The Effect of Functional Diversity on Team Creativity - Behavioral and fNIRS Evidence

Yasheng Chen[*], Adam Presslee[†], Sue Yang[‡]
* School of Management, Xiamen University, China † School of Accounting and Finance, University of Waterloo, Canada ‡ Lyons School of Accounting, Texas Tech University, US

Content:

Online Supplemental Appendix A: Participant Instructions

Study Overview

Thank you for attending this study. You have been recruited for this study because you are trained to be **a business professional [an engineer]**, and business professionals [engineers] pay special attention to the **practical, cost-effective nature [novel, unique nature]** of a project.

Participants in today's session have been assigned to work in a group with one other person (i.e., a two-person group) on a proposal for a real on-campus project. Thus, your proposals will help the University plan for the related project. The other person in your group has been recruited because he/she is trained to be [**a business professional/ an engineer]**. Group members will communicate with one another solely through Zoom video meeting. There is no internet browsing during today's session and Zoom meetings will be recorded.

Participants must keep all the information in this study confidential. Please wait until the administrator gives you additional instructions.

Task Description

Your group's task in today's session is to submit a proposal for the **creative use (i.e., a use that is original, innovative, practical, and implementable within a reasonable budget)** of the top floor (7th floor) of The [ ] Building (The Management School) at [ ] University. The mission statement of the Management School of [ ] University is "[ ]" The top floor (7th floor) of The [ ] Building has around 8,450 sqft (65*130 ft) that is unfurnished and unoccupied, and the budget for the project is $ 240,000.

Please note that throughout the process, you and the other person in your group can **freely express your opinions through Zoom. You may speak as little or as much as you like, and you do not need to wait for the other person in your group to speak.**

Your group's task is broken out into three parts. In Part 1, your group will spend 15 minutes developing specific proposals for the creative use of the space. In Part 2, your group will then spend 5 minutes (1) deciding which proposal to submit as the most creative, and (2) discussing why the submitted proposal is a creative use of the campus space. In Part 3, your group will spend 1 minute to summarize and submit the proposal selected in Part 2 in Zoom Chat. The administrator will give you instructions at the beginning of each part.

Additional Compensation

A panel of professionals and doctoral students will evaluate the creativity of each group's submitted proposal and assign it a rating. In addition to the flat show-up fee paid to you, the group receiving the highest group creativity rating out of all groups will receive $78 reward, which will be evenly divided among the two group members.

Online Supplemental Appendix B: The MNI Coordinates of CHs of a Typical Participant

CHs	MNI coordinates (mm)			Brodmann Areas (Percentage)
	x	y	z	
CH1	*73*	*-25*	*-5*	21-Middle temporal gyrus (96%)
CH2	*69*	*-50*	*0*	37-Fusiform gyrus (46%)
CH3	*53*	*-78*	*-5*	19-V3 (83%)
CH4	*71*	*-13*	*8*	22-Superior temporal gyrus (92%)
CH5	*71*	*-39*	*12*	22-Superior temporal gyrus (98%)
CH6	*62*	*-62*	*14*	37-Fusiform gyrus (54%)
CH7	*46*	*-87*	*6*	19-V3 (76%)
CH8	*70*	*-23*	*26*	2-Primary somatosensory cortex (68%)
CH9	*66*	*-50*	*27*	22-Superior temporal gyrus (61%)
CH10	*53*	*-75*	*24*	39-Angular gyrus (84%)
CH11	*68*	*-6*	*32*	43-Subcentral area (72%)
CH12	*68*	*-34*	*40*	40-Supramarginal gyrus (73%)
CH13	*59*	*-60*	*38*	39-Angular gyrus (84%)
CH14	*43*	*-81*	*35*	19-V3 (70%)
CH15	*65*	*-20*	*46*	1-Primary somatosensory cortex (57%)
CH16	*61*	*-43*	*50*	40-Supramarginal gyrus (100%)
CH17	*48*	*-68*	*49*	39-Angular gyrus (82%)
CH18	*57*	*-2*	*51*	6-Pre-motor cortex (86%)
CH19	*60*	*-29*	*54*	1-Primary somatosensory cortex (50%)
CH20	*51*	*-54*	*56*	40-Supramarginal gyrus (84%)
CH21	*31*	*-70*	*56*	7-Somatosensory association cortex (100%)
CH22	*52*	*-11*	*59*	4-Pre-motor cortex (54%)
CH23	*45*	*-34*	*66*	3-Primary somatosensory cortex (42%)
CH24	*32*	*-59*	*68*	7-Somatosensory association cortex (96%)

Online Supplemental Appendix C: IBS

C1. IBS, IBS Increments, and How They Are Calculated

IBS fundamentally represents the simultaneous patterns of brain oxygen levels between two or more individuals. The analysis of brain oxygen levels includes two critical aspects: the temporal sequence (*time*) and the occurrence rate of repetitive patterns (*frequency*) (Buzsáki and Draguhn 2004, Cohen 2014, Cui et al. 2012). First, *Time* in this context refers to the progression of data points in a sequence. For example, if daily temperature readings were documented for a year, the "time" aspect refers to the days over which the readings were taken. Second, *Frequency* alludes to the repetition rate of the data within a given time frame, typically measured in cycles per second or Hertz (Hz). Higher frequencies indicate rapid repetition and lower frequencies denote fewer repetition (Cohen 2014).

The Wavelet Transform Coherence (WTC) technique is utilized for assessing the correlation of both *time* and *frequency* of multiple time-series. Specifically, WTC can be dissected into two foundational concepts: (1) Wavelet Transform, and (2) Coherence (Lachaux et al. 1999, Mallat 1999, Percival and Walden 2000, Torrence and Compo 1998). Imagine we have a song, and we want to understand its different pitches and rhythms. A wavelet transform is like a sophisticated tool that helps us break down this song into its various components (such as different instruments or vocals). It is more advanced than just breaking a song into bass and treble; it can isolate very specific sounds and frequencies. Whereas coherence, in this context, is the degree of similarity or synchronicity between two phenomena (e.g., musical pieces or signal segments). High coherence signifies strong similarity or synchronization, while low coherence indicates disparity or asynchronicity. Therefore, WTC is a method that leverages this intricate wavelet transform to analyze two signals (for instance, musical tracks or data sets) and evaluate their synchronicity or similarity across time.

When we apply WTC to analyze time series data from two participants within a single group, we examine the behavior of these two datasets over *Time* and across *Frequency*. For instance, consider two financial time series such as the stock prices of two different companies. These series are spread over time, and their analysis via WTC involves examining their day-to-day movements (*Time*) and identifying recurring patterns in stock price fluctuations (*Frequency*), which could range from short-term variations (high frequency, e.g., daily) to long-term trends (low frequency, e.g., yearly or multiple years). Two stocks could share similar long-term trends, e.g., the stock price kept increasing in the past five years. However, their short-term pattern could be very different, e.g., monotonic increase vs. fluctuation throughout the day with ups and downs. In this example, these two stocks have higher coherence at low frequencies, but lower coherence at higher frequencies.

In the fNIRS studies, WTC is a prevalent metric for quantifying inter-brain synchrony (IBS), as it assesses the correlation between two brain activity time series, considering both *Time* and *Frequency* (e.g., Cui et al. 2012, Jiang et al. 2012, Lu et al. 2021). A high IBS value means there is a consistent phase relationship and similar power in the fluctuations of their brain signals at specific *Time* and *Frequency*. This consistent relationship can facilitate coordinated actions, shared attention, or similar cognitive or emotional states, which are critical in cooperative or competitive interactions (e.g., Cui et al. 2012, Czeszumski et al. 2020, Wang et al. 2018).

Drawing on fNIRS research (e.g., Cui et al. 2012, Lu et al. 2021, Nozawa et al. 2016), below is a diagram demonstrating the data collection and WTC analysis of Channel 10 using one randomly selected participating team in our study.

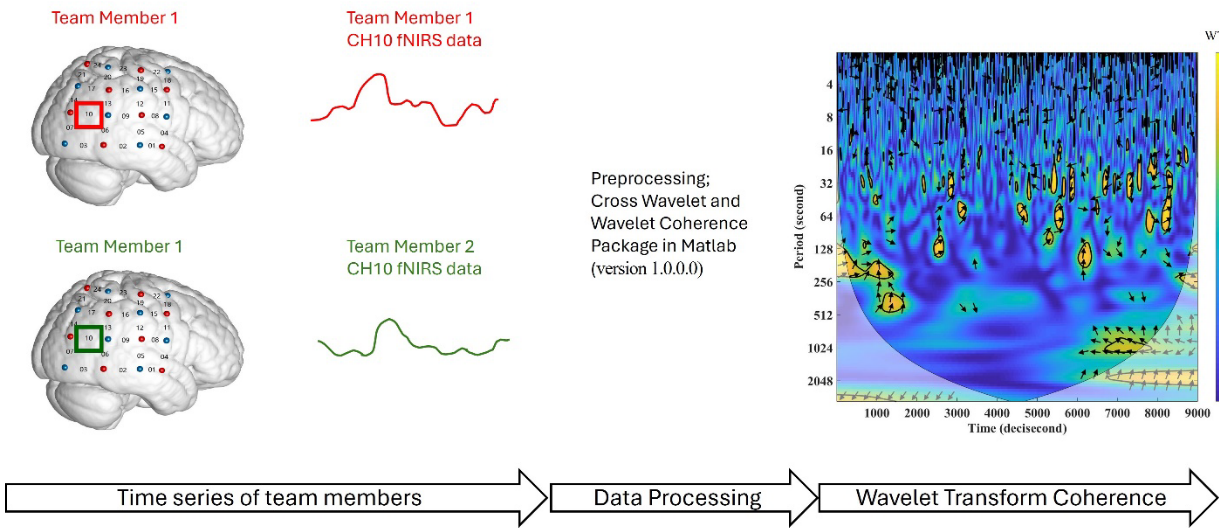


Figure E-1. Data processing and the estimation of WTC. WTC ranges between 0 and 1, with 0 (lowest coherence) in blue and 1 (highest coherence) in yellow. The x-axis is *Time* (desisecond is the unit), and the y-axis is *Period* (second is the unit). *Period* is the other aspects of waves and is the inverse of *Frequency*. While it can be confusing, neuroimaging studies report WTC diagrams using period and time instead of frequency and time. While *Frequency* refers to the number of times a repeating event occurs per unit time, *Period* is the duration of time it takes to complete one cycle of a repeating event. For example, if the frequency is 2 per second, then the period is 0.5 second. In the existing neuroimaging literature, *Period* is normally the y-axis in a WTC plot. This plot demonstrates how the coherence between the two time series changes at different frequencies. A point at a lower period (higher up on the y-axis) means it's analyzing the coherence of higher frequency components of the data. Conversely, a point at a higher period (lower down on the y-axis) is analyzing the coherence of lower frequency components.

In fNIRS hyperscanning studies, IBS is normally averaged across the task period at various brain activity frequencies. Some frequencies represent real brain activities, while others may be attributed to noise and artifacts. The more advanced fNIRS devices, such as the model used in our study, efficiently filter out most noise. In addition, following the best practices in fNIRS literature (e.g., Cui et al. 2012, Jiang et al. 2012, Lu et al. 2020, 2021), brain activity data below 0.015 Hz and above 0.700 Hz are excluded from our analysis such data is likely due to noise. To further isolate the specific effects of a task or intervention on IBS, baseline levels of IBS that exist independently of the task are usually subtracted from the IBS value measure during the task. Baseline measurements capture the state of brain synchrony before the introduction of any task or stimulus. This state can vary significantly between different pairs of participants due to individual differences in resting-state brain activity, mood, familiarity with each other, or even the experimental setting. This approach generates a new variable, IBS increment, which provides a clearer understanding of how inter-brain synchrony changes in response to specific stimuli and thus is normally the focus in fNIRS hyperscanning studies rather than raw IBS data (e.g., Cui et al. 2012, Czeszumski et al. 2020, Lu et al. 2021, Reinero et al. 2021, Wang et al. 2018)

C2. IBS Data Processing and Identification of Frequency Band of Interest

As previously mentioned, some brain activity frequencies captured by fNIRS technique may be the result of noise. Pre-processing of data (e.g., retain only frequencies within a specific range) alone cannot accurately identify the frequencies that correspond to genuine brain activities (Cui et al. 2012, Lu et al. 2021, Reinero et al. 2021, Wang et al. 2018). Using IBS increments allows for a standardized approach to measuring and reporting changes in inter-brain synchrony, facilitating comparison across different studies and experimental conditions. When the IBS during the task is *higher than during the resting period*, it indicates an increase in synchronization. This could be due to the task requiring coordinated or synchronized cognitive processes between the team members. Conversely, if the IBS during the task is *lower than during the resting period*, it indicates a decrease in synchronization. This could occur if the task promotes more independent or divergent thinking, leading to less synchronized brain activities (e.g., Cui et al. 2012, Czeszumski et al. 2020, Lu et al. 2021, Reinero et al. 2021, Wang et al. 2018).

The human brain is never completely inactive, even during rest. The default mode network (DMN) is a well-known brain network that remains active during rest, supporting self-referential thoughts, mind wandering, and consolidation of memories. The spontaneous activity of the DMN or other resting-state networks can contribute to observed non-zero IBS during resting periods (e.g., Buckner et al. 2008, Greicius and Menon 2004, Raichle et al. 2001). In addition, participants in a hyperscanning study, even during rest, share the same environment. This shared context can lead to similar sensory inputs (e.g., ambient sounds, room temperature), which can synchronize neural responses across individuals even in the absence of a direct interaction or task (e.g., Schmidt and Richardson 2008, Tognoli et al. 2007, Tognoli and Kelso 2015). Further, individuals might have intrinsic connectivity patterns that are naturally more aligned with each other, possibly due to similar brain structures, functional architectures, or psychological states. These intrinsic similarities can lead to non-zero IBS values even during resting states (e.g., Fransson and Marrelec 2008, Hasson et al. 2004, 2012). All these highlight the importance of controlling the baseline IBS in fNIRS studies. As the IBS value at resting period could be non-zero, IBS increments *could be both positive and negative* (e.g., Chang and Glover 2010, Pan et al. 2018). Negative IBS increments indicate that the team members are engaged in brain activities that are very different with synchronization, even lower than the resting period's synchronization.

To identify the real brain activity variations corresponding to the intervention introduced in the study, following prior literature, we run a one-way ANOVA using *Functional_Diversity* on IBS increments of each channel (24 channels in our study, and each channel corresponds to a brain area) along the full brain activity frequency range (0.015-0.70 Hz). There were 67 total frequencies along the full frequency range and 24 channels resulting in 1,608 (24x67) *p* values.

Given the large number of tests, all resulting *p* values (generated by ANOVAs) are adjusted using the **false discovery rate (FDR)** method, which is standard practice (Cui et al. 2012, Czeszumski et al. 2020, Lu et al. 2021, Reinero et al. 2021, Wang et al. 2018). Specifically, neuroimaging technologies, such as fNIRS and fMRI, involve the simultaneous testing of numbers of voxels (for fMRI) or channels (for fNIRS) across the brain. Each voxel or channel is tested for statistically significant activity, leading to thousands of statistical tests being performed in a single study. With each additional test, the chance of incorrectly rejecting the null hypothesis (type I error) for at least one test increases. Without appropriate correction, this could lead to many false positives, where random fluctuations in the data are mistakenly interpreted as significant findings. Therefore, in neuroimaging studies, FDR *p*-value correction is employed to effectively manage the multiple comparisons problem inherent in the analysis of high-dimensional neuroimaging data (Genovese et al. 2002, Nichols and Hayasaka 2003). The steps of FDR correction are as follows:

a) Rank the *p* values obtained from the ANOVA tests in ascending order. Let these ordered p values be $p_1, p_2, \ldots p_m$, where m is the total number of tests.
b) Calculate the critical value. For each ranked *p* value p_i, calculate its critical value using the formula: Critical $\text{Value}_i = (i/m) \times Q$, where i is the rank of the *p* value, m is the total number of tests, and Q is the chosen FDR level (e.g., 0.05 for 5%).
c) Find the largest *p* value (p_k) that meets the criteria (FDR corrected *p* value no larger than 5%).
d) Accept all hypotheses up to p_k. All hypotheses corresponding to *p* values up to and including p_k are considered significant. This means you reject the null hypothesis for all tests from 1 through k.

FDR *p* value correction ensures that the findings from these neuroimaging studies are reliable and scientifically valid, minimizing the risk of false-positive results while preserving the ability to detect true effects. In our study, after FDR correction, there are only 12 *p* values from the above-mentioned ANOVA tests that meet the 5% criteria in Task Part I (15 minutes divergent thinking phase), and 12 *p* values in Task Part II (5 minutes convergent thinking phase). Table F-1 summarizes these FDR corrected *p* values.

As discussed in our manuscript, the human brain is the most complex object in the known universe (Ackerman 1992). Any physiological or cognitive function is associated with a network of brain areas, and creativity is no exception. Drawing on neuroscientific theories, we select right hemisphere angular gyrus (r-AG) and right hemisphere superior temporal gyrus (r-STG) as our brain regions of interest as these two brain regions have consistently been shown as being associated with creativity (Berkowitz and Ansari 2010, Fink et al. 2010, 2012, Jung et al. 2010, Jung-Beeman et al. 2004, Lu et al. 2021, Seghier 2013, Shen et al. 2017). Therefore, it is not surprising that other than r-AG (channel 10) and r-STG (channel 5), there are only a few other brain areas (covered by Channel 1, 7, 11, and 12) showing significant variations. CH1 covers the right hemisphere middle temporal gyrus, which has the main function of language processing, visual perception, and multimodal sensory integration (Onitsuka et al. 2004). CH 7 covers the right hemisphere V3 area, which is one of the main components of visual neural systems (Arcaro and Kastner 2015). CH11 covers the right hemisphere subcentral area, which has the main function to respond to pressure on the eardrum and to eating and drinking (Job et al. 2011). CH 12 covers the right hemisphere supramarginal gyrus, which is mainly involved in phonological and articulatory processing of words and emotional processing (Kropf et al. 2019, Silani et al. 2013). These channels detect brain areas that correspond to fundamental human physiological functions. Therefore, we believe they are not related to our research questions and are not included in the data analysis.

Task Part I			**Task Part II**		
HZ	**Channel**	**FDR *p* value**	**HZ**	**Channel**	**FDR *p* value**
0.258	11	0.0082752	0.290	1	0.0380519
0.290	**10**	**0.0103088**	0.307	1	0.0385870
0.274	**10**	**0.0121188**	0.038	12	0.0433330
0.307	**10**	**0.0124994**	**0.038**	**5**	**0.0433330**
0.258	**10**	**0.0168429**	**0.036**	**5**	**0.0449356**
0.326	**10**	**0.0221823**	**0.034**	**5**	**0.0461327**
0.274	11	0.0224237	0.290	7	0.0488154
0.244	11	0.0265252			
0.307	1	0.0356175			
0.326	1	0.0371268			
0.345	1	0.0494100			
0.345	**10**	**0.0494100**			

Table F-1: FDR corrected *p* values that meet the 5% criteria in Task Part I and Part II.

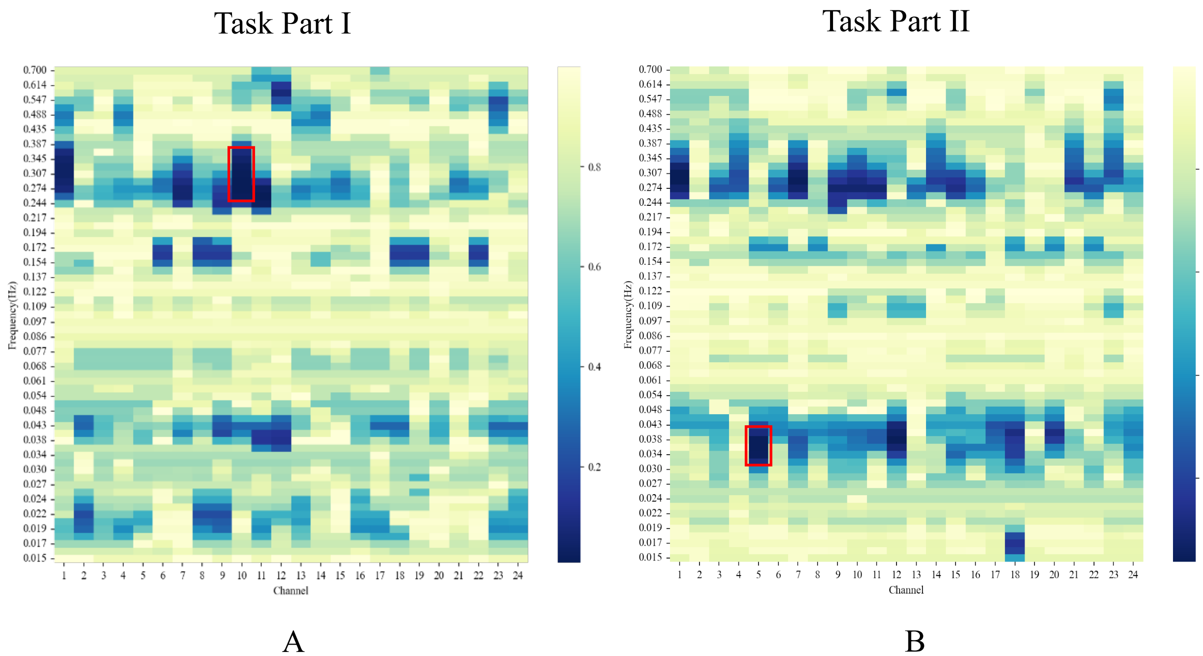
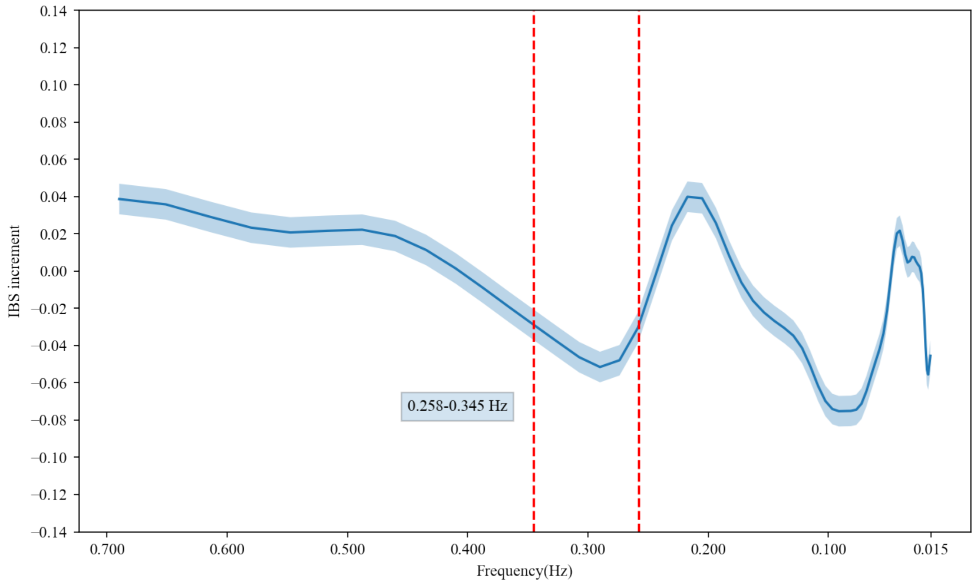


Figure F-1: Visual demonstration of the filtering process. Heat map of FDR corrected *p* values from the one-way ANOVA using *Functional_Diversity* on IBS increments from each channel and at each brain activity frequency. Blue corresponds to lower *p* values, while yellow corresponds to higher *p* values. (A) Task Part I *p* value heat map. As this map shows, *p* values from tests using IBS increments captured at Channel 10 at frequency range of 0.258-0.345 Hz are among the most significant. (B) Task Part II *p* value heat map. As this map shows, *p* values from tests using IBS increments captured at Channel 5 at frequency range of 0.034-0.038 Hz are among the most significant.

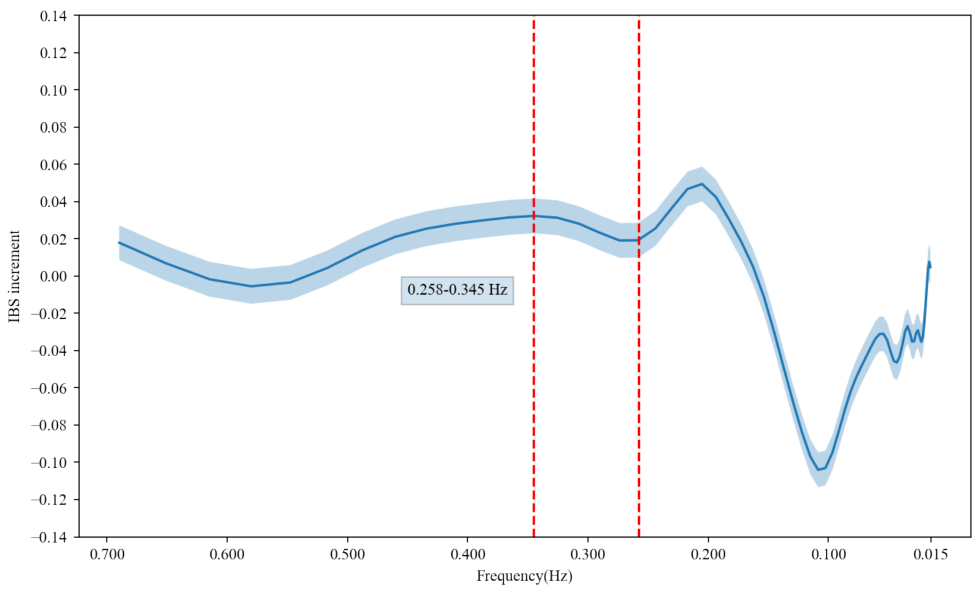
Figure F-1 offers a detailed visual representation of our strategy for selecting channels and frequency bands, providing a supplemental view of Figure 4 in our manuscript. While Figure 4 presents heat maps of *average* IBS increments (transformed into z values) for CH10 and CH5 within their respective frequency bands of interest, Figure F-1 describes the rationale behind choosing CH10 and CH5 as our focal regions in this study. Furthermore, it explains our selection of the frequency bands 0.258-0.345 Hz and 0.034-0.038 Hz for these channels, offering insights into their significance and relevance to our research objectives.

Figure F-2 below further demonstrates the magnitude of IBS increment changes in these two brain regions. Figures F-2A through F-2D collectively reinforce the justification for our chosen brain regions of interest (r-AG and r-STG) and the brain activity frequency band of interest for these regions.

A. IBS Increments from Channel 5 - Functionally Homogeneous Teams



B. IBS Increments from Channel 5 - Functionally Heterogenous Teams



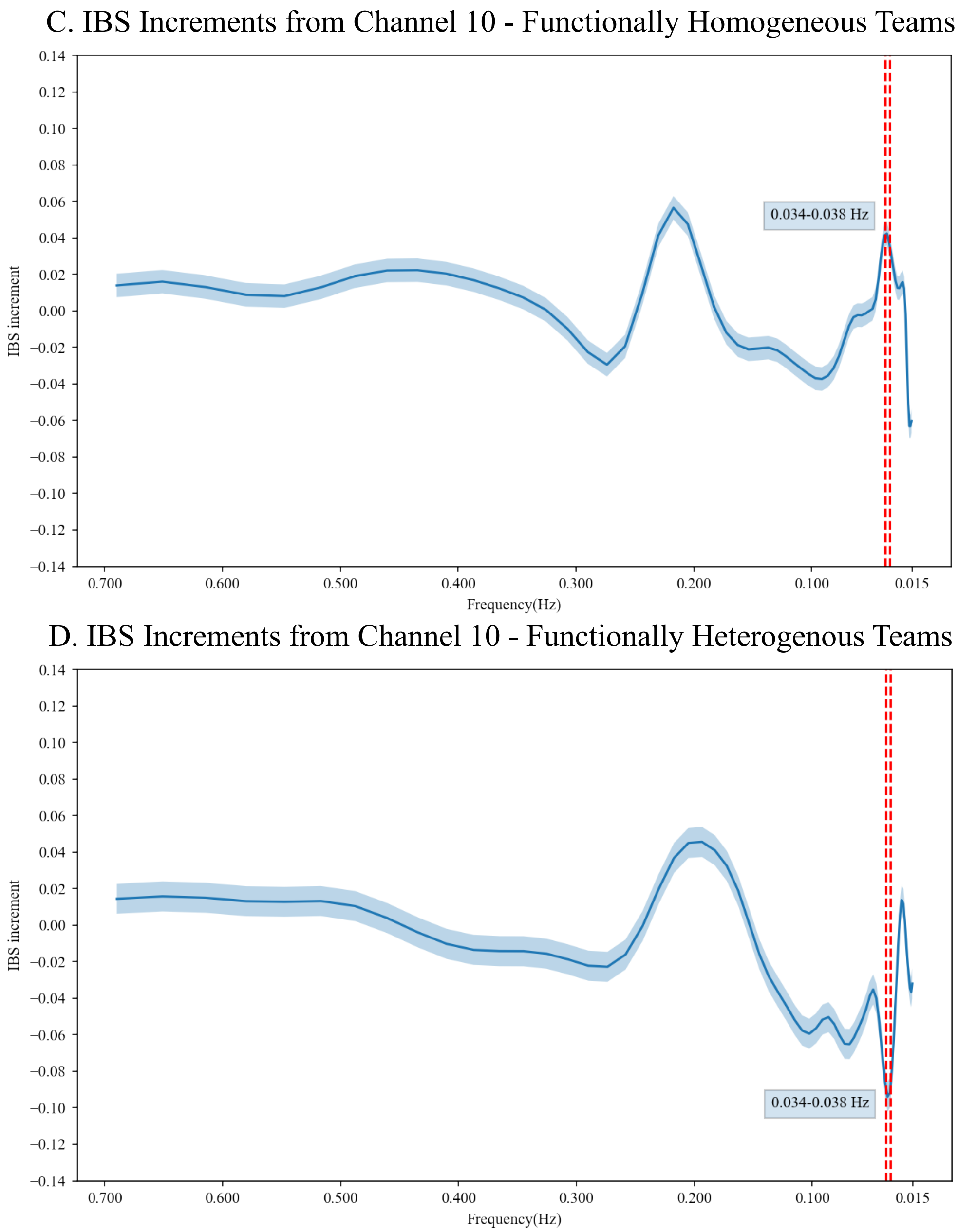


Figure F-2. The IBS increments of CH10 and CH5 are plotted against the frequencies (x-axis) for homogenous and heterogeneous groups, respectively (shaded areas indicates 95% confidence interval). (A) IBS increment plot of CH10 for the control group. (B) IBS increment plot of CH10 for the treated group. (C) IBS increment plot of CH5 for the control group. (D) IBS increment plot of CH5 for the treated group.

Online Supplemental Appendix D: Rater Instructions

D1. Instructions for Research Assistants to Rate Uniqueness

Recently, we conducted an experiment in which we recruited two-person teams of student participants. These teams were asked to develop a specific proposal for the use of the 7th floor of the Management School of Xiamen University. We need you to serve as an independent, unbiased rater of the **original and innovative** quality of each team's proposal.

When evaluating the **original and innovative** quality of each team's proposal, please consider proposals **that are particularly unusual or rare such that few groups come up with them** as being more original and innovative.

Your job as an independent, unbiased rater is to read through each submitted proposal and provide a rating. Please rate each proposal **on a scale from 1 (unoriginal and not innovative) to 7 (original and innovative)**.

Please record all ratings in the Excel spreadsheet provided to you. **Please provide a brief explanation for each rating in the "Comment/Explanation" column of the spreadsheet**.
After you finish rating the proposals, please complete **a short questionnaire** which is on the second tab of the scorecard excel document.

Thank you very much for your assistance with our experiment!

D2. Instructions for Professionals to Rate Usefulness

Recently, we conducted an experiment in which we recruited two-person teams of student participants. These teams were asked to develop a specific proposal for the use of the 7th floor of the Management School of [ ] University. We need you to serve as an independent, unbiased rater of the **practical and implementable** quality of each team's proposal.

When evaluating the **practical and implementable** quality of each team's proposal, we ask you to equally consider three aspects of the proposal:

1. <u>Mission Statement</u>. More practical and implementable proposals are those that fit with the Management School's mission statement of "[ ]" The Management School's mission statement narrowly focuses on the activities of university faculty and students.
2. <u>Available Space</u>. More practical and implementable proposals are those that effectively and efficiently use the available space. The available space is the top floor (7th floor) of The [ ] Building, which has around 8,450 sqft (65*130 ft) of unfurnished and unoccupied space. For proposals that contain multiple uses of the space, please consider whether there is sufficient space for each functional use contained within the proposal.
3. <u>Budget.</u> More practical and implementable proposals are those that effectively and efficiently use the Management School's budget. The budget for the project is $240,000. For proposals that contain multiple uses of the space, please consider whether there are sufficient funds to pay for each functional use contained within a proposal.

Your job as an independent, unbiased rater is to read through each submitted proposal and provide a rating. Please rate each proposal **on a scale from 1 (impractical and unimplementable) to 7 (practical and implementable)**.

Please record all ratings in the Excel spreadsheet provided to you. **Please provide a brief explanation for each rating in the "Comment/Explanation" column of the spreadsheet**.
After you finish rating the proposals, please complete **a short questionnaire** which is on the second tab of the scorecard excel document.

Thank you very much for your assistance with our experiment!

D3: Instructions for Research Assistants to Rate Creativity

Recently, we conducted an experiment in which we recruited two-person teams of student participants. These teams were asked to develop a specific proposal for a creative use for the 7th floor of the Management School of [ ] University. We define **creativity** as proposals that are **original**, **innovative**, **practical, and implementable within a reasonable budget**. We need you to serve as an independent, unbiased rater of the creativity of each team's proposal.

To assist your creativity evaluations, please review the same instructions we gave to the participants (see page 2). As described in these instructions, each team interacted solely through a Zoom video meeting. Team members talked with each other in the Zoom video meeting for 20 mins to discuss their ideas. At the end of the meeting, each team submitted their final proposal by typing it in the Zoom chat panel.

Your job as an independent, unbiased rater is to read through each submitted proposal and provide a creativity rating. Please rate the creativity of each proposal **on a scale from 0 (lowest creativity) to 10 (highest creativity)**. Importantly, **your average rating across all 40 proposals must be "5" Please rate the proposals relative to each other and not according to some absolute standard. To establish a range and a feel of the relative quality of the proposals, please first review them all before assigning specific ratings.**

Again, we define creativity as ideas that are **original**, **innovative**, **practical, and implementable within a reasonable budget**.

Please record all ratings in the Excel spreadsheet provided to you. **Please provide a brief explanation for each rating in the "Comment/Explanation" column of the spreadsheet**.
After you finish rating the proposals, please complete **a short questionnaire** which is the second tab of the scorecard document.

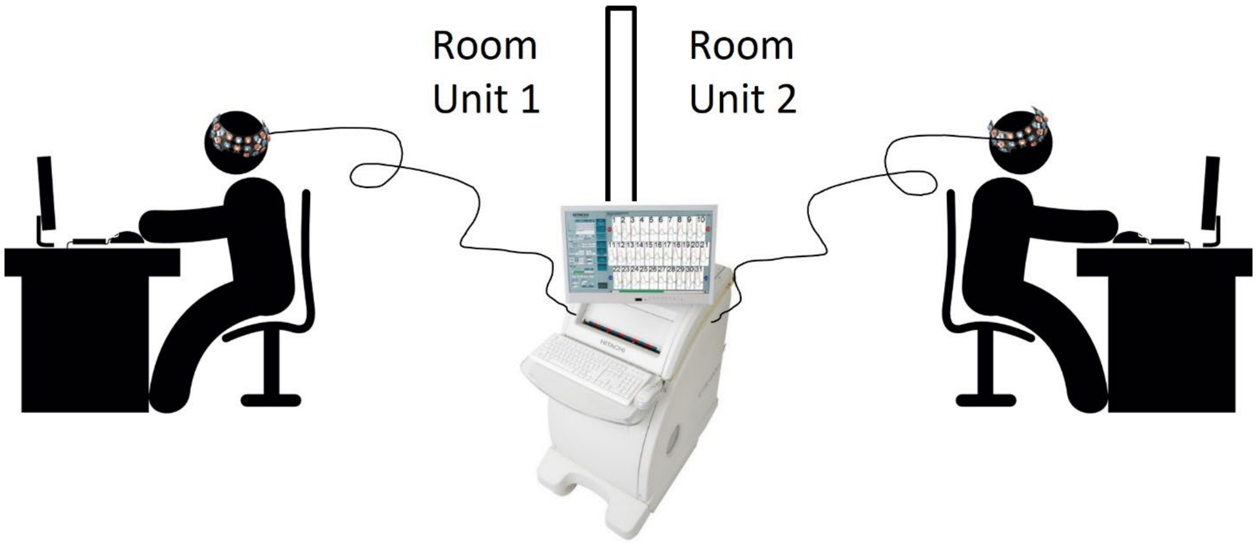
Thank you very much for your assistance with our experiment!

Online Supplemental Figure 1: fNIRS Instrument

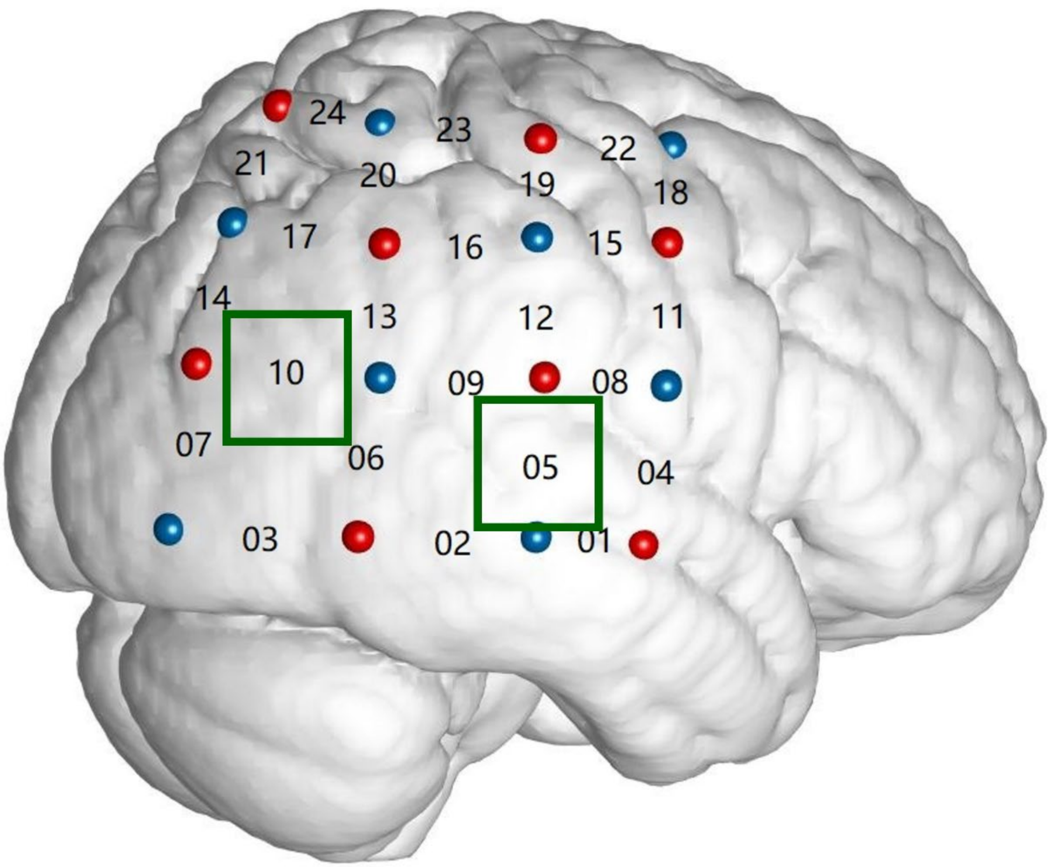
Panel A: Hitachi ETG 4100 functional near-infrared spectroscopy used in this study

Panel B: An individual wearing the fNIRS cap (from: www.hitachi-medical.com)

Panel C: Experiment setting: Two participants on the same team sit in separate room units, wearing fNIRS caps that are connected to the fNIRS machine.



Online Supplemental Figure 2: Location of All fNIRS Channels in the Experiment[1]



[1] This image shows the location of the 24 channels used in the study. Red dots are emitters, and blue dots are receivers. Channels are the brain areas located between emitters and receivers. The channel 10 and 5 are marked with a green rectangle box.