Report

Current Biology

Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom

Highlights

- We report a real-world group EEG study, in a school, during normal class activities
- EEG was recorded from 12 students simultaneously, repeated over 11 sessions
- Students' brain-to-brain group synchrony predicts classroom engagement
- Students' brain-to-brain group synchrony predicts classroom social dynamics

Authors

Suzanne Dikker, Lu Wan, Ido Davidesco, ..., Jay J. Van Bavel, Mingzhou Ding, David Poeppel

Correspondence

sdikker@gmail.com (S.D.),
david.poeppel@gmail.com (D.P.)

In Brief

Dikker, Wan, et al. follow a group of high school seniors for a semester and record their brain activity during their regular biology class. They find that students' brainwaves are more in sync with each other when they are more engaged during class. Brain-to-brain synchrony is also reflective of how much students like the teacher and each other.

Dikker et al., 2017, Current Biology *27*, 1375–1380

http://dx.doi.org/10.1016/j.cub.2017.04.002

Current Biology
Report

Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom

Suzanne Dikker,[1,2,7,8,*] Lu Wan,[3,7] Ido Davidesco,[1] Lisa Kaggen,[1] Matthias Oostrik,[5] James McClintock,[6] Jess Rowland,[1] Georgios Michalareas,[4] Jay J. Van Bavel,[1] Mingzhou Ding,[3] and David Poeppel[1,4,*]
[1]Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA
[2]Department of Language and Communication, Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, the Netherlands
[3]J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, 1275 Center Drive, Gainesville, FL 32611, USA
[4]Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany
[5]Leidekkerssteeg 1, 1012 GH Amsterdam, the Netherlands
[6]New York, NY 10024, USA
[7]These authors contributed equally
[8]Lead Contact
*Correspondence: sdikker@gmail.com (S.D.), david.poeppel@gmail.com (D.P.)
http://dx.doi.org/10.1016/j.cub.2017.04.002

SUMMARY

The human brain has evolved for group living [1]. Yet we know so little about how it supports dynamic group interactions that the study of real-world social exchanges has been dubbed the "dark matter of social neuroscience" [2]. Recently, various studies have begun to approach this question by comparing brain responses of multiple individuals during a variety of (semi-naturalistic) tasks [3–15]. These experiments reveal how stimulus properties [13], individual differences [14], and contextual factors [15] may underpin similarities and differences in neural activity across people. However, most studies to date suffer from various limitations: they often lack direct face-to-face interaction between participants, are typically limited to dyads, do not investigate social dynamics across time, and, crucially, they rarely study social behavior under naturalistic circumstances. Here we extend such experimentation drastically, beyond dyads and beyond laboratory walls, to identify neural markers of group engagement during dynamic real-world group interactions. We used portable electroencephalogram (EEG) to simultaneously record brain activity from a class of 12 high school students over the course of a semester (11 classes) during regular classroom activities (Figures 1A–1C; Supplemental Experimental Procedures, section S1). A novel analysis technique to assess group-based neural coherence demonstrates that the extent to which brain activity is synchronized across students predicts both student class engagement and social dynamics. This suggests that brain-to-brain synchrony is a possible neural marker for dynamic social interactions, likely driven by shared attention mechanisms. This study validates a promising new method to investigate the neuroscience of group interactions in ecologically natural settings.

RESULTS AND DISCUSSION

The classroom is an ideal starting point for real-world neuroscience: it provides a practically important and ecologically naturalistic context but also a semi-controlled environment, governed by a sequence of activities led by a teacher. This allowed us to measure brain activity and behavior in a systematic fashion over the course of a full semester as students engaged in a series of predetermined class activities (repeated across 11 50-min classes, students followed lectures, watched instructional videos, and participated in group discussions). We explored the hypothesis that synchronized neural activity across a group of students predicts (and possibly underpins) classroom engagement and social dynamics. When students feel connected or engaged with the material or each other, are their brains in fact "in sync" in a formal, quantifiable sense? To investigate these questions, we used low-cost portable electroencephalogram (EEG) systems ([16]; Supplemental Experimental Procedures, section S2) paired with a novel analysis technique to characterize the synchronization of brain activity between individuals: total interdependence (TI; [17]; Supplemental Experimental Procedures, section S3). Figures 1C and 1D lay out how TI is operationalized.

We focused on the relationship between TI and classroom engagement, on the one hand, and social dynamics, on the other—both of which are critical for student learning [18]. Classroom engagement was quantified as student appreciation ratings of different teaching styles (Figure 1B) and student day-by-day self-reported focus. Classroom social dynamics were quantified in terms of socially relevant personality traits (group affinity [19, 20] and empathy [21]) and as social closeness during class interactions (between students and with the teacher; see Supplemental Experimental Procedures, section S1 for details).

Brain-to-Brain Synchrony and Class Engagement

We first examined the relationship between brain-to-brain synchrony (indexed by TI) and student ratings of four different

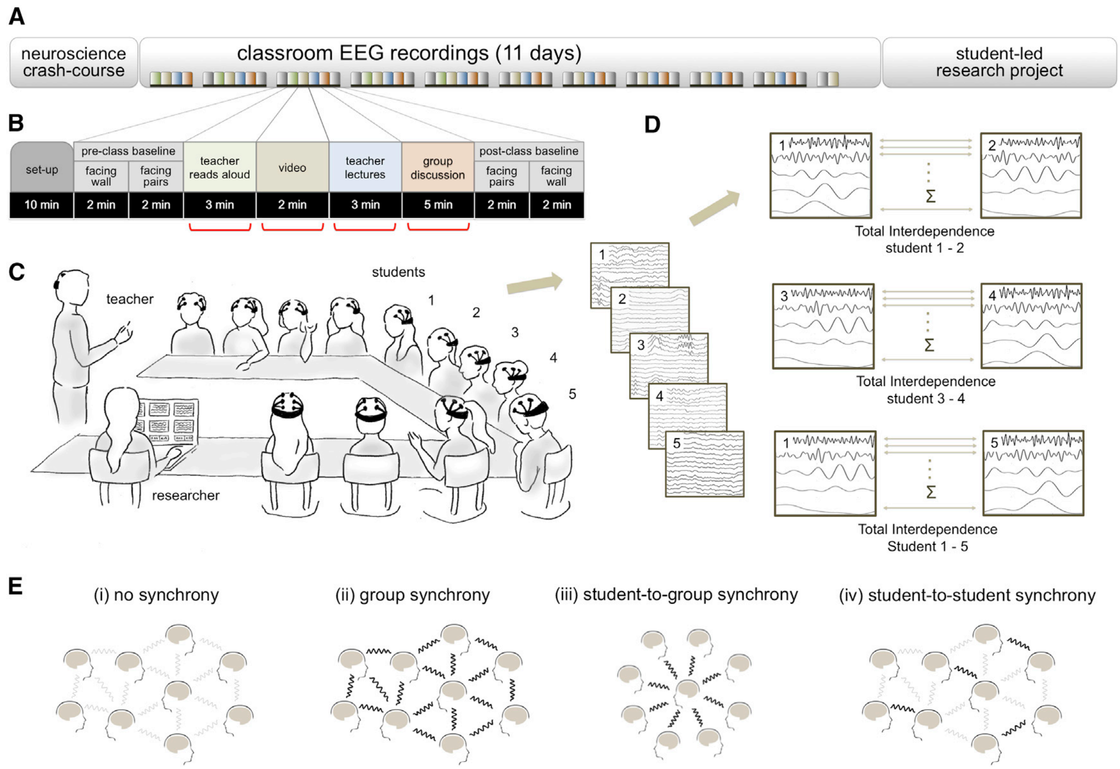

Figure 1. Experimental Setup, Procedure, and Rationale

(A) Timeline of the experiment. The fall semester started with a crash course in neuroscience, followed by 11 recording days distributed over a 3-month period. In the spring semester, students designed, executed, and carried out their own original research projects (see Supplemental Experimental Procedures, section S1).

(B) Sample experimental procedure of a typical recording day: EEG activity was recorded during video, lecture, and discussion teaching styles separately, which were consistently carried out across all 11 recording days. Other tasks were alternated (Supplemental Experimental Procedures, section S1). TI values were averaged for each teaching style separately (marked in red; Supplemental Experimental Procedures, section S3).

(C) Illustration of experimental setup in the classroom with 12 students wearing the EMOTIV EPOC headset (Supplemental Experimental Procedures, section S2). These portable devices offer a rich opportunity to involve students both as participants and as experimenters (Supplemental Experimental Procedures, section S1).

(D) Brain-to-brain synchrony (TI) was computed by taking each student's raw EEG signal, decomposing it into frequency bins (1–20 Hz, 0.25 Hz resolution), and calculating the sum of the inter-brain coherence between pairs of students for each bin. Thus, TI quantifies the inter-brain coherence across the frequency spectrum, allowing a data-driven identification of the brain signals of interest (see Figure S3 for further details).

(E) TI enables us to analyze brain-to-brain synchrony at multiple socially relevant levels of investigation: group synchrony (averaging TI values across all possible pairs within a group) (i); student-to-group synchrony (averaging TI values between a given student and each of his/her peers) (ii); and student-to-student synchrony (TI values between pairs of students) (iii).

See also Figure S1.

teaching styles over time. Students rated each segment after every recording and were also asked to provide overall ratings of each teaching style after the semester was over (Figures 1A and 1B; Supplemental Experimental Procedures, section S1). Significant main effects of teaching style were observed on both student ratings (repeated-measures two-way ANOVA with teaching style and time as main factors [see Supplemental Experimental Procedures, section S4 for details]: day-by-day ratings: $F(3,24) = 16.85$; $p < 10^{-5}$; post-semester ratings: $F(3,27) = 33.29$; $p < 10^{-8}$) and brain-to-brain synchrony (group synchrony: $F(3,12) = 5.93$; $p < 0.0005$; student-to-group synchrony: $F(3,27) = 5.94$; $p < 0.005$; see Supplemental Experimental Procedures, section S2). Overall, students preferred watching videos and engaging in group discussions over listening to the teacher reading aloud or lecturing (Figure 2A, left panel), an effect that was even more pronounced in the post-semester ratings (Figure 2A, right panel). A strikingly similar pattern was observed for group synchrony (Figure 2B, left) as well as student-to-group synchrony (Figure 2B, right; see Table S2 for detailed statistics). Student-to-group synchrony exhibited a strong positive correlation with student ratings: the higher the post-semester student ratings, the stronger the student-to-group synchrony averaged across days ($r = .61$, $p < 0.0001$; Figure 2C; Figure 2A, right shows the same data, separated by condition and averaged across subjects). Day-by-day ratings and group synchrony were not correlated.

Is Brain-to-Brain Synchrony Purely Stimulus Driven?

How much of brain-to-brain synchrony is explained by "mere" stimulus attributes (i.e., teaching style; cf. [6]), and how much do individual differences (cf. [7]) contribute to synchrony? To explore this, we performed a number of multiple regression

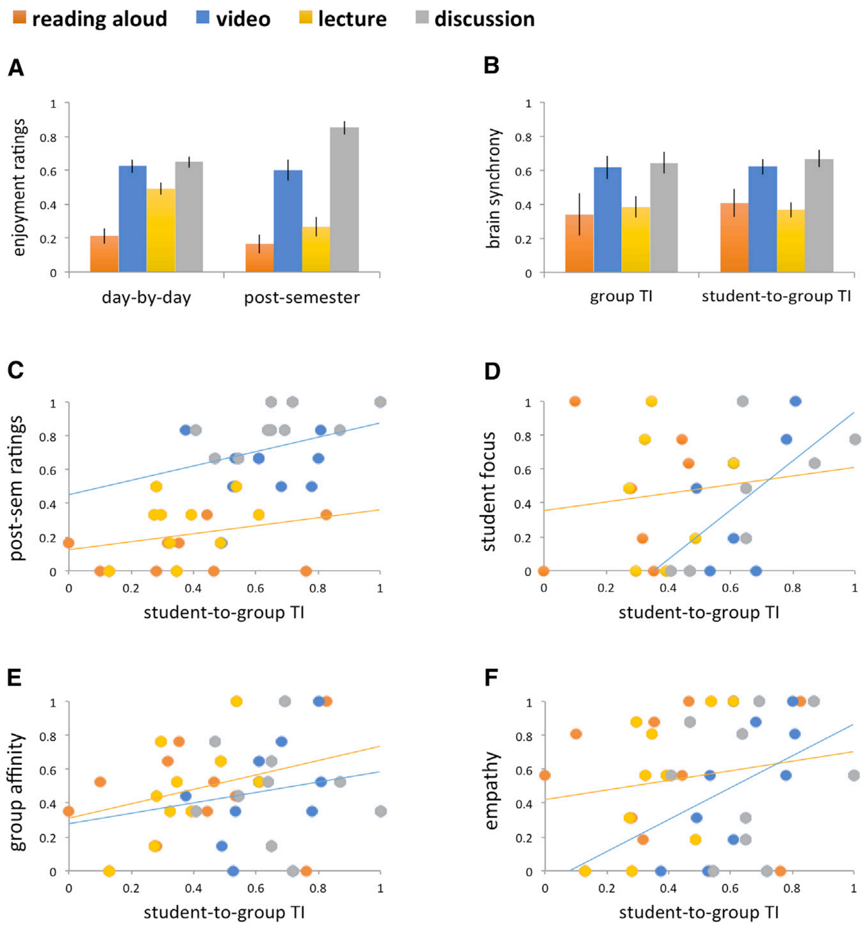

Figure 2. Independent Contributions of Teaching Style and Individual Differences to Brain-to-Brain Synchrony

(A) Average day-by-day (left) and post-semester (right) student appreciation ratings for four teaching styles: reading aloud, video, lecture, and discussion sessions. Error bars reflect standard errors over students.

(B) Average group TI (left) and student-to-group TI (right) for four teaching styles. Error bars reflect standard errors over days (left) and students (right).

(C) Post-semester ratings, while exhibiting a main effect on student-to-group synchrony, did not independently predict student-to-group TI over teaching style.

(D–F) Student focus (D), group affinity (E), and empathy (F) did each predict student-to-group TI in addition to teaching style.

Trend lines are displayed by teaching style (blue: discussion and video; yellow: reading aloud and lecture). All values were normalized to a 0–1 scale (max-min) for presentation purposes, and each dot reflects one student's TI in one of four teaching styles averaged across days (see Figure S4 for data further separated by days).

See also Figures S2, S3, and S4 and Tables S1 and S2.

analyses to assess the relationship between TI and a number of individual variables (ratings, focus, group affinity, and empathic disposition), with teaching style included as a factor representing the stimulus attribute (see Supplemental Experimental Procedures, section S3).

Post-semester ratings, while exhibiting a main effect on student-to-group synchrony ($F(1,220) = 20.79$, $p < 0.0001$), did not independently predict synchrony over teaching style (post-semester ratings: $F(1,210) = 2.28$, $p = 0.1327$ and teaching style: $F(1,9) = 2.37$, $p = 0.1581$; Figure 2C). Student focus, in contrast, did predict student-to-group synchrony independent of teaching style: students who were more focused on a given day also showed higher synchrony for that day (focus: $F(1,126) = 4.64$, $p = 0.0331$ and teaching style: $F(1,9) = 29.23$, $p = 0.0004$; Figure 2D).

Next, we examined the relationship between brain-to-brain synchrony and students' personality traits, in particular their group affinity and empathic disposition ([20]; see Supplemental Experimental Procedures, section S1 for details). Both group affinity and empathy predicted student-to-group synchrony independently of teaching style (group affinity: $F(1,115) = 5.95$, $p = 0.0163$ and teaching style: $F(1,9) = 12.73$, $p = 0.0060$; empathy: $F(1,115) = 5.71$, $p = 0.0185$ and teaching style: $F(1,9) = 13.53$, $p = 0.0062$).

Together, these findings demonstrate that individual factors (focus and personality traits) contribute to synchrony above and beyond the nature of the stimulus itself.

Brain-to-Brain Synchrony and Classroom Social Dynamics

Our findings suggest that brain-to-brain synchrony is driven by a combination of stimulus properties (teaching styles) and individual differences (student focus, teaching style preferences, teacher likeability, and personality traits). However, none of these factors speak directly to whether the presence of others had an effect on synchrony during class. For example, empathic disposition affects brain-to-brain similarities even in the absence of others [14].

To address classroom social dynamics directly, we collected social closeness ratings from students both toward the teacher and to the other students (Supplemental Experimental Procedures, section S1) and introduced manipulations that either did or did not involve direct social interaction. To investigate the effect of the teacher on student-to-group synchrony, we compared the two teaching styles in which the teacher was minimally involved (videos) and maximally involved (lectures). Figure 2D illustrates that, while students varied with respect to their overall student-to-group synchrony, synchrony was consistently higher for video than lecture sessions across students ($p = 0.007$; see Table S1). This difference was correlated with students' evaluations of the teacher: the more favorable a student's rating of the teacher, the smaller that student's difference in synchrony between video (where the teacher played no role) and lecture sessions (where the teacher played an integral role; Figure 2E; $r = 0.72$, $p = 0.018$ for data averaged across days).

We then tested whether pairwise student-to-student synchrony varied as a function of the classroom configuration (in each class, students were randomly assigned seats by the experimenters; see Supplemental Experimental Procedures,

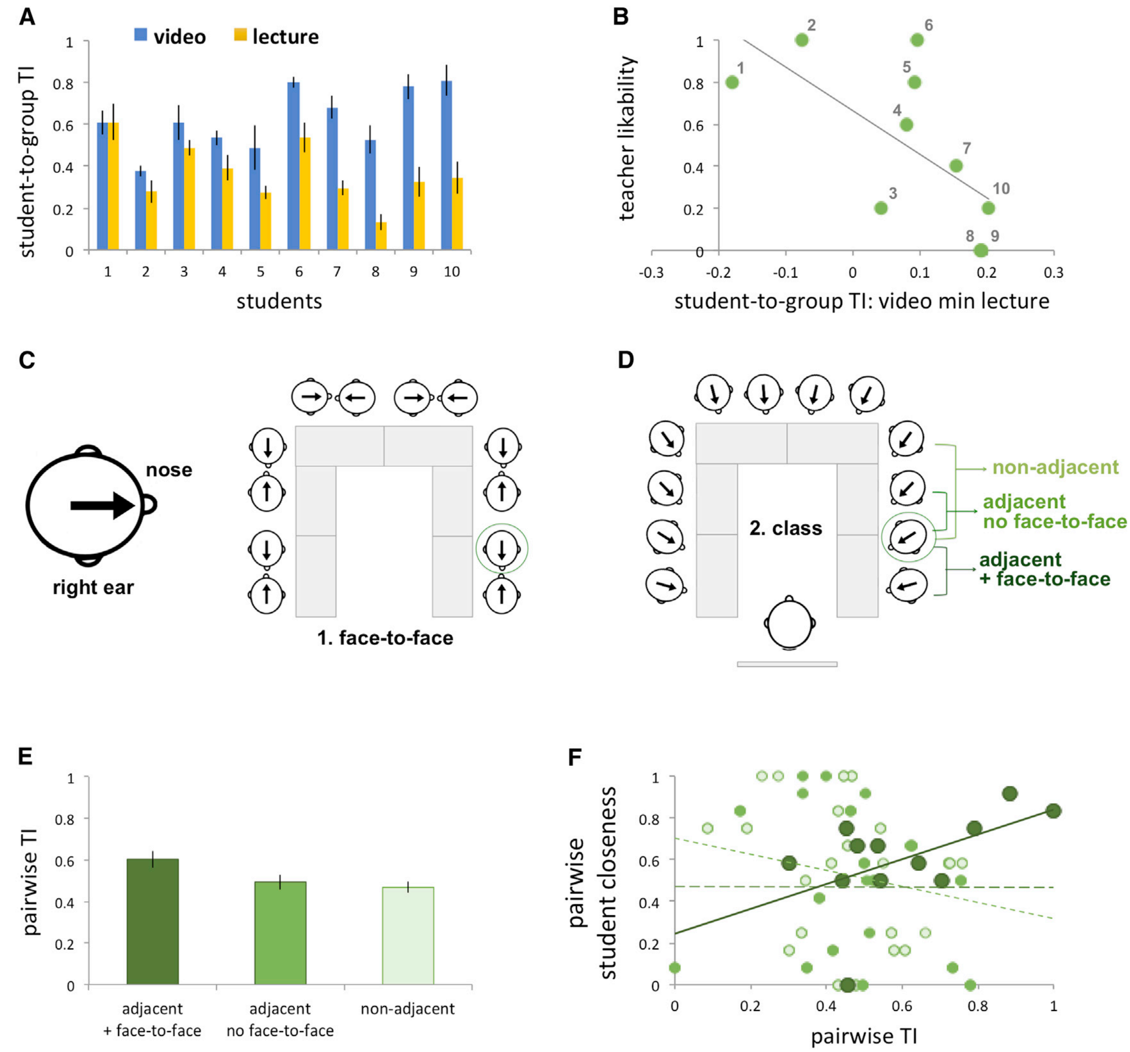

Figure 3. Brain-to-Brain Synchrony Predicts Classroom Social Dynamics

(A and B) The difference in student-to-group TI between video and lecture sessions across students (A) (error bars reflect standard errors over days) was negatively correlated with their ratings of the teacher (B) ($r = -.72$, $p = 0.018$; each dot represents one student; TI values are averaged across days; teacher likeability was recorded once for each student, after the semester was over).

(C) Before class, students sat face-to-face, engaging in eye contact for 2 min with one peer (Supplemental Experimental Procedures, section S1).

(D) An illustration for one student (green circle) of how the face-to-face baseline allowed a comparison of pairwise TI for three types of students: students who sat adjacent to each other and had engaged in silent eye contact prior to class (adjacent + face-to-face), students who sat next to each other but had not participated in a face-to-face baseline together (adjacent, no face-to-face), and students who were not sitting next to each other (non-adjacent).

(E) Students showed the highest pairwise synchrony during class with their face-to-face partner compared to the other two student pairings (error bars reflect standard errors over student pairs).

(F) Pairwise TI is correlated with mutual closeness ratings for adjacent + face-to-face pairs (solid dark green), but not for adjacent, no face-to-face pairs (solid light green) or non-adjacent pairs (no fill green). Each dot represents one student pair, averaged across teaching styles. All values were normalized to a 0–1 scale (max-min) for presentation purposes.

See also Figures S2, S3, and S4 and Table S1.

section S1) and student interaction: as illustrated in Figures 1B and 3C, students engaged in eye contact (face-to-face) with an assigned peer for 2 min prior to class (see Supplemental Experimental Procedures, section S1 for details). This allowed us to compare the relationship between pairwise synchrony and students' self-reported closeness to each other for three types of student pairs: students who sat adjacent to each other and had engaged in silent eye contact prior to class (adjacent + face-to-face), students who sat next to each other but had not participated in a face-to-face baseline together (adjacent, no face-to-face), and students who were not sitting next to each other (non-adjacent; illustrated in Figure 3D). Students showed the highest pairwise synchrony during class with their face-to-face partner compared to the other two student pairings (Figure 3E; one-way ANOVA: $F(2,102) = 5.66$, $p = 0.0047$). In addition, brain-to-brain synchrony was correlated with students'

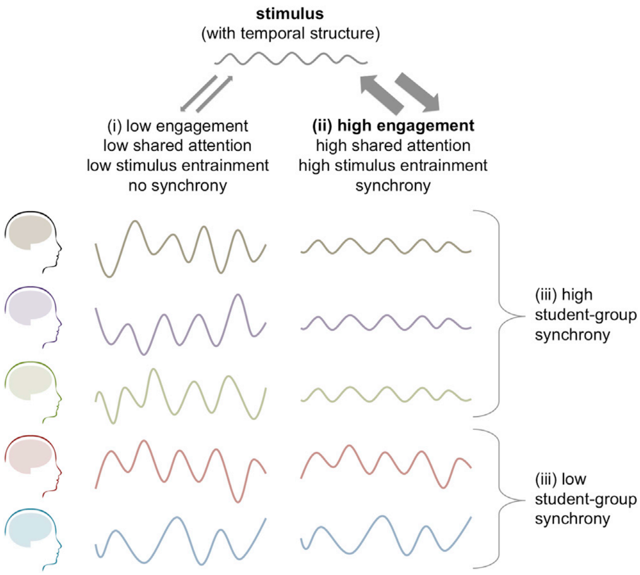

Figure 4. Shared Attention as a Possible Account of Brain-to-Brain Synchrony

Schematic illustration of a possible joint attention account of brain-to-brain synchrony. Neural entrainment to an external stimulus (video, teacher, or each other) is driven by a combination of stimulus properties (shown as arrows flowing down from "stimulus") and attention (arrows flowing up to the stimulus). Under "low attention" conditions, students' neural oscillations are not entrained to an external stimulus (video, teacher, or each other) (i). Under "shared attention" conditions, students' alpha oscillations are attenuated and entrained with an engaging external stimulus: a video, the teacher, or each other (ii). Some students are in a more attentive state, have more socially engaged personality traits, or have directly interacted, modulating the extent to which their neural oscillations are entrained with the stimulus (the teacher, a video, or each other) (iii).

mutual closeness ratings, but exclusively for adjacent + face-to-face pairs: student pairs who reported higher social closeness to each other exhibited stronger pairwise brain-to-brain synchrony during class activities, only if they had engaged in eye contact prior to class ($r = 0.5265$, $p = 0.0082$; solid green dots and solid line in Figure 3F; note that there was only a marginal main effect of condition on the TI × closeness correlation: $F(2,75) = 2.83$, $p = 0.0654$). In sum, face-to-face interaction prior to class not only increased brain-to-brain synchrony during class but also seemed to serve as an "activator" for interpersonal relationship features: actual joint attention, and not passive co-presence, predicted student-to-student synchrony.

Shared Attention as a Likely Source of Brain-to-Brain Synchrony

It is important to emphasize that brain-to-brain synchrony is not a mechanism in itself. Instead, neural synchrony across participants is a measurable reflection of the underlying neural computations that underpin some of the psychological processes under investigation. To better understand the synchronization effects we observe, mental constructs like focus, empathy, and closeness need to be decomposed into basic psychological processes that provide more suitable linking hypotheses to neural metrics. As already briefly discussed above, the finding that student-to-student synchrony is correlated with mutual closeness ratings during class—but only for pairs of students who had engaged in eye contact prior to class—aligns with research suggesting that eye contact sets up a context for joint attention [22]. Joint attention (shared intentionality) has been proposed to form a scaffold for social cognition in a range of social-psychological contexts, including development [21, 23], and provides a plausible account for prior findings showing an increase in brain-to-brain synchrony during laboratory tasks that required dyads to coordinate visual attention (e.g., [3, 5, 8, 11]).

We speculate that stimulus properties (teaching style [13]), individual differences (focus, engagement, and personality traits [14]), and social dynamics (social closeness and social interaction) each mediate attention at the neural level. This, in turn, affects students' neural entrainment to their surrounding sensory input: the teacher, a video, or each other [24]. This ties directly to behavioral evidence showing that people physically (and typically subconsciously) entrain to each other when engaging in tasks that require joint attention (pupil dilation, gestures, walking; e.g., [25]). More broadly, student-to-group synchrony as a function of shared attention follows directly from a range of electrophysiological results showing that brain rhythms lock to the rhythms of auditory and audiovisual input, which is amplified when the input is attended [24, 26, 27].

To provide additional evidence that speaks to a shared attention account, we examined the relationship between student-to-group synchrony and alpha band power—a well-characterized index of attention [28, 29]. As predicted, a reduction in a student's alpha oscillatory activity was accompanied by an increase in student-to-group alpha coherence ($r = -0.64$, $p = 0.0044$).

In sum, this study suggests that brain-to-brain synchrony increases as shared attention modulates entrainment by "tuning" neural oscillations to the temporal structure of our surroundings. Individuals who are less engaged with the stimulus show lower brain-to-brain synchrony levels with the rest of the group (Figure 4), and people who have interacted face-to-face show increased entrainment to each other.

Simultaneously recording EEG data from a group of teenagers under naturalistic circumstances presents obvious challenges when compared to laboratory-generated EEG experiments. Although we could not attain the level of experimental rigor that characterizes laboratory studies, we imposed as much structured design as possible, while minimally limiting students to engage with each other and with the class content, as they would under normal circumstances. Second, we carried out EEG recordings on 11 different days with the same series of experimental conditions, essentially replicating the same experiment 11 times on the same group of students (Figure 1A). Finally, we carried out a series of experiments to verify that we obtained interpretable recordings and that TI reliably indexes the synchronization of the neural signal across individuals in both the laboratory and in a classroom context (Figure S2).

Conclusions

We repeatedly recorded brain activity from a group of 12 students simultaneously as they engaged in natural classroom activities and social interactions. Over the course of 11 different school days distributed over one semester, we found that brain-to-brain synchrony between students consistently predicted class engagement and social dynamics. These findings suggest that brain-to-brain synchrony is a sensitive marker that can predict dynamic classroom interactions, and this relationship may be driven by shared attention within the group. The approach we describe provides a promising new avenue to investigate the neuroscience of group interactions under ecologically natural circumstances.

ACCESSION NUMBERS

The raw data reported in this paper have been deposited in the Open Science Framework under ID code OSF: 10.17605/OSF.IO/NSUHJ.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.04.002.

AUTHOR CONTRIBUTIONS

S.D. and D.P. conceptualized the research. L.W., L.K., J.M., M.D., and D.P. designed the research. S.D., L.K., J.R., and I.D. performed the research. M.O. and S.D. designed custom software. L.W., S.D., L.K., J.R., and G.M. analyzed data. S.D., D.P., L.W., I.D., G.M., J.J.V.B., and M.D. wrote the paper.

ACKNOWLEDGMENTS

This research was supported by NSF INSPIRE Track 1 Award 1344285 and Netherlands Organisation for Scientific Research Award 275-89-018. We thank the school staff and especially the Advanced Biology students for generously granting access to the school and for donating their time and resources (especially M. Schaffer and S. Dhanesar); M. Westerlund, S. Ashrafi, and M. Rabadi for co-facilitating the educational portion of the project; K. Du, G. Mackellar, and the rest of the EMOTIV team for hardware support; A. Flinker for technical consultation; and B. Tuller for comments. The research was approved by New York University's Committee on Activities Involving Human Subjects.

Received: October 20, 2016
Revised: February 27, 2017
Accepted: April 4, 2017
Published: April 27, 2017

REFERENCES

1. Dunbar, I.R. (2002). The social brain hypothesis. In Foundations in Social Neuroscience, J.T. Cacioppo, ed. (MIT Press), pp. 69–88.
2. Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (2013). Toward a second-person neuroscience. Behav. Brain Sci. *36*, 393–414.
3. Babiloni, F., and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. Neurosci. Biobehav. Rev. *44*, 76–93.
4. Dmochowski, J.P., Bezdek, M.A., Abelson, B.P., Johnson, J.S., Schumacher, E.H., and Parra, L.C. (2014). Audience preferences are predicted by temporal reliability of neural processing. Nat. Commun. *5*, 4567.
5. Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. PLoS ONE *5*, e12166.
6. Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. Trends Cogn. Sci. *16*, 114–121.
7. Hari, R., Himberg, T., Nummenmaa, L., Hämäläinen, M., and Parkkonen, L. (2013). Synchrony of brains and bodies during implicit interpersonal interaction. Trends Cogn. Sci. *17*, 105–106.
8. Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., and Lu, C. (2015). Leader emergence through interpersonal neural synchronization. Proc. Natl. Acad. Sci. USA *112*, 4274–4279.
9. Pfeiffer, U.J., Timmermans, B., Vogeley, K., Frith, C.D., and Schilbach, L. (2013). Towards a neuroscience of social interaction. Front. Hum. Neurosci. *7*, 22.
10. Stephens, G.J., Silbert, L.J., and Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. Proc. Natl. Acad. Sci. USA *107*, 14425–14430.
11. Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., Rossi, S., Bruni, I., Rossini, P.M., and Perani, D. (2012). Brains "in concert": frontal oscillatory alpha rhythms and empathy in professional musicians. Neuroimage *60*, 105–116.
12. Duan, L., Dai, R.N., Xiao, X., Sun, P.P., Li, Z., and Zhu, C.Z. (2015). Cluster imaging of multi-brain networks (CIMBN): a general framework for hyperscanning and modeling a group of interacting brains. Front. Neurosci. *9*, 267.
13. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. Science *303*, 1634–1640.
14. Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I.P., Hari, R., and Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. Proc. Natl. Acad. Sci. USA *109*, 9599–9604.
15. Dikker, S., Silbert, L.J., Hasson, U., and Zevin, J.D. (2014). On the same wavelength: predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. J. Neurosci. *34*, 6267–6272.
16. Debener, S., Minow, F., Emkes, R., Gandras, K., and de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? Psychophysiology *49*, 1617–1621.
17. Wen, X., Mo, J., and Ding, M. (2012). Exploring resting-state functional connectivity with total interdependence. Neuroimage *60*, 1587–1595.
18. Reyes, M.R., Brackett, M.A., Rivers, S.E., White, M., and Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. J. Educ. Psychol. *104*, 700–712.
19. Balliet, D., Wu, J., and De Dreu, C.K. (2014). Ingroup favoritism in cooperation: a meta-analysis. Psychol. Bull. *140*, 1556–1581.
20. Saarela, M.V., Hlushchuk, Y., Williams, A.C.D.C., Schürmann, M., Kalso, E., and Hari, R. (2007). The compassionate brain: humans detect intensity of pain from another's face. Cereb. Cortex *17*, 230–237.
21. Mundy, P., and Newell, L. (2007). Attention, joint attention, and social cognition. Curr. Dir. Psychol. Sci. *16*, 269–274.
22. Farroni, T., Csibra, G., Simion, F., and Johnson, M.H. (2002). Eye contact detection in humans from birth. Proc. Natl. Acad. Sci. USA *99*, 9602–9605.
23. Tomasello, M. (1995). Joint attention as social cognition. In Joint Attention: Its Origins and Role in Development, C. Moore, and P. Dunham, eds. (Erlbaum), pp. 103–130.
24. Ki, J.J., Kelly, S.P., and Parra, L.C. (2016). Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. J. Neurosci. *36*, 3092–3101.
25. Knoblich, G., Butterfill, S., and Sebanz, N. (2011). Psychological research on joint action: theory and data. In The Psychology of Learning and Motivation, B. Ross, ed. (Academic Press), pp. 59–101.
26. Zion Golumbic, E.M., Ding, N., Bickel, S., Lakatos, P., Schevon, C.A., McKhann, G.M., Goodman, R.R., Emerson, R., Mehta, A.D., Simon, J.Z., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". Neuron *77*, 980–991.
27. Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science *320*, 110–113.
28. Haegens, S., Händel, B.F., and Jensen, O. (2011). Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. J. Neurosci. *31*, 5197–5204.
29. Palva, S., and Palva, J.M. (2007). New vistas for α-frequency band oscillations. Trends Neurosci. *30*, 150–158.

Current Biology, Volume *27*

Supplemental Information

Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom

Suzanne Dikker, Lu Wan, Ido Davidesco, Lisa Kaggen, Matthias Oostrik, James McClintock, Jess Rowland, Georgios Michalareas, Jay J. Van Bavel, Mingzhou Ding, and David Poeppel

A

Number of channels / Channel names (Int.10-20 locations)	14 (plus CMS/DRL references. P3/P4 locations) / AF3, AF4, F3, F4, F7, F8, FC5, FC6, P3 (CMS), P4 (DRL), P7, P8, T7, T8, O1, O2
Sampling method / sampling rate	Sequential sampling, Single ADC / 128 SPS (2048 Hz internal)
Resolution / Bandwidth	16 bits (14 bits effective) 1 LSB = 0.51μV / 0.2 - 45Hz, 50 Hz & 60 Hz digital notch filters
Dynamic range (input referred)	256mVpp
coupling mode	AC coupled
Connectivity / battery life	Proprietary wireless, 2.4GHz band / 12 hours (typical)

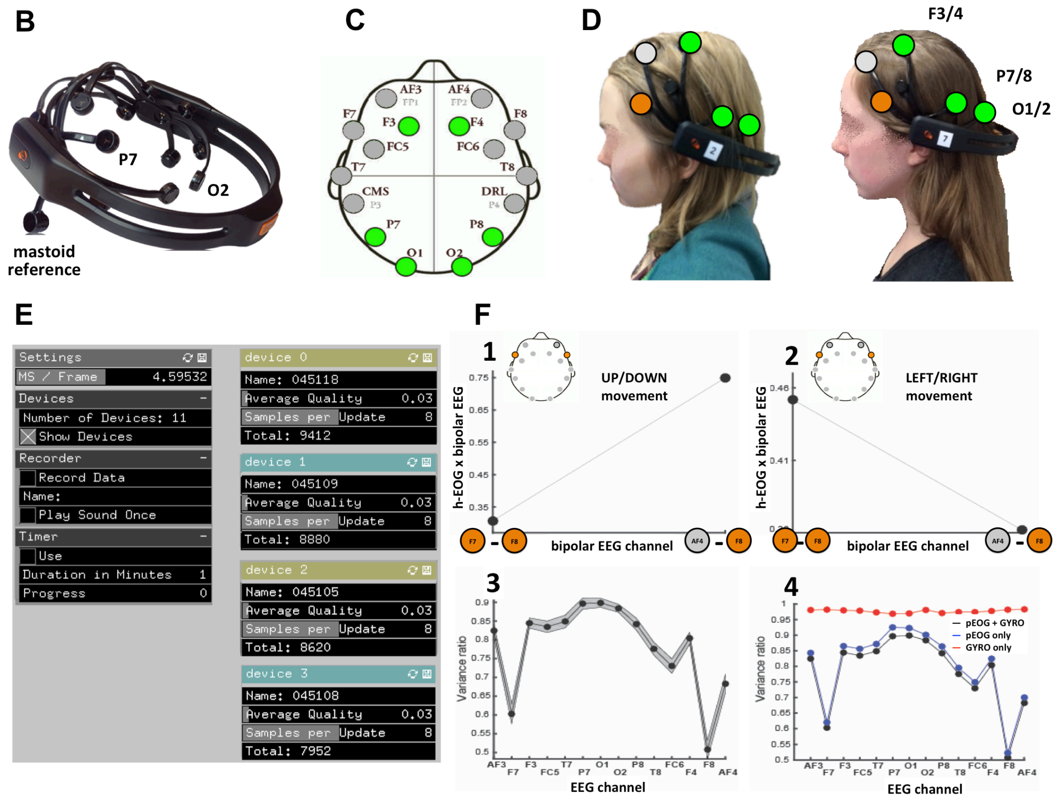

Figure S1. EEG data collection equipment (Related to Figure 1)

A. Hardware specifications of the emotiv EPOC EEG headset; **B.** Image of the emotiv EPOC headset (side view); **C**. Top-view of the electrode locations with those included in analysis marked in green (see text and S3 for details); **D**. Side-view of two sample students wearing the EEG headset. Green electrodes were included in the synchrony analysis and electrodes used as h-EOG and v-EOG proxies are marked in orange and grey respectively. **E.** Screenshot of the recording software, which was developed specifically for the purposes of the present study. **F.** Eye & head-motion artifact regression analysis: **F-1.** Correlation between v-EOG and bipolar EEG channels during UP-DOWN eye movements; **F-2.** Correlation between h-EOG and bipolar EEG channels during LEFT-RIGHT eye movements; **F-3.** Ratio of variance after vs. before regression; **F-4.** Ratio of variance after vs. before regression using different parts of the model.

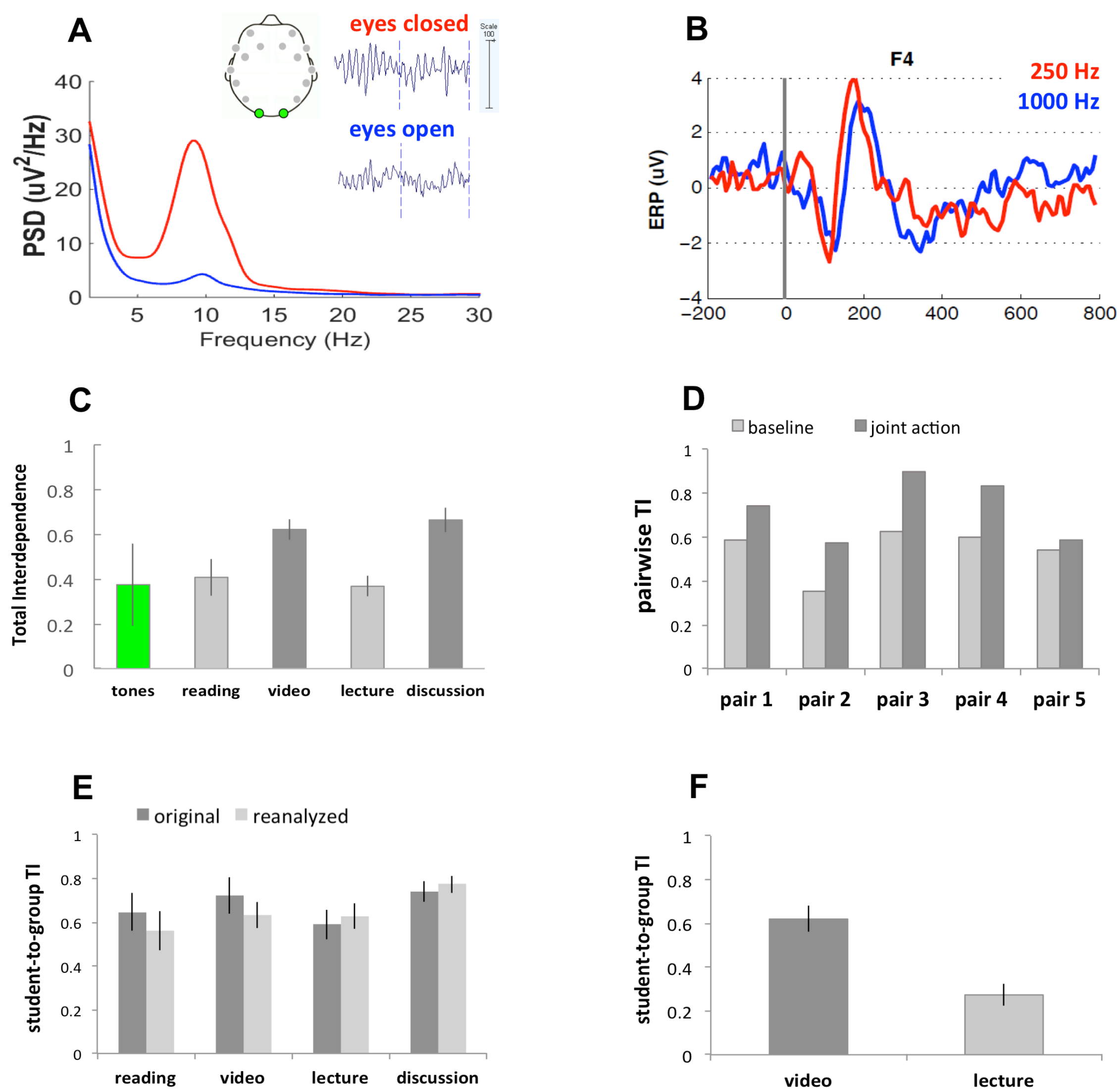

Figure S2. Control experiments (Related to Figures 2-3).

A-right. 2-second raw EEG trace recorded from posterior electrodes of a sample student with eyes closed vs. eyes open; **A-left**. Average power spectra comparing eyes open vs. eyes closed combining two occipital channels at the group level (all 10 students); **B**. Event-related potentials to 250 Hz (blue) and 1000 Hz (red) tones recorded from a representative electrode (F4); **C.** Total Interdependence to tones, compared to teaching styles (See S2). **D.** Data from two joint attention control experiments, showing that joint action conditions yielded consistently higher pairwise TI values across the five pairs (pairs 1-3 are taking from Control Experiment 1; pairs 4-5 participated in control experiment 2, see text). **E.** For one representative day, student-to-group TI values were recomputed after regressing out head motion and residual eye artifacts (see text). No significant changes in student-to-group synchrony were observed. **F.** Data from 9 senior biology students at a separate high school also exhibited consistently higher student-to-group TI for video than lecture teaching styles, replicating the pattern reported in the study.

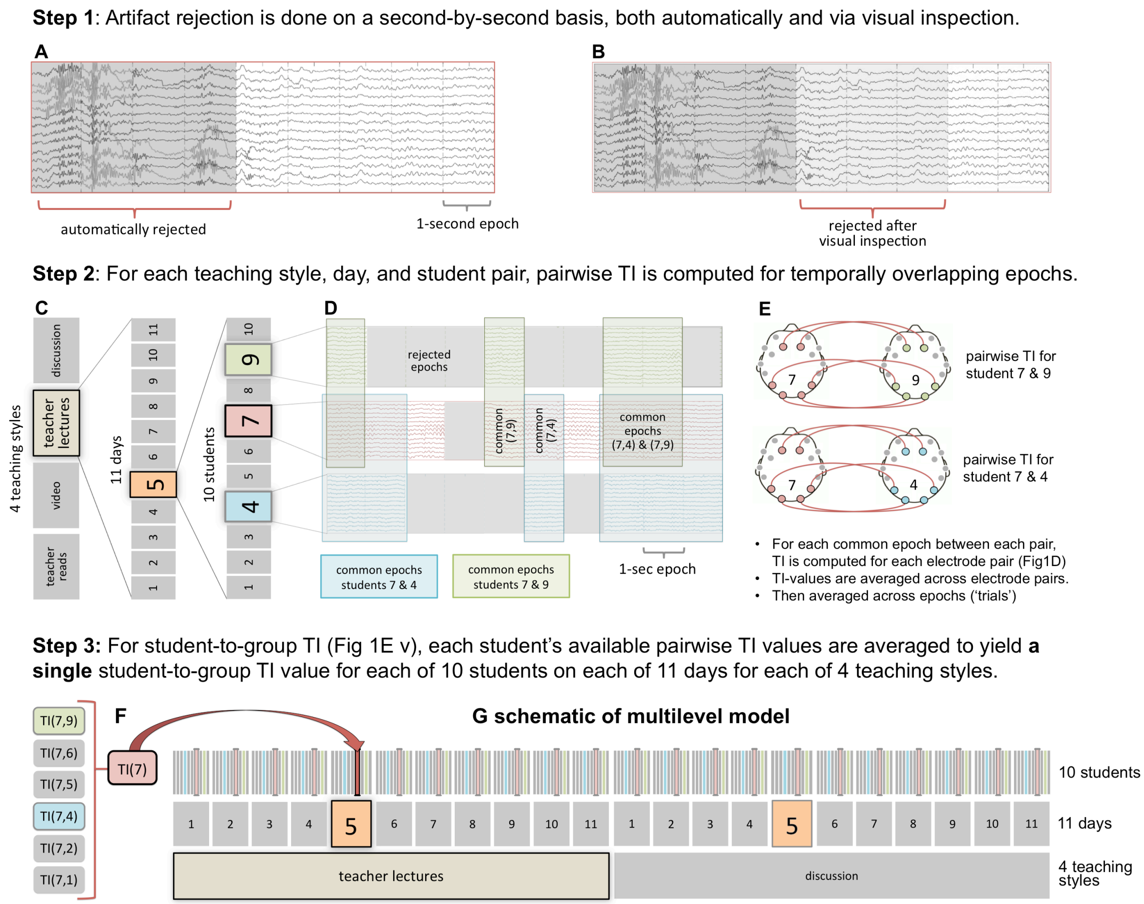

Figure S3. Preprocessing and Analysis Pipeline (Related to Figures 2-3)

A. Following band-pass filtering (0.5-35 Hz), the continuous EEG data was divided into 1-second epochs for artifact rejection. Epochs with movement and ocular artifacts were first rejected automatically according to a 100 μV rejection threshold, followed by manual rejection through visual inspection (**B**). **C**. Class content was presented using four teaching styles (with the exception of reading aloud, see text) on 11 separate days to a group of 12 students (only 10 students were included in the analysis). **D**. Pairwise TI was computed for each student with each other student, on each day and for each teaching style. For each pair of students, only epochs that were accepted in both students (common epochs) were included in the analysis. **E**. TI was computed for each electrode pair, and then averaged across electrode pairs and across common epochs. **F.** For group TI, all available pairwise TI values were averaged to generate a single value per day and teaching style (also see Figure 1E(ii)); For student-to-group TI, each student's available pairwise TI values were averaged to yield one single synchrony value for each student on each day and for each teaching style (also see Figure 1E(iii)). **G.** This value was then included in a multilevel model (10 students for 11 days and 4 teaching styles) for analysis.

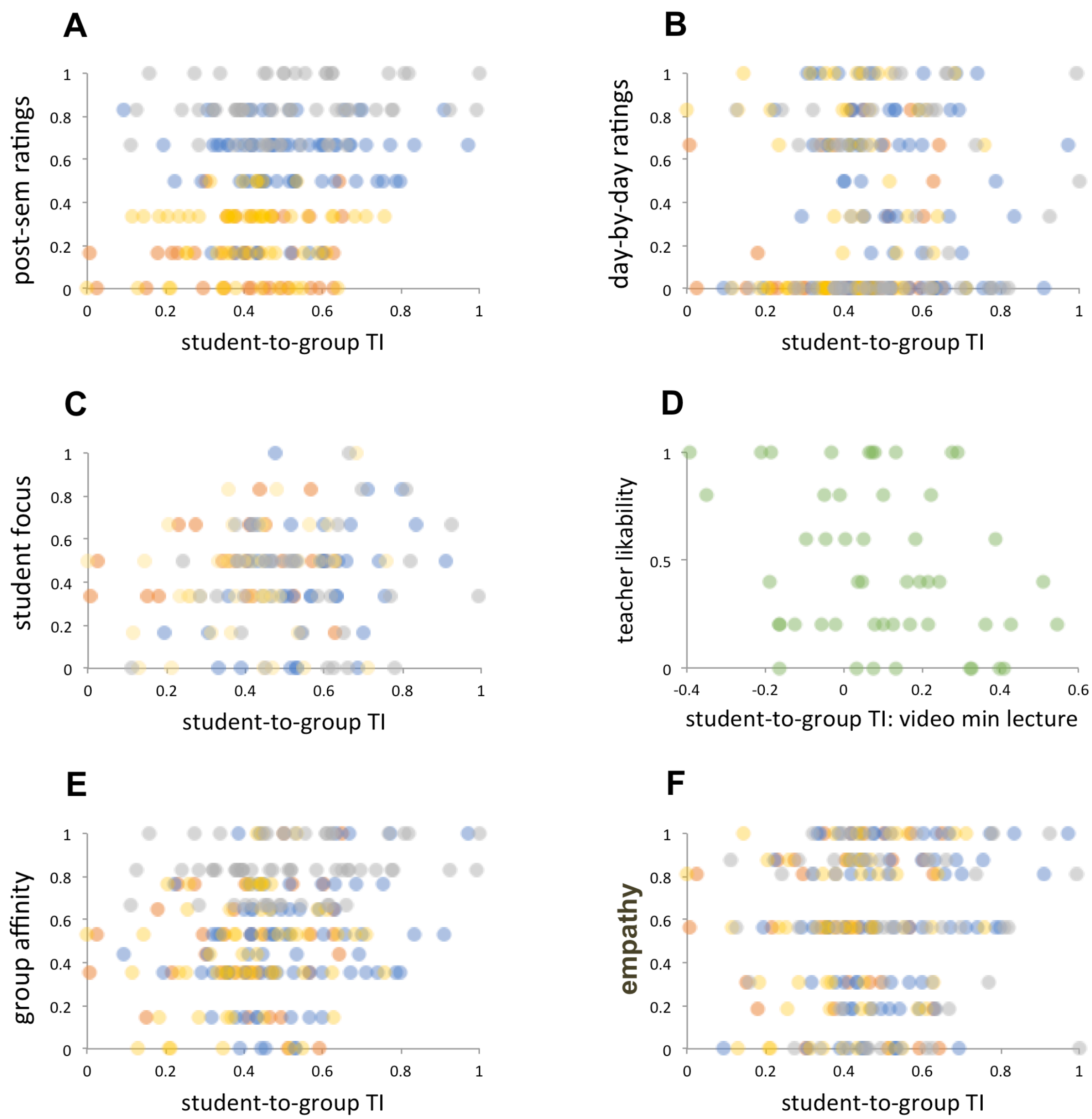

Figure S4. Correlation scatterplots of individual difference and student ratings with student-to-group synchrony during class (Related to Figures 2 & 3).

Individual student-to-group TI values for each student on each day for each variable under investigation: A. Post-Semester Ratings; B. Day-By-Day Ratings; C. Student Self-Reported Focus; D. Teacher Likeability; E Group Affinity; and F. Empathy. Statistics were conducted over these values, which are max-min normalized for presentation purposes.

Table S1: Percentage of data accepted by teaching style, across days and subjects (Related to Figures 2 and 3; Also see S3)

Teaching Style	% Epochs Accepted	StdErr
Reading Aloud	59 %	2 %
Video	69 %	2 %
Lecture	58 %	3 %
Discussion	38 %	2 %

Note that these values correspond to the % epochs accepted for individual students. Since the TI analysis is based on pairs of students, only epochs that were accepted in both students were included in the analysis (see Figure S3).

Table S2: Post-hoc Tukey-Kramer test for dependent variables (Related to Figure 2)
Highlighted rows mark statistically significant differences ($p<0.05$).

Day-by-day student ratings

Teaching style #1	Teaching style #2	Avg. diff. (#1 minus #2)	StdErr of diff.	P-val
Discussion	Lecture	1.79	0.57	0.054
Discussion	Reading	3.32	0.45	0.000
Discussion	Video	1.60	0.47	0.038
Lecture	Reading	1.53	0.55	0.089
Lecture	Video	-0.19	0.43	0.969
Reading	Video	-1.72	0.29	0.002

Post-semester student ratings

Teaching style #1	Teaching style #2	Avg. diff. (#1 minus #2)	StdErr of diff.	P-val
Discussion	Lecture	3.50	0.50	0.000
Discussion	Reading	4.10	0.48	0.000
Discussion	Video	1.50	0.54	0.086
Lecture	Reading	0.60	0.31	0.269
Lecture	Video	-2.00	0.49	0.013
Reading	Video	-2.60	0.40	0.001

Group synchrony

Teaching style #1	Teaching style #2	Avg. diff. (#1 minus #2)	StdErr of diff.	P-val
Discussion	Lecture	0.0112	0.0036	0.114
Discussion	Reading	0.0130	0.0060	0.278
Discussion	Video	-0.0040	0.0040	0.747
Lecture	Reading	0.0018	0.0051	0.983
Lecture	Video	-0.0152	0.0035	0.039
Reading	Video	-0.0170	0.0061	0.150

Student-to-group synchrony

Teaching style #1	Teaching style #2	Avg. diff. (#1 minus #2)	StdErr of diff.	P-val
Discussion	Lecture	0.0099	0.0026	0.019
Discussion	Reading	0.0083	0.0032	0.109
Discussion	Video	-0.0018	0.0032	0.940
Lecture	Reading	-0.0016	0.0037	0.971
Lecture	Video	-0.0117	0.0026	0.007
Reading	Video	-0.0101	0.0047	0.203

Supplemental Experimental Procedures

S1. Detailed description of the procedure

Overall procedure. The experiment took place between September 2014 and January 2015. We partnered with a senior biology class at a New York City high school. In September-October 2014, NYU investigators taught the students a neuroscience crash course, including the basics of signal processing and experimental design. Students were further familiarized with the EEG equipment. This was both educationally and practically motivated: each class was only 50 minutes, and the set-up time could be minimized by letting students help out with the application of the EEG headsets. From October 2014-January 2015, we visited the class eleven times to record the students' EEG as they engaged in semi-regular classroom activities. Class content followed the regular biology curriculum. In the spring semester of 2015, students conducted their own EEG experiment, under the guidance of NYU researchers.
For a video showing students setting up the emotiv EEG headsets and engaging in a series of test runs, see https://vimeo.com/212150060).

Participants. The subjects of this study were 12 healthy high school students in their senior year (9 females and 3 males, age 17-18) with no known history of neurological disease. Students were enrolled in an Advanced Placement Biology class. All participants provided written informed consent after receiving a detailed explanation of the experimental procedures. The Institutional Review Board of New York University approved all experimental procedures for this study. Two girls were excluded from EEG data analysis because they consistently had poor data quality due to their hair volume and head shape.

Set-up. During setup, NYU researchers helped students (and students helped each other) fit the EEG headsets, and students filled out a pre-class questionnaire.

Teaching styles. As can be seen in Figure 1, the classroom activities included four main teaching styles. For three minutes, the teacher read to the students from his lecture notes. Next, students watched a two-minute instructional video related to the class' topic. The teacher then proceeded to lecture to the students for three minutes. Finally, students engaged in a group discussion about the class topic for five minutes. The Reading Aloud activity was terminated after five classes because it did not serve the students' learning goals. Video, lecture, and discussion were included in all 11 recording sessions. Data analyses were performed for each student on each day and teaching style separately.

Baseline activities. Three different baseline activities were conducted at the beginning and end of each class: facing the wall (2 minutes; post-class recording was skipped for classes 2-4 and 8 due to time constraints), facing the group (2 minutes for classes 4, 6, 8, and 10), or facing each other in pairs (2 minutes before class for classes 5, 7, 9, and 11). Pairs were determined each class based on where students sat in class (See Figure 3C for an example class configuration), which was randomized each time: The experimenters, and not the students, determined the seat assignment. During these baseline activities, students were instructed to sit still, refrain from talking, and focus on the wall, the group, or each other.

Questionnaires. Prior to the eleven EEG-recording sessions, information about students' social traits was collected, including group affinity (see Figure 2E) and the Personal Distress Scale of the Interpersonal Reactivity Index as an index of empathy ([S1, S2]; Figure 2F)).
Second, students filled out pre- and post-class questionnaires (during nine out of eleven classes) asking them how focused they were (student focus; Figure 2F) and how much they enjoyed different segments of class (day-by-day ratings; Figure 2A) on a scale from 1 to 7 (max-min normalized in all figures for presentation purposes). These questionnaires also included a number of items proposed by the students (e.g., asking how hungry or caffeinated they were) that were not analyzed for the purposes of the present study (also see Analysis Strategy below).
Third, after the completion of all eleven EEG-recording sessions, students filled out one additional questionnaire, again asking them how much they enjoyed the four teaching styles overall (post-semester ratings; Figure 2A), how much they liked the teacher (teacher likeability; Figure 2E), and how close they felt to each of their peers (closeness ratings; Figure 3C-D).

S2. Data collection & preprocessing

EEG data collection. We recorded simultaneous EEG activity from 12 students in their regular classroom as they were following their regular AP Biology curriculum (adapted to the teaching formats indicated in Figure 1) during 11 classes, which took place at 8:30am, 10:40am, or 2:20pm (EEG data available online at https://osf.io/nsuhj). Students had been made aware of movement artifacts during the neuroscience crash-course they received prior to the recording sessions and were instructed to minimize overt movement during the EEG recordings. We used 14-electrode emotiv EPOC wireless EEG headsets (mastoid reference locations), paired with custom software built using the OpenFrameworks software package (www.openframeworks.com) capable of recording EEG data from twelve students simultaneously onto a single computer (MacBook Pro). See Figure S1A for the emotiv EPOC hardware specifications, Figure S1B for a picture of the headset, and Figure S1E for a screenshot of the recording software Graphical User Interface.

EEG data quality. Figure S1C shows a schematic top-view of the electrode locations, with the electrodes included for analysis marked in green. Although electrode locations are mapped onto a standard 10/20 layout, it is important to note that the actual electrode placement was inconsistent across students since by design the headset was placed slightly differently on each student's head (Figure S1D). Therefore, we did not differentiate between electrodes in our analysis (see below): For each subject pair, the unit of analysis was the TI computed for each electrode pair (one from each subject, marked green in Figure S1C) and then averaged across all available electrode pairs between subjects (see Electrode Selection and Figure S3E). Figure S3A-B show raw EEG traces from a representative headset recorded on a typical recording day for all 14 electrodes.

Control Experiments.
We carried out multiple control experiments and 'sanity checks' to verify that (a) our recordings indeed consisted of reliable, interpretable EEG data (Figure S1F, S2A-B), (b) Total Interdependence is a valid method to capture synchronized neural activity across students as a function of shared attention (Figure S2C-E), and (d) our findings are not only replicable across days, but also across groups and schools (Figure S2F). These are discussed in turn below.

S1F. Regressing out the effect of eye and head movements from the EEG data. When inter-brain synchrony is computed from EEG measurements during naturalistic experimental conditions, one of the main questions that arise is whether it is driven by synchronous non-brain signals present in the EEG measurements. In the context of the current experiment, where students were attending to various forms of classroom interactions, some of such non-brain sources could be eye or head movements modulated by the stimuli that the students were presented with or from the surrounding environment. For example, the estimated interbrain synchrony could be driven by synchronous eye movements between students simultaneously tracking salient moving stimuli. Similar non-brain-related synchrony could result from synchronous head movements, driven by the environment or presented stimuli, which stress or strain similar muscles across students and create similar artifacts in the EEG recordings. In order to estimate the effect of eye and head movements in the inter-brain synchrony, four explanatory variables representing these artifacts were regressed out of the EEG recordings and the TI metric was recomputed.
Head movements were assessed using a two-axis gyroscope built in the emotiv EEG headset and sampled synchronously to the EEG channels. The output of the gyroscope provides two variables, G_x and G_y, representing angular velocity with a fixed resolution ($1\ °/s$) around the X and Y axes respectively. The X-axis is the "horizontal axis" running through the bases of the headset arm, over the ears of the subject (Figure S1B). The Y-axis is the "vertical axis" running along the azimuthal direction in right angles to the headset arm and the X-axis. These two variables were used as regressors in order to remove the effect of head movements from the EEG data. Because the gyroscopes in the emotiv headset have a fixed resolution of $1\ °/s$, the data is not smooth but it has a step-wise behavior. In order to avoid contaminating the EEG data with such step-wise patterns after the regression, the gyroscope variables were low-pass filtered at 30 Hz. This filtering converted the step-wise values into smooth and continuous variables suitable for regression of the EEG data.
Eye movements are typically assessed using Electro-oculography(EOG), with bipolar electrodes placed over and below the eye, for detecting vertical eye movements(v-EOG), and on the left and right outer canthi for horizontal eye movements (h-EOG). Since no bipolar EOG electrodes were deployed in the current experiment, two proxy EOG variables were devised from the recorded EEG channels. From the

spatial layout of the emotiv headset electrodes (see Figure S1C-D), it is evident that the closest proxy to a h-EOG measurement is the difference between the electrodes F7 and F8. As these electrodes are located almost on a horizontal line in parallel to horizontal eyes movements and are the nearest electrodes to the eye muscles, it was expected that they would be mostly sensitive to horizontal and much less sensitive to vertical eye movements. The most suitable identified candidates for a v-EOG proxy variable were the bipolar derivations AF4-F8 and AF3-F7. As these bipolar derivations are identical with respect to the eye in measurement, only one of them was selected for the rest of the analysis, namely the derivation AF4-F8 (Figure S1F-1). The spatial locations of the electrodes in this proxy v-EOG variable are significantly different from what would be expected for an actual v-EOG measurement. However, it is expected that this derivation can still capture vertical eye movements. This is based on the fact that the electric field diminishes with the inverse of the square of distance. Thus, although the electric field measurements from vertical eye movements will have the same polarity in electrodes AF4 and F8 (they are both located above the eye muscle), the field in F8 will be much stronger than in electrode AF4. As a result, the difference will still capture the electric field from the vertical eye movements. Since the electrodes F8 and AF4 also have a horizontal offset, we expected that their bipolar derivation would also capture horizontal eye movements. Using single EEG electrodes as proxies to v-EOG and h-EOG was not considered a possible solution because single electrodes carry a mixture of electric fields from the entire brain, and regressing them out would remove a large amount of brain activity. Bipolar derivations of electrodes, close to the eyes like the ones described above, eliminate such common signals and are more sensitive to strong local electric field gradients from the eyes.

In order to test the suitability of the proxy EOG variables F7-F8 and AF4-F8, two surrogate experiments were performed with a single participant, during which EEG (with an emotiv headset) and EOG (using separate electrodes) were recorded simultaneously. In the first surrogate experiment, the participant was instructed to move her eyes up and down for one minute. In the second surrogate experiment, the participant was instructed to move her eyes left and right, also for one minute. The electrodes for the actual v-EOG were placed above and below the right eye. The electrodes for the actual h-EOG were placed on the outer canthi of the eyes. Both the EOG and EEG signals were pass-band filtered from 1 to 10 Hz which typically contains the main envelope of eye movement artifacts (High gamma band artifacts from eye movements were not accessible as the effective frequency range of the emotiv headset goes only up to 43 Hz [S3]. First, the Pearson correlation coefficient between v-EOG channels and the proxy EOG bipolar variables, F7-F8 and AF4-F8, was computed during up-down eye movements (surrogate experiment 1). As can be seen in figure S1F-1 the bipolar vertical proxy AF4-F8 had a much higher correlation with the actual v-EOG channel and confirmed the expectation that is a suitable candidate for capturing vertical eye movements. Similarly, the correlation between the actual h-EOG and the two proxy variables was computed during left-right eye movements, shown in figure S1-A2. In this case the horizontal proxy variable F7–F8, as expected, had the highest correlation and this confirmed its selection for capturing horizontal eye movements. Here it must be mentioned that the correlation of v-EOG with AF4-F8, $(0.749, p < 10^{-20})$, was notably higher than the correlation of h-EOG with F7-F8 $(0.452, p < 10^{-20})$. One possible explanation for this higher correlation is that, in addition to vertical eye movements, the bipolar variable AF4-F8 is also sensitive to eye blinks, which create strong electric fields. These additional events increase the correlation with the actual v-EOG signal. The variables h-EOG and F7-F8 are much less sensitive to eye blinks.

Together, the above surrogate experiments verified that the bipolar EEG variables F7-F8 and AF4-F8 can be used as proxies to h-EOG and v-EOG and can be used to regress residual eye-movement related artifacts out of the EEG data. For the scope of the regression analysis presented below the proxy variable F7-F8 will be termed P_h and AF4-F8 will be termed P_v. In addition to these two variables, the two gyroscope output variables described at the beginning of this session were used as regressors of the EEG channels in a multivariate linear model.

In the linear multivariate approach described below, instead of building a model directly on the EEG data, its principal components were used. This is also the case for the two proxy EOG regressor variables and the two gyroscope variables. The reason for using Principal Component Analysis (PCA) here is that there is a high degree of collinearity between the EEG channels, between the two proxy regressors and to a much lesser degree between the gyroscope regressors. By applying PCA on the data and on each group of regressors, collinearity within them is removed. The validity of similar modeling approaches, based on orthogonal decompositions of correlated measurements into their uncorrelated components, has already been demonstrated in the analysis of EEG and MEG data [S4, S5].

The linear regression has the form:

$$Y_{pca} = A + B \cdot X + E_{pca}$$

where Y_{pca} are the principal components of the EEG data Y. Each of the 14 rows corresponds to a principal component and each column to a sample.

$$Y_{pca} = C \cdot Y$$

C is the principal component coefficient matrix. It maps EEG channels to principal components. Each row corresponds to a principal component and each column to an EEG channel.

$$X = \begin{bmatrix} G_{pca} \\ P_{pca} \end{bmatrix}$$

X contains the 2 groups of regressor variables.
G_{pca} contains the principal components of the two gyroscope variables G_x and G_y.
P_{pca} contains the principal components of the proxy EOG variables P_h and P_v.
A contains the costant terms of the linear model for each of the EEG channels.
B contains the slope coefficients for each of the regressor variables and each modelled EEG channel. Each row corresponds to a channel and each column to a regressor.
E_{pca} contains the residual EEG data that could not be modelled by the given regressors. This residual is at the principal component space. Each row corresponds to a principal component and each column to a sample. This residual data can be then back-projected to the EEG sensor space simply by multiplying it with the inverse of the principal component coefficient matrix.

$$E = C^{-1} \cdot E_{pca}$$

E is the final residual of the EEG data after the modulation by the regressor variables has been removed. This is the residual data, based on which the inter-brain synchrony (TI) was re-computed and compared to the originally calculated metric before the regression.

Prior to applying the above modeling to the data, the EEG data was high-pass filtered at 0.5 Hz and demeaned. The gyroscope data was low-pass filtered at 30 Hz (for converting its step-wise fluctuations to continuous) and demeaned. Before the data was entered into PCA and the multivariate modeling process, all data segments flagged as noisy were removed (exactly the same as in the data used for the original TI computation). The multivariate model was estimated by the MATLAB "*mvregress*" routine, using the covariance-weighted("*cwls*") least-squares method [S6, S7] . Although initially the maximum-likelihood method was employed, the multivariate distribution of the model residuals was not normal, as indicated by the Royston's Multivariate Normality Test, and so the least-squares method was finally used [S8, S9].

The above modeling process was applied to each student's recorded data during each of 11 different experimental sessions spanning three days (reading, video, lecture and discussion for each day except the third day, for which reading aloud data was not available). In total, there were 90 EEG datasets for which a multivariate model was estimated and the modulation of eye and head movements was regressed out, leaving a residual dataset.

In order to assess the effect of the regression on the original data, the ratio of the variance after versus before the regression was estimated for each of the 90 datasets and each sensor. Figure S1F-3 presents the mean variance ratio values across for each EEG sensor. The gray-shaded area represents the standard error of the mean (SE) across all the used datasets. The three channels with the highest variance reduction were F8, F7 and AF4. This was expected, as they are the channels that constructed the two proxy EOG regressors. Channel F8, which was a constituent of both regressors, exhibited the highest variance reduction. An interesting fact is that the variance reduction in channel AF3 was not as high as for its mirror channel AF4. This probably is due to the fact that the vertical proxy variable P_v captured not only vertical eye movements but also local brain activity on the right side, which was not present in the symmetric location on the left side of the head (in contrast to what should be expected in the case of eye movements). This is also supported by the fact that the variance reduction for lateral right side sensors F4, FC6, T8 and P8 was higher than for their symmetric left-side sensors (F3, FC5, T7 and P7). Nevertheless, even these sensors showed significant variance reduction, between 10.3% and 16.5%. In overall it must be noted that the variance reduction showed a descending gradient from anterior towards posterior and from lateral towards medial EEG sensors, with occipital and parietal sensors showing the lowest variance reduction.

The final assessment of the effect of regression on the data was to examine how much of the variance reduction was attributed to eye and how much to head movements, as represented by the regressor variables. In order to pursue this, for each of the 90 datasets the model coefficients for the gyroscope regressor variables, representing head movements, were omitted. Based on this reduced model, the residual and the variance reduction ratio for it were computed. This ratio represented how much of eye movement artifacts were present in the original EEG data. The exact same procedure was then repeated but with omitting the eye movement regressors and keeping the head-movement ones. This indicated how much of head movement artifacts were present in the original EEG data. These results are shown in figure S1F-4. It is obvious that most of the variance reduction is attributed to the eye-movement regressors (blue line). The variance reduction from the gyroscope regressors (red) remained low, below 5%, and quite evenly distributed across the sensor array. This indicates that head movements during the experimental sessions were limited, and the EEG data was not contaminated by accompanying artifacts such as neck muscle and sensor movement artifacts. On the same figure the variance ratio from the full model is presented again (black line), for comparison purposes, which as expected is equal to 1 minus the sum of the variance ratios from the two reduced models.

In summary, the procedures described so far in this section showed that:

- Eye movement artifacts in the data can be captured using proxy EOG variables from frontal EEG channels.
- After regressing out of the EEG data the eye and head movement artifacts, the residual variance was significantly reduced in all sensors, with the least reduction in occipital and parietal areas.
- Most of this variance reduction was due to the regressors that represented eye movements. (This showed that students followed instructions and refrained from large head movements during the experiment).

After the above evaluation was applied to each dataset, the resulting EEG residual for one representative day was used in the computation of student-to-group synchrony (TI), in place of the originally used non-regressed data. This analysis is discussed below and results are shown in Figure S2E.

S2A. Data quality assessment of the emotiv system. As a rudimentary 'sanity check' test of the data quality of the emotiv system, EEG was contrasted between eyes-open (EO) and eyes-closed (EC; 2 minutes each). Figure S2A-right shows a sample 2-seconds EEG trace recorded from a posterior electrode of an example subject, comparing eyes-open vs. eyes-open resting. As clearly visible in the raw EEG from the sample subject, EC resting was accompanied by large alpha rhythm, which was significantly attenuated during EO resting. Figure S2A-left shows power spectra averaged over all students (data obtained on the first recording day). EO-EC power difference in the alpha band is again apparent. This observation confirms a well-established phenomenon sometimes referred to as alpha blocking [S9].

S2B. Two-tone ERPs. In a second test, students were instructed to passively listen to 100 instances of a 250 Hz tone and 100 instances of a 1000 Hz tone alternating at a stimulus-onset-asynchrony (SOA) of 1000 ms (tone duration 400 ms; played via an external speaker that was set up on the teacher's desk). The EEG signals were epoched between -200 ms before the onset of each stimulus and 800 ms after the onset of the same stimulus. Figure S2B shows a canonical ERP response to the two tones recorded from electrode location F4: a clear N1-P2 complex is visible, with the 250 Hz tones (P2 peak: $M=194.5$, $SD=13.4$) generating a slightly earlier peak latency in both components than the 1000 Hz (P2 peak: $M=218.5$, $SD=13.4$) tones ($t(3)=-3.67$, $p=0.017$; data from channels F3 and F4 was included in this analysis). This observation is again consistent with prior reports using well-established equipment [S10].

S2C. Validating TI as a measure of shared stimulus entrainment. To validate the TI measure, we first computed student-to-group Total Interdependence for each student who participated in the tone experiment (combining the 250 Hz and 1000 Hz stimuli). If Total Interdependence reflects (attention-modulated) entrainment to external stimuli, as we argue in Figure 4, TI values should be comparable to those teaching styles with a single-source auditory input, namely when the teacher was either reading aloud or lecturing. As can be seen in Figure S2C, student-to-group TI values in response to tones were indeed numerically similar to those obtained during Lecture and Reading Aloud teaching styles (note that for the Tones, we only have available data from five students, recorded once, while the Lecture and Reading Aloud data is averaged across multiple sessions). For additional discussions of the EEG data collected by the emotiv system, including comparisons to other EEG systems, see [S11, S12].

S2D. Validating TI as a measure of joint attention. In one control experiment, EEG data of three pairs of participants were collected from a finger-moving experiment, replicating [S13]. In the baseline condition, two subjects sat face-to-face while eye contact was not mandatory. In the joint action session, one subject was randomly selected to serve as a leader whose finger movements were tracked and mimicked by the follower, and then they switched their roles for another session. Both baseline and joint action session last 2 minutes.

A second control experiment compared collaborative vs. competitive contexts in the same task. As mentioned previously, following the conclusion of data collection, in the spring semester of 2015, the students conducted their own EEG experiment, under the guidance of NYU researchers. In this experiment, students simultaneously recorded EEG from two subjects during a block design task. Subjects had to rearrange colored blocks to match a pattern. The two subjects were either competing against each other or collaborating with each other. In the competition condition, each subject worked independently, trying to complete the task ahead of the other subject. In the collaboration condition, the two subjects worked together on the same puzzle, trying to complete the task together at the shortest time.

Together, these findings validate Total Interdependence as a valid method to capture differences in synchrony between dyads as a function of joint synchrony, in line with findings showing an increase in brain-to-brain synchrony during laboratory tasks that required dyads to coordinate visual attention (e.g., [S14, S15, S16, S17]).

S2E. Validating TI as a *neural* measure of synchrony as a function of shared attention. To ensure that synchrony values were not in part explained by head motion or residual eye movement, we reanalyzed data for one recording day after regressing out residual motion artifacts (details are described above). We reanalyzed data for Day 5, selected because of two criteria: (1) all four teaching styles were carried out on that day, (2) most TI values were available for that day relative to other days. As can be seen in Figure S2G, no consistent differences between the original and reanalyzed data were observed across or within teaching styles ($t(1,30) = 0.8602$, $p = .38$). In addition, the difference in student-to-group synchrony for Video and Discussion teaching styles on the one hand and Reading Aloud and Lecture styles on the other were virtually identical (Original: $t(1,11) = 3.98$, $p = 0.0021$, difference: 0.112; Reanalyzed: $t(1,11) = 3.91$, $p = .0024$, difference: 0.108). Thus, we can conclude that estimated interbrain synchrony was not significantly driven by synchronous eye or head movements between participants, something that would be manifested as a reliable TI reduction when the regressed data was used.

S2F. Replicating classroom student-to-group TI in a second school. To assess the robustness of our findings across time and contexts, in addition to replicating the same experiment 11 times, we replicated our study at a different high school. Figure 2SF shows student-to-group TI values for 9 senior biology students of a second New York City high school receiving the regular class content either via video or through their teacher lecturing (we only included these two teaching styles because these teaching styles exhibited the cleanest data in the first group of students). Student-to-group TI was significantly higher for video than lecture sessions ($F(1,5) = 21.59$, $p < .006$), replicating findings from the first group of students reported in the rest of this paper.

S3. Analysis. Quantifying brain-to-brain synchrony.

Figure S3 summarizes our preprocessing and analysis procedure.

Step 1: EEG preprocessing.
First, EEG segments corresponding to each teaching style were extracted. The signals were then band-pass filtered between 0.5 Hz and 35 Hz, and divided into 1-second epochs for artifact rejection and EEG analysis. Epochs with large movement artifacts were removed by setting a rejection threshold in EEGLAB [S18] of +/- 100 μV for any of the 14 channels at any time within the epoch (Figure S3A). Then, epochs with blinks and muscle artifacts were manually removed by visually inspecting each of approximately 52800 epochs of data collected across days, students, and teaching styles (Figure S3B). Any participant for whom more than 50 percent of epochs had to be excluded based on these criteria was excluded from analysis for that particular teaching style. Average acceptance rates for each teaching style (before excluding students) are included in Table S1.
Although the numbers in Table S1 may seem low, it is important to emphasize that since we were in a naturalistic environment, we were unable to minimize subjects' eye movements and body movements, resulting in a much larger proportion of artifacts than those observed in conventional laboratory or virtual-reality-based experiments.
After the removal of eye-movement and other movement artifacts, those channels whose average amplitude exceeded the mean channel amplitude by four standard deviation were excluded from analysis, and this process was iterated three times [S19].

Step 2: Computing brain-to-brain synchrony: Total Interdependence (TI)

To compute brain-to-brain synchrony between students for each teaching style on each day and each student pair (Figure S3C), we employed the method of *total interdependence* (TI; S20) to compute brain-to-brain synchrony among multiple students during a given teaching style. Total Interdependence is defined in terms of spectral coherence. In this study spectral coherence was computed based on the Welch method which, according to [S21], controls bias in coherence estimation.
For a pair of simultaneously acquired time series: $(x_1, y_1),\ (x_2, y_2),\ (x_3, y_3), \dots, (x_n, y_n)$, TI as defined by [S22] was computed according to:

$$TI_{x,y} = -\frac{1}{2\pi}\int_{-\pi}^{\pi} ln\left(1 - C_{xy}^2(\lambda)\right) d\lambda, \qquad (1)$$

where $C_{xy}(\lambda)$ is the coherence between the two signals, x and y, at frequency $f = \lambda/2\pi$. For two Gaussian processes, this formula was shown to measure the total amount of mutual information between them. [S23] further demonstrated that TI captures the total linear relationship between x and y time series. Numerically, for a given sampling frequencyf_s, Eq. (1) can be recast into an implementable form:

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

where $\Delta f = \frac{f_s}{2(N-1)}$ is the frequency resolution and N is the number of desired frequency points in the interval between 0 and the Nyquist frequency $f_s/2$.

In this study, TI was estimated by computing the magnitude squared coherence using the Welch method for 6 one-on-one paired combinations of electrodes from two subjects (Figure 1D and S3E) for which data was most often free of noise across students (also see above): two occipital channels (O1, O2), two frontal channels (F3, F4), and two parietal channels (P7, P8) were included for analysis (excluding any electrodes that may have been rejected following the procedure outlined above).

The magnitude squared coherence was calculated for the frequency range between 1 to 20 Hz by tapering non-overlapping 1s epochs (zero-padded [S24, S25] to 4s) with a Hanning window and performing the Fourier transform with 0.25 Hz frequency resolution. A minimum of 60 artifact-free common epochs for paired subjects was used for analysis. The common epochs were selected by taking the overlapping epochs that were accepted with the aforementioned preprocessing procedure between pairwise subjects for each teaching style on each day. For example, for the Lecture session on one recording day, epochs 1-10 and 12-18 were accepted for student 1, and 2-13 and 16-18 were accepted for student 2, then common epochs for student 1 and 2 were selected to be 2-10, 12, 13, 16-18 (and so on) which were next used for TI computation. This is illustrated in Figure S3D.
TI for one pair of subjects was obtained by averaging TI values across all paired electrodes (Figure S3E) and student-to-group TI was obtained by averaging TI values over all possible pairwise combinations between that one student and the rest of the group (Figure S3F). These values were then entered in a multilevel model for statistical analysis (Figure S3G; see Analysis Strategy below).

Alpha coherence was computed by summing the coherence values within 8-12 Hz. Student-to-group TI for student *i* was obtained by averaging all pairwise TI values where student *i* was a member of the dyad (i.e., $(i, j), i \neq j$, where *i* and *j* are indices of available students in each teaching strategy on each day; Figure S3E).Student-to-group alpha coherence was obtained by using the same averaging procedure as described for student-to-group TI (Figure S3E-G). Both TI and alpha band coherence values were normalized across conditions and recordings using a min-max transformation. Group alpha coherence explains 33 % of group TI (R^2 = .33)**.**

S4. Analysis Strategy. In order to test whether student ratings (day-to-day ratings and post-semester ratings) and brain synchrony (student-to-group synchrony and group synchrony) varied significantly between teaching styles, repeated-measures two-way ANOVAs were conducted with teaching style and day of recording as main factors. In the case of post-semester student ratings, only teaching style was considered as a factor (as these ratings were obtained only once). The ANOVAs were followed by post-hoc Tukey-Kramer tests to assess the significance of pairwise comparisons (e.g. lecture compared to video; See Tables S2).

Next, to investigate the relationship between student-to-group TI and questionnaire metrics across days we created multilevel models [S26] with days nested within students. Multilevel models were implemented in

the SAS PROC MIXED procedure (all random effect were modeled wherever possible; [S27]). To assess the independent contribution of the stimulus nature (teaching style) and individual factors respectively (listed below), we first created a binary factor Teaching Style (stimulus attribute) by reorganizing the data into two main 'conditions', grouping together Discussion and Video sessions on the one hand, and Reading Aloud and Lecture sessions on the other (previous analyses showed no difference between students ratings of these styles). As shown in Figure 2B, there was a main effect of Teaching Style on student-to-group synchrony, with significantly lower student-to-group TI for Reading Aloud and Lecture sessions than Discussion and Video sessions ($F(1,9) = 20.90$, $p = .0013$).
We then conducted a series of repeated measures analyses on student-to-group TI: Teaching Style & Post-Semester Ratings, Teaching Style & Self-Reported Focus, Teaching Style & Group Affinity, Teaching Style & Empathy (see the main text for statistics).
Finally, to assess any effects of the class configuration on TI, we ran a model comparing pairwise TI values and pairwise closeness ratings across three categories: pairs of students that were seated next to each other and had engaged in eye contact prior to class (adjacent + face-to-face pairs), students who sat next to each other but had *not* engaged in eye contact before class (adjacent + no face-to-face pairs), and students who were not seated next to each other (non-adjacent pairs; see main text and S1 above).

As already mentioned above, we collected a number of exploratory individual measures via questionnaires, many of which were designed by the student participants rather than the experimenters (e.g., "How much caffeine did you consume?" "How distracted are you by school work?"). These were largely for educational, rather than scientific purposes (i.e., to educate students about how to construct their own hypotheses; see SI for more information). To help minimize Type I errors, we focused on the subset of metrics mentioned above: self-reported focus, group affinity, empathy, and social closeness. These are listed and motivated in the Analysis Plan below.

Analysis Plan

Analysis		**Motivation/Research Question**
Main effect of stimulus & ratings: group vs. individual & day-by-day vs. semester		
(1) 2-way repeated measures ANOVAs on day-by-day & post-semester ratings	planned	How do students' self-reports of the teaching styles vary by condition?
(2) 2-way repeated measures ANOVAs on group TI & student-to-group TI	planned	Does a group metric or an individual metric better capture differences in our data?
(3) correlations: student ratings x TI	follow-ups	Are student ratings and brain-to-brain synchrony related? [S28]

Independent effects of stimulus vs. individual differences on brain-to-brain synchrony: 2 'state' variables and 2 'trait' variables		
Individual 'state' differences		
(1) repeated-measures multilevel regression analysis assessing effects of Teaching Style x ratings on student-to-group TI	planned	Does student engagement with the material independently predict student-to-group synchrony, above and beyond the nature of the stimulus (teaching style)?
(2) repeated-measures multilevel regression analysis assessing effects of Teaching Style x focus on student-to-group TI	planned	Does focus, a proxy for attention [S29] independently predict student-to-group synchrony?
Individual 'trait' differences		
(1) repeated-measures multilevel regression analysis assessing effects of Teaching Style x group affinity on student-to-group TI	planned	Do students with a natural affinity for joining and affiliating with groups experience greater neural synchrony?
(2) repeated-measures multilevel regression analysis assessing effects of Teaching Style x empathy on student-to-group TI	planned	Do student with greater empathy (indexed by the Personal Distress Scale [S1]) experience greater neural synchrony with the rest of the group?

Effects of co-presence: teacher likeability and student closeness (joint attention vs. class configuration)		
Teacher likeability		
(1) teacher likeability x student-to-group TI correlation	planned	Are student evaluations of their teacher associated with greater neural synchrony?
Student closeness		
(1) 1-Way ANOVA comparing pairwise TI for adjacent + face-to-face pairs, adjacent no face-to-face pairs, and non-adjacent pairs	follow-up	Does co-presence (class configuration) or social interaction (face-to-face baseline) better predict synchrony during class?
(2) correlation: student closeness x pairwise TI for adjacent + face-to-face pairs	planned	Does face-to-face interaction (eye contact) modulate whether closeness and synchrony are correlated during class?
(3) correlation: student closeness x pairwise TI for adjacent + not face-to-face pairs	follow-up	Does physical proximity modulate whether closeness and synchrony are correlated during class?
(4) correlation: student closeness x pairwise TI for non-adjacent pairs	follow-up	Are closeness and synchrony are correlated during class irrespective of direct interaction or physical proximity?

Post-hoc attention analysis		
(1) correlation: alpha coherence x alpha power correlation	follow-up	This was a post-hoc analysis providing additional, indirect evidence in support of an attention account of student-to-group synchrony.

Supplemental References

S1. Mundy, P., and Newell, L. (2007). Attention, joint attention, and social cognition. Curr. Dir. Psychol. 16, 269-274.

S2. Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. J. Pers. Soc. Psychol. 44, 113-126.

S3. Aydemir, O., Pourzare, S. and Kayikcioglu, T. (2012). Classifying various EMG and EOG artifacts in EEG signals. Przegląd Elektrotechniczny 88, 218-222.

S4. Klados, M. A., Papadelis, C., Braun, C. and Bamidis, P. D. (2011). REG-ICA: a hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. Biomed. Signal Process. Control 6, 291-300.

S5. Michalareas, G., Schoffelen, J.M., Paterson, G. and Gross, J. (2013). Investigating causality between interacting brain areas with multivariate autoregressive models of MEG sensor data. Hum. Brain Mapp. 34, 890-913.

S6. Little, R. J. A. and Rubin, D. B. (2002). Statistical analysis with missing data, Second Edition, (Hoboken, NJ: Wiley).

S7. The MathWorks, I. (2015). MATLAB and Statistics Toolbox Release 2015b, Natick, Massachusetts, United States.

S8. Trujillo-Ortiz, A., Hernandez-Walls, R., Barba-Rojo, K. and Cupul-Magana, L. (2007). Roystest:Royston's Multivariate Normality Test. A MATLAB file. [WWW document]. http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=17811.

S9. Berger, H. (1929). Uber das Elektroencephalogramm des Menschen (On the electroencephalogram of man). Arch. Psychiatr. Nervenkr. 87, 527–570.

S10. Roberts, T. P., and D. Poeppel (1996). Latency of auditory evoked M100 as a function of tone frequency. Neuroreport 7, 1138-1140.

S11. Debener, S., Minow, F., Emkes, R., Gandras, K., and De Vos, M. (2012) How about taking a low-cost, small, and wireless EEG for a walk? Psychophysiology 49, 1617-1621.

S12. Badcock, N.A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J. and McArthur, G. (2013). Validation of the emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs. PeerJ 1, e38.
S13. Yun, K., Watanabe, K., and Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. Sci. Rep. 2, 959.
S14. Babiloni, F., and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. Neurosci. Biobehav. Rev. 44, 76-93.
S15. Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. PLoS One 5, e12166.
S16. Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., and Lu, C. (2015). Leader emergence through interpersonal neural synchronization. Proc. Natl. Acad. Sci. USA 112, 4274-4279.
S17. Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., Rossi, S., Bruni, I., Rossini, P.M. and Perani, D. (2012). Brains "in concert": frontal oscillatory alpha rhythms and empathy in professional musicians. Neuroimage 60, 105-116.
S18. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9-21.
S19. Snyder, A. C., and Smith, M. A. (2014) Stimulus-dependent spiking relationships with the EEG. J. Neurophys. 114, 1468-1482.
S20. Wen, X., Mo, J., and Ding, M. (2012). Exploring resting-state functional connectivity with total interdependence. Neuroimage 60, 1587-1595.
S21. Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note. Front. Human. Neurosci. 7, 881.
S22. Gel'fand, I. M., and Yaglom, A. M. (1959). Calculation of the amount of information about a random function contained in another such function. Am. Math. Soc. Trans. 2, 199–246.
S23. Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. J. Amer. Statist. Assoc. 77, 304-313.
S24. Lalor, E. C., Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B. and Mcdarby, G. (2005). Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. EURASIP J. Adv. Signal Process., 2005, 3156-3164.
S25. Mäki, H., and Ilmoniemi, R. J. (2010). EEG oscillations and magnetically evoked motor potentials reflect motor system excitability in overlapping neuronal populations. Clin Neurophysiol., 121(4), 492-501.
S26. Goldstein, H. (1995). Multilevel statistical models, Second Edition (London: Arnold).
S27. Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. J. Educ. Behav. Stat. 23, 323-355.
S28. Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., and Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. Nat. Commun. 5, e4567.
S29. Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320, 110-113.