Leader emergence through interpersonal neural synchronization

Jing Jiang (蒋静)[a,b,c,d], Chuansheng Chen (陈传升)[e], Bohan Dai (代博涵)[a,b,f,g], Guang Shi (时光)[a,b], Guosheng Ding (丁国盛)[a,b], Li Liu (刘丽)[a,b], and Chunming Lu (卢春明)[a,b,1]

[a]State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China; [b]Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China; [c]Berlin School of Mind and Brain, Humboldt University, 10117 Berlin, Germany; [d]Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; [e]Department of Psychology and Social Behavior, University of California, Irvine, CA 92697; [f]Max Planck Institute for Psycholinguistics, Nijmegen 6500 AH, The Netherlands; and [g]International Max Planck Research School for Language Sciences, Nijmegen 6500 AH, The Netherlands

The neural mechanism of leader emergence is not well understood. This study investigated (i) whether interpersonal neural synchronization (INS) plays an important role in leader emergence, and (ii) whether INS and leader emergence are associated with the frequency or the quality of communications. Eleven three-member groups were asked to perform a leaderless group discussion (LGD) task, and their brain activities were recorded via functional near infrared spectroscopy (fNIRS)-based hyperscanning. Video recordings of the discussions were coded for leadership and communication. Results showed that the INS for the leader–follower (LF) pairs was higher than that for the follower–follower (FF) pairs in the left temporo-parietal junction (TPJ), an area important for social mentalizing. Although communication frequency was higher for the LF pairs than for the FF pairs, the frequency of leader-initiated and follower-initiated communication did not differ significantly. Moreover, INS for the LF pairs was significantly higher during leader-initiated communication than during follower-initiated communications. In addition, INS for the LF pairs during leader-initiated communication was significantly correlated with the leaders' communication skills and competence, but not their communication frequency. Finally, leadership could be successfully predicted based on INS as well as communication frequency early during the LGD (before half a minute into the task). In sum, this study found that leader emergence was characterized by high-level neural synchronization between the leader and followers and that the quality, rather than the frequency, of communications was associated with synchronization. These results suggest that leaders emerge because they are able to say the right things at the right time.

leader emergence | neural synchronization | babble hypothesis | quality of communication | communication skill

Leadership is a ubiquitous feature of all social species, including human and nonhuman animals (1, 2). However, the neural mechanism of leader emergence is still not well-understood. Evolutionary theories suggest that, whereas both human and nonhuman animals have evolved tendencies to compete for dominance over access to survival-related resources (3–5), human leaders also play an important role in maintaining group cohesion (6). Thus, human leaders need to take into account not only their own needs but also the needs of their followers to facilitate cooperation among group members (7–9). Interestingly, recent imaging evidence indicates that the neural activities of two individuals are more synchronized when they perform a cooperative rather than a competitive task (10). Moreover, the level of interpersonal neural synchronization (INS) is closely associated with the level of understanding between partners (11). It is unknown, however, whether INS is involved in leader emergence.

Previous evidence has shown that communication plays an important role in the increase of INS (12). However, the role of communication in leader emergence has been extensively debated. On the one hand, the so-called "babble" hypothesis postulates that the most talkative member of a group often becomes the group's leader (13, 14). Indeed, there is evidence that the frequency of communication (regardless of its usefulness) is a better predictor of leader emergence than other factors such as the quality of communication (15). It is suggested that communication frequency is probably one of the main factors that increase the probability for initiating group action (16).

On the other hand, various recent studies have suggested that the quality of communication is a more important predictor of leader emergence than is the frequency of communication (17–20). Consistent with this "quality-of-communication" hypothesis, evidence shows that the frequency of communication has no real effect on leader emergence (20). Although the frequency of communication boosts leadership ratings, it does so only when the content is of high quality (17). Furthermore, in task-oriented groups, the quality rather than the quantity of communication is a better predictor of leader emergence (18, 19). Research has also suggested that high-quality communication tends to involve a high level of mentalizing: i.e., the ability to read social situations and to alter one's own behavior to fit in and act appropriately (21). Indeed, communication skills have been considered to be an important part of leader competence in modern societies (22). It is likely that leaders emerge when they possess tactful communication skills and competence: i.e., being able to say the right things at the right time.

Research is needed to investigate how communications are related to INS, which in turn may be related to leader emergence. Considering the interactive nature of leader emergence,

Significance

Great leaders are often great communicators. However, little is known about the neural basis of leader–follower communication. Only recently have neuroscientists been able to examine interpersonal neural synchronization (INS) between leaders and followers during social interactions. Here, we show that INS is significantly higher between leaders and followers than between followers and followers, suggesting that leaders emerge by synchronizing their brain activity with that of the followers. Moreover, the quality rather than frequency of the leaders' communications makes a significant contribution to the increase of INS. This result supports the "quality of communication" hypothesis in leader emergence. Finally, our results show that leadership can be predicted shortly after the onset of a task based on INS as well as communication behaviors.

it is imperative to adopt the "second-person approach": i.e., measuring two or more persons' brain activities simultaneously (23). This approach is also termed "hyperscanning" and has proven to be promising in the field of social neuroscience (23–25). By using an EEG-based hyperscanning approach, recent evidence showed that, during guitar playing, the a priori-assigned leaders showed higher levels of delta-phase locking than did the followers and that INS from the leaders to the followers was stronger than that from the followers to the leaders (26, 27). Evidence further showed implicit synchronization of both body movements and neural activity between a priori-assigned leaders and followers during social interactions (28). However, previous hyperscanning studies did not examine the neural mechanism of leader–follower (LF) communications and did not compare the INS between the LF and the follower–follower (FF) pairs. Thus, it is still unknown whether and how INS is involved in leader emergence. In addition, EEG is sensitive to motor artifacts and suffers from poor spatial resolution. In contrast, functional near infrared spectroscopy (fNIRS) is more tolerant of movements and is able to measure local hemodynamic effect. These advantages make it particularly suitable for testing the role of communication in leader emergence in a realistic situation.

This study examined whether and how INS was involved in leader emergence by using the fNIRS-based hyperscanning approach. During the experiment, three-person groups were recruited to perform a leaderless group discussion (LGD) task. This task has been used successfully in many studies to induce a discussion-oriented, problem-solving situation (19). INS of neural activity was computed. It was hypothesized that INS of the LF pairs would be higher than that of the FF pairs. Based on the babble hypothesis, it was expected that (*i*) leaders would initiate more communications than the followers and (*ii*) the increased INS for the LF pairs would be mainly due to the emerging leaders' communication frequency and would occur in language-related brain areas. Alternatively, based on the quality-of-communication hypothesis, leaders would not initiate more communications than the followers, and INS for the LF pairs would be associated with the emerging leaders' communication skills and competence, rather than the frequency, and would occur in brain areas associated with social mentalizing. Finally, using a Fisher linear discrimination analysis, we investigated how early during the LGD session the INS data and communication behaviors could predict the emergence of leaders.

Results

Interpersonal Neural Synchronization. The experimental setup is illustrated in Fig. 1*A*. For each session, three participants sat face-to-face in a triangle and were given a topic for an LGD (see *Materials and Methods* for details). Their brain activities were simultaneously recorded with an fNIRS system (Fig. 1*B*). The discussion was video-taped and coded by independent judges for leadership, communication skills and competence, initiation of communications, and frequencies of verbal and nonverbal communications.

For the LF pairs, a significant INS increase compared with the resting-state condition was identified at channel 6 (CH6), which roughly covered the left temporo-parietal junction (TPJ) [$t_{(10)} = 4.62$, $P = 0.001$, false discovery rate (FDR) correction] (Fig. 1*C*). No INS increase was found for any channel of the FF pairs (Fig. 1*D*). Group differences between the LF and FF pairs were significant for CH6 ($t_{(20)} = 3.51$, $P = 0.002$), but not for any other CHs.

To validate that the above results could not have been obtained by chance, we assessed the likelihood of obtaining significant INS increases for any random pairings of the participants. Specifically, we reanalyzed the data after randomizing the LF pairing both within and between discussion groups. The first was the within-group permutation: Each of the two followers was assigned to be the "leader" and the INS data were reanalyzed. The second approach was the between-group permutation: All 33 participants were randomly assigned to 11 three-member
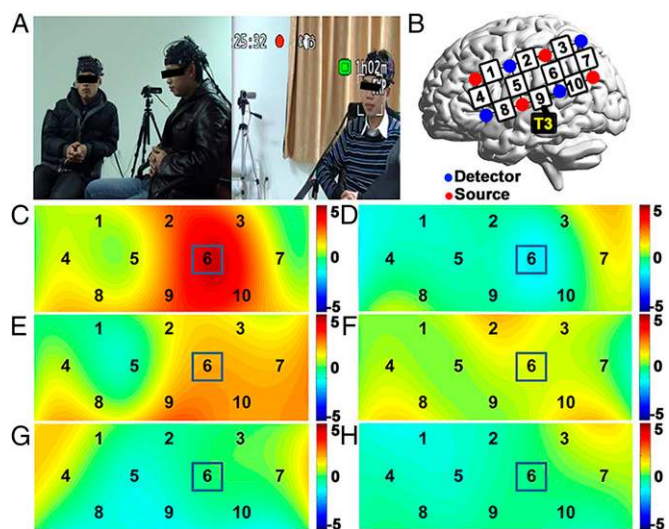

Fig. 1. Experimental procedure and the increase of interpersonal neural synchronization (INS). (*A*) For each group, three persons sat in a triangle. Two cameras were placed in opposite positions. The figure shows two sample frames from the cameras in the opposite directions. Participants were asked to discuss a topic for 5 min and then to choose a leader to report their conclusion. (*B*) The optode probe set was placed on the left frontal, temporal, and parietal cortices. T3 corresponds to a position in the international 10–20 system. (*C* and *D*) Shown are *t* maps for results of the original pairs (i.e., real data). (*E* and *F*) Shown are *t* maps for the permutation results of pairs with a follower from the same group randomly assigned as the leader. (*G* and *H*) Shown are *t* maps for the permutation results of randomized pairs from across groups. [*C, E,* and *G* are *t* maps for averaged leader–follower (LF) pairs; *D, F,* and *H* are *t* maps for the follower–follower (FF) pairs.]

groups, and the INS analysis was then reconducted. This permutation was conducted 1,000 times. Both approaches showed no significant INS increases for any CHs. Fig. 1 *E* and *F* and *G* and *H* shows the results of typical within- and between-group validation analyses, respectively. Complete results for CH6 from 1,000 permutations of between-group validation analyses are shown in Fig. S1. These results suggested that the significant INS increase in the left TPJ was specific to the particular LF relationship in the LGD context.

Who Synchronized with Whom? Granger causality analysis (GCA) was conducted on the time series of CH6 to determine whether it was the leader who synchronized with the followers or whether it was the other way around. One-sample *t* tests on the pairwise-conditional causalities showed that the mean causalities of both directions were significantly higher than zero: from the leaders to the followers ($t_{(10)} = 10.001$, $P < 0.001$) and from the followers to the leaders ($t_{(10)} = 7.272$, $P < 0.001$). However, two-sample *t* test showed that the mean causality from the leaders to the followers was significantly higher than that from the followers to the leaders ($t_{(10)} = 2.177$, $P = 0.027$). These results indicated a more important role of the leaders than the followers in the INS increase in the LF pairs at CH6.

Communication Behaviors and INS. Both verbal and nonverbal communication frequencies were significantly higher for the LF pairs than for the FF pairs: ($t_{(20)} = 3.873$, $P = 0.001$) for verbal and ($t_{(20)} = 4.565$, $P < 0.001$) for nonverbal communications (Fig. 2*A*). However, the leaders did not differ significantly from the followers in the frequency of communication initiation ($t_{(10)} = -1.602$, $P = 0.125$). To investigate whether the role of leaders in communication initiation might have changed as the discussion progressed, the initiation data were reanalyzed by the first and the second halves of the LGD session. Still, no differences were found between the LF and FF pairs: ($t_{(10)} = -0.433$, $P = 0.674$)

for the first half and ($t_{(10)} = -0.858$, $P = 0.411$) for the second half of the LGD session (Fig. 2B). These results suggested that, although communication frequency was higher for the LF pairs than for the FF pairs, leaders and followers contributed equally throughout the LGD session.

We next examined INS that accompanied different types of communications (verbal, nonverbal, and no communications). For the LF pairs, INS during verbal communications (INS-V) differed significantly from both INS during nonverbal communication (INS-NV) ($t_{(10)} = 2.951$, $P = 0.015$) and INS when no communications occurred (INS-NC) ($t_{(10)} = 2.758$, $P = 0.02$) (Fig. 2C). Fig. 3 shows the correspondence between INS (coherence value) and video frame for a typical LF pair at CH6. No significant results were found for the FF pairs. Group difference in INS-V between the LF and FF pairs was also significant ($t_{(20)} = 3.178$, $P = 0.005$). No significant group differences were found for INS-NV ($t_{(20)} = -0.24$, $P = 0.813$) and INS-NC ($t_{(20)} = 0.982$, $P = 0.338$). These results indicated that the INS difference was specific for verbal communication between the leaders and the followers.

In terms of the role of communication initiation, leader-initiated communications induced a higher level of INS than the ones initiated by the followers ($t_{(20)} = 2.176$, $P = 0.042$) (Fig. 2D). This result suggested that leader-initiated communications were likely to be of higher quality (and thus led to increased INS). This conjecture was further supported by two other results. First, leaders' communication skills and competence were more highly rated (M = 25.279, SD = 0.800) than those of the followers (M = 22.020, SD = 1.112) ($t_{(20)} = 7.894$, $P < 0.001$) (Fig. 2E). Second, there was a significant correlation between INS during leader-initiated communications and judge-rated leaders' communication skills and competence ($r = 0.697$, $P = 0.017$) (Fig. 2F). The correlation between INS during leader-initiated communications

Fig. 3. The correspondence between INS at CH6 and coded communication behaviors. (A) A time course of INS for one randomly selected LF pair. (B) The corresponding communication behaviors coded from video frames. Blue points, follower-initiated verbal communications; green points, nonverbal communications; red points, leader-initiated verbal communications. The sections of the line without color points represent no communications. The numbers 1, 2, and 3 in A highlight time points that correspond to video-frame examples in B.

and the leaders' initiation frequency was not significant ($r = 0.247$, $P = 0.465$). This difference in correlation coefficients was in favor of the quality-of-communication hypothesis over the babble hypothesis although a larger sample of leaders would be needed to allow for a statistical test of the difference.

Prediction of Leadership. To investigate how early the leaders emerged during the LGD, Fisher linear discrimination analyses were conducted. Fig. 4A shows the time course of the prediction accuracy in the discriminant analysis based on the INS data, which differentiated the LF pairs from the FF pairs. The analysis included three indexes: sensitivity (percentage of LF pairs correctly predicted, red line), specificity (percentage of FF pairs correctly predicted, blue line), and the generalization rate of accuracy (overall proportions of LF and FF pairs correctly predicted, green line). A moving-window analysis (window size = 9 s) revealed that the prediction accuracy was sporadic during the initial period, but the prediction accuracy of all three indexes was stably higher than the chance level starting at 23 s ($P < 0.05$, corrected by FDR) [see the purple section above the chance-level (0.50) line in Fig. 4A]. A similar discriminant analysis was conducted based on the communication frequency (Fig. 4B). The results showed that the prediction accuracy of all three indexes was stably higher than the chance level starting at 29 s ($P < 0.05$, corrected by FDR) (see the purple section above the chance-level line in Fig. 4B). In sum, the INS and communication frequency data were able to discriminate the leaders from the followers less than half a minute into the LGD task.

Discussion

This study used an fNIRS-based hyperscanning approach to test the hypothesis that INS was involved in leader emergence. The results demonstrated that INS increased from the baseline more significantly for the LF pairs than for the FF pairs. Further analysis revealed that, although the communication initiation frequency of leaders and followers did not differ significantly, leader-initiated communication induced greater INS than did follower-initiated communication. The INS increase during leader-initiated communications was also associated with leaders' communication skills and competence. These results suggest that quality rather than quantity (or frequency) of communication is more important in leader emergence. These results are discussed sequentially below.

First, results of this study confirmed our hypothesis that the LF relationship in the LGD context would be characterized by a high level of INS. We derived our hypothesis from integrating
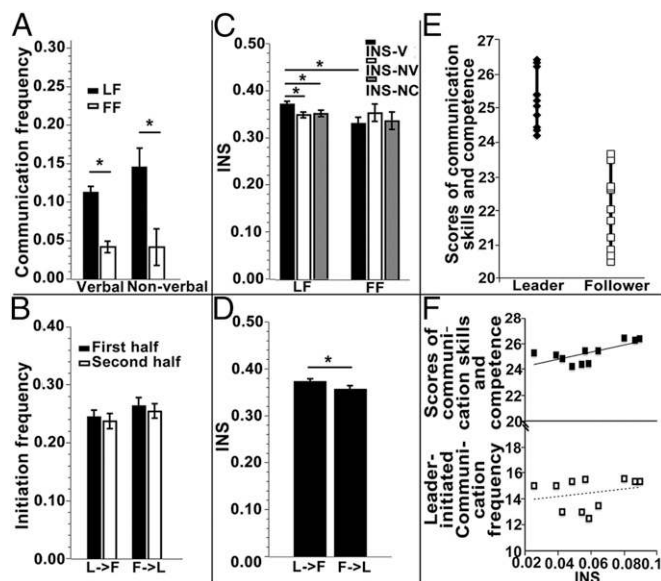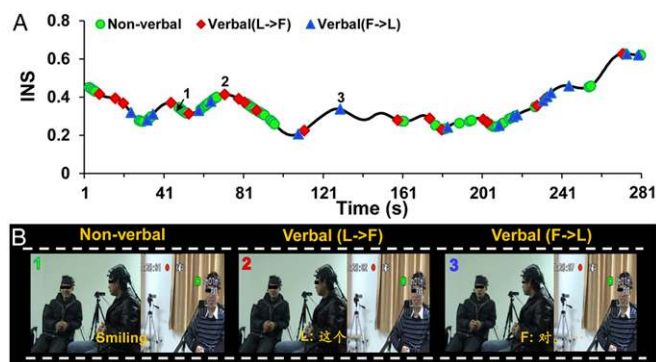

Fig. 2. (A) Verbal and nonverbal communication frequencies during the task. The averaged frequency of the two leader–follower (LF) pairs (black) was higher than the frequency of the follower–follower (FF) pairs (white). (B) There were no significant differences in leader-initiated (L→F) vs. follower-initiated (F→L) verbal communications. (C) LF pairs' INS during verbal communication (INS-V) was higher than INS for all other situations. NC, no communication occurred; NV, nonverbal communication; V, verbal communication. (D) INS during leader-initiated communication was higher than that during follower-initiated communication. (E) Leaders' communication skills and competence were more highly rated than those of the followers. (F) INS during leader-initiated communication was positively associated with ratings of communication skills and competence (Upper), but not with leader-initiated communication frequency (Lower). *$P < 0.05$.
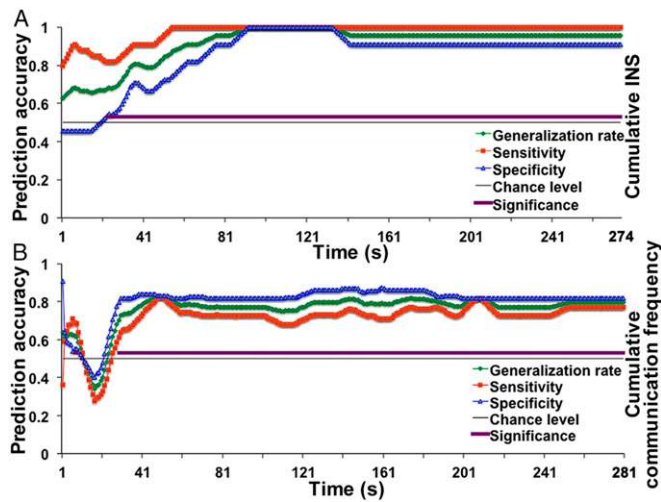
Fig. 4. Time course of prediction accuracy. (*A*) Prediction results based on the cumulative INS data. (*B*) Prediction results based on cumulative communication frequency. There were a total of 274 time points for *A* after shifting 6 s toward the left due to fNIRS signal delay (*Materials and Methods*) and 280 time points for *B*. The time courses were smoothed by using a moving average method (span = 9 s). The purple line above the chance-level line indicates the time points where all three accuracy indexes were significantly higher than the chance level (0.50).

recent imaging evidence that cooperation between persons led to a high level of INS (10, 28) with recent perspectives about human leaders' role as the coordinators who help their groups to solve various tasks, including resource sharing and decision making (9, 29). According to the service-for-prestige theory of leadership (9), human leaders and followers are involved in reciprocal exchange: Leaders may incur costs to provide followers with public goods, and, in return, followers incur costs to provide leaders with prestige, particularly in a relatively small group. We interpret the higher INS for the LF pairs as a reflection of their closer cooperation and social exchange.

Second, we found that the level of INS was increased specifically during verbal communications between the leaders and followers, not during nonverbal or no communications, nor for any type of communications involving the FF pairs. This result was consistent with previous studies showing that verbal communication was one of the main factors that affected leader emergence (13, 17–19). The present results further suggest that verbal communication affects leader emergence by modulating the neural synchronization. Because of the importance of verbal communications in INS, this particular route of leader emergence may be specific to humans (e.g., the service-for-prestige theory) (9). Nonhuman animals typically establish leadership via dominance (e.g., displays of physical strength), so it would be interesting to investigate whether they also show INS.

Third, although the leaders and followers contributed equal numbers of communications, leader-initiated verbal communications were found to lead to higher INS than did follower-initiated ones. Moreover, the GCA results showed that INS was bidirectional but was significantly stronger from the leaders to the followers than the other direction. These results suggested that dynamic social interactions played an important role in leader emergence. Indeed, as Schilbach et al. (23) suggested, dynamic social interaction is a key constituent of grasping the minds of others. An action by an "initiator" may lead to closer monitoring of the outcome of the interaction, including the responses by other individuals (23). In our study, the leaders initiated the communications, monitored the followers' responses, and closely synchronized their brain activities with those of the followers. This speculation was further supported by the significant correlation between communication skills and competence

and INS. It seems that a leader is someone who would say the right things at the right time to increase neural synchronization with the followers.

Fourth, the increased INS for the LF pairs was found in the left TPJ, but not in the language area [i.e., left inferior frontal cortex (IFC)]. This result was consistent with previous evidence that high quality of communication is associated with high-level mentalizing (21), which was partly subserved by the left TPJ. Specifically, previous evidence has shown that interpersonal coordination or communication is facilitated by the mutual abilities to predict each other's subsequent action (i.e., high-level mentalizing) (30). Researchers have debated about which specific parts of the left or right TPJ or both are involved in mentalizing and understanding and reasoning about the beliefs and intentions of others (31–33). In one study, a lesion in the left TPJ was found to affect the representations of someone else's beliefs (33). In another study, the posterior part of the right TPJ and the parietal cortex were found to be involved in social cognition and memory retrieval whereas the anterior part of the right TPJ as well as the motor cortex and insula were involved in attention (32). Although the poor spatial resolution of fNIRS did not allow us to precisely locate the position of the INS increase, the most likely area would be the posterior part of the left TPJ (for high-level mentalizing) because no motor cortex was involved in this study.

Finally, discriminant analyses showed that, shortly after the start of the LGD task, the INS data and communication behaviors could successfully distinguish the LF from the FF pairs. These results further supported the quality-of-communication hypothesis by suggesting that the communication frequency matters when the quality is of high level (17). These results also confirmed previous findings (26–28, 34) that neural activity (as well as interactive communication behaviors) could be used to differentiate reliably the leaders from the followers. It is worth noting that different studies have found different earliest time points for successful discrimination based on neural activity: before the onset of the interactions in Sänger et al. (26, 27) and Konvalinka et al. (34) and about half a minute into the interaction in our study. One possible explanation of these variations is that the time point for successful discrimination depends on how the leaders emerge. In Sänger et al. (26, 27), leaders were assigned a priori; in Konvalinka et al. (34), leaders emerged through a number of repeated trials; and, in the present study, leaders emerged during a single LGD task. Future research should specifically examine the role of neural activity or INS in predicting different types of leader emergence.

Several limitations of this study need to be noted. First, our findings from the LGD task may not be generalized to other types of situations for leader emergence. The process of leader emergence from a free discussion among equals (all college students) may be different from one involving members who are of different ages, genders, social status, etc. In addition, the phenomenon of INS may also be different for leader emergence than for situations with a leader assigned a priori, as discussed earlier. Second, our sample size was adequate for the examination of group differences, but not as satisfactory for individual differences in leaders. Similarly, the statistical power was limited when we tested the babble hypothesis because of both the small sample size and the somewhat limited verbal behaviors from the short period of the LGD task. Third, we did not measure other important characteristics of leadership, such as charisma (35), which should be considered in future research for their role in INS. Finally, because of the poor spatial resolution of fNIRS, it was difficult to identify exactly which brain areas were responsible for the responses at CH6.

In summary, leadership is an important feature of human society, but little is known about the neural basis of leader emergence. Using the fNIRS-based hyperscanning approach in a realistic interpersonal-communication context, the current study found evidence that human leaders cooperated with their followers to achieve group decision by synchronizing their brain

PSYCHOLOGICAL AND COGNITIVE SCIENCES

activities with those of the followers through their tactful communication skills and competence. We further found that it was possible to predict leadership based on the INS data as well as communication behaviors early in their interactions. These findings contribute to the theoretical discussion about the importance of communications in leader emergence and advance our understanding of the neural mechanism of leader emergence. The results also potentially may be used in neuro-feedback or neuro-intervention during leadership training.

Materials and Methods

Participants. Thirty-six healthy adults (mean age 22 ± 2 y) participated in this study. They were pseudorandomly split into 12 three-person groups. For each group, the members had to be of the same sex (to avoid a potential confound of intergender interactions) and were total strangers to one another. There were 6 female groups and 6 male groups. One female group was excluded because of data collection failure.

Written informed consent was obtained from all participants. The study protocol was approved by the Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University.

Tasks and Procedures. For each group, an initial resting-state session of 5 min served as the baseline. During this session, the participants were required to remain as motionless as possible with their eyes closed and mind relaxed (36).

After the resting-state session, each group was instructed to perform the LGD task. Two additional 30-s resting-state periods (one at the initial phase and the other at the ending phase of the LGD) were used to allow the imaging instrument to reach a steady state.

During the LGD, the three participants of each group sat face-to-face in a triangle. Two digital video cameras were placed at opposite positions so that all three participants could be recorded (see Fig. 1A for two sample frames). Participants received the following topic for discussion: "An airplane crash-landed on a deserted island. Only 6 persons survived: A pregnant woman, an inventor, a doctor, an astronaut, an ecologist, and a vagrant. Whom do you think should be given the only one-person hot-air balloon to leave the island?" The participants were asked to read and think about the topic for 5 min without interacting with one another. Afterward, each group was instructed to discuss the topic for 5 min. Each group was then required to choose a member to report their conclusion to the experimenter. The reporting session lasted 1 min. The whole procedure was video recorded for subsequent coding.

Determination of the Leaders and Evaluation of Communication Skills and Competence. After the experiment, an additional group of eight graduate students was recruited to view the video recordings of the discussion session and to judge who the leader was for each group. Judges were asked to use their own criteria to make the judgment. For each group, the member with votes from more than half of eight judges was defined as the leader. The average vote for the leaders was $77.3 \pm 15.6\%$. The intraclass reliability (ICC) among judges was 0.874 ($P < 0.001$). For 9 of the 11 groups, the judges' choice of the leader agreed with the group members' own choice (i.e., the person who gave the report). For subsequent analyses, we used the more objective choices by the judges.

Judges were also asked to evaluate the communication skills and competence of each group member on a 5-point scale (Table S1). There were seven aspects of communication skills and competence (group coordination, active participation, new perspectives, input quality, logic and analytic ability, verbal communication, and nonverbal communication). Judges were given explanations of the above categories and a scoring guide (see Table S1 for details). Interjudge reliability was determined by ICC, and it was satisfactory to high (ranging from 0.773 to 0.926) for all but one item (new perspectives, ICC = 0.412). Possible reasons for the judges' lack of consensus on "new perspectives" might be the low frequencies of relevant behavior or ambiguity of this construct. This item was removed from further analyses. For the remaining items, ratings from the eight judges were averaged for each item. The final scale of communication skills and competence included six items with high internal consistency (Cronbach alpha = 0.930).

Coding of Communication Behaviors. Two additional coders, who were not involved in the voting of leaders and the evaluations of communication skills and competence, coded communication behaviors. We used new coders to avoid the leader voting's potential contamination of behavior coding. Communication behaviors included verbal communications, such as turn-taking and interjections, and nonverbal communications, such as orofacial

movements, facial expressions, and sign gestures. Each of the 280 s during the LGD was coded as having either verbal communication, nonverbal communication, or no communications. If both verbal and nonverbal behaviors occurred for a given second, the dominant behavior was coded. The frequencies of verbal and nonverbal communications were calculated as the proportions of time (out of the 280 s) when verbal and nonverbal communications occurred, respectively. The intercoder reliability (based on ICC) was 0.930 for verbal communications (vs. no communications) and 0.952 for nonverbal communications (vs. no communications).

In addition, the initiator of each occurrence of verbal communication was also coded. The frequency of initiations for each member was calculated as the ratio of time points where a member initiated a communication over the total number of that member's verbal communications (ICC = 0.949).

FNIRS Data Acquisition. During the experiment, the participants sat in a quiet room. An ETG-4000 optical topography system (Hitachi Medical Company) was used to collect imaging data from the three participants of each group simultaneously. Three sets of the same customized optode probes were used. The probe was placed on the left hemisphere so as to cover both the left inferior frontal cortex (an area important for language) (37) and the temporal-parietal junction (TPJ) (an area closely associated with social mentalizing) (31, 33).

The optode probes consisted of 10 measurement channels (four emitters and four detectors, 30 mm optode separation). CH9 was placed just at T3 in accordance with the international 10–20 system (Fig. 1B). The probe set was examined and adjusted to ensure consistency of the positions among the participants of each group and across the groups.

The absorption of near-infrared light at two wavelengths (695 and 830 nm) was measured with a sampling rate of 10 Hz. The changes in the oxy-hemoglobin (HbO) and deoxy-hemoglobin (HbR) concentrations were recorded in each channel based on the modified Beer–Lambert law. This study focused only on the changes in the HbO concentration, which was demonstrated to be the most sensitive indicator of changes in the regional cerebral blood flow in fNIRS measurements (38).

Imaging-Data Analysis.

Interpersonal neural synchronization. Data collected during the resting-state and LGD sessions were entered into the analysis. During preprocessing, data in the initial and ending periods (30 s resting state plus 10 s LGD, respectively) were removed, leaving 280 s of data for each session. Wavelet transform coherence (WTC) was used to assess the cross-correlation between two fNIRS time series generated by pairs of participants as a function of frequency and time (39). The wavelet coherence MatLab package was used (40) [for more thorough information, please see Grinsted et al. (40) and Chang and Glover (41)]. Briefly, three HbO time series were obtained simultaneously for each CH from the three participants of each group. WTC was applied to each pair of the time series to generate 2D coherence maps. According to previous studies (10, 12), the coherence value increases when there are interactions between persons, compared with that during the resting state. Based on the same rationale, the average coherence value between 0.02 and 0.2 Hz was calculated. This frequency band also excluded the high- and low-frequency noises, such as those associated with respiration (about 0.2–0.3 Hz) and cardiac pulsation (about 1 Hz), all of which would lead to artificial coherence. Finally, the coherence value was time-averaged.

The averaged coherence value of the resting-state session was subtracted from that of the LGD session, and the difference was used as an index of the INS increase between two persons. Because each group had two LF pairs and only one FF pair, the INS increases for the two LF pairs were averaged for matched-sample t tests (SI Text and Fig. S2). For each channel, after converting the INS increase into a z value, a one-sample t test was performed on the z value across the participant pairs, and two t maps of the INS increase ($P < 0.05$, corrected by FDR) were generated, one for the LF pairs and the other for the FF pairs. The t maps were smoothed using the spline method.

Validation by randomizing the data. To verify that the INS increase was specific to the LF relationship that emerged during the LGD, two validation approaches were applied. The first was the within-group permutation: Each of the two followers was assigned to be the "leader," and the INS data were reanalyzed. The second approach was the between-group permutation: All 33 participants were randomly assigned to 11 three-member groups, and the INS analysis was then reconducted. This permutation was conducted 1,000 times.

Who synchronized with whom? For CHs that showed significant INS increases, GCA was conducted to determine the direction of synchronization (i.e., whether it was the leaders who synchronized with the followers or the other way around). GCA is a method that uses vector autoregressive models to measure the causal relationship (i.e., pairwise-conditional causalities from the

source to the target) between time series such as the fNIRS data (42). We computed the pairwise-conditional causalities of both directions: from the leaders to the followers and from the followers to the leaders. These two causality indices were statistically tested to see whether they differed from zero and from each other.

Communication Behaviors and INS. To confirm the contribution of communication to the INS increase during the LGD, the CHs that showed significantly greater INS increases for the LF pairs than for the FF pairs were selected. First, the time courses of INS in the selected CHs were downsampled to 1 Hz to obtain point-to-frame correspondence between the signal's time course and video recordings. Second, the time points of the video were marked as having either verbal or nonverbal or no communications. Third, the corresponding INSs were separately averaged to obtain three indexes: i.e., INS-V, INS-NV, and INS-NC, for INS during verbal, nonverbal, and no communications, respectively. The INS data were adjusted for the delay-to-peak effect in the fNIRS signal (about 6 s) (43). Finally, these indexes were statistically compared for the LF and FF pairs separately (using a paired two-sample *t* test), as well as between the LF and FF pairs (using an independent two-sample *t* test).

To examine the role of the leaders, further analyses were conducted to clarify whether the results were driven by leader-initiated or follower-initiated communications and whether the increase of INS was associated with the leaders' communication skills and competence or communication frequency. The results were threshholded at $P < 0.05$ level (FDR-corrected).

Prediction of Leadership. The time course of INS for the LF and FF pairs during the LGD session was baseline-corrected by subtracting their respective averaged INS during the resting state. Cumulative INS across the time was calculated and then used as the neural-classification feature to classify the LF and FF pairs: i.e., the type of relationship (i.e., LF or FL) was the classification label. The cumulative INS at time point n was computed as a sum of the INS at time points from 1 to $n − 1$. The discriminant analysis was conducted for each time point. A leave-one-out cross-validation method was used to obtain the prediction accuracy. Time courses were generated for three indexes of prediction accuracy: sensitivity, specificity, and the generalization rate of accuracy. Because the fNIRS signal needs 6 s to reach a peak value after the presentation of a stimulus (43), the recorded time points did not match the brain-activity time points (or behavioral time points, such as communications). To adjust for the delay, we deleted the first 6 time points, yielding a total of 274 time points for INS. Then, a moving window of 9 s was used to identify the time points when the prediction accuracy differed significantly from a chance level (0.50). Similar analyses were conducted based on the communication frequency at each time point. Finally, a moving average method (span = 9 s) was used to smooth the time courses of prediction accuracy.

The prediction results based on moment-to-moment INS data and communication frequency are provided in Fig. S3, which suggested that the cumulative data provided more stable prediction accuracy than the moment-to-moment data.

1. King AJ, Johnson DD, van Vugt M (2009) The origins and evolution of leadership. *Curr Biol* 19(19):R911–R916.
2. van Vugt M, Ahuja A (2010) *Naturally Selected: The Evolutionary Science of Leadership* (HarperBusiness, New York).
3. Dyer JRG, Johansson A, Helbing D, Couzin ID, Krause J (2009) Leadership, consensus decision making and collective behaviour in humans. *Philos Trans R Soc Lond B Biol Sci* 364(1518):781–789.
4. Johnstone RA, Manica A (2011) Evolution of personality differences in leadership. *Proc Natl Acad Sci USA* 108(20):8373–8378.
5. Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. *Nature* 433(7025):513–516.
6. van Vugt M (2014) On faces, gazes, votes, and followers: Evolutionary psychological and social neuroscience approaches to leadership. *New Frontiers in Social Neuroscience*, Research and Perspectives in Neurosciences, eds Decety J, Christen Y (Springer, Cham, Switzerland), Vol 21, pp 93–110.
7. Krajbich I, Camerer C, Ledyard J, Rangel A (2009) Using neural measures of economic value to solve the public goods free-rider problem. *Science* 326(5952):596–599.
8. Gillet J, Cartwright E, van Vugt M (2011) Selfish or servant leadership? Evolutionary predictions on leadership personalities in coordination games. *Pers Individ Dif* 51(3):231–236.
9. Price ME, van Vugt M (2014) The evolution of leader-follower reciprocity: The theory of service-for-prestige. *Front Hum Neurosci* 8:363.
10. Cui X, Bryant DM, Reiss AL (2012) NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59(3):2430–2437.
11. Stephens GJ, Silbert LJ, Hasson U (2010) Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci USA* 107(32):14425–14430.
12. Jiang J, et al. (2012) Neural synchronization during face-to-face communication. *J Neurosci* 32(45):16064–16069.
13. Mullen B, Salas E, Driskell J (1989) Salience, motivation, and artifact as contributors to the relation between participation rate and leadership. *J Exp Soc Psychol* 25:545–559.
14. Wickham K, Walther J (2007) Perceived behaviors of emergent and assigned leaders in virtual groups. *Int J e-Collaboration* 3(1):1–17.
15. Sorrentino RM, Boutillier RG (1975) The effect of quantity and quality of verbal interaction on ratings of leadership ability. *J Exp Soc Psychol* 11:403–411.
16. van Vugt M (2006) Evolutionary origins of leadership and followership. *Pers Soc Psychol Rev* 10(4):354–371.
17. Jones EE, Kelly JR (2007) Contributions to a group discussion and perception of leadership: Does quantity always count more than quality? *Group Dyn* 11(1):15–30.
18. Balthazard PA, Waldman DA, Warren JE (2009) Predictors of the emergence of transformational leadership in virtual decision teams. *Leadersh Q* 20(5):651–663.
19. Riggio RE, Riggio HR, Salinas C, Cole EJ (2003) The role of social and emotional communication skills in leader emergence and effectiveness. *Group Dyn* 7(2):83–103.
20. Sarker S, Grewal R, Sarker S (2002) Emergence of leaders in virtual teams: What matters? *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (IEEE, Los Alamitos, CA).
21. Dobbins GH, Long WS, Dedrick EJ, Clemons TC (1990) The role of self-monitoring and gender on leader emergence: A laboratory and field study. *J Manage* 16(3):609–618.
22. Price ME, van Vugt M (2014) The service-for-prestige theory of leader–follower relations: A review of the evolutionary psychology and anthropology literatures.

Biological Foundations of Organizational Behavior, eds Colarelli SM, Arvey RD (Univ of Chicago Press, Chicago), pp 169–201.
23. Schilbach L, et al. (2013) Toward a second-person neuroscience. *Behav Brain Sci* 36(4):393–414.
24. Babiloni F, Astolfi L (2014) Social neuroscience and hyperscanning techniques: Past, present and future. *Neurosci Biobehav Rev* 44:76–93.
25. Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C (2012) Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends Cogn Sci* 16(2):114–121.
26. Sänger J, Müller V, Lindenberger U (2012) Intra- and interbrain synchronization and network properties when playing guitar in duets. *Front Hum Neurosci* 6:312.
27. Sänger J, Müller V, Lindenberger U (2013) Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Front Hum Neurosci* 7:234.
28. Yun K, Watanabe K, Shimojo S (2012) Interpersonal body and neural synchronization as a marker of implicit social interaction. *Sci Rep* 2:959.
29. Liden RC, Wayne SJ, Zhao H, Henderson D (2008) Servant leadership: Development of a multidimensional measure and multi-level assessment. *Leadersh Q* 19(2):161–177.
30. Konvalinka I, Vuust P, Roepstorff A, Frith CD (2010) Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Q J Exp Psychol (Hove)* 63(11):2220–2230.
31. Saxe R (2006) Uniquely human social cognition. *Curr Opin Neurobiol* 16(2):235–239.
32. Bzdok D, et al. (2013) Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. *Neuroimage* 81:381–392.
33. Samson D, Apperly IA, Chiavarino C, Humphreys GW (2004) Left temporoparietal junction is necessary for representing someone else's belief. *Nat Neurosci* 7(5):499–500.
34. Konvalinka I, et al. (2014) Frontal alpha oscillations distinguish leaders from followers: Multivariate decoding of mutually interacting brains. *Neuroimage* 94:79–88.
35. Dinh JE, et al. (2014) Leadership theory and research in the new millennium: Current theoretical trends and changing perspectives. *Leadersh Q* 25(1):36–62.
36. Lu C-M, et al. (2010) Use of fNIRS to assess resting state functional connectivity. *J Neurosci Methods* 186(2):242–249.
37. Holland R, et al. (2011) Speech facilitation by left inferior frontal cortex stimulation. *Curr Biol* 21(16):1403–1407.
38. Hoshi Y (2007) Functional near-infrared spectroscopy: Current status and future prospects. *J Biomed Opt* 12(6):062106.
39. Torrence C, Compo GP (1998) A practical guide to wavelet analysis. *Bull Am Meteorol Soc* 79(1):61–78.
40. Grinsted A, Moore J, Jevrejeva S (2004) Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process Geophys* 11(5/6):561–566.
41. Chang C, Glover GH (2010) Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50(1):81–98.
42. Im C-H, et al. (2010) Estimation of directional coupling between cortical areas using near-infrared spectroscopy (NIRS). *Opt Express* 18(6):5730–5739.
43. Cui X, Stetson C, Montague PR, Eagleman DM (2009) Ready...go: Amplitude of the FMRI signal encodes expectation of cue arrival time. *PLoS Biol* 7(8):e1000167.