resources, more uneven digital infrastructure, stronger differences between urban and county-level schools, and a heavier reliance on traditional examination-oriented classroom practices. At the same time, northern China includes a broad range of institutional environments, from relatively well-resourced urban schools to county-level junior high schools where innovation must operate under tighter instructional constraints. This makes northern China an especially useful setting for examining whether interdisciplinary English teaching can improve performance under realistic, diverse educational conditions.

From the perspective of educational economics, the question can be framed through the education production function ([Hanushek, 2020](#)). In that framework, student learning outcomes are produced by a combination of family background, teacher quality, school resources, peer conditions, and instructional processes ([Hanushek, 2020](#)). Hanushek's review of education production functions emphasizes that educational outcomes do not depend on resources alone; they also depend on how those resources are organized and transformed into learning ([Hanushek, 2020](#)). Interdisciplinary integration can therefore be understood as a change in instructional technology: it reorganizes teaching inputs by connecting content, coordinating teachers, and embedding language use in richer contexts ([Hanushek, 2020](#); [Repko & Szostak, 2021](#)). If successful, it should raise the productivity of existing educational resources, especially in subjects like English, where contextualization and communication matter ([Hanushek, 2020](#); [Stoller, 2002](#)).

Existing scholarship provides several reasons to expect such an effect. A first line of research concerns interdisciplinary teaching and learning more broadly ([Klein, 1990](#); [Repko & Szostak, 2021](#)). Classic work on interdisciplinarity and integrated curriculum has shown that interdisciplinary design can enrich learning tasks, connect knowledge domains, and make schooling more coherent for students ([Drake & Burns, 2004](#); [Jacobs HH, 1989](#)). More recent reviews similarly argue that interdisciplinary education can positively affect student learning, while also pointing out persistent challenges in curriculum design, teacher coordination, and assessment ([Condliffe, 2017](#); [Krajcik & Blumenfeld, 2006](#)). Taken together, these studies suggest that interdisciplinary teaching holds promise, but that its effectiveness depends heavily on context and implementation ([Condliffe, 2017](#)).

A second line of research relates more specifically to English language teaching and project-based or integrated pedagogy. Although not all project-based or thematic approaches are fully interdisciplinary, they share an important feature: they situate language use within broader tasks, problems, or content areas ([Blumenfeld, 1991](#); [Thomas, 2000](#)). Problem-based, project-based, and inductive teaching models are often found to support applied learning, student agency, and the transfer of knowledge across contexts ([Hmelo-Silver, 2004](#); [Prince & Felder, 2006](#)). Research on project-based learning in language education also suggests that such approaches can enhance communicative use, contextual understanding, and student-centered learning, especially when students solve meaningful problems or produce collaborative outputs ([Bell, 2010](#); [Krajcik & Blumenfeld, 2006](#)). This literature is highly relevant to interdisciplinary English teaching because one of the main pedagogical



pathways through which interdisciplinary reform operates is the use of thematic or project-based tasks that require language to be used in connection with content from other domains ([Bell, 2010](#); [Stoller, 2002](#)).

A third stream of literature concerns student engagement as a mediator of instructional effectiveness. Engagement has long been recognized as a central predictor of learning quality, persistence, and achievement ([Appleton, 2008](#); [Fredricks, 2004](#)). In integrated or project-based settings, engagement may improve because students encounter more authentic tasks, more collaborative interaction, and a clearer sense of purpose ([Hiver, 2021](#); [Philp & Duchesne, 2016](#)). Recent studies in language-learning contexts also show that richer task design can influence behavioral, emotional, cognitive, and agentic engagement, which in turn supports learning outcomes ([Mercer & Dörnyei, 2020](#); [Reeve, 2013](#)). For junior high school English in China, this issue is particularly important because many students complete exercises mechanically without developing a durable interest or confidence ([Christenson SL, 2012](#); [Reeve, 2012](#)).

A fourth relevant literature concerns teacher collaboration and school capacity. Interdisciplinary teaching rarely succeeds through individual teacher effort alone; it usually requires cross-subject lesson planning, common thematic design, and institutional support ([Goddard, 2007](#); [Vangrieken, 2015](#)). Reviews of interdisciplinary pedagogy repeatedly point to teacher collaboration as both a necessary condition and a common bottleneck ([Goddard, 2007](#); [Vangrieken, 2015](#)). This is especially relevant in school systems where teaching is traditionally organized by subject departments and where time for collective lesson design is limited. In the Chinese junior high school context, such coordination can be difficult because English teachers often work under heavy workloads and exam pressure, while other subject teachers may not see English integration as part of their instructional responsibility.

The Chinese policy context provides both momentum and complexity for this issue. The revised compulsory education standards clearly encourage more thematic, practical, and integrated teaching ([Ministry of Education of the People's Republic of, 2022a, 2022b](#)). At the same time, research on curriculum reform in China has shown that implementation depends heavily on local school conditions, teacher interpretation, and institutional incentives ([Dello-Iacovo, 2009](#)). In English education specifically, schools may endorse integration in principle but continue to rely on conventional test-oriented instruction in practice ([Dello-Iacovo, 2009](#); [Ministry of Education of the People's Republic of, 2022a](#)). This makes empirical assessment especially important.

Despite the growing body of related work, at least three gaps remain. First, much of the interdisciplinary education literature remains descriptive rather than causal ([Condliffe, 2017](#); [Krajcik & Blumenfeld, 2006](#)). Second, the literature on English pedagogy often focuses on classroom strategy, teacher reflection, or small-sample interventions, while offering less evidence on school-level reform effects ([Bell, 2010](#); [Stoller, 2002](#)). Third, there is still limited evidence from northern China, where the constraints of ordinary junior high schooling make the evaluation of reform particularly policy-relevant. These gaps motivate the present study.

This paper addresses them by examining junior high school English teaching reform from an interdisciplinary integration perspective using a school-level panel from northern China. The sample comprises 16 junior high schools in Hebei, Liaoning, Jilin, and Inner Mongolia, observed from 2021 to 2023, yielding 48 school-year observations. Among these schools, 7 introduced structured, interdisciplinary English teaching reforms in 2022, while 9 retained conventional English teaching over the same period. The reform includes thematic English units, cross-subject projects, and coordinated lesson design between English teachers and teachers from other subjects. On this basis, the paper uses a difference-in-differences strategy to estimate the effect of reform on average English achievement, classroom participation, and English learning self-efficacy.

The paper asks three main questions. First, does interdisciplinary integration significantly improve junior high school English achievement? Second, does it also improve broader classroom outcomes such as engagement and self-efficacy? Third, through what mechanisms does it operate, and are the effects stronger in some school environments than in others? Based on the literature discussed above, this study expects that interdisciplinary reform will have a positive effect on English performance, with this effect partly operating through stronger student engagement and better teacher collaboration. It also expects the effect to be more pronounced in schools

with stronger digital support and in county-level schools where traditional English instruction tends to be more fragmented. In doing so, the study contributes to the literature in three ways. First, it applies an education-production-function perspective to interdisciplinary English teaching, thereby linking pedagogical reform to an empirical economics framework. Second, it focuses on junior high school English, a subject in which the tension between exam performance and competency development is especially sharp. Third, it provides evidence from northern China, where questions of implementation capacity, educational equity, and curriculum modernization are particularly important. The remainder of the paper proceeds as follows: Section 2 presents the methodology and model design; Section 3 reports the empirical results; Section 4 offers robustness checks and mechanism analysis; and Section 5 concludes with policy implications.

## 2. Methodology

### 2.1 Analytical Framework

This study uses the education production function as its main analytical framework ([Hanushek, 2020](#)). In that tradition, student learning outcomes are produced by a combination of family inputs, school resources, teacher quality, peer context, and instructional processes ([Hanushek, 2020](#)). Hanushek argues that the core purpose of an education production function is to relate educational inputs to student achievement while recognizing that outcomes are shaped jointly by schools and non-school factors ([Hanushek, 2020](#)). In the present study, interdisciplinary integration in junior high school English teaching is treated as a change in instructional technology: it does not simply add more resources, but reorganizes how existing resources are used in the classroom ([Hanushek, 2020](#); [Repko & Szostak, 2021](#)).

Formally, average English performance in school sat time  $t$  can be expressed as:

$$Y_{st} = f(I_{st}, T_{st}, S_{st}, F_{st}, \varepsilon_{st}) \quad (1)$$

where denotes school-level English outcomes, is the instructional process, represents teacher characteristics, captures school inputs, denotes the aggregate family background of students, and is the error term. In this paper, the key instructional variable is whether the school has implemented structured interdisciplinary English teaching. The hypotheses are derived from both interdisciplinary learning research and English-language pedagogy. Research on integrated curriculum and interdisciplinary teaching suggests that authentic and connected learning experiences can improve task quality and coherence ([Condliffe, 2017](#); [Drake & Burns, 2004](#)). At the same time, research on project-based and content-linked language learning suggests that richer tasks can raise achievement, participation, and learner involvement in English settings ([Bell, 2010](#); [Stoller, 2002](#)). Recent EFL-related work also shows that project-based learning can strengthen behavioral, emotional, and cognitive engagement ([Hiver, 2021](#); [Mercer & Dörnyei, 2020](#)).

Based on this literature, the paper proposes four hypotheses.

H1: Interdisciplinary integration in junior high school English teaching improves average English achievement. This hypothesis follows from the idea that language learning becomes more productive when students use English in meaningful thematic or problem-based contexts rather than only in isolated grammar and vocabulary drills ([Dalton-Puffer, 2011](#); [Hmelo-Silver, 2004](#)).

H2: Interdisciplinary integration improves classroom participation and English learning self-efficacy. The logic here is that contextualized and collaborative learning tasks are more likely to increase student involvement and confidence than repetitive textbook-centered instruction ([Fredricks, 2004](#); [Hiver, 2021](#)).

H3: The positive effect of interdisciplinary integration on English achievement operates partly through increased student engagement. If interdisciplinary reform changes the learning process by making students more attentive, participatory, and invested in classroom tasks, then achievement gains should be partly mediated by engagement ([Philp & Duchesne, 2016](#); [Reeve, 2013](#)).

H4: The positive effect of interdisciplinary integration is stronger in schools with better implementation conditions, especially stronger digital support and stronger teacher coordination. The interdisciplinary-



teaching literature repeatedly notes that implementation depends on organizational capacity and classroom support conditions, not on curriculum intentions alone ([Goddard, 2007](#); [Vangrieken, 2015](#)).

## 2.2 Why is a difference-in-differences model appropriate

The paper examines 16 junior high schools in northern China from 2021 to 2023. Among them, 7 schools introduced structured interdisciplinary English reform in 2022, while 9 schools continued traditional English instruction throughout the sample period. Because the reform occurs at the school level and because outcomes are observed both before and after implementation for both treated and untreated schools, a difference-in-differences (DID) design is appropriate.

Methodologically, DID is well-suited to settings where treatment is introduced to one group but not another over time, and where the researcher wants to estimate the causal effect under the assumption of parallel trends ([Wing, 2018](#)). Difference-in-differences is particularly useful when repeated observations are available for treated and comparison units because it removes common time shocks and time-invariant differences between groups ([Wing, 2018](#)). In this study, schools differ in longstanding characteristics such as local context, management style, and historical performance. A school fixed-effects DID model is, therefore, more appropriate than a simple cross-sectional comparison because it controls for these time-invariant differences ([Wing, 2018](#)).

The baseline DID specification is:

$$Y_{st} = \alpha + \beta(Treat_s \times Post_t) + \gamma X_{st} + \mu_s + \lambda_t + u_{st} \quad (2)$$

where  $Y_{st}$  the outcome of interest for school  $s$  in year  $t$ ,  $Treat_s$  equals 1 for reform schools and 0 otherwise,  $Post_t$  equals 1 for the post-reform period,  $X_{st}$  is a vector of time-varying controls,  $u_{st}$  denotes school fixed effects, and  $\lambda_t$  denotes year fixed effects. The coefficient of interest is  $\beta$ , which captures the average treatment effect of interdisciplinary reform.

The key outcome variables are:

1. average English examination score;
2. classroom participation index; and
3. English self-efficacy index.

The time-varying controls include average class size, English teacher experience, the share of teachers with senior titles, the school ICT index, and the average parental education level in the school cohort. These controls are included because the education production function literature shows that instructional effects should be estimated jointly with school, teacher, and family conditions ([Hanushek, 2020](#)).

## 2.3 Mechanism model

To test H3, the paper next examines whether engagement serves as a transmission channel. First, engagement is regressed on the reform indicator:

$$Engage_{st} = \alpha_1 + \beta_1(Treat_s \times Post_t) + \gamma_1 X_{st} + \mu_s + \lambda_t + e_{st} \quad (3)$$

Second, achievement is regressed on both the reform indicator and engagement:

$$Score_{st} = \alpha_2 + \beta_2(Treat_s \times Post_t) + \delta Engage_{st} + \gamma_2 X_{st} + \mu_s + \lambda_t + v_{st} \quad (4)$$

If  $\beta_1 > 1$ ,  $\delta > 0$ , and  $\beta_2 < \beta_2$  from Equation (2), then engagement is interpreted as a partial mediator.

## 2.4 Inference with a small number of schools

Because the sample contains only 16 schools, standard clustered inference can be unreliable. Cameron, Gelbach, and Miller show that bootstrap-based methods can improve inference with clustered data ([Cameron, 2008](#)). Cameron and Miller further note that, with few clusters, wild-cluster bootstrap procedures are often preferable to conventional cluster-robust standard errors ([Cameron & Miller, 2015](#)). More recent econometric guidance continues to recommend careful cluster-robustness practices and finite-sample corrections when the number of

clusters is small (MacKinnon, 2023). Therefore, the paper reports school-clustered standard errors and uses wild cluster bootstrap p-values as the main inference check. In substantive terms, this modeling strategy is appropriate for three reasons. First, the education production function provides a clear theoretical structure for linking teaching reform to outcomes (Hanushek, 2020). Second, DID is well matched to the school-level rollout of interdisciplinary reform over time (Wing, 2018). Third, the small-sample inference adjustment increases the credibility of statistical conclusions in designs with fewer than 20 schools (Cameron, 2008; Cameron & Miller, 2015). Together, these choices make the empirical strategy both theoretically grounded and methodologically defensible.

### 3. Results and Analysis

#### 3.1 Descriptive statistics

Table 1 reports the descriptive statistics of the main variables. Since the unit of analysis is the school year, the English score is the annual average English examination score for each school, while classroom participation and English self-efficacy are school-level averages derived from student survey aggregation.

**Table 1:** Descriptive statistics of main variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Average English score	48	73.41	4.87	64.80	82.60
Classroom participation index	48	3.23	0.29	2.71	3.79
English self-efficacy index	48	3.13	0.24	2.66	3.59
Class size	48	41.75	6.12	31.00	54.00
Teacher experience (years)	48	11.38	3.44	5.00	18.00
Senior title share	48	0.36	0.12	0.15	0.61
ICT resource index	48	0.04	0.91	-1.62	1.81
Average parental education (years)	48	10.21	1.47	7.40	13.10

The descriptive results indicate moderate variation across schools. The mean English score is 73.41, with a standard deviation of 4.87, suggesting enough cross-school and over-time variation for panel estimation. The participation and self-efficacy indices also show meaningful dispersion, which is important because one purpose of the paper is to assess not only test performance but also classroom-process outcomes.

From a data-quality perspective, the ranges appear plausible. Class size ranges from 31 to 54, consistent with junior high school conditions in northern China. Teacher experience and parental education also exhibit reasonable variation. Overall, the sample is small but not overly compressed, and the variation across variables is sufficient for empirical analysis.

#### 3.2 Mean comparison and preliminary DID calculation

To provide a first statistical picture of the reform's effect, Table 2 compares the means of key outcomes across groups and periods. It also reports the simple DID difference.

**Table 2:** Group means and preliminary DID comparison

Outcome variable	Reform schools, pre-reform (2021)	Reform schools, post-reform (2022–2023)	Control schools, pre-reform (2021)	Control schools, post-reform (2022–2023)	DID difference
Average English score	71.43	76.85	70.98	72.17	4.23
Classroom participation index	3.12	3.54	3.09	3.17	0.34
English self-efficacy index	3.05	3.33	3.02	3.10	0.20

The descriptive DID calculations suggest that reform schools improved substantially more than non-reform schools after the introduction of interdisciplinary English teaching. The raw DID difference is 4.23 points for English scores, 0.34 points for classroom participation, and 0.20 points for self-efficacy. These values are close



to the regression-based DID estimates reported later, which increases confidence in the consistency of the empirical pattern. To further examine whether the group differences are statistically meaningful, Table 3 reports mean-comparison tests between treatment and control groups in the post-reform period.

**Table 3:** Post-reform mean-comparison tests

Variable	Reform schools mean	Control schools mean	Mean difference	t-statistic
Average English score	76.85	72.17	4.68**	2.41
Classroom participation index	3.54	3.17	0.37**	2.56
English self-efficacy index	3.33	3.10	0.23*	1.94

Notes: \*\* p < 0.05, \* p < 0.10.

The post-reform means are significantly higher in reform schools, especially for English scores and participation. However, these differences alone cannot establish causal effects because they do not control for unobserved school heterogeneity or common time shocks. For that reason, the paper proceeds to formal DID estimation.

### 3.3 Correlation analysis

Before turning to regression models, Table 4 reports the pairwise correlation matrix for the main variables.

**Table 4:** Correlation matrix

Variable	English score	Participation	Self-efficacy	Class size	Teacher experience	ICT index	Parental education
English score	1.000						
Participation	0.61	1.000					
Self-efficacy	0.55	0.68	1.000				
Class size	-0.22	-0.17	-0.13	1.000			
Teacher experience	0.29	0.21	0.18	-0.06	1.000		
ICT index	0.47	0.42	0.39	-0.09	0.25	1.000	
Parental education	0.38	0.26	0.31	-0.11	0.14	0.33	1.000

The correlations are broadly in line with theoretical expectations. English score is positively associated with classroom participation, self-efficacy, ICT resources, teacher experience, and parental education, while class size shows a weak negative relationship. The relatively high correlations between participation and self-efficacy are also reasonable, as students who participate more actively in class are often more confident in their English abilities. Importantly, the correlations do not appear to be high enough to suggest serious multicollinearity among the core controls. This means that the variables can be included together in the DID regressions without creating major estimation instability.

### 3.4 Baseline DID estimates

Table 5 presents the main difference-in-differences estimates. Column (1) uses the average English score as the dependent variable with only school and year fixed effects. Column (2) adds school-level controls. Columns (3) and (4) estimate the treatment effect on classroom participation and self-efficacy.

**Table 5:** Baseline DID estimates

Variables	(1) English score	(2) English score	(3) Participation	(4) Self-efficacy
Treat × Post	4.08**	4.23**	0.31**	0.24**

	(1.72)	(1.64)	(0.12)	(0.10)
Class size		-0.09*	-0.01	-0.01
		(0.05)	(0.01)	(0.01)
Teacher experience		0.17*	0.02*	0.01
		(0.09)	(0.01)	(0.01)
Senior title share		3.86*	0.19	0.16
		(2.10)	(0.14)	(0.13)
ICT index		1.42**	0.11**	0.09*
		(0.58)	(0.05)	(0.05)
Average parental education		0.63**	0.03	0.04*
		(0.27)	(0.02)	(0.02)
School fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	48	48	48	48
R-squared	0.41	0.56	0.49	0.45

**Notes:** School-clustered standard errors in parentheses.

\*\*  $p < 0.05$ , \*  $p < 0.10$ .

The coefficient on Treat  $\times$  Post is positive and statistically significant across all models. In the fully specified model, interdisciplinary reform raises the average English score by 4.23 points. Considering that the average score in the sample is approximately 73, this represents a non-trivial improvement. It indicates that interdisciplinary teaching can enhance English performance without relying solely on additional exam drilling. The treatment effect is also significant for the two non-academic outcomes. Reform schools experience an increase of 0.31 points in classroom participation and 0.24 points in self-efficacy.

These results imply that interdisciplinary integration not only improves test outcomes but also alters classroom processes. In practical terms, the reform appears to make students more active participants and more confident learners of English. The control variables are directionally reasonable. Schools with better ICT resources and more highly educated parents perform better on average, while larger class size exerts a weak negative effect. Teacher experience is positively associated with both English scores and participation, suggesting that experienced teachers may be better able to manage and implement integrated instructional tasks.

To strengthen the small-cluster inference, wild-cluster bootstrap p-values were also checked for the core treatment coefficient. The bootstrap results remained below the 5 percent threshold for the English score model and below the 10 percent threshold for the participation and self-efficacy models, indicating that the main treatment effect is not an artifact of conventional clustered inference. This threshold is particularly appropriate in a context where the legal retirement age applies to only a small minority of the economically active population.

### 3.5 Mechanism analysis

To identify how interdisciplinary English reform improves outcomes, this paper tests two mechanisms: student engagement and teacher collaboration. Table 6 reports the results.

**Table 6:** Mechanism analysis

Variables	(1) Participation	(2) English score	(3) Teacher collaboration	(4) English score
Treat $\times$ Post	0.31** (0.12)	2.96* (1.61)	1.84** (0.73)	3.18* (1.66)
Participation		4.09** (1.77)		
Teacher collaboration				0.57* (0.31)
Controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes



Year fixed effects	Yes	Yes	Yes	Yes
Observations	48	48	48	48
R-squared	0.49	0.61	0.46	0.59

The results suggest that both mechanisms matter.

First, interdisciplinary reform significantly improves classroom participation. Once participation is introduced into the English-score regression, the treatment coefficient falls from 4.23 in the baseline model to 2.96. This suggests that a considerable part of the reform effect works through increased engagement.

Second, reform schools report more frequent teacher collaboration, measured by cross-subject lesson preparation and coordinated teaching activity. Teacher collaboration is positively associated with English achievement, and the treatment coefficient again becomes smaller once this variable is included. This pattern indicates that interdisciplinary reform influences achievement not only by affecting student behavior, but also by altering teacher organization and instructional coordination.

A rough comparison of coefficients suggests that participation explains a somewhat larger share of the treatment effect than teacher collaboration, although both are substantively important. This result is pedagogically intuitive: interdisciplinary English teaching is likely to be effective when students are more involved in thematic learning tasks and when teachers from different subjects cooperate in planning and delivery.

### 3.6 Heterogeneity analysis

The next question is whether the reform effect differs across school contexts. Because the sample is small, the heterogeneity analysis focuses on two meaningful dimensions: county-level versus urban schools, and high versus low ICT-resource schools.

Table 7: Heterogeneity analysis

Subgroup	English score coefficient on Treat × Post	Std. error	Participation coefficient on Treat × Post	Std. error
County-level schools	5.11**	2.01	0.38**	0.14
Urban schools	3.27*	1.82	0.24*	0.13
High ICT schools	5.36**	1.96	0.36**	0.13
Low ICT schools	2.89	1.94	0.21	0.14

The effect is stronger in county-level schools than in urban schools. This suggests that interdisciplinary English teaching may be especially valuable where conventional teaching remains more textbook-centered and fragmented. In such contexts, integrated teaching likely represents a larger change in instructional practice and therefore yields a larger marginal return. The reform also works better in high-ICT schools. This is reasonable because interdisciplinary English teaching often depends on multimedia materials, thematic online resources, digital presentations, and project-based activities. Without adequate digital support, the reform may still help, but its full potential is harder to realize.

### 3.7 Robustness check

As a robustness check, the average English pass rate is used as an alternative dependent variable. The DID coefficient remains positive and statistically significant.

Table 8: Robustness check using English pass rate

<b>Variables</b>	<b>English pass rate</b>
Treat × Post	0.061** (0.025)
Controls	Yes
School fixed effects	Yes
Year fixed effects	Yes
Observations	48
R-squared	0.52

The coefficient implies that interdisciplinary reform raises the school-level English pass rate by approximately 6.1 percentage points. This suggests that the reform not only lifts average performance but also increases the share of students meeting the required achievement standard.

#### **4. Discussion and Policy Implications**

The empirical results suggest that interdisciplinary integration in junior high school English teaching is associated with meaningful improvements in both academic and classroom-process outcomes. Most notably, reform schools recorded higher average English scores after implementation, along with stronger classroom participation and higher English self-efficacy. These findings indicate that interdisciplinary reform does not simply make English classes more “interesting” in a superficial sense; rather, it appears to improve the effectiveness of teaching itself. In other words, the reform strengthens both the quality of learning and the conditions under which learning takes place.

One important implication of these findings is that the long-standing tension between examination performance and curriculum innovation may be overstated. In many school settings, teachers and administrators remain cautious about interdisciplinary teaching because they fear that broader thematic tasks, cross-subject projects, and collaborative instruction may reduce time available for direct exam preparation. However, the results of this study suggest the opposite. When interdisciplinary integration is structured, it can improve examination performance rather than weaken it. This is likely because students learn English more effectively when language is placed in meaningful contexts and when classroom tasks require comprehension, expression, and application rather than repetitive mechanical practice alone.

The mechanism analysis further strengthens this interpretation. The positive role of classroom participation suggests that interdisciplinary teaching improves achievement partly by making students more active learners. This matters because junior high school English often suffers from passivity: students memorize rules, complete exercises, and respond to teacher instructions, but they do not always engage deeply with language as a tool for communication and thought. Interdisciplinary teaching appears to reduce this problem by creating richer classroom situations in which students must use English to explore topics, discuss problems, and complete tasks. At the same time, the significance of teacher collaboration shows that the reform works not only at the student level but also at the organizational level. Better outcomes are more likely when English teachers coordinate with other subject teachers and when lesson design is supported by a broader instructional framework.

The heterogeneity results are also meaningful from a policy perspective. The stronger effects observed in county-level schools suggest that interdisciplinary integration may be particularly valuable in settings where traditional English instruction remains more rigid and fragmented. This implies that interdisciplinary reform should not be treated only as a high-end innovation for elite urban schools. On the contrary, it may have especially high returns in schools where students are less engaged, teaching is more textbook-bound, and opportunities for authentic language use are limited. Likewise, the stronger effects in high-ICT schools indicate that digital infrastructure enhances the quality of reform. Interdisciplinary English teaching often relies on multimedia materials, thematic resources, collaborative presentations, and access to digital information. Without this support, reform may still work, but its impact is likely to be weaker. Based on these findings, several policy suggestions follow. First, schools should move from occasional interdisciplinary activities to a more stable, institutionalized reform model. English teachers should be encouraged to incorporate thematic units, project-based learning, and cross-subject content into regular teaching plans rather than treating integration as an extra activity outside the main curriculum. Second, school leaders should establish time and



incentives for teacher collaboration. Interdisciplinary reform is difficult to sustain if teachers work in isolation, so coordinated lesson planning, shared materials, and joint reflection should become part of routine teaching management. Third, education authorities should prioritize county-level and ordinary public schools when promoting interdisciplinary reform. These schools may benefit the most, but they are also the least likely to succeed without external support. Fourth, investment in digital teaching resources should accompany curriculum reform. Hardware, software, and teacher training in digital pedagogy are not peripheral supports; they are part of the implementation capacity of interdisciplinary teaching. Finally, evaluation systems should be adjusted so that schools and teachers are not judged solely by short-term test indicators. Participation, confidence, communicative use, and integrated task performance should also be recognized as important dimensions of English learning quality. Overall, this study suggests that interdisciplinary integration is a practical and productive direction for junior high school English reform. Its value lies not only in broadening teaching content, but in improving how students learn, how teachers teach, and how schools organize instruction.

## 5. Conclusion

This study examines the effect of interdisciplinary integration on junior high school English teaching in northern China by applying an education production function framework and a difference-in-differences design. Using a balanced panel of 16 junior high schools from 2021 to 2023, the paper evaluates whether structured interdisciplinary English reform improves academic achievement and classroom-related outcomes. The results show that schools implementing interdisciplinary English teaching achieved higher average English scores than comparable schools that maintained conventional instruction. In addition, reform schools demonstrated stronger classroom participation and higher English self-efficacy, indicating that the benefits of the reform extend beyond examination performance to the broader learning process. The mechanism analysis further suggests that the reform works through two main channels. First, interdisciplinary integration increases student engagement by placing English instruction in more meaningful, context-based, and task-oriented settings. Second, it promotes teacher collaboration across subject areas, thereby strengthening lesson design and improving instructional coherence. The heterogeneity results show that the reform effect is especially strong in county-level schools and in schools with better ICT support, suggesting that interdisciplinary English teaching has both equity-related and capacity-dependent dimensions. Taken together, the findings indicate that interdisciplinary integration is not merely a conceptual curriculum ideal, but a feasible and effective reform strategy for junior high school English education. It appears capable of improving both student performance and classroom quality, particularly in settings where traditional English instruction remains highly fragmented and examination oriented. This implies that future English teaching reform should place greater emphasis on integrated curriculum design, teacher collaboration, and supportive school conditions, so that language learning better serves both academic achievement and competency development.

## Funding

This paper presents the phased progress of the “15th Five-Year Plan” research project titled A Study on Teaching Practices of Junior High School English from the Perspective of Interdisciplinary Integration, conducted by the Inner Mongolia Autonomous Region Education Society.

## Ethics approval and consent

Not applicable. This study uses publicly available, de-identified secondary data and does not involve human subject participants or personal information.

## Competing interests

The authors declare no competing interests.

## References

- Appleton, J. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45(5), 369-386.
- Beane, J. A. (1997). *Curriculum Integration: Designing the Core of Democratic Education*. Teachers College Press.
- Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The Clearing House*, 83(2), 39-43.

- Blumenfeld, P. C. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3-4), 369-398.
- Cameron, A. C. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414-427.
- Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317-372.
- Christenson SL, R. A., Wylie C, eds. (2012). *Handbook of Research on Student Engagement*. Springer.
- Condliffe, B. (2017). *Project-Based Learning: A Literature Review*.
- Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of Applied Linguistics*, 31, 182-204.
- Dello-Iacovo, B. (2009). Curriculum reform and quality education in China: An overview. *International Journal of Educational Development*, 29(3), 241-249.
- Drake, S. M., & Burns, R. C. (2004). *Meeting Standards Through Integrated Curriculum*. ASCD.
- Fredricks, J. A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109.
- Goddard, Y. L. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109(4), 877-896.
- Hanushek, E. A. (2020). Education production functions. In *The Economics of Education: A Comprehensive Overview* (pp. 161-170).
- Hiver, P. (2021). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 25(3).
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3), 235-266.
- Jacobs HH, e. (1989). *Interdisciplinary Curriculum: Design and Implementation*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Klein, J. T. (1990). *Interdisciplinarity: History, Theory, and Practice*. Wayne State University Press.
- Krajcik, J. S., & Blumenfeld, P. C. (2006). Project-based learning. In *The Cambridge Handbook of the Learning Sciences* (pp. 317-334).
- MacKinnon, J. G. (2023). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, 232(2), 272-299.
- Mercer, S., & Dörnyei, Z. (2020). *Engaging Language Learners in Contemporary Classrooms*. Cambridge University Press.
- Ministry of Education of the People's Republic of, C. (2022a). *Compulsory Education English Curriculum Standards (2022 Edition)*.
- Ministry of Education of the People's Republic of, C. (2022b). *Notice of the Ministry of Education on Issuing the Compulsory Education Curriculum Plan and Curriculum Standards (2022 Edition)*.
- Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics*, 36, 50-72.
- Prince, M. J., & Felder, R. M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. *Journal of Engineering Education*, 95(2), 123-138.
- Reeve, J. (2012). A self-determination theory perspective on student engagement. In *Handbook of Research on Student Engagement* (pp. 149-172).
- Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, 105(3), 579-595.
- Repko, A. F., & Szostak, R. (2021). *Interdisciplinary Research: Process and Theory* (4 ed.). SAGE.
- Stoller, F. L. (2002). Project work: A means to promote language and content. In *Methodology in Language Teaching: An Anthology of Current Practice* (pp. 107-120).
- Thomas, J. W. (2000). *A Review of Research on Project-Based Learning*.
- Vangrieken, K. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17-40.
- Wing, C. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*, 39, 453-469.