Syncing Minds for Learning: Student-Teacher Interaction Forecasts Academic Improvements Over Semesters

Abstract

Background: Student performance continuously fluctuates throughout the learning process, with teachers playing a pivotal role in shaping these changes. While previous research has established the influence of student-teacher interactions on learning outcomes, how such interactions drive variations in performance across different students remains an open question. In practice, teachers are often unclear about which students interact well with them, which may hinder the improvement of teaching strategies.

Aims: Leveraging advancements in hyperscanning technology, this study utilizes longitudinal EEG data collected in real classrooms to explore the impact of student-teacher inter-brain coupling on academic improvements.

Methods: The study employs a novel dynamic model and two static models to explore the relationship between teacher-student inter-brain coupling and students' academic performance and progress, while controlling for the effects of teacher self-selection.

Results: The findings reveal that enhanced coupling predominately at the high-beta frequency band serves as a positive predictor of academic improvements in both Chinese and math subjects.

Conclusions: This study highlights the critical role of inter-brain coupling in learning, providing robust evidence of ecological validity and confirming its practical potential.

Keywords: student-teacher interaction, hyperscanning, inter-brain coupling, longitudinal research, real classroom

Introduction

In almost all educational settings cross cultures, student performance is always differentiating and fluctuating over time. Students' learning trajectories often diverge as they progress. Teachers, as the essential facilitators and guides in the learning process, play a pivotal role in shaping these outcomes (Hamre & Pianta, 2007; Holper et al., 2013; Pianta et al., 2012; Saha & Dworkin, 2009). In today's academically diverse classrooms, how teachers interact with different students and whether student-teacher interactions influence the variation of student performance remains an open question. Moreover, do these interactions differ in their impact across diverse student groups? Addressing this question is crucial for advancing our understanding of student differentiation, the evolving role of teachers, and for improving both the quality and equity of education. Exploring how student-teacher interactions contribute to performance variations could provide insights into optimizing teaching practices to better support students at different ability levels. This understanding may lead to more

effective, tailored pedagogical strategies that promote an inclusive learning environment and ensure opportunities for all students to succeed.

However, effectively measuring the level of student-teacher interactions presents significant methodological challenges. Classroom interactions are complex, encompassing both observable behaviors, such as questioning and responding, and more implicit cognitive-emotional exchanges (Ellis, 1993; Ellis, 2005; Hamre & Pianta, 2005). For instance, experienced teachers can intuitively tailor their instruction through these interactions to align with students' zones of proximal development, crafting learning experiences that resonate with students' current abilities and potential for growth (Shabani et al., 2010; Vygotsky, 1978). Therefore, dynamically assessing the quality of student-teacher interactions is crucial for gaining a deeper understanding of student learning outcomes. However, objectively and accurately quantifying these interactions remains a difficult task. Traditionally, assessments have relied on subjective self-reporting via questionnaires, which are susceptible to biases from social desirability and large measurement error (Assor & Connell, 1992; Miller, 2011; Nederhof, 1985; Pace, 1985).

The advancement of hyperscanning technology in the field of neuroscience has introduced a more objective approach to measuring these interactions. Hyperscanning involves the simultaneous recording of brain activities from multiple individuals, allowing for the observation of inter-brain coupling, which reflects the degree of neural synchronization during social interactions (Konvalinka & Roepstorff, 2012; Liu et al., 2019; Zheng et al., 2018; Zheng et al., 2020). Studies have indicated that stronger inter-brain coupling could be associated with improved learning outcomes, suggesting that this measure could serve as a potent indicator of educational success (Dikker et al., 2017; Zheng et al., 2018). However, most research utilizing hyperscanning has been confined to laboratory settings, with limited external validity and application in real-world classroom environments (Goudeau et al., 2024; Sonkusare et al., 2019). Furthermore, there is a dearth of longitudinal studies that could provide a deeper understanding of the dynamics of student-teacher interactions and their impact on learning over time, with higher internal validity (Goudeau et al., 2024). With recent advancements in wearable neurotechnology, the improved practicality and high temporal resolution of EEG devices now enable the dynamic tracking and comprehensive assessment of student-teacher interactions and learning processes in real-world classroom settings (Chen et al., 2023; Dikker et al., 2017).

In real classrooms, student-teacher interaction is a far more complex and dynamic process compared to controlled laboratory environments. Even when we can represent their interaction levels through neurophysiological measures, individual differences among students and teacher's self-selection regarding how to interact with different students make it challenging to detect the causal relationship between interaction and academic performance (Angrist & Pischke, 2009). For instance, when instruction is tailored to an average level, students performing at or below this level may show a higher student-teacher interaction, while higher achievers may show lower

student-teacher interaction. This scenario could lead to an apparent negative correlation between interaction intensity and achievement. However, as other patterns of self-selection coexist—such as instruction being tailored to a top level and only high-achieving students focusing diligently or lower-performing students disengaging—these static associations become convoluted, making it difficult to disentangle specific effects (Pashler et al., 2008; Schmeck, 2013).

To truly appreciate the impact of student-teacher interactions on learning, longitudinal studies are essential; tracking the dynamic evolution of interaction quality over time allows for a nuanced understanding of how these relationships influence academic trajectories and meet diverse student needs. Leveraging longitudinal data with econometric modeling techniques, particularly the methods of difference, can address biases from unobserved correlated variable (Gandolfo, 1971; Liker et al., 1985; Tuma, 1984). By calculating the change in inter-brain coupling values across two time points and using this to forecast changes in scores, this approach effectively controls for biases arising from teachers' self-selection behaviors and individual differences among students. Additionally, because initial interaction levels may influence academic improvements, and since the initial level and the change over time is always negatively associated statistically (See Table S1), the baseline inter-brain coupling levels must be included as a predictor variable, in order to avoid the omitted-variable bias and improve internal validity (Angrist & Pischke, 2009).

The study is based on a longitudinal investigation conducted in middle school, a representative context of basic education, focusing on real classroom learning environments in Chinese and math subjects. Wearable EEG devices were used to collect neurophysiological data from both teachers and students over multiple semesters. Given the distinct functional significance of different EEG frequency bands, separate analyses were conducted for each band. For each participant, we calculated the average levels of student-teacher inter-brain coupling, as well as standardized scores representing students' learning outcomes, for each half-semester (i.e., one *period*). Leveraging the longitudinal design, the study employed a dynamic model to forecast the relationship between the growth in learning outcomes and changes in inter-brain coupling between consecutive periods. As a comparison, two other static models were also developed, which used inter-brain coupling levels to forecast learning outcomes in the same period. The study aims to provide a clearer understanding of how student-teacher interactions, as measured by inter-brain coupling, contribute to student learning outcomes over time.

Results

In this study, forty-eight first-year middle school students (22 males; aged 12-14) and four teachers (two teaching Chinese and two teaching math) underwent EEG recordings during Chinese and math classes spanning three semesters, each lasting four months. EEG recordings were carried out for one week each month, capturing 1-2 daily sessions of both Chinese and math classes, with each session lasting 40 minutes. Prior

to each session, students were equipped with EEG headbands, participated in usual class activities (Fig. 1A, B).

[Figure 1 About Here]

Students' learning outcomes were assessed through half-semester examinations, with scores provided by the school administration. Each half-semester was operationally defined as a single period. Students were stratified into quartiles based on their grade rankings: 0–25%, 25–50%, 50–75%, and 75–100%, with the 75–100% range representing the top 25%. Fig. 1C and 1D illustrate the temporal changes in Chinese and math scores, respectively. Notable fluctuations in student rankings were observed across performance categories over time. Scores were then normalized against the entire student cohort for further analysis, with higher normalized scores indicating better performance.

The research employed the total interdependence (TI) value to quantify the student-teacher inter-brain coupling, conducting analyses across multiple frequency bands (Chen et al., 2023; Dikker et al., 2017; Xu et al., 2024). For each student, the student-teacher TI value was calculated between that student and their respective teacher during each class session (Fig. 1E). The changes in student-teacher TI across periods are illustrated in *Fig. S1*. These TI values were then averaged and standardized for each period. The differences in TI values and scores were calculated between two consecutive periods, and this analysis was extended across various frequency bands including broadband (1-40 Hz), delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), low-beta (13-18 Hz), high-beta (18-30 Hz), and gamma (30-40 Hz) frequency bands. Fig. 1F illustrates the scores over periods and corresponding TI values in the high-beta frequency band for selected example students, as shown in the left two columns. The trends of students' TI values are quite different from their learning outcomes. To further explore the potential causal relationship between student-teacher inter-brain coupling and learning outcomes, the differences between successive periods for both scores and TI values were calculated, resulting in ΔTI and Δscore. The results are presented in the right two columns of Fig. 1F. Students' ΔTI values and Δscores exhibited a similar trend of variation.

Changes in student-teacher inter-brain coupling forecast students' academic improvement

The study first employed a dynamic model to investigate whether increases in student-teacher inter-brain coupling could forecast improvements in student performance. The model used the changes in learning outcomes (Δscores) between the two adjacent *periods* as the outcome variable, with the baseline student-teacher coupling values (TI values) and changes in coupling values (ΔTI values) as predictors. After controlling for the baseline TI values, the analysis showed that ΔTI values significantly impacted the improvement in both Chinese (Fig. 2A) and math scores (Fig. 2B). For Chinese, ΔTI in broadband and high-beta frequency bands can positively forecast the improvements in Chinese scores ($\beta_{\Delta broadbandTI} = 0.374, p =$

0.040; $\beta_{\Delta highbetaTI} = 0.457, p = 0.008$), whereas for math, ΔTI in high-beta frequency bands positively forecasted the improvements in math scores ($\beta_{\Delta highbetaTI} = 0.187, p = 0.043$). These findings suggest that more pronounced increases in student-teacher TI values can forecast greater academic improvements. TI values also possess a predictive effect on academic improvement. For Chinese, TI values in the broadband and high-beta frequency bands similarly exhibit a significant positive predictive effect on score increases ($\beta_{broadbandTI} = 0.461, p = 0.019$; $\beta_{highbetaTI} = 0.542, p = 0.001$). For math, TI values in the broadband, delta, and high-beta bands are all capable of positively forecasting increases in scores ($\beta_{broadbandTI} = 0.464, p = 0.001$; $\beta_{deltaTI} = 0.291, p = 0.020$; $\beta_{highbetaTI} = 0.218, p = 0.039$). This implies that higher coupling values with teachers can also account for the progress in learning outcomes. The results of the fitted value are depicted in Fig. 2C (Chinese) and Fig. 2D (math). To explore the lagged effect of student-teacher coupling on Δscores across multiple subsequent periods, we employed TI and ΔTI to forecast Δscores with lags ranging from 0 to 4 *periods*. For instance, if ΔTI represents the change in TI values from *period* 1 to *period* 2, then with a lag period of 0, the corresponding dependent variable would be the change in scores from *period* 1 to *period* 2, yielding results shown in Fig. 2A and 2B. When the lag period is set to 1, the dependent variable shifts to the change in scores from period 2 to period 3. The outcomes are depicted in Fig. 2E (Chinese) and 2F (math), indicating that student-teacher coupling does not have a predictive effect on subsequent changes in academic performance. The model outcomes and the results of the permutation tests are presented in *Table S2* and *Fig. S2* of the *Supplementary Information*.

[Figure 2 About Here]

Student-teacher inter-brain coupling forecast students' academic achievement

As a comparison, the study implemented a *static model I* to forecast the coefficients between learning outcomes within the same period and student-teacher inter-brain coupling (TI values). The results, as shown in Fig. 3, indicate no significant association between inter-brain coupling and student performance in Chinese classes. However, in math classes, there was a significant negative correlation between student-teacher coupling in the delta frequency band and student performance ($\beta_{deltaTI} = -0.315, p < 0.001$). Fig. 3C and 3D present the linear regression results for the significant frequency bands identified in the dynamic model. The model outcomes are presented in *Table S3*.

[Figure 3 About Here]

Static model II was then developed as a supplementary analysis. This model utilized the learning outcome (Scores) of each period as the outcome variable, and the student-teacher coupling values (TI values) and changes in student-teacher coupling values (ΔTI values) as predictors. For both Chinese (Fig. 4A) and math (Fig. 4B), the student-teacher coupling can forecast learning outcomes. Broadband TI and high-beta

TI positively forecast Chinese scores ($\beta_{broadbandTI} = 0.428, p = 0.037$; $\beta_{highbetaTI} = 0.435, p = 0.047$), while broadband TI and Δgamma TI significantly forecast math scores ($\beta_{broadbandTI} = 0.404, p = 0.021$; $\beta_{\Delta gammaTI} = 0.369, p = 0.016$). The fitted values of significant frequency bands are presented in Fig. 4C and 4D. These findings demonstrate that higher TI values and their growth correlate with better academic achievements. Subsequently, we performed forecasts using TI and ΔTI for scores with lags ranging from 0 to 4 *periods*. The significant predictive effect diminishes after one lagged period, with high-beta TI and Δhigh-beta TI demonstrating a significant predictive effect on Chinese scores for the one lagged period ($\beta_{highbetaTI} = 0.468, p = 0.037$; $\beta_{\Delta highbetaTI} = 0.556, p = 0.008$). These results reveal that high-beta inter-brain coupling can effectively forecast Chinese scores for the next period, but fails to forecast performance in periods beyond that. The model outcomes are presented in *Table S4*. The results of the permutation tests are presented in *Fig. S2* in *Supplementary Information*.

The study found either no correlation or a negative correlation between student-teacher inter-brain coupling and students' learning outcomes within the same session. However, significant positive correlations were identified between the increments of these variables. This suggests that teachers may adopt varying instructional strategies to selectively interact with students of different achievement levels.

[Figure 4 About Here]

Model comparison

Subsequently, the study compared the model performance through model fit. By comparing the Wald chi-squared results of the models (see Table 1), we found that the dynamic model yielded higher Wald chi-squared values compared to the static models, indicating that the dynamic model demonstrated significantly better predictive performance. Combined with the preceding theoretical analysis and model comparison results, our findings suggest that the dynamic model is more appropriate than the two static models.

Table 1. **The model fit of *dynamic* model and *static* models.**

	Chinese			**Math**			
	N	**Broadband Wald chi2(2)**	**High-beta band Wald chi2(2)**	**N**	**Broadband Wald chi2(2)**	**Delta Wald chi2(2)**	**High-beta band Wald chi2(2)**
Dynamic model	**37**	**6.68**	**13.17**	**54**	**11.00**	**12.82**	**5.48**
Static model I	37	0.35	0.04	54	2.72	11.62	0.56
Static model II	37	4.75	4.49	54	5.32	11.68	2.55

Forecasting learning outcomes across different student samples

Given the individual differences among students, the benefits derived from student-teacher interactions likely vary across performance groups. To further explore this, the model's fit was examined after excluding students from different performance groups, as shown in Table 2. The results indicate that in general the model fit improved the most after excluding the top 10% of students (i.e., students with scores in the 0–90% range).

Table 2. The model fit for students at different performance levels.

	Chinese			**Math**			
	N	**Broadband Wald chi2(2)**	**High-beta band Wald chi2(2)**	**N**	**Broadband Wald chi2(2)**	**Delta Wald chi2(2)**	**High-beta band Wald chi2(2)**
ALL	37	6.68	13.17	54	11.0	12.82	5.48
0%-90%	34	**7.41**	13.39	49	**16.76**	**15.69**	**6.52**
0%-80%	30	6.78	**13.96**	44	10.71	7.42	4.67
0%-70%	26	5.26	6.34	38	7.06	6.89	4.66
0%-60%	23	5.03	6.39	33	1.97	4.89	1.48
40% - 100%	22	5.32	5.07	32	12.41	12.94	4.89
30% - 100%	25	6.69	9.42	37	**13.46**	15.26	**6.65**
20% - 100%	29	6.87	9.74	43	8.64	**16.67**	6.12
10% - 100%	33	**7.76**	**11.70**	48	8.88	15.11	5.08

Discussion

This study investigates the impact of student-teacher inter-brain coupling values on learning outcome in real Chinese and math classes over six periods (i.e., three semesters) at a middle school. By utilizing longitudinal data, we identified a positive impact of student-teacher inter-brain coupling on learning outcomes. Specifically, the student-teacher ΔTI and TI values within the broadband and high-beta frequency bands were found to positively forecast the improvement in Chinese academic achievement, while Δhigh-beta TI along with TI values across broadband, delta, and high-beta bands, positively forecast the improvement in math scores. Furthermore, in static model I, no significant coefficient was observed between student-teacher inter-brain coupling and students' learning outcomes in Chinese classes, while in math classes, student-teacher inter-brain coupling in the delta frequency band was negatively correlated with students' learning outcomes within the same period. In static model II, higher broadband and high-beta TI values were predictive of Chinese

scores, while increased broadband TI and Δgamma TI were linked to better math performance. The predictive capacity of student-teacher coupling extended into subsequent semesters, suggesting a sustained impact on academic success. Finally, the study found that the model fit improved when the top 10% of students were excluded. These findings underscore the importance of the dynamic model in forecasting the relationship between student-teacher inter-brain coupling and academic performance, offering valuable insights into the neurophysiological mechanisms that support effective student-teacher interactions in real classroom scenario.

This study, through the application of dynamic modeling, reveals that students who experience enhanced inter-brain coupling with their teacher during class tend to achieve higher academic gains. Specifically, those students who benefit the most are those exhibiting elevated levels of neural synchrony with their teachers, particularly when this synchrony shows significant enhancement. This finding underscores the practical significance of Δinter-brain coupling values as a valuable metric in real-world educational settings. By providing feedback to teachers and students regarding the changes in inter-brain coupling, this measure can offer timely, accurate insights into instructional efficacy, facilitating real-time adjustments. Fig. 5A and 5B illustrate the corresponding TI values and ΔTI values, respectively, when ranked by score and Δscore. Specifically, as shown in Fig. 5B, teachers can more intuitively observe that students who achieved greater improvements in their scores consistently exhibited higher ΔTI values. For novice teachers, who may find it challenging to assess the benefits received by diverse student groups, the increase in inter-brain coupling offer an efficient way to obtain feedback on teaching efficacy. Additionally, the increase in inter-brain coupling can provide students with personalized feedback on their learning progress, empowering them to better understand and reflect on their performance, and adapt their study strategies accordingly.

[Figure 5 About Here]

The dynamic model employs the rise in student-teacher inter-brain coupling to forecast score improvements, effectively mitigating biases arising from teacher's self-selection behaviors. In real classroom settings, teachers often adopt diverse instructional strategies, with some prioritizing a broader reach that includes students with mid or lower performance levels, potentially leading to a negative correlation between student-teacher coupling and learning outcomes. Furthermore, students with different baseline performance levels may exhibit varying degrees of improvement through student-teacher interaction. Consequently, by controlling for baseline student performance, the model more accurately captures the advantages gained through student-teacher interactions. Thus, after accounting for individual factors of both teachers and students, the positive association observed in coupling increments provides compelling evidence that groups with higher student-teacher interaction quality exhibit greater academic improvement. The comparative analysis of model fit between dynamic and static models further indicates that, unlike static measures, increases in inter-brain coupling may serve as a more robust predictor of student

learning outcomes. Prior research has similarly indicated that inter-brain coupling values alone are insufficient for reliably forecasting performance outcome (Bevilacqua et al., 2019), suggesting that static models may encounter uncontrolled confounding variables in predictive contexts. In contrast, the dynamic model emerges as a potentially more effective choice, offering a more robust metric for accurate forecasting.

This study consistently demonstrated a positive predictive effect of TI and ΔTI on both learning outcomes and their improvements. Specifically, in the Chinese subject, the broadband and high-beta frequency band were consistently significant predictors of both learning outcomes and improvements, as well as notable forecasters of enhancements in math scores. The discovery of broadband coupling signifies a relationship between the neural activity similarities of teachers and students at a holistic level and learning outcomes. Previous research on inter-brain coupling between students has similarly affirmed the predictive value of broadband similarity in academic performance (Dikker et al., 2017). The neural signatures observed in the high-beta frequency bands corroborate previous findings in both educational contexts and broader cognitive neuroscience (Baceviciute et al., 2021; Chikhi et al., 2022), solidifying and expanding these observations. Extensive research has linked the beta frequency band to various cognitive functions, including working memory (Deiber et al., 2007; Schapkin et al., 2020), language processing (Weiss & Mueller, 2012), reward processing and long-term memory (Hanslmayr et al., 2014). The high-beta frequency, in particular, is believed to play a more significant role in top-down information processing (Buschman & Miller, 2007; Kornblith et al., 2016; Marco-Pallarés et al., 2015). Furthermore, the immature brain development of early adolescents' results in distinct patterns of brain activity compared to late adolescents and adults. Specifically, the brain activity of middle school students is more vibrant, exhibiting higher power across all frequency bands, particularly in fast wave activity such as the high-beta bands (Lüchinger et al., 2011; Segalowitz et al., 2010). This suggests that the high-beta frequency band plays a critical role in the sophisticated cognitive activities typical of early adolescents. Additionally, our exploration of the lagged effects of this predictive capability reveals that it accurately forecasts performance only for the most recent 1-2 periods (i.e., two to four months), implying a closed relationship between coupling and academic achievement within a reasonable duration.

Several limitations of this study warrant discussion. First, to minimize disruption to real classroom environments, we employed wearable, portable devices for data collection. However, these devices have limitations, particularly in the number of available channels. Future studies may benefit from using multimodal wearable equipment, integrating additional physiological signals such as skin conductance and heart rate to enhance model development and predictive accuracy. Second, variations in classroom structures—shaped by factors such as national, cultural, and grade-level

differences–may also impact results, underscoring the need for more diverse measurements in future research.

In conclusion, this study explored the profound directional impact of student-teacher interactions on learning outcomes within real classroom environments, offering neurophysiological insights. Our findings highlight the pivotal role of these interactions in enhancing effective learning, as analyzed from a neuroscientific perspective. By affirming the critical importance of increased student-teacher inter-brain coupling in forecasting academic improvement, this research further elucidates the intricate mechanisms underlying human learning, thereby enriching our understanding of the cognitive foundations of education.

Materials and Methods

Participants.

Forty-eight first-year middle school students (22 males; 12-14 years old) and four teachers (two Chinese teachers and two math teachers) from two separate classes (23 students from class 1; 25 students from class 2) from a middle school participated in the study.

Three students were discarded from the analysis due to the unsatisfactory EEG retention rates (>80% of epochs were removed, see more details in the “Data preprocessing” section). As a result, 45 participants were included in the analyses.

Data Collection.

EEG data.

The EEG data collection took place in real classroom environment and spanned a total duration of 12 months, covering three academic semesters. Data acquisition extended over four months per semester, with each half-semester identified as a distinct 'period'. Within these periods, data collection took place for a week every month, during which one to two sessions were recorded daily from two separate student classes. Specifically, for Class 1, 62 math and 59 Chinese sessions were recorded, while Class 2 accounted for 55 math and 50 Chinese sessions. Each lesson had a duration of 40 minutes, during which the teachers conducted their planned teaching activities, and students engaged in their regular learning tasks. Both students and teachers equipped themselves with the EEG headbands before each session and removed them subsequently, as directed by the research team.

EEG data were captured using a dual-channel headband featuring dry electrodes, specifically positioned at the Fp1 and Fp2 regions of the frontal cortex. The data recording occurred at a sampling rate of 250 Hz, employing the Brainno device

(SOSO H&C, South Korea). The reference electrode was clipped to the right ear, serving as the ground for Fpz.

Student's learning outcomes.

Data on the students' academic achievement were gathered through their midterm and final examination scores in Chinese and math over three semesters, encompassing a total of six examinations, provided by the school administration.

Data Processing.

Student's learning outcome preprocessing.

We utilized standardized ranking information to represent students' learning outcomes, enabling a consistent and equitable comparison of student performance across various exams and cohort sizes. Students were first ranked on a scale from 1 to *N*, where *N* represents the total number of students in the cohort. The highest-performing student was assigned a rank of 1, progressing sequentially to the student with the lowest score, who received the rank of *N*. This approach effectively mitigating the issues of score saturation and truncation and highlights the relative performance among students, enhancing the robustness of the model. This initial ranking was subsequently standardized using a specific formula:

$$Score = \frac{(N+1) - Initial\ Rank}{N} \quad \#(1)$$

Where $Score$ denotes students' standard score of students, representing their academic achievement. N denotes the total number of students, and the $Initial\ Rank$ indicates the ranking of individual student within the cohort. The $Score$ ranges between 0 and 1, with higher values signifying better performance.

EEG data preprocessing.

The EEG data for each student were preprocessed through a five-step procedure for both channels, consistent with methodologies employed in previous studies (Chen et al., 2023; Xu et al., 2024). The preprocessing steps included the identification of missing data, robust detrending, band-pass filtering, waveform segmentation, and ocular artifact mitigation. Subsequently, the data from each channel were segmented into 30-second epochs. For a single 40-minute class session, this resulted in a total of 160 epochs across both channels for every student (*Fig. S3*).

Additionally, 33 of the participants were also participates in a separate research project, during which eyes-open and eyes-closed EEG data in resting states were collected. The spectral analysis was conducted on resting state data. The spectrum of

the eye-closed data prominently featured a peak in the alpha frequency band, thereby affirming the reliability of the collected EEG data (*Fig. S4*).

Student-teacher Inter-brain Coupling Calculation.

This study employs the total interdependence values (TI) for assessing student-teacher, student-class and student-excellence inter-brain coupling. This approach relies on the coherence of two signals within the frequency domain, which has been validated for its efficacy (Xu et al., 2024). The utility of TI values as a metric for inter-brain coupling has been increasingly recognized in a variety of real-world contexts, including concert hall, and classroom (Bevilacqua et al., 2019; Chabin et al., 2022; Chen et al., 2023; Dikker et al., 2017). The computation of TI values is as follow:

$$TI_{xy} = -\frac{2}{f_s}\sum_{i=1}^{N-1} \ln\left(1 - C_{xy}^2(i\Delta f)\right)\Delta f \quad \#(2)$$

where C_{xy} denotes the spectral coherence of the signal at the frequency f, computed by the Welch method. Here, f_s denotes the sampling rate, and N represents the number of data points between 0 and the Nyquist sampling rate $f_s/2$. The frequency resolution Δf is computed employing the relation $\frac{f_s}{2(N-1)}$. In this analysis, TI values were computed across the EEG's broad frequency spectrum, ranging from 1 to 40Hz. Calculations also incorporated individual frequency bands: delta (1-4Hz), theta (4-8Hz), alpha (8-13Hz), low beta (13-18Hz), high-beta (18-30Hz), and gamma (30-40Hz).

Within each session, the TI value for each student-teacher pair was determined by averaging the TI values across all epochs that successfully matched between the student and teacher, considering data from both recording electrodes (as shown in Fig. 1C). A 'segment' was defined as a matched epoch, recognized as such only when it included artifact-free and valid data from both the teacher and the student. A minimum threshold was established, requiring that only student-teacher pairs with at least six valid segments were to be included (Chen et al., 2023; Dikker et al., 2017). For Chinese classes, the average proportion of valid student-teacher coupling segments for each session was 50.2%, while for math classes, it was 46.6% (*Fig. S5*).

Following this, the inter-brain coupling values within each period (i.e., half-semester) were averaged for each student. To ensure the quality of the data, student samples with valid sessions constituting less than 20% within a period were omitted from the analysis. These averaged TI values were then normalized to a range of [0, 1] for each frequency band.

Regression Analysis.

Linear mixed-effects (LME) models were fitted to the data to investigate the relationship between student-teacher TI and academic achievement, which allow to effectively partition the overall variation of the dependent variable into components corresponding to different levels of data hierarchy (Gałecki et al., 2013). According to the dependent variables, two kinds of models are designed. The *dynamic model* (equations 5-6) uses the difference in test score between *period* t and the previous *period* as the dependent variable, representing the improvement of learning achievement, while the *static model* (equations 7-8) uses standardized test score as dependent variable. Given the potential non-independence of EEG data within each individual students or time *period*, the two categorical variables are employed to classify the observations into separate groups, to account for correlated observations in the data (Pinheiro & Bates, 2000).

The intraclass correlation analysis.

The intraclass correlation coefficient (ICC) was first calculated for each categorical variable, in order to determine if it is appropriate to include the corresponding random effect. The formula of ICC is as follows:

$$ICC = \frac{Var_{between}}{Var_{between} + Var_{within}} \quad \#(3)$$

where $Var_{between}$ denotes the between-group variation of the dependent variable, and Var_{within} denotes the within-group variation. The ICC value is between 0 and 1, where a higher ICC value indicates larger distinctions between groups. An ICC value above 0.05 highlights noteworthy intraclass correlation that should be considered (Bujang & Baharum, 2017).

Forecasting of academic improvement.

The *dynamic* model is applied in order to investigate whether student-teacher coupling can forecast student academic achievement. The investigation aimed to explore whether the increase in TI values could dynamically affect changes in academic achievement after controlling for the initial impact of TI. In pursuit of this, the model integrated both the baseline TI values and their incremental growth as independent variables, with the growth of standardized scores as the dependent variable in the analysis.

The *dynamic models* are:

$$\Delta Score_{it} = Score_{it} - Score_{i(t-1)}$$
$$= \beta_0 + \beta_1 frequencyTI_{i(t-1)} + \beta_2 \Delta frequencyTI_{it} + u_i + \mathrm{u_t} + \varepsilon_{it} \ \#(4)$$

where $\Delta Score_{it}$ denotes the improvement of students' academic achievement;

$\mathrm{Score_{it}}$ represents the standardized score of student i in period t, and $\mathrm{Score_{i(t-1)}}$

denotes the standardized score of the same student i in the preceding period, $t-1$; $frequencyTI_st_{i(t-1)}$ indicate the vector of student-teacher TI values of student i in period $t-1$ in different frequency bands (i.e., broadband, delta, theta, alpha, low-beta, high-beta, and gamma bands); $\Delta frequencyTI_{it}$ is the difference between $frequencyTI_{it}$ and $frequencyTI_{i,t-1}$; u_i is the individual fixed ability (or say, random effect of student), u_t is the period fixed characteristics (or say, random effect by period), and ε_{it} denotes the random error term.

ICC was calculated first. For Δscore of Chinese, the ICC for student individual (ICC_{stu_id}) was 0.000 and for period factor (ICC_{period}) was 0.000. In the case of Δscore of math, ICC_{stu_id} was 0.000, and ICC_{period} was 0.137. As a result, in the forecasting of Δscore of Chinese, the individual factor was treated as the random effect. For the forecasting of Δscore of math, the period factor was treated as the random effect.

Forecasting of student's academic achievement.

The study then investigates the influence of student-teacher inter-brain coupling on the students' learning outcome. The inter-class correlation analysis (ICC) was first conducted to refine the regression models. For Chinese score, the ICC for individual student (ICC_{stu_id}) was 0.796, and for *period* (ICC_{period}) it was 0.000; For math score, ICC_{stu_id} was 0.625, and ICC_{period} was 0.000. Thus, only the individual level was treated as random effect in the analysis.

The *static model I* was designed as:

$$Score_{it} = \beta_0 + \beta_1 frequencyTI_{st\,it} + u_i + \varepsilon_{it} \quad \#(5)$$

The *static model II* was designed as:

$$Score_{it} = \beta_0 + \beta_1 frequencyTI_{st\,i(t-1)} + \beta_2 \Delta frequencyTI_{st\,it} + u_i + \varepsilon_{it} \quad \#(6)$$

where all the symbols are consistent with those defined before (*See formula 4*).

Forecasting of student's academic achievement over multiple lagged periods.

To investigate whether the student-teacher coupling has a long-term predictive effect on score and Δscore, this study utilized current TI and ΔTI to forecast scores over multiple lagged periods. The formula is as follow:

$$\Delta Score_{i(t+lag\ period)} = \beta_0 + \beta_1 frequencyTI_{i(t-1)} + \beta_2 \Delta frequencyTI_{it} + u_i + \varepsilon_{it} \quad \#(7)$$

$$Score_{i(t+lag\ period)} = \beta_0 + \beta_1 frequencyTI_{i(t-1)} + \beta_2 \Delta frequencyTI_{it} + u_i + \varepsilon_{it} \quad \#(8)$$

where $lag\ period$ denotes the number of periods by which the score lags behind ΔTI, $Score_{i(t+lag\ period)}$ is the achievement of student i in period $t + lag\ period$; and other symbols are consistent with before (*See formula 4, 5*). For example, if the $lag\ period$ is set to 2, and t equals 3, this implies forecasting the score for the 5th period using the TI from the 2nd period and the ΔTI from the 2nd to 3rd period.

We conducted nonparametric permutation tests to further verify the correlation between inter-brain coupling and learning outcomes. To assess the significance of our findings, we generated 5,000 random permutations of the independent and dependent variables and calculated the distribution of forecasted coefficients from these shuffled datasets. By comparing the original coefficient values with those generated from the 5,000 shuffled versions, we evaluated the robustness of our results. Results with p-values less than 0.05 from both parametric and nonparametric tests are considered significant (Hou et al., 2020).

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Assor, A., & Connell, J. P. (1992). The validity of students' self-reports as measures of performance affecting self-appraisals. *Student perceptions in the classroom*, *13*, 25-47.

Baceviciute, S., Terkildsen, T., & Makransky, G. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality. *Computers & Education*, *164*, 104122.

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., & Dikker, S. (2019). Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: Evidence from a real-world classroom electroencephalography study. *Journal of cognitive neuroscience*, *31*(3), 401-411.

Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofacial Science*, *12*(1).

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *science*, *315*(5820), 1860-1862.

Chabin, T., Gabriel, D., Comte, A., Haffen, E., Moulin, T., & Pazart, L. (2022). Interbrain emotional connection during music performances is driven by physical proximity and individual traits. *Annals of the New York Academy of Sciences*, *1508*(1), 178-195.

Chen, J., Qian, P., Gao, X., Li, B., Zhang, Y., & Zhang, D. (2023). Inter-brain coupling reflects disciplinary differences in real-world classroom learning. *npj Science of Learning*, *8*(1), 11.

Chikhi, S., Matton, N., & Blanchet, S. (2022). EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology*, *59*(6), e14009.

Deiber, M.-P., Missonnier, P., Bertrand, O., Gold, G., Fazio-Costa, L., Ibanez, V., & Giannakopoulos, P. (2007). Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics. *Journal of cognitive neuroscience*, *19*(1), 158-172.

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., & Ding, M. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current biology*, *27*(9), 1375-1380.

Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, *5*(3), 289-318.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, *27*(2), 305-352.

Gałecki, A., Burzykowski, T., Gałecki, A., & Burzykowski, T. (2013). *Linear mixed-effects model*. Springer.

Gandolfo, G. (1971). *Economic dynamics: Methods and models* (Vol. 16). Elsevier.

Goudeau, S., Easterbrook, M. J., & Fayant, M.-P. (2024). How to do research in classroom settings. *Nature Human Behaviour*.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child development*, *76*(5), 949-967.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms.

Hanslmayr, S., Matuschek, J., & Fellner, M.-C. (2014). Entrainment of prefrontal beta oscillations induces an endogenous echo and impairs memory formation. *Current biology*, *24*(8), 904-909.

Holper, L., Goldin, A. P., Shalom, D. E., Battro, A. M., Wolf, M., & Sigman, M. (2013). The teaching and the learning brain: A cortical hemodynamic marker of teacher–student interactions in the Socratic dialog. *International Journal of Educational Research*, *59*, 1-10.

Hou, Y., Song, B., Hu, Y., Pan, Y., & Hu, Y. (2020). The averaged inter-brain coherence between the audience and a violinist predicts the popularity of violin performance. *NeuroImage*, *211*, 116655.

Konvalinka, I., & Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Frontiers in human neuroscience*, *6*, 215.

Kornblith, S., Buschman, T. J., & Miller, E. K. (2016). Stimulus load and oscillatory activity in higher cortex. *Cerebral Cortex*, *26*(9), 3772-3784.

Liker, J. K., Augustyniak, S., & Duncan, G. J. (1985). Panel data and models of change: A comparison of first difference and conventional two-wave models.

Social Science Research, *14*(1), 80-101.

Liu, J., Zhang, R., Geng, B., Zhang, T., Yuan, D., Otani, S., & Li, X. (2019). Interplay between prior knowledge and communication mode on teaching effectiveness: interpersonal neural synchronization as a neural marker. *NeuroImage*, *193*, 93-102.

Lüchinger, R., Michels, L., Martin, E., & Brandeis, D. (2011). EEG–BOLD correlations during (post-) adolescent brain maturation. *NeuroImage*, *56*(3), 1493-1505.

Marco-Pallarés, J., Münte, T. F., & Rodríguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. *Neuroscience & Biobehavioral Reviews*, *49*, 1-7.

Miller, A. L. (2011). Investigating Social Desirability Bias in Student Self-Report Surveys. *Association for Institutional Research (NJ1)*.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European journal of social psychology*, *15*(3), 263-280.

Pace, C. R. (1985). The credibility of student self-reports.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological science in the public interest*, *9*(3), 105-119.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In *Handbook of research on student engagement* (pp. 365-386). Springer.

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: basic concepts and examples. *Mixed-effects models in S and S-Plus*, 3-56.

Saha, L. J., & Dworkin, A. G. (2009). *International handbook of research on teachers and teaching* (Vol. 21). Springer Science & Business Media.

Schapkin, S., Raggatz, J., Hillmert, M., & Böckelmann, I. (2020). EEG correlates of cognitive load in a multiple choice reaction task. *Acta neurobiologiae experimentalis*, *80*(1), 76-89.

Schmeck, R. R. (2013). *Learning strategies and learning styles*. Springer Science & Business Media.

Segalowitz, S. J., Santesso, D. L., & Jetha, M. K. (2010). Electrophysiological changes during adolescence: a review. *Brain and Cognition*, *72*(1), 86-100.

Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English language teaching*, *3*(4), 237-248.

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends in cognitive sciences*, *23*(8), 699-714.

Tuma, N. B. (1984). *Social dynamics models and methods*. Elsevier.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.

Weiss, S., & Mueller, H. M. (2012). “Too many betas do not spoil the broth”: the role of beta brain oscillations in language processing. *Frontiers in Psychology*, *3*, 201.

Xu, X., Kong, Q., Zhang, D., & Zhang, Y. (2024). An evaluation of inter-brain EEG coupling methods in hyperscanning studies. *Cognitive Neurodynamics*, *18*(1), 67-83.

Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., Zhang, Z., Han, Z., Liu, L., & Guo, T. (2018). Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Human brain mapping*, *39*(7), 3046-3057.

Zheng, L., Liu, W., Long, Y., Zhai, Y., Zhao, H., Bai, X., Zhou, S., Li, K., Zhang, H., & Liu, L. (2020). Affiliative bonding between teachers and students through interpersonal synchronisation in brain activity. *Social Cognitive and Affective Neuroscience*, *15*(1), 97-109.

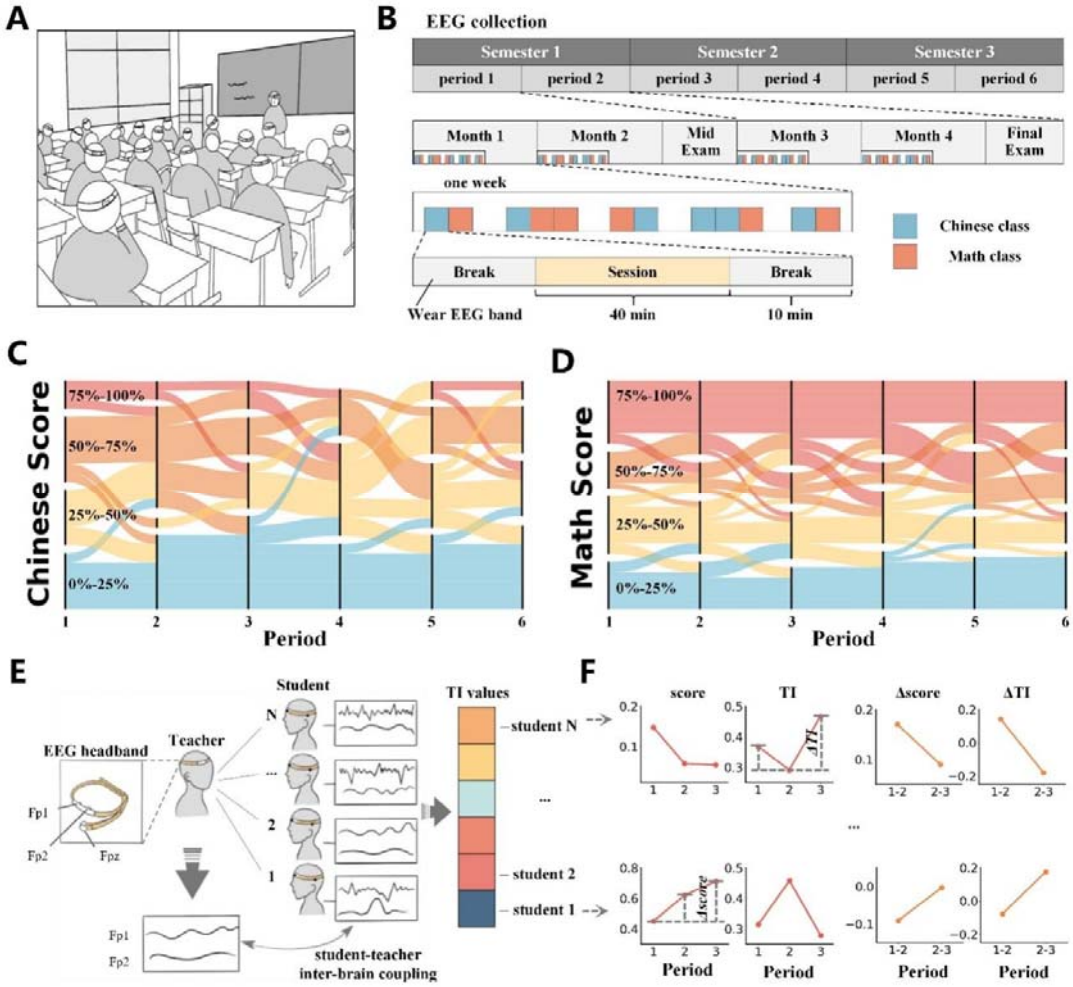

Fig. 1. An illustration of the data collection process. (A), Schematic representation of data acquisition in a real classroom setting. **(B),** Outline of the data collection process. **(C),** Variation in Chinese scores across different periods. **(D),** Variation in math scores across different periods. **(E),** An illustration of the calculation of student-teacher inter-brain coupling of each session. The total interdependence (TI) value is calculated for each student per session, then averaged across all sessions within each *period*. **(F),** Student examples with Chinese scores for each period (first column on the left) and corresponding TI values (second column on the left), along with the differences, i.e. **Δ** Score (second column on the right) and **Δ**TI values (first column on the right).

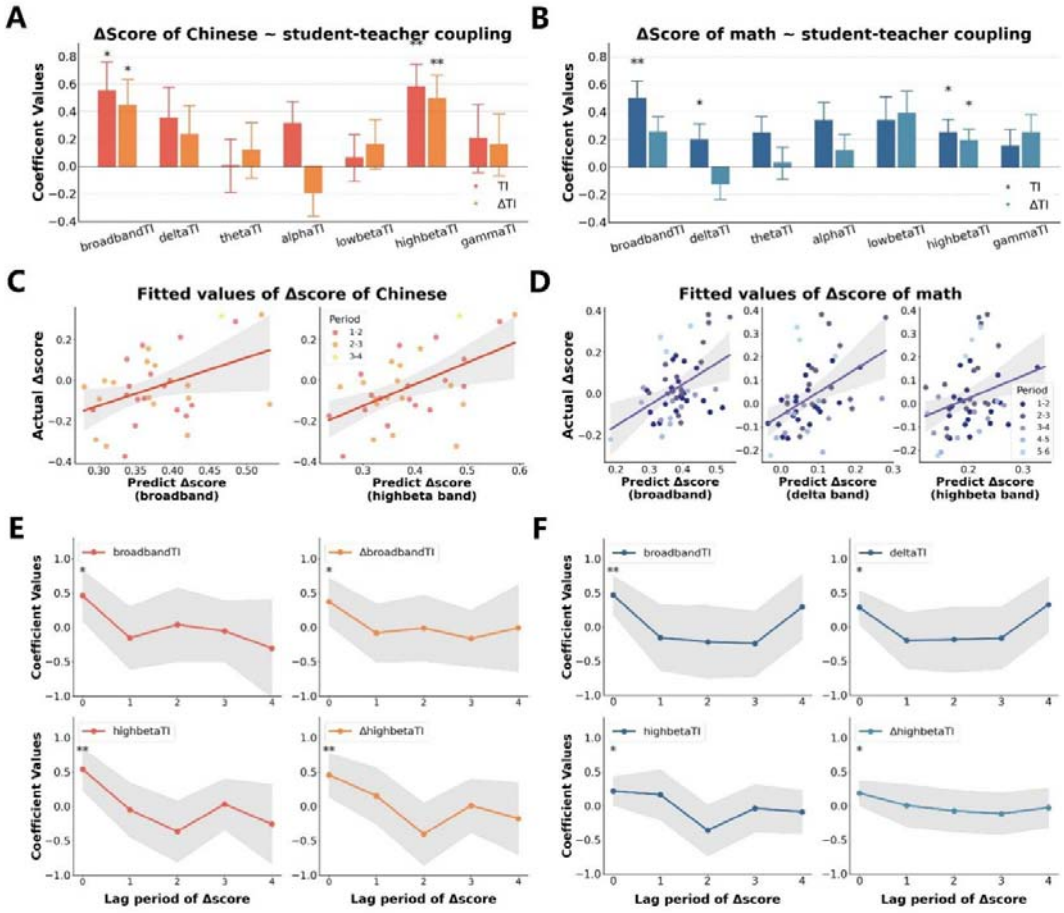

Fig. 2. Coefficients of dynamic model. (A), Coefficients between Chinese improvement and student-teacher coupling, N=37. The horizontal axis represents TI values and ΔTI values across different frequency bands, while the vertical axis corresponds to the associated coefficient values. Asterisks denote significant correlation, *p□<□0.05, **p□<□0.01. **(B),** Coefficients between math improvement and student-teacher coupling, N=54. **(C),** The fitted values of forecasting Δscore of Chinese. The horizontal axis denotes forecasted Δscore, while the vertical axis represents the actual values of Δscores. Data points from different *periods* are denoted by distinct colors. **(D),** The fitted values of forecasting Δscore of math. **(E),** Coefficients of lagged models in Chinese. The horizontal axis indicates the lagged periods, and the vertical axis represents the coefficient values between TI/ΔTI and Δscore in Chinese. The shaded areas represent the 95% confidence intervals. **(F),** Coefficients of lagged models in math.

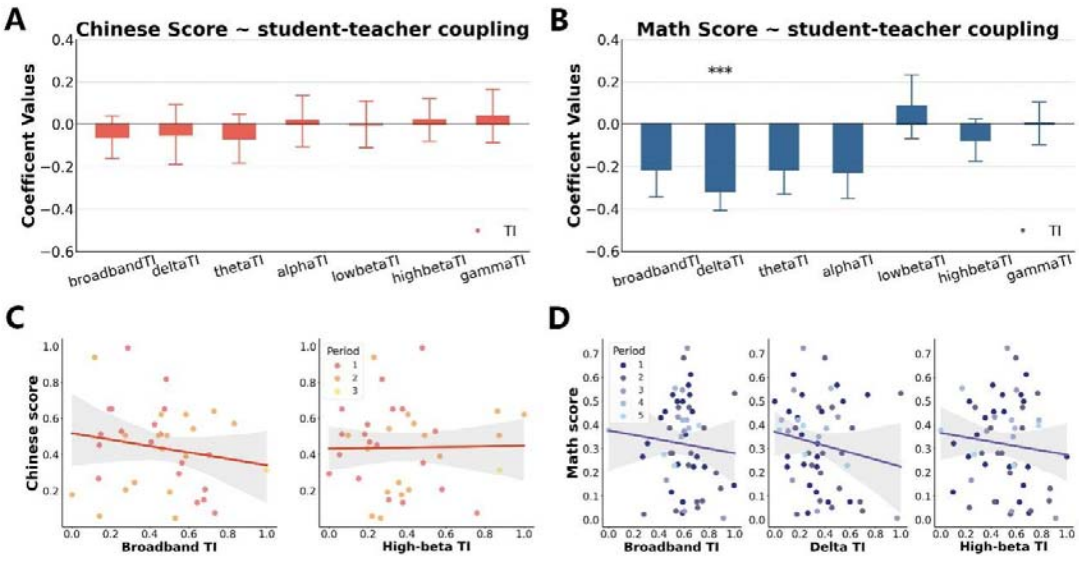

Fig. 3. Coefficients of static model I. (A), Coefficients between Chinese scores and student-teacher coupling, N=37. The horizontal axis represents TI values across different frequency bands, while the vertical axis corresponds to the associated coefficient values. **(B),** Coefficients between math scores and student-teacher coupling, N=54. Asterisks denote significant correlation, *p□<□0.05, **p□<□0.01, ***p<0.001. **(C),** The fitted values for forecasting Chinese score. The horizontal axis denotes broadband TI and high-beta band TI, while the vertical axis represents the values of scores. Data points from different *periods* are denoted by distinct colors. **(D),** The fitted values for forecasting math score.

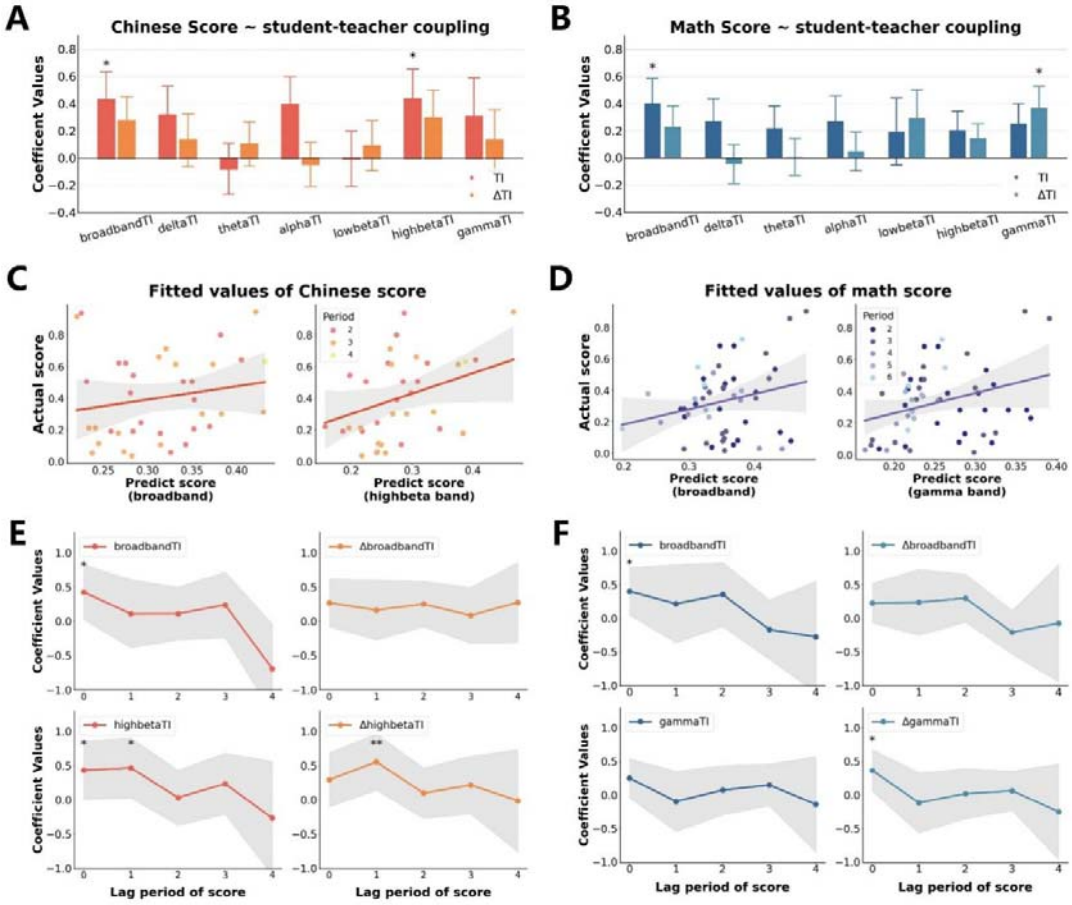

Fig 4. Coefficients of static model II across different frequency bands. (A), Coefficients between Chinese score and student-teacher coupling, N=37. **(B),** Coefficients between math learning outcomes and student-teacher coupling, N=54. **(C),** The fitted values for forecasting Chinese score. **(D),** The fitted values for forecasting math score. **(E),** Coefficient values for forecasting multi-period Chinese score. **(F),** Coefficient values for forecasting multi-period math score.

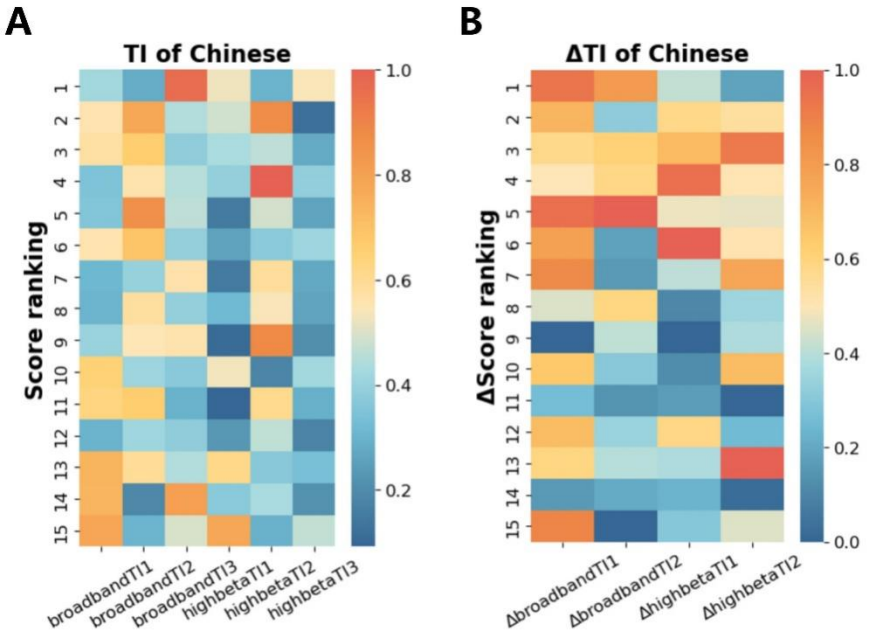

Fig 5. Visualization of student-teacher inter-brain coupling values across different score distributions. (A), Heatmap of TI values sorted by Chinese score ranking. The color gradient represents the magnitude of TI values, with darker shades indicating higher values. Students are arranged from top to bottom in descending order of their Chinese scores, and the x-axis corresponds to broadband TI values and high-beta TI values across periods 1, 2, and 3. **(B),** Heatmap of ΔTI values sorted by predicted Δscore of Chinese. The color gradient represents the magnitude of ΔTI values, with darker shades indicating higher values. Students are arranged from top to bottom in descending order of their score improvements, and the x-axis corresponds to Δbroadband TI values and Δ high-beta TI values across periods 1 and 2. Only students who have data for all the periods are included.