npj | science of learning

Article

Published in partnership with The University of Queensland

https://doi.org/10.1038/s41539-024-00286-y

Structured interaction between teacher and student in the flipped classroom enhances learning and interbrain synchrony

Check for updates

Qi Li[1], Die Wang[1], Weilong Xiao[1,2,3], Yingying Tang[4], Qi Sun[1,2] ✉, Binghai Sun[1,2,3] ✉ & Zhishan Hu[4,5] ✉

Studies have found that flipped classroom teaching (FT) improves learning compared to lecture-based teaching (LT). However, whether the structured teacher–student interaction—the key feature of FT—plays an essential role in enhancing learning remains unclear, as do its neural underpinnings. Here, we compared three teaching conditions: FT with a video lecture and structured interaction, LT with a face-to-face lecture and spontaneous interaction, and control teaching (CT) with a video lecture and spontaneous interaction. The fNIRS-based hyperscanning technique was used to assess the interbrain synchrony (IBS) from teacher-student dyads. Results showed that the learning was significantly improved in FT than in LT and CT, and FT significantly increased teacher–student IBS in left DLPFC. Moreover, the IBS and learning improvements were positively correlated. Therefore, these findings indicate that the structured teacher–student interaction is crucial for enhancing learning in FT, and IBS serves as its neural foundation.

Teaching, a cornerstone of knowledge transfer in human society, has facilitated the sharing of experience and wisdom across generations. In the traditional lecture-based teaching (LT) model, teachers deliver lectures directly, with students primarily engaging as passive and note-taking listeners, although they can ask questions and receive immediate responses[1–3]. The flipped classroom teaching (FT) model, leveraging video lectures before class to make room for interactive discussions during class, is on the rise in modern educational settings[4–6]. Research has demonstrated that the FT improves students' learning in subject areas such as humanities, mathematics, and engineering[3,7]. Moreover, numerous studies have found that FT enhances learning compared to LT[5,7–9]. Compared with the LT, the FT includes video lectures and structured teacher–student interaction phases. However, it is yet to be elucidated whether the structured teacher–student interaction plays a key role in improving learning outcomes in the FT and its underlying neural basis.

The FT model generally consists of two phases: the lecture phase and the knowledge consolidation phase (also known as the teacher–student interaction phase)[10–12]. In the lecture phase, students engage with lecture videos to acquire foundational knowledge. Subsequently, in the teacher–student interaction phase, they delve deeper into the material through face-to-face discussions with the teacher. Additionally, before attending the FT class, the teacher generally arranges the interaction process step by step, generating a structured teacher–student interaction phase. In contrast, the teacher–student interaction is spontaneous or voluntary in the LT model, which happens in the process of lecturing. If students have no questions and the teacher does not initiate queries, interaction between teacher and students would not occur. Hence, it is reasonable to conclude that the FT will bolster communication during the interaction phase, thanks to prior knowledge acquisition and the well-planned schedule.

According to the interactive-constructive-active-passive theory (ICAP)[13], the interaction (e.g., teacher–student interaction), with a sufficient degree of turn-taking and predominantly constructive statements by interactants, promotes students' learning. During the interaction, teachers' explanations and feedback help to complete students' knowledge schemas by filling in the gaps in their understanding. Moreover, teachers prompt students to reevaluate their existing knowledge, correct misconceptions, and infer new insights. Empirical studies have found that well-arranged teacher–student interactions are effective in improving learning[14,15]. For

[1]School of Psychology, Zhejiang Normal University, Jinhua, P. R. China. [2]Intelligent Laboratory of Zhejiang Province in Mental Health and Crisis Intervention for Children and Adolescents, Jinhua, P. R. China. [3]Research Center of Tin Ka Ping Moral Education, Zhejiang Normal University, Jinhua, P. R. China. [4]Neuroimaging Core, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China. [5]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, P. R. China. ✉e-mail: sunqi_psy@zjnu.edu.cn; jky18@zjnu.cn; huzhishan@sjtu.edu.cn

instance, a previous study found that students' learning performance was significantly improved when teachers used scaffolding strategies (e.g., asking guiding questions or providing hints) to individualize instruction for students[15]. Therefore, the structured teacher–student interaction in the FT model may play a vital role in enhancing learning.

Moreover, the educational benefits of teacher–student interaction can be due to the enhanced interbrain synchrony (IBS), which is defined as the temporal coupling of neural systems between two individuals engaged in social interaction[16]. The burgeoning of "second-person neuroscience"[17] has prompted an increasing number of researchers to use neuroimaging techniques (e.g., fNIRS) to simultaneously measure the brain activity of two or more individuals during interaction (known as "hyperscanning")[18]. Studies have revealed that the teacher–student interaction enhances the IBS in the prefrontal cortex (PFC) and the temporal-parietal junction (TPJ)[15,19,20]. Importantly, students' learning outcomes were improved with the increase in the IBS[15,19–21]. For example, a previous study found that when teachers adopted a face-to-face communication mode to teach students with prior knowledge, the IBS in the left PFC between teachers and students was enhanced, and the larger the IBS, the better the student's academic performance[22]. Hence, the IBS may be the neural basis of the FT to enhance students' learning.

In summary, the current study aimed to investigate whether the structured teacher–student interaction is crucial for enhancing learning and to explore the neural basis for improving students' learning in the FT model. It is expected that the FT will enhance learning given its superior teacher–student interaction compared to the LT. Additionally, we expect that the IBS in the FT will be significantly higher than that in the LT and the IBS is positively correlated with the student's learning. Taken together, it can be supposed that the structured interaction in the FT model played a key role in improving learning.

To examine the aforementioned proposals, an experiment was conducted in which a teacher taught three groups of students using three types of teaching models: the FT condition, which contained a video-based lecture phase and a structured teacher–student interaction phase; the LT condition, which included a face-to-face lecture phase and a spontaneous teacher–student interaction phase; and the control teaching (CT) condition, which comprised a video-based lecture phase and a spontaneous interaction phase. The CT condition was designed to directly compare the impact of structured and spontaneous interactions. Moreover, we recorded the brain activity of teacher–student dyads using the functional near-infrared spectroscopy (fNIRS)-based hyperscanning technique during both phases. We compared the learning improvements and the IBS across the three groups and further examined the relationship between learning improvements and IBS. This study seeks to elucidate the important role of structured teacher–student interactions in improving students' learning outcomes, offering empirical evidence to inform future educational reforms.

Results

Behavioral results

Pre- and post-test scores of students in three conditions are displayed in Fig. 1. A two-way ANOVA analysis was performed on these scores. We found a significant main effect of the test phases (F (1,61) = 126.83, $p < 0.001$, $\eta_p^2 = 0.68$, observed power = 1.00), and the lecture types (F (2,61) = 4.75, $p = 0.012$, $\eta_p^2 = 0.13$, observed power = 0.78). Additionally, the interaction between the test phases and lecture types was significant (F (2,61) = 7.74, $p = 0.0010$, $\eta_p^2 = 0.20$, observed power = 0.95). Then, the simple-effect test with Bonferroni correction showed that the post-test scores were higher than the pre-test scores in all conditions (ps < 0.001, Cohen's ds > 1.00). Meanwhile, for the pre-test score, there was no significant difference among the FT, LT, and CT conditions (ps > 0.10, Cohen's ds < 0.23). In contrast, for the post-test score, the FT condition (M = 9.90, SD = 2.68) was significantly higher than the LT condition (M = 7.71, SD = 2.93, $p = 0.049$, Cohen's d = 0.78) and the CT condition (M = 6.75, SD = 2.99, $p = 0.0019$, Cohen's d = 1.11); there was no significant difference between the latter two conditions ($p = 0.83$, Cohen's d = 0.32).

In addition, one-way ANOVAs were performed on likability and appearance attractiveness. No significant main effect of the lecture types was found on likability (teacher: F (2,61) = 0.38, $p = 0.69$, $\eta_p^2 = 0.012$; student: F (2,61) = 1.42, $p = 0.25$, $\eta_p^2 = 0.045$) and appearance attractiveness (teacher: F (2,61) = 0.23, $p = 0.80$, $\eta_p^2 = 0.0075$; student: F (2,61) = 0.76, $p = 0.47$, $\eta_p^2 = 0.024$). Due to the fact that the data did not meet the prerequisite assumption of equal variance, we used the Kruskal–Wallis test to calculate the perceived quantity of teacher–student interaction and found a significant main effect of the lecture types ($\chi^2(2) = 18.95$, $p < 0.001$, $\eta^2 = 0.28$). The perceived quantity of teacher–student interaction was higher in the FT condition (M = 7.24, SD = 1.04) than in the LT (M = 4.33, SD = 2.24, $p < 0.001$, Cohen's d = 1.67, Bonferroni corrected) and CT conditions (M = 4.45, SD = 2.60, $p < 0.001$, Cohen's d = 1.41, Bonferroni corrected), with no significant difference between the LT and CT conditions ($p > 0.10$). Therefore, we suggest that the implementation of the experiment was in line with our expectations. That is, there were more interactive discussions between teacher and student in the FT condition, indicating the implementation of the experiment with fidelity.

GLM tests showed that no significant linear relationships were identified between perceived quantity of interaction and learning improvements (FT condition: $t = -0.84$, $R^2 = 0.036$, $\beta = -0.47$, $p = 0.41$; LT condition: $t = 0.64$, $R^2 = 0.021$, $\beta = 0.17$, $p = 0.53$; CT condition: $t = 1.16$, $R^2 = 0.063$, $\beta = 0.23$, $p = 0.26$).

Interbrain synchrony results

Within the FOI, one-sample t-tests against zero (FDR corrected) revealed the enhanced class-related IBS during the FT condition (Fig. 2a) in the left dorsolateral prefrontal cortex (DLPFC, CH09, M = 0.038, SE = 0.011, t (20) = 3.37, $p = 0.019$, Cohen's d = 0.73, observed power = 0.99; CH11, M = 0.052, SE = 0.014, t (19) = 3.73, $p = 0.017$, Cohen's d = 0.84, observed power = 0.99), superior temporal gyrus (STG, CH13, M = 0.037, SE = 0.012, t (18) = 3.14, $p = 0.027$, Cohen's d = 0.73, observed power = 0.99), and primary somatosensory cortex (S1, CH14, M = 0.038, SE = 0.011, t (19) = 3.63, $p = 0.017$, Cohen's d = 0.81, observed power = 0.99). In contrast, no significant enhanced class-related IBS was observed for the LT (ts < 2.60, ps > 10, Cohen's ds < 0.57, Fig. 2b) and the CT (ts < 2.02, ps > 0.10, Cohen's ds < 0.44, Fig. 2c) conditions.

For the channels showing significant enhanced class-related IBSs, a one-way ANOVA was conducted to compare the class-related IBS across the lecture types. The main effect of the lecture types was significant in the left DLPFC (CH11, F (2,56) = 5.30, $p = 0.031$, $\eta_p^2 = 0.16$, observed power = 0.85, FDR corrected). It should be noted that 5 dyads were excluded from the analysis as the CH11 in them was determined to be a bad channel (see "Methods" for bad channel identification methods). Class-related IBS in the FT condition (M = 0.052, SE = 0.014) was higher than that in the LT condition (M = -0.012, SE = 0.012, $p = 0.012$, Cohen's d = 1.09, Bonferroni corrected) and the CT condition ($M = -0.0029$, SE = 0.019, $p = 0.040$, Cohen's d = 0.76, Bonferroni corrected); however, the latter two were not significantly different ($p > 0.10$, Fig. 3a). No significant main effect was found in other channels (ps > 0.10, FDR corrected; see Supplementary Table 1 for exact p values).

We further examined the lecture and interaction phases. We found that the main effect of the lecture types was significant in the interaction phase (F (2,56) = 6.00, $p = 0.017$, $\eta_p^2 = 0.18$, observed power = 0.89, FDR corrected, Fig. 3c) but not in the lecture phase (F (2,56) = 2.69, $p = 0.19$, $\eta_p^2 = 0.088$, FDR corrected, Fig. 3b). Also, it was not significant for other channels (ps > 0.10, FDR corrected; see Supplementary Table 1 for exact p values).

Correlation between behavioral indices and IBS

The IBS results have highlighted the important role of the left dorsolateral prefrontal cortex (CH11) in the FT condition. Next, we examined whether there was a linear relationship between the IBS in the left DLPFC and the learning improvements in each teaching model. GLM tests showed that the IBS showed a strong linear relationship with learning improvements in both FT and CT conditions, regardless of whether the IBS was class-

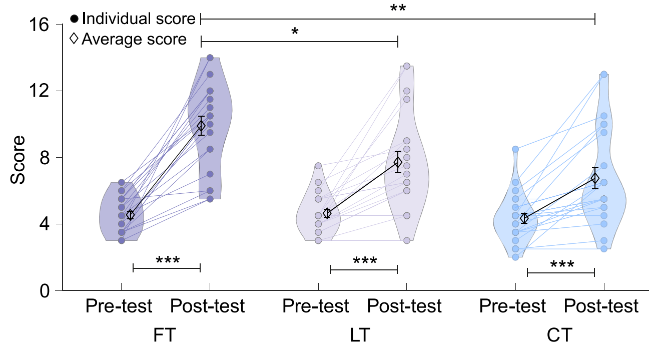

Fig. 1 | Pre- and post-test scores of students in three conditions (FT, LT, and CT). Each dot represents a student's score. Diamonds indicate the mean score across all participants in each group. The error bars indicate the standard error of the mean. Significance levels are marked as follows: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; Abbreviations: FT flipped classroom teaching condition, LT lecture-based teaching condition, CT control teaching condition.

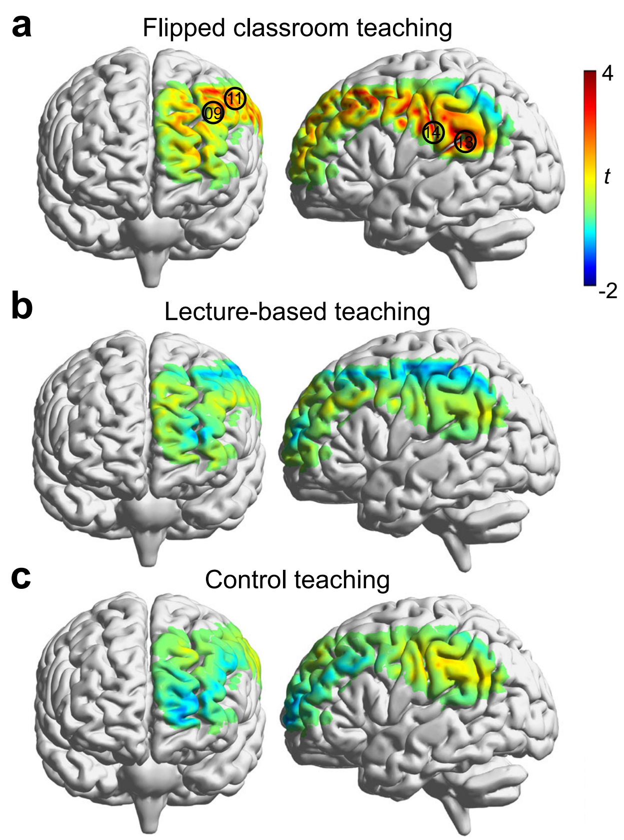

Fig. 2 | T-map from one-sample t-test against zero on class-related IBS. Each row represents a teaching model: **a** flipped classroom teaching, **b** lecture-based teaching, and **c** control teaching, with brighter color representing a higher *t* value. Channels with significant results are labeled accordingly.

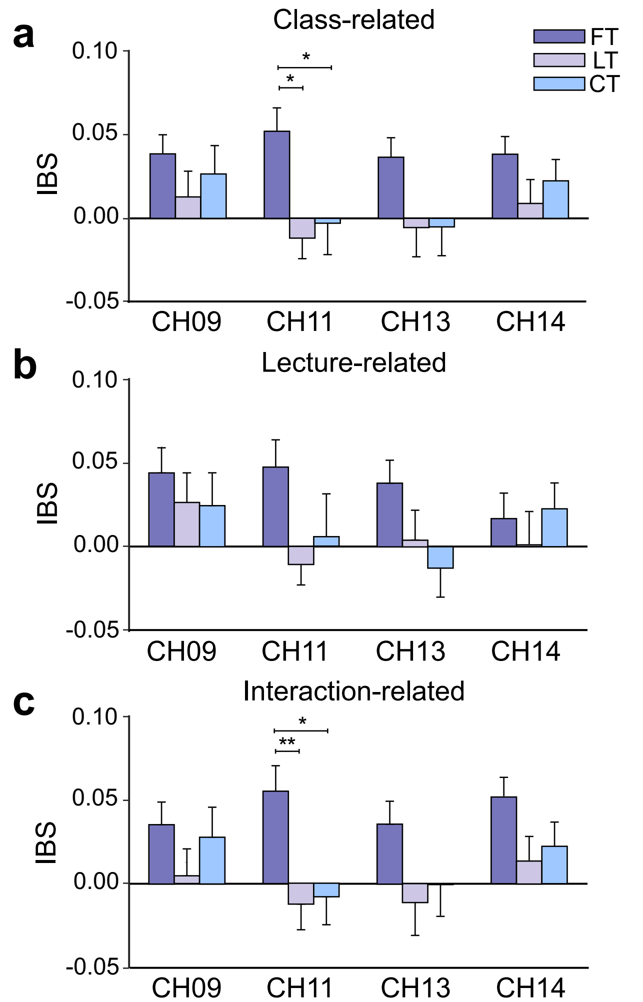

Fig. 3 | Task-related IBS at CH09, CH11, CH13, and CH14 among three conditions (FT, LT, and CT). Class-related (**a**), lecture-related (**b**), and interaction-related (**c**) IBS at CH09, CH11, CH13, and CH14 among three conditions. The error bars indicate the standard error of the mean. Significance levels are marked as follows: *$p < 0.05$; **$p < 0.01$; Abbreviations: FT flipped classroom teaching condition, LT lecture-based teaching condition, CT control teaching condition.

related, lecture-related, or interaction-related. However, this relationship was not identified in LT. The statistical results are displayed in Table 1 and Fig. 4.

In addition, we analyzed the relationships between the perceived quantity of teacher–student interaction and interaction-related IBS for each teaching model. GLM tests revealed no significant linear relationship between them (FT condition: $t = 0.57$, $R^2 = 0.018$, $\beta = -0.0082$, $p = 0.58$; LT condition: $t = -0.53$, $R^2 = 0.015$, $\beta = -0.0036$, $p = 0.61$; CT condition: $t = 1.32$, $R^2 = 0.093$, $\beta = 0.0090$, $p = 0.20$).

Discussion

Since the earliest days of chalk on slate, teaching has relied on the dynamic interaction between the teachers and students. Technological advances have enhanced this interaction. The flipped classroom model (FT), emerging from such developments, encourages students to explore new materials through lecture videos before class. As a result, students can have deeper and more informed interactions with their teachers in the following class session[6,10]. The current study considers that the FT enhances the quality of teacher–student interactions and their neural synchrony, thereby leading to better learning outcomes. Moreover, we suggest that a well-designed, structured interaction is essential for augmenting this advantage.

In the present exploratory study, we conducted a comparative analysis among the FT condition, the traditional lecture-based teaching (LT) condition with a face-to-face lecture and spontaneous interaction, and a control teaching (CT) condition including a video lecture and spontaneous interaction. Through this approach, the study comprehensively examined the impact of structured interaction on improving learning outcomes. Importantly, we employed the fNIRS-based hyperscanning technology to uncover the role of interbrain synchrony (IBS) in the flipped classroom setting.

The behavioral results showed that the FT significantly improved students' learning outcomes compared to both LT and CT, while the latter

Table 1 | Results of the GLM test between task-related IBS at CH11 and learning improvements

	Flipped classroom teaching				Lecture-based teaching				Control teaching			
	t	R^2	β	p	t	R^2	β	p	t	R^2	β	p
Class-related	2.98	0.33	24.42	0.0080	0.41	0.0093	4.42	0.69	2.91	0.33	17.48	0.0098
Lecture-related	3.60	0.42	24.52	0.0020	1.06	0.059	11.51	0.30	2.61	0.29	11.96	0.018
Interaction-related	2.18	0.21	18.30	0.043	0.039	0.001	0.34	0.97	2.71	0.30	18.79	0.015

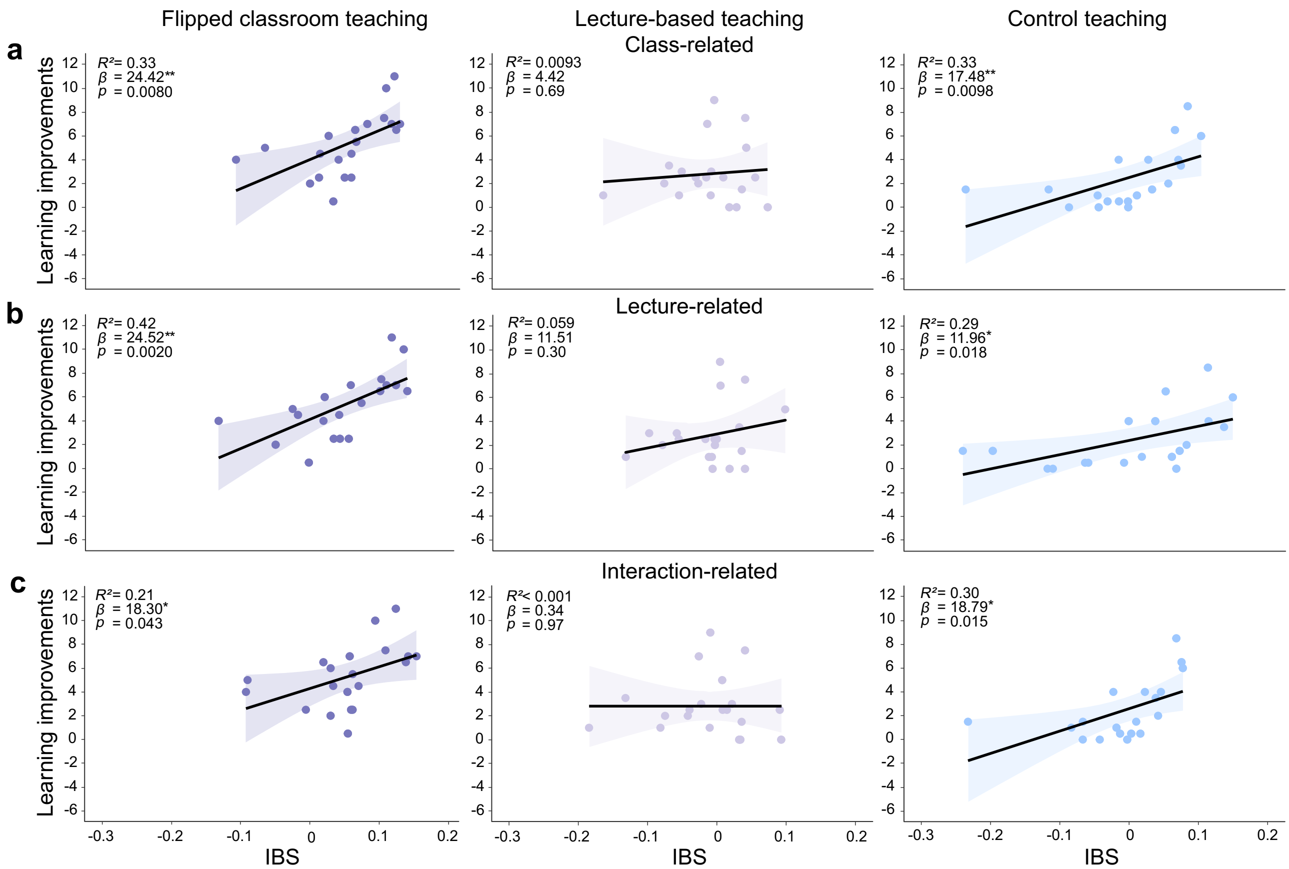

Fig. 4 | Linear regression results between task-related IBS at CH11 and learning improvements. Linear relationship between the class-related (**a**), lecture-related (**b**), and interaction-related (**c**) IBS in left dorsolateral prefrontal (CH11) and learning improvements in flipped classroom teaching (left), lecture-based teaching (middle), and control teaching (right) conditions. *x*-Axis indicates the IBS value, and *y*-axis indicates the learning improvements. Shaded areas indicate 95% confidence interval. Significance levels are marked as follows: $^*p < 0.05$; $^{**}p < 0.01$.

two conditions showed no significant difference. The distinction between LT and CT conditions lay solely in the lecture phase; LT involved an in-person lecture, whereas CT featured a video lecture by the same instructor. This variation in lecture format did not result in different learning outcomes, consistent with a prior study[20], suggesting that the video lecture is comparable to the in-person lecture. Additionally, the FT condition differed from the CT condition solely in the interaction phase, with the FT containing a structured interaction and the CT allowing a spontaneous interaction. This distinction led to significantly enhanced learning in the FT condition, supporting the hypothesis that structured interaction plays a key role in improving learning. These findings indicate that spontaneous interaction offers lower educational benefits than structured interaction. Overall, a well-planned, structured interaction is crucial for achieving positive teaching outcomes, which supports and expands the ICAP theory[13].

Another interesting finding was that the perceived quantity of teacher–student interaction was higher in the FT condition compared to the LT and CT conditions, but GLM tests revealed no significant linear relationship to learning improvements. These results suggest that the interaction frequency had a minor impact on students' learning enhancement in the FT condition. However, previous evidence indicated that the greater the quantity of teacher–student interaction, the better the students learned[23,24]. This inconsistency could be explained by the differences in measurement methods and learning content. Specifically, Pan and his colleagues quantified interactions by coding videos[23], while we relied on participants' subjective reports. Additionally, the previous study focused on skill learning (a song)[23], whereas our study centered on theoretical knowledge. Skill learning, which requires more practice[25], benefits from frequent interactions that provide more practice opportunities. In contrast, theoretical knowledge learning benefits more from the depth or quality of interaction, where people share their insights revolving around a certain knowledge, facilitating each other's thinking and thus promoting deeper learning. Hence, the key to enhancing student learning outcomes in FT may lie in the quality of structured interaction rather than the frequency alone.

The fNIRS results indicated that the neural synchrony captured by fNIRS-based hyperscanning appears to underlie the enhanced communication in the flipped classroom. The FT alone showed significantly enhanced IBS in the left dorsolateral prefrontal cortex (DLPFC), superior temporal gyrus (STG), and primary somatosensory cortex (S1). These findings align with previous hyperscanning studies[15,22,26,27], which also observed enhanced IBS in these cortical areas during teacher–student instruction, collaborative learning, and group engagement in creative problem-solving. Specifically, the superior temporal cortex, a critical area for theory of mind and mentalizing[28], is involved in processing syntactic and semantic information[29,30]. The lecture given in the current study required deep semantic processing of management psychology knowledge. Teacher and students in the FT condition formed a shared knowledge representation reflected by the increase in IBS[15,22] in the STG. Furthermore, the primary somatosensory cortex is widely recognized as a key region for processing sensory and motor signals[27,31], and the left DLPFC plays important roles in various cognitive functions, such as mentalizing[32,33], information integration[34,35], joint attention[36], and monitoring and feedback[37,38]. The increased IBS in these regions indicated intense interaction dynamics between teacher and students, along with a joint attention to educational materials[39,40].

We found that the FT condition exhibited significantly higher IBS in the left DLPFC compared to the LT and CT conditions in the interaction phase. However, no significant difference in the IBS was identified between the LT and CT conditions. These patterns were consistent with the behavioral results regarding learning improvement, implying that increased IBS underlies improved learning outcomes from structured interaction in the flipped classroom. These findings provide neurobiological evidence for the ICAP theory[13].

Importantly, GLM analysis supported the linear relationship between learning improvements and IBS in left DLPFC in the structured interaction phase of FT. Specifically, a well-planned interaction schedule enables progressively deepening engagement between teacher and students, leading to a convergence in their understanding of knowledge. This alignment, driven by extensive cognitive involvement, is reflected by enhanced IBS and improved learning outcomes[19]. In addition, we observed that the perceived amount of teacher–student interaction did not significantly predict the IBS, which might suggest that the increased IBS mainly depended on high-quality teacher–student interaction[41]. In sum, our findings emphasize the critical role of structured designs in the interaction phase for enhancing IBS and learning.

Additionally, we observed that in the spontaneous interaction phase of the CT, IBS in the left DLPFC was correlated with learning improvements, an association not found in the same phase of the LT condition. Combined with the findings about the lecture phase, lecture-related IBS was linked to learning improvements in CT, but not in LT. We conjecture that the effects of the lecture phase might continue into the interaction phase; that is, a kind of "aftereffect" is generated. Specifically, the lecture and interaction phases of each teaching model are closely related. Students attending the video lecture may have a better learning experience than those in the face-to-face lecture[39,40]. The immersive experience of students during the video lecture (owing to minimal distractions) persists into the following interaction phase in CT. Likewise, the distracted state of students during the in-person lecture (due to unpredictable disruptions) continues into the interaction phase in LT. Consequently, the interaction phases show a similar correlation pattern to the lecture phases. Moreover, the spontaneous interaction could be subject to various potentially unpredictable factors[42], which might lead to an unstable link between IBS and learning outcomes. However, empirical evidence is needed to further support this supposition; future studies could track the dynamics and changes in IBS and behavior over time and link coded behavioral labels with IBS point-by-point[23,27] to explore this effect.

Regarding the IBS in the lecture phase, no significant difference in IBS in the left DLPFC was found among the FT, LT, and CT conditions. GLM analysis revealed significant links between lecture-related IBS and learning improvements in FT and CT, but not in LT. Specifically, the lecture videos in FT and CT conditions are prerecorded, allowing for careful control prior to the lecturing, and students watch them alone. As a result, they are exposed to fewer potential distractions, which enables the IBS to be determined by the joint attention on teaching materials[39,40]. In contrast, the on-site lecture in the LT condition exposes both the instructor and students to more potential distractions and additional sensory inputs[43], which might account for the lack of a relationship between the IBS and learning improvements.

Moreover, the current study opens several avenues for future studies. First, this study highlights the benefits of the structured interaction of the flipped classroom. To determine whether a structured interaction following a face-to-face lecture can enhance learning as effectively as that following a video lecture, the former needs to be tested. Second, our study involves one-on-one teaching conducted in a controlled laboratory setting, which is a simulated and "micro" flipped class. It would be valuable to employ portable neuroimaging techniques such as EEG or fNIRS to explore the effects of FT in real-world classroom settings, as in previous studies[39,40]. Third, more types of knowledge (e.g., procedural knowledge) should be explored rather than declarative knowledge alone. Fourth, this study did not measure the number of questions asked by students across conditions. Future studies could collect such data to gain a more comprehensive understanding of the instructional process. Finally, given that this study is a preliminary attempt to explore the neural foundation of FT using fNIRS hyperscanning technology, it is encouraged that future studies increase the sample size to consolidate the current findings.

In conclusion, this study provides empirical evidence that the flipped classroom teaching model, characterized by pre-recorded lectures and structured interaction phases, improves learning outcomes significantly more than traditional lecture-based teaching or the flipped teaching model controlled with spontaneous interactions. This improvement is underpinned by increased teacher–student IBS in key brain regions, such as the left DLPFC, STG, and S1, which are associated with cognitive processing and teacher–student interaction dynamics. Such evidence highlights the crucial role of interaction in the learning process, which is promoted by the structured approach of flipped classrooms. The findings suggest that carefully planned interaction schedules can significantly benefit educational content delivery.

The current study, therefore, offers valuable insights for the optimization of flipped classroom methodologies, potentially guiding future educational strategies and policies to better meet the needs of both teachers and students in achieving optimal learning experiences. For example, in future teaching activities, teachers should design their interaction processes with students carefully and scientifically and strive to catch students' attention and engage them, thereby fostering synchronization in learning behaviors and neural activities between teachers and students.

Methods

Participants

A previous study using a comparable experimental design reported large effects[44]. We thus performed a power analysis for large effects using G* Power 3.1. The analysis identified the minimum number of participants as 60, with a power $(1 - \beta) = 0.80$, an error probability of $\alpha = 0.05$, and a large effect size of $f = 0.42$. Sixty-six university students (27 males, 39 females; $M = 20.96$, $SD = 2.23$) and 1 teacher (male, 23 years old, a postgraduate student with one year of teaching experience) were recruited through public announcements posted at our university to participate in this experiment. Note that the sample size is consistent with previous studies[20,22,44]. All participants were healthy and right-handed, had normal hearing and vision, and were naive to the purposes of the experiment. The students did not major in psychology. The teacher, professionally trained in psychology, was thoroughly familiar with the educational content. Each student was paired with the teacher, generating 66 teacher–student dyads. The members of each dyad were not familiar with each other. Note that the same teacher was involved in every dyad to maintain a consistent teaching style[23].

The students were randomly assigned to three conditions, each with twenty-two students. The first group (8 males, $M = 21.59$, $SD = 3.25$; 14 females, $M = 21.84$, $SD = 2.01$) attended the FT condition. The second group (10 males, $M = 19.91$, $SD = 2.03$; 12 females, $M = 21.66$, $SD = 1.87$) attended

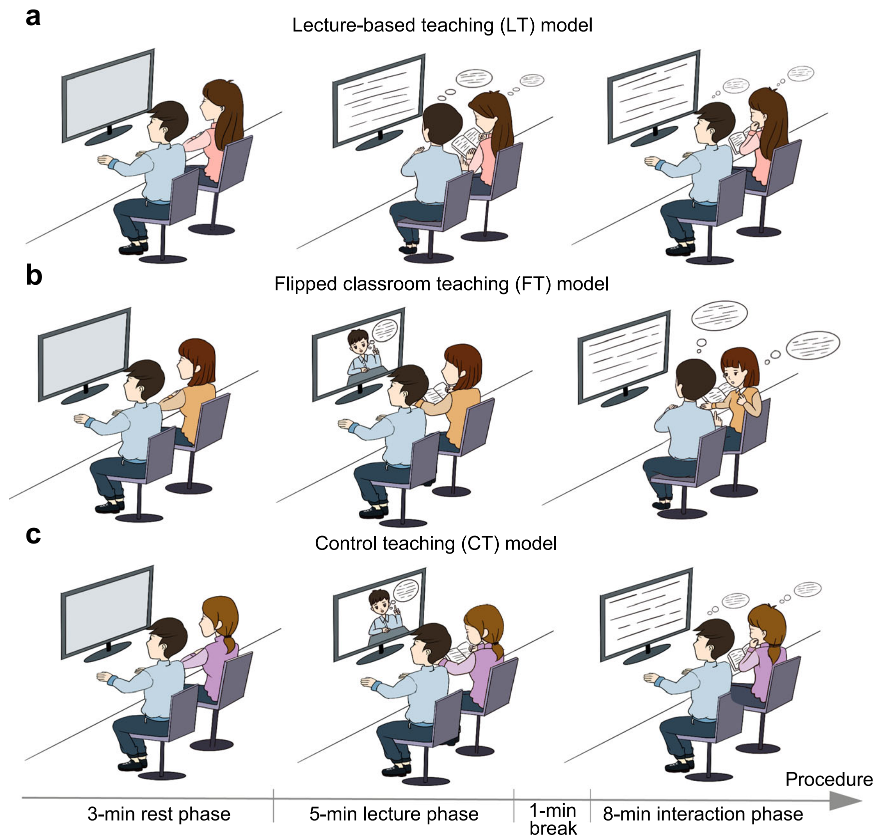

Fig. 5 | Experimental design. Students were randomly assigned to three conditions with different teaching models: **a** lecture-based teaching (LT), **b** flipped classroom teaching (FT), and **c** control teaching (CT). All models encompassed three phases: rest, lecture, and teacher–student interaction.

the LT condition. The third group (9 males, $M = 20.34$, $SD = 2.39$; 13 females, $M = 20.19$, $SD = 1.65$) attended the CT condition. Note that two female students, one each from the FT and LT conditions, were removed because their post-test scores were lower than the pre-test scores and the learning improvements (the difference between post- and pre-test scores) deviated from the mean by more than three standard deviations, leaving 21 students in both FT and LT conditions. The aim of the study was to improve students' learning outcomes through effective teaching. The decrease in the two participants' learning outcomes suggests that they did not take the experiment seriously. Therefore, we removed them from the participant pool.

We got the written consent form from all participants before the experiment. The study was in accordance with the Declaration of Helsinki and supported by the Ethics Committee of Zhejiang normal university (ZSRT2022020).

Experimental materials

Students were asked to study two theories selected from "Management psychology": the "achievement need theory"[45] and the "ERG theory"[46].

To ensure consistent teaching content and strategy across all dyads, the teacher underwent standardized training before attending the experiment. Specifically, a teaching script containing teaching contents and procedures was provided to the teacher. He was then asked to rehearse and prepare at home. Three days later, the teacher demonstrated his teaching to the experimenter in the lab. The main experiment started only after the teacher demonstrated consistency in the duration of instruction, speech rate, instructional content, and methodology[15,22]. Additionally, after training, we invited him to pre-record a 5-min lecture video that would be used in the FT and CT conditions in the lecture phase.

Two sets of parallel tests regarding the two theories were prepared for pre- and post-tests, respectively. They were two short case analysis questions with 15 points for each test. The tests were adapted from the textbook *Human Resource Management: Theory, Practice and Art*[47].

To assess the difficulty between the pre- and post-tests, 12 new participants (6 males, 6 females; $M = 19.45$, $SD = 0.98$) were recruited to complete the two tests. Two raters who were naive to the experimental purposes were invited to score the tests. The Pearson correlation coefficients between the two raters for the pre- and post-tests were 0.72 ($p = 0.0085$) and 0.69 ($p = 0.013$), respectively, indicating that inter-rater reliability was acceptable. The final scores of each student were the average of the two raters' scores. A paired t-test on the final score revealed no significant difference between the pre- and post-tests (t (11) = −0.95, $p = 0.36$, Cohens' $d = 0.28$), suggesting that the difficulties of the two tests were comparable. In addition, a Pearson correlation analysis showed that the correlation coefficient between the two tests was 0.70 ($p = 0.011$), indicating that the two tests were reliable.

Experimental tasks and procedures

At the beginning of the experiment, all students completed the pre-test within 15 min. Then, they rested for 3-min (left panels in Fig. 5), with no communication allowed. Students were asked to relax and remain still.

After the rest, a 5-min lecture about the above-mentioned theories was given in either the LT, FT or CT condition. Note that the CT condition was included to underscore the impact of structured interaction on improving learning outcomes. In the LT condition (middle panel in Fig. 5a), the teacher taught the theories. The students listened and took notes, during this phase the students were allowed to raise questions and the teacher would give feedback immediately. In contrast, in the FT and CT conditions (middle panel in Fig. 5b, c), the students learned the theories by watching the lecture video and taking notes. The lecture video could not be paused, fast-forward, fast-backward, and replayed.

After a 1-min break, an 8-min teacher–student interaction phase was followed. In the LT and CT conditions (spontaneous interaction, right panel in Fig. 5a, c), the students could ask questions freely, and the teacher would give feedback immediately. If the students did not have any questions, they were instructed to review their notes. In the FT condition (structured interaction, right panel in Fig. 5b), the students were asked to summarize the

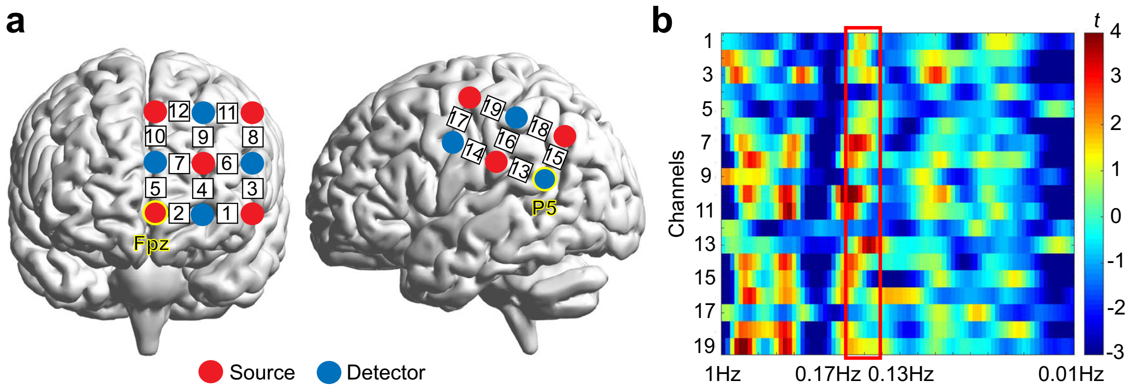

Fig. 6 | Optode probe patch and frequency band of interest. a The patches were placed over the left prefrontal and temporal-parietal junction regions. **b** *t*-value map. *x*-Axis denotes the frequency ranging from 0.01 to 1 Hz. *y*-Axis denotes the *t*-value resulting from the comparison between lecture and rest phases in each channel. The frequency band of interest, 0.13 to 0.17 Hz, is highlighted by a red rectangle.

teaching contents, and the teacher would give feedback. Next, the teacher encouraged the students to ask questions and provided answers. Following this, the teacher asked questions and the students responded, receiving feedback accordingly. Finally, the teacher summarized the knowledge concisely.

After the interaction phase, the students completed the post-test and reported their perceived quantity of teacher–student interaction. Additionally, both teachers and students were required to complete their subjective ratings of the likability and appearance attractiveness of the interacting partners. All ratings were on a 9-point Likert scale, with 1 denoting "not very much" and 9 denoting "very much". No communication was allowed during the assessment.

During the experiment, the brain activities of the teacher and student were recorded using an fNIRS device. Note that, for the lecture phase in the CT and FT conditions, we collected the teacher's brain activity while he recorded the lecture video before the main experiment began.

Behavioral data analyses

Each student's pre- and post-tests were scored by the two raters. Pearson correlation analysis showed that the two raters' scores were significantly correlated for the pre-test ($r = 0.65$, $p < 0.001$) and the post-test ($r = 0.79$, $p < 0.001$), indicating that the inter-rater reliabilities were acceptable.

For each student, the final score of each test, reflecting the learning outcome, was the average score of the two raters. To examine the effects of test phases and lecture types on learning, an ANOVA with test phases (pre vs. post) as the within-subjects factor and lecture types (FT, LT, vs. CT) as the between-subjects factor was conducted. If the FT can improve students' learning outcomes better than the LT and CT conditions, then we anticipated a significant interaction effect. In specific, we expected no significant difference in the pre-test score among the three lecture types, but that the post-test score for the FT condition should be significantly higher than that of the LT and CT conditions.

In addition, we conducted one-way ANOVAs to analyze likability, appearance attractiveness, and students' perceived quantity of teacher–student interaction among the three lecture types. We then calculated whether the perceived quantity of interaction could predict students' learning improvements—the difference between post- and pre-test scores—with a general linear model (GLM).

fNIRS data acquisition

The brain activity of both the teacher and student was measured by the ETG-4000 (Hitachi Medical Corporation, Japan) fNIRS system with a sampling rate of 10 Hz. This system contains eight near-infrared light emitters at two wavelengths (695 and 830 nm) and seven detectors, which formed 19 measurement channels. This layout covered the left prefrontal cortex (PFC) and temporal-parietal junction (TPJ). Both brain areas have been recognized as playing a key role in teacher–student interactions[15,20–22,34]. A 3×3 probe patch (five emitters and four detectors, resulting in 12 measurement channels with 30 mm optode separation) was used to cover each participant's left PFC and a 2×3 probe patch (three emitters and three detectors, resulting in 7 measurement channels with 30 mm optode separation) was placed over each participant's left TPJ. According to the international 10–20 system, the rightmost optode of the lowest probe row of the 3×3 patch was placed at Fpz, and the leftmost optode of the lowest probe row of the 2×3 patch was placed at P5 (Fig. 6a). We applied a virtual registration method to obtain correspondence between the measured channels and the measured points on the cerebral cortex[48,49].

Preprocessing of fNIRS data

The fNIRS data were preprocessed with the NIRS-KIT toolbox[50], based on MATLAB. In line with previous studies, the quality of fNIRS data was checked by visual inspection. All channels that did not show a clear heart band at around 1 Hz in the wavelet transform plot were identified as bad channels and were excluded from all further analysis (approximately 11% of the channels)[51–54]. Subsequently, the first and last 10 s of data in each phase were removed to ensure the measurement of stable brain activity. The remained data were first converted to changes in oxyhemoglobin (HbO) and deoxyhemoglobin (HbR), which were calculated according to the modified Beer-Lambert law. Since the HbO concentration is more sensitive to changes in cerebral blood flow than the HbR concentration[55], this study focused on the HbO concentration, which aligns with previous fNIRS-based hyperscanning studies[20,21,56,57]. Then, motion artifacts were corrected by using a "correlation-based signal improvement" (CBSI) approach[58]. Finally, a bandpass filtering procedure (0.01–1 Hz) was performed to reduce low-frequency drifts and high-frequency noise[21,59].

Interbrain synchrony (IBS) calculation

After preprocessing, we employed wavelet transform coherence (WTC) analysis to estimate the neural synchrony between analogous channels of each dyad[60]. In specific, we estimated the time-averaged (also averaged across all channels in each dyad) IBS across the frequencies from 0.01 to 1 Hz (frequency step = 0.058, after logarithmic transformation)[61,62] and then a series of paired sample *t*-tests were conducted to compare the IBS between the rest phase and the whole teaching process (i.e., lecture phase plus interaction phase). The resulting *p*-values were corrected using false discovery rate (FDR) method[63]. The IBS was significantly increased during the entire teaching process compared to the rest phase within the frequencies between 0.13 Hz and 0.17 Hz (i.e., period 5.89–7.42 s). Consequently, as depicted in Fig. 6b, this frequency range was selected as the frequency of interest (FOI) for subsequent analyses, and the IBS values within this FOI were averaged.

We quantified enhancements in IBS throughout the class (class-related IBS), as well as during the lecture (lecture-related IBS) and interaction phases (interaction-related IBS), relative to the baseline established during the rest phase. Consequently, we derived three distinct categories of task-related IBS, which we then transformed into Fisher *z*-scores for further analysis[64]. Subsequently, we conducted one-sample *t*-tests against zero for each channel to identify the channels showing significant class-related IBS with FDR correction ($p < 0.05$). For all significant channels, one-way ANOVAs were performed to assess differences in class-related, lecture-related, and interaction-related IBS across three types of teaching models. The statistical results were adjusted using FDR correction. For results with a significant main effect of group, post hoc multiple comparisons with Bonferroni correction were performed.

Relationship between IBS and behavioral indices

We first calculated each student's learning improvements (post-test score minus pre-test score). Next, we used GLM to predict learning improvements from IBS measures and analyzed whether the perceived quantity of teacher–student interaction predicted IBS. We then assessed the significance of the regression coefficient (β) to confirm the predictive power of IBS and the perceived quantity of interaction.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability

The underlying codes for this study are available on request by contacting the corresponding author.

Received: 10 April 2024; Accepted: 26 November 2024;
Published online: 03 December 2024

References

1. Charlton, B. G. Lectures are such an effective teaching method because they exploit evolved human psychology to improve learning. *Med. Hypotheses* **67**, 1261–1265 (2006).
2. King, A. From sage on the stage to guide on the side. *Coll. Teach.* **41**, 30–35 (1993).
3. Strelan, P., Osborn, A. & Palmer, E. The flipped classroom: a meta-analysis of effects on student performance across disciplines and education levels. *Educ. Res. Rev.* **30**, 100314 (2020).
4. Cheng, L., Ritzhaupt, A. D. & Antonenko, P. Effects of the flipped classroom instructional strategy on students' learning outcomes: a meta-analysis. *Educ. Technol. Res. Dev.* **67**, 793–824 (2019).
5. Cho, H. J., Zhao, K., Lee, C. R., Runshe, D. & Krousgrill, C. Active learning through flipped classroom in mechanical engineering: Improving students' perception of learning and performance. *Int. J. STEM Educ.* **8**, 1–13 (2021).
6. Lin, G. Y., Wang, Y. S. & Lee, Y. N. Investigating factors affecting learning satisfaction and perceived learning in flipped classrooms: The mediating effect of interaction. *Interact. Learn. Environ.* **31**, 5759–5780 (2023).
7. Güler, M., Kokoç, M. & Önder Bütüner, S. Does a flipped classroom model work in mathematics education? A meta-analysis. *Educ. Inf. Technol.* **28**, 57–79 (2023).
8. Fadol, Y., Aldamen, H. & Saadullah, S. A comparative analysis of flipped, online and traditional teaching: a case of female middle Eastern management students. *Int. J. Manag. Educ.* **16**, 266–280 (2018).
9. Martínez-Jiménez, R. & Ruiz-Jiménez, M. C. Improving students' satisfaction and learning performance using flipped classroom. *Int. J. Manag. Educ.* **18**, 100422 (2020).
10. Akçayır, G. & Akçayır, M. The flipped classroom: a review of its advantages and challenges. *Comput. Educ.* **126**, 334–345 (2018).
11. Vitta, J. P. & Al-Hoorie, A. H. The flipped classroom in second language learning: a meta-analysis. *Lang. Teach. Res.* **27**, 1268–1292 (2023).
12. Zhang, J., Wang, Y. & Zhang, B. Introducing a new teaching model: flipped classroom. *J. Distance Educ.* **30**, 46–51 (2012).
13. Chi, M. T. & Wylie, R. The ICAP framework: linking cognitive engagement to active learning outcomes. *Educ. Psychol.* **49**, 219–243 (2014).
14. Chiu, Y. C., Jen, C. H., Chang, C. Y., Lee, P. L. & Yeh, T. K. Learning seismology through inquiry: Structured, guided, or both? *Seismol. Res. Lett.* **87**, 998–1007 (2016).
15. Pan, Y. et al. Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *Neuroimage* **211**, 116657 (2020).
16. Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S. & Keysers, C. Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn. Sci.* **16**, 114–121 (2012).
17. Schilbach, L. et al. Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414 (2013).
18. Czeszumski, A. et al. Hyperscanning: a valid method to study neural inter-brain underpinnings of social interaction. *Front. Hum. Neurosci.* **14**, 39 (2020).
19. Zhang, L., Xu, X., Li, Z., Chen, L. & Feng, L. Interpersonal neural synchronization predicting learning outcomes from teaching-learning interaction: a meta-analysis. *Front. Psychol.* **13**, 835147 (2022).
20. Zheng, L. et al. Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Hum. Brain Mapp.* **39**, 3046–3057 (2018).
21. Zhu, Y. et al. Instructor–learner neural synchronization during elaborated feedback predicts learning transfer. *J. Educ. Psychol.* **114**, 1427–1441 (2022).
22. Liu, J. et al. Interplay between prior knowledge and communication mode on teaching effectiveness: interpersonal neural synchronization as a neural marker. *Neuroimage* **193**, 93–102 (2019).
23. Pan, Y., Novembre, G., Song, B., Li, X. & Hu, Y. Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *Neuroimage* **183**, 280–290 (2018).
24. Bjorklund, S. A., Parente, J. M. & Sathianathan, D. Effects of faculty interaction and feedback on gains in student skills. *J. Eng. Educ.* **93**, 153–160 (2004).
25. Lin, C. H. et al. Interleaved practice enhances skill learning and the functional connectivity of fronto-parietal networks. *Hum. Brain Mapp.* **34**, 1542–1558 (2013).
26. Lu, K., Qiao, X., Yun, Q. & Hao, N. Educational diversity and group creativity: evidence from fNIRS hyperscanning. *Neuroimage* **243**, 118564 (2021).
27. Pan, Y., Cheng, X. & Hu, Y. Three heads are better than one: cooperative learning brains wire together when a consensus is reached. *Cereb. Cortex* **33**, 1155–1169 (2023).
28. Baker, J. M. et al. Sex differences in neural and behavioral signatures of cooperation revealed by fNIRS hyperscanning. *Sci. Rep.* **6**, 26492 (2016).
29. Correia, J. et al. Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe. *J. Neurosci.* **34**, 332–338 (2014).
30. Friederici, A. D., Rüschemeyer, S. A., Hahne, A. & Fiebach, C. J. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb. Cortex* **13**, 170–177 (2003).
31. Tamè, L. et al. The contribution of primary and secondary somatosensory cortices to the representation of body parts and body sides: an fMRI adaptation study. *J. Cogn. Neurosci.* **24**, 2306–2320 (2012).
32. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).

33. Dixon, M. L. et al. Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proc. Natl Acad. Sci. USA* **115**, E1598–E1607 (2018).
34. Takeuchi, N., Mori, T., Suzukamo, Y. & Izumi, S. I. Integration of teaching processes and learning assessment in the prefrontal cortex during a video game teaching-learning task. *Front. Psychol.* **7**, 2052 (2017).
35. Raposo, A., Vicens, L., Clithero, J. A., Dobbins, I. G. & Huettel, S. A. Contributions of frontopolar cortex to judgments about self, others and relations. *Soc. Cogn. Affect. Neurosci.* **6**, 260–269 (2011).
36. Grossmann, T. & Johnson, M. H. Selective prefrontal cortex responses to joint attention in early infancy. *Biol. Lett.* **6**, 540–543 (2010).
37. Anticevic, A. et al. The role of default network deactivation in cognition and disease. *Trends Cogn. Sci.* **16**, 584–592 (2012).
38. Mansouri, F. A., Buckley, M. J. & Tanaka, K. Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science* **318**, 987–990 (2007).
39. Bevilacqua, D. et al. Brain-to-Brain synchrony and learning outcomes vary by student-teacher dynamics: evidence from a real-world classroom electroencephalography study. *J. Cogn. Neurosci.* **31**, 401–411 (2019).
40. Dikker, S. et al. Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Curr. Biol.* **27**, 1375–1380 (2017).
41. Pi, Z., Chai, H., Li, L., Zhang, X. & Li, X. High task motivation learners co-viewing video lectures facilitates learning. *J. Comput. Assist. Learn.* **40**, 588–599 (2024).
42. Raj Prabhu, N., Raman, C. & Hung, H. Defining and quantifying conversation quality in spontaneous interactions. *ICMI 2020 Companion—Companion Publ. 2020 Int. Conf. Multimodal Interact*. 196–205 (2020).
43. Novembre, G. & Iannetti, G. D. Hyperscanning alone cannot prove causality. Multibrain stimulation can. *Trends Cogn. Sci.* **25**, 96–99 (2021).
44. Zheng, L. et al. Affiliative bonding between teachers and students through interpersonal synchronisation in brain activity. *Soc. Cogn. Affect. Neurosci.* **15**, 97–109 (2020).
45. McClelland, D. C. *Human Motivation* (Cambridge University Press, 1987).
46. Alderfer, C. *Existence, Relatedness, & Growth* (New York: Free Press, 1972).
47. Li, C. *Human Resource Management: Theory, Practice and Art* (Peking University Press, Bei Jing, China Agricultural University Press, 2017).
48. Singh, A. K., Okamoto, M., Dan, H., Jurcak, V. & Dan, I. Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *Neuroimage* **27**, 842–851 (2005).
49. Tsuzuki, D. et al. Virtual spatial registration of stand-alone fNIRS data to MNI space. *Neuroimage* **34**, 1506–1518 (2007).
50. Hou, X. et al. NIRS-KIT: a MATLAB toolbox for both resting-state and task fNIRS data analysis. *Neurophotonics* **8**, 010802 (2021).
51. Lu, K., Yu, T. & Hao, N. Creating while taking turns, the choice to unlocking group creative potential. *Neuroimage* **219**, 117025 (2020).
52. Shao, C., Zhang, X., Wu, Y., Zhang, W. & Sun, B. Increased interpersonal brain synchronization in romantic couples is associated with higher honesty: an fNIRS hyperscanning study. *Brain Sci.* **13**, 833 (2023).
53. Zhao, H. et al. Inter-brain neural mechanism underlying turn-based interaction under acute stress in women: A hyperscanning study using functional near-infrared spectroscopy. *Soc. Cogn. Affect. Neurosci.* **17**, 850–863 (2022).
54. Nguyen, T. et al. The effects of interaction quality on neural synchrony during mother-child problem solving. *Cortex* **124**, 235–249 (2020).
55. Lindenberger, U., Li, S. C., Gruber, W. & Müller, V. Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* **10**, 1–12 (2009).
56. Sun, B. et al. Behavioral and brain synchronization differences between expert and novice teachers when collaborating with students. *Brain Cogn.* **139**, 105513 (2020).
57. Sun, B. et al. Cooperation with partners of differing social experience: an fNIRS-based hyperscanning study. *Brain Cogn.* **154**, 105803 (2021).
58. Cui, X., Bray, S. & Reiss, A. L. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage* **49**, 3039–3046 (2010).
59. Liu, J. et al. Shared intentionality modulates interpersonal neural synchronization at the establishment of communication system. *Commun. Biol.* **6**, 832 (2023).
60. Grinsted, A., Moore, J. C. & Jevrejeva, S. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* **11**, 561–566 (2004).
61. Nozawa, T., Sasaki, Y., Sakaki, K., Yokoyama, R. & Kawashima, R. Interpersonal frontopolar neural synchronization in group communication: an exploration toward fNIRS hyperscanning of natural interactions. *Neuroimage* **133**, 484–497 (2016).
62. Xue, H., Lu, K. & Hao, N. Cooperation makes two less-creative individuals turn into a highly-creative pair. *Neuroimage* **172**, 527–537 (2018).
63. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
64. Cui, X., Bryant, D. M. & Reiss, A. L. NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* **59**, 2430–2437 (2012).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 32200842) to Q.S. and (No. 62007002) to Z.H. The funders were not involved in the study design, data collection and analysis, publication decisions, or manuscript preparation. We thank Chong Shao for his assistance in data analysis and Xuecheng Zhang and Xiumei Gong for their help in data collection.

Author contributions

Conceptualization: Q.L., Q.S., D.W., and Z.H. Methodology: Q.L., D.W., W.X., and Q.S. Investigation: Q.L. and D.W. Data curation: Q.L. Formal analysis: Q.L. and W.X. Visualization: Q.S. and Q.L. Validation: Q.S., Q.L., and Z.H. Supervision: Q.S. Project administration: Q.S. Funding acquisition: Q.S., and Z.H. Writing—original draft: Q.L. and Q.S. Writing—review & editing: Q.S., Z.H., D.W., B.S., and Y.T. All authors have read and approved the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41539-024-00286-y.

Correspondence and requests for materials should be addressed to Qi Sun, Binghai Sun or Zhishan Hu.