In-person, Video Conference, or Audio Conference? Examining Individual and Dyadic Information Processing As a Function of Communication System

Supplementary Section 1 – The Dynamical Interpersonal Communication Systems Model

Table S1

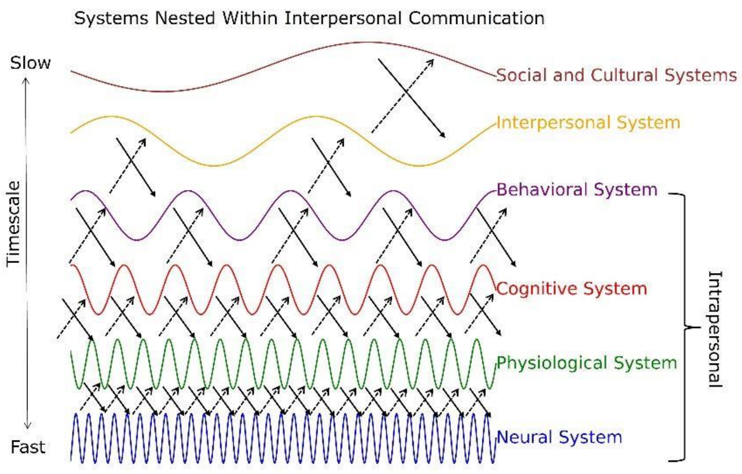
Interpersonal Communication Systems

Interpersonal Communication Systems (ICS)	Encoding System Involved in ICS	Examples of ICS	Same Location	Types of Directly Perceivable Information	Amount of Directly Perceivable Information
Evolved	Evolved Symbolic	In-person Conference	Yes	Visual Auditory	High
Representational	Representational Symbolic	Audio/Video Conference	No	Visual Auditory	Medium High
		Audio Conference	No	Auditory	Medium Low
Symbolic	Symbolic	Messaging	No	Pictorial	Low

Note. All communication systems involve the use of language which is encoded symbolically.

Figure S1

Systems Nested Within Interpersonal Communication



Note. The figure illustrates the systems nested within a specific interpersonal communication system. The slower systems restrict the faster ones, while the faster ones also influence the slower ones. The figure was adapted from Eiler et al. (2013).

Supplementary Section 2 – Physiological Data Preparation and Calculation

Physiological Data Preparation

Participants were instructed to place three pre-gelled disposable ECG electrodes on their bodies using the three-lead system. To allow participants' arms to move freely (e.g., gestures), two electrodes were placed just below the collarbones and the other on the waist. Prior to attaching the electrodes, participants were instructed to clean the designated skin areas using an alcohol pad. To ensure privacy, the experimenters left the lab during the ECG preparation process, and for the dyads in the in-person condition, separate spaces were provided for each individual if necessary.

After attaching the ECG electrodes, the experimenters returned to the lab and began preparations for the EMG and EDA electrodes. For the facial EMG signals, the skin areas were first wiped using a paper towel dampened with distilled water, lightly abraded with an exfoliating abrasion pad, and a small amount of conductive gel was applied to improve the conductivity of the signal. Two reusable 4 mm Ag/AgCl electrodes filled with an isotonic gel were placed just above the left eyebrow to measure corrugator supercilii muscle activity, and two same electrodes were attached to the skin under the left lower eyelid for orbicularis oculi data recording. To ensure the inter-electrode impedances of the facial EMG signals were within an acceptable range, a CheckTrode electrode impedance meter was used, with a mean impedance of 12.66 kΩ across all participants. For the EDA preparation, the skin was cleaned using a paper towel dampened with distilled water. Two pre-gelled disposable EDA electrodes were attached to the inner arch of the left foot. To ensure data quality, the experimenters carefully monitored all physiological signals during recording. All physiological signals were recorded with BIOPAC MP150 systems at a sampling rate of 2,000Hz. The physiological data preparation procedures were also detailed in Han (2020) and Han et al. (2022).

Note, the placement of ECG and EDA was intentionally selected to allow for free movement of the upper body, in order to enhance the ecological validity of this study. Nonetheless, it is possible that different communication conditions may be affected differently by the invasive equipment. Specifically, the audio-only condition may have been less influenced because participants could not see the sensors on their partner, thereby not being constantly reminded of these devices. On the other hand, participants in visual-based communication might have been periodically reminded, possibly affecting their communication behavior. Researchers could consider some wearable (and thus smaller) devices such as the Apple Watch for heart rate, which might be less noticeable, and try to replicate some findings of this study. In addition, future research can collect naturalistic conversation recordings and examine facial expressions using facial recognition software (which is less invasive) to identify and analyze real-time facial expressions to understand emotional processes involved in interpersonal communication (though facial recognition software cannot detect non-visible activities). These two methodologies may remove the impact of invasive equipment on the processes we studied in the paper. Nonetheless, it should be noted that as of now there are no wearable devices or facial recognition tools that can match the accuracy of the physiological measurements we employed in the study.

Physiological Data Frequency

The pre-processed data, obtained from the previous study, had varying frequencies, including 5Hz RSA, 1Hz EDA, and 10Hz fEMG. The reasons why different frequencies were extracted for different signals are as follows. First, 5Hz RSA time series were created using the dynamic RSA method developed by Abney et al. (2021). Parameters proposed by Abney et al. including the 5Hz frequency were followed to best preserve the dynamics of RSA as tested by Abney et al. (2021). Second, for EDA, because EDA runs on a second-based scale, 1Hz of skin conductance data was obtained as the result of the pre-processing process (see Han et al., 2022). Lastly, facial EMG data were initially exported with two data sets, a 1Hz dataset and a 10Hz dataset. Nevertheless, the 1Hz data set contained many files with inaccurate data from AcqKnowledge, which resulted from an error during the batch process. Therefore, 10Hz data were used for the present study.

RSA Calculation

First, inter-beat intervals (IBIs) were extracted from ECG signal using AcqKnowledge 5.1. The software's script was used to automatically identify each R peak from the QRS complex. Trained research assistants then visually reviewed the data and corrected any incorrectly labeled R-peaks. The time between two adjacent R peaks was calculated as the inter-beat interval.

Second, following the method outlined by Abney et al. (2021), the IBI time series was resampled to create a time-based 5Hz IBI time series. This time-based time series allows for real-time matching of the RSA data with behavioral data. The remaining steps in the RSA calculation were carried out using the Porges-Bohrer method as described in Abney et al. (2021), including a final transformation step that takes a natural logarithm of RSA values to ensure a normal distribution (Porges, 1985; Riniolo & Porges, 2000). It should be noted that to match the RSA time series with behavioral data, the data stretching step involved in Abney et al.'s method was omitted in this study, with the agreement of Abney and his team.

Facial EMG Data Calculation

Facial EMG data were pre-processed following the guidelines provided by BIOPAC. Specifically, a spectral analysis was applied to identify any 60Hz noise present in the data, and if yes, a 60Hz comb band filter was used to remove this noise. After that, EMG data were run through a 28-500Hz bandpass filter and then rectified and integrated (Read, 2020). Finally, 10Hz data were extracted from the AcqKnowledge software as per the previous study, from which the data for this study was obtained.

EDA Data Calculation

Please see the detailed procedures described in Han et al. (2022).

Supplementary Section 3 – Participants and Lab Setting

Participants Information

Most participants identified themselves as being White (n = 100, 66.67%), followed by Asian (n = 19, 12.67%), more than one race (n = 6, 4.00%), Black or African Americans (n = 4, 2.67%), and Native Hawaiian or other Pacific Islanders (n = 1, 0.67%). Four participants preferred not to answer the racial identity question. The race identity of 16 participants was not collected due to technical failure.

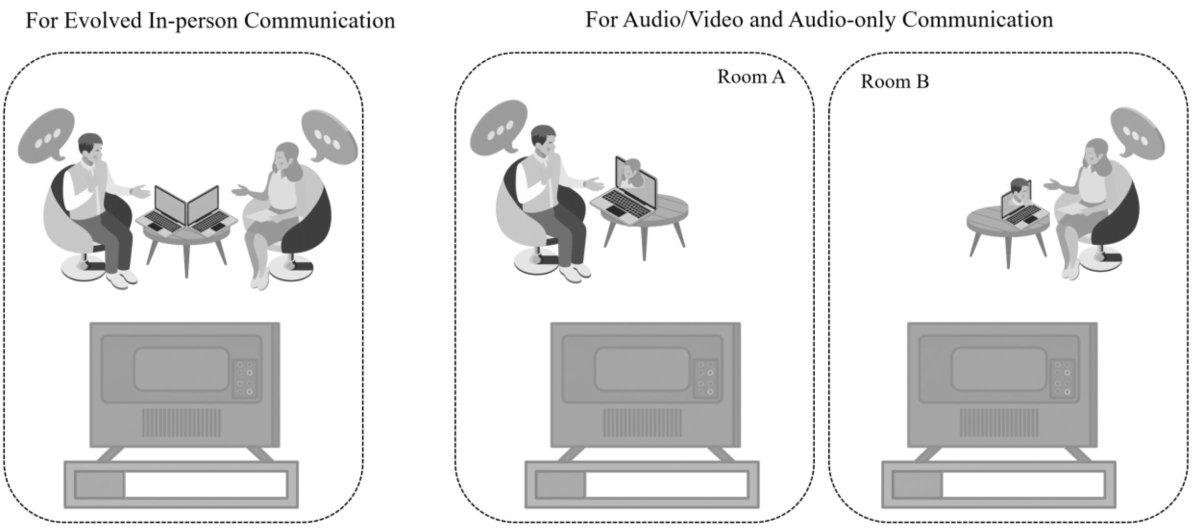
Lab Setting

Dyads were seated in front of a big television screen where a static basketball image was presented. For the video and audio conference groups, dyads were guided to two separate rooms with the same setting as the in-person condition. Cameras for the video conference group were placed on the participants' laptop which was placed on a small coffee table next to participants. Participants were not allowed to adjust the settings, and in, fact none of them had the intention to do so.

Figure S2 below illustrates the lab setting used for this study. The in-person communication condition involved two conversants sitting in the same room and engaging in an unstructured conversation for a duration of two minutes. The audio/video and audio-only communication conditions, on the other hand, required participants to be in different rooms and engage in their conversation using a video or audio conference tool. In audio/video communication, participants were able to perceive visual cues of their conversing partner, while the audio-only condition did not provide any visual information.

Figure S2

Lab Setting



Supplementary Section 4 – Data Description

Missing Data

Four dyads' data were missing (one in the in-person condition and two in the audio conference condition) due to the failure of participants to execute the conversation (during which they thought they should keep silent). For RSA, one additional dyad's data in the video conferencing condition was deleted because data for their silence phase were missing, which is needed for calculating the tonic resting RSA level in the analysis. For EDA, eight dyads' data were very noisy and thus removed. For fEMG data, one participant's data were deleted because the AcqKnowledge 5.0.1 failed to rectify the filtered data. Data distribution for the communication conditions for each level of analysis is summarized in Table S2 of this document.

Table S2

Data Distribution for the Three Communication Systems

	In-person Conference	Video Conference	Audio Conference
Individual level analysis (data unit: individual)			
RSA	54	46	40
Skin Conductance	44	45	36
Facial EMG-Corrugator	54	45	41

Facial EMG-OO	54	45	41
Dyadic level analysis (data unit: dyad)			
RSA	27	23	20
Skin Conductance	22	22	18
Facial EMG-Corrugator	27	22	20
Facial EMG-OO	27	22	20

Supplementary Section 5 – Description of Variables

Friendship Closeness and Satisfaction

The strength of the dyads' friendship was measured with two 7-point scales: one Likert scale on psychological closeness (6 items, Vangelisti & Caughlin, 1997) and one semantic differential scale (10 items) on relational satisfaction (Huston et al., 1986). Cronbach's alphas for the closeness and satisfaction scales were .88 and .90 respectively.

Big Five Personality

The five personality traits including agreeableness, conscientiousness, extroversion, emotional stability, and openness to experience, were assessed by the ten-item personality inventory (TIPI, Gosling et al., 2003). Agreeableness, as described in the main text, is the tendency to cooperate with others. Conscientiousness is the degree of self-discipline and self-control. Extroversion describes one's tendency on emotional expressiveness and sociability. Openness to experience describes one's tendency to seek novel experiences and intellectual development. Emotional stability, oftentimes called neuroticism, is associated with mood swings, anxiety, sadness, etc. TIPI was not developed for traditional reliability measures (see the scale developers' note at https://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/a-note-on-alpha-reliability-and-factor-structure-in-the-tipi/).

RSA Introduction

Respiratory sinus arrhythmia (RSA) refers to a high frequency band (0.12-0.40Hz, Porges, 1985) of heart rate variability and has been used to measure the parasympathetic influence on the heart, indicative of changes in mental effort, cognitive workload, emotional regulation, anxiety, and other psychological processes (Beauchaine, 2001; Porges, 1995). RSA reflects the body's response to environmental challenges by modulating the vagal break which controls the metabolic resources required by the body and the environment. Specifically, when one gets challenged by their surroundings, either physically, mentally, or emotionally, the body needs more metabolic resources to deal with the outside stress in order to maintain its homeostasis. To meet the increased metabolic demands, the vagal break is released (vagal withdrawal), lowering RSA and activating the sympathetic nervous system, which results in a faster heart rate. On the other hand, in a peaceful environment, vagal break is maintained for social engagement behaviors, resulting in increased RSA and decreased heart rate (Porges, 1991, 1995, 2007). In the context of this study, lower RSA is interpreted as indicating more cognitive effort during interpersonal communication.

Supplementary Section 6 –
Cross Recurrence Quantification Analysis (CRQA)

Parameters

To calculate synchrony levels using CRQA, several parameters including time lag and embedding dimension were determined by the CRQA MATLAB toolbox (Kay & Richardson, 2015). Following the same procedures demonstrated in Han et al. (2022), a time lag of 50, an embedding dimension of 10, and a radius of 8 were used for the 10Hz corrugator data, a time lag of 40, an embedding dimension of 10, and a radius of 10 for the 10Hz orbicularis oculi, and lastly a time lag of 40, an embedding dimension of 6, and a radius of 25 for the 5Hz RSA. These parameters were chosen such that most of the %REC fell into the preferred 0.5% – 5% range (Richardson et al., 2007; Shockley, 2005). All other parameters were the same as those reported by Han et al. (2022), including those for skin conductance.

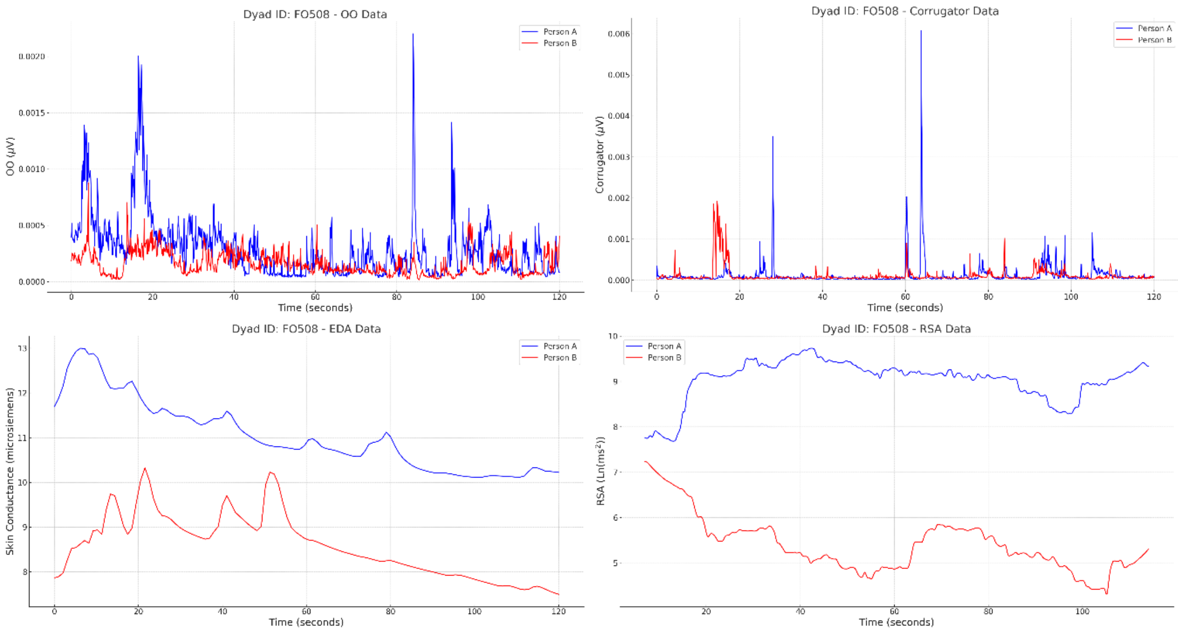
Data Demonstration

The figures below depict individual data from a single dyad and their corresponding recurrence plots derived from CRQA.

Figure S3

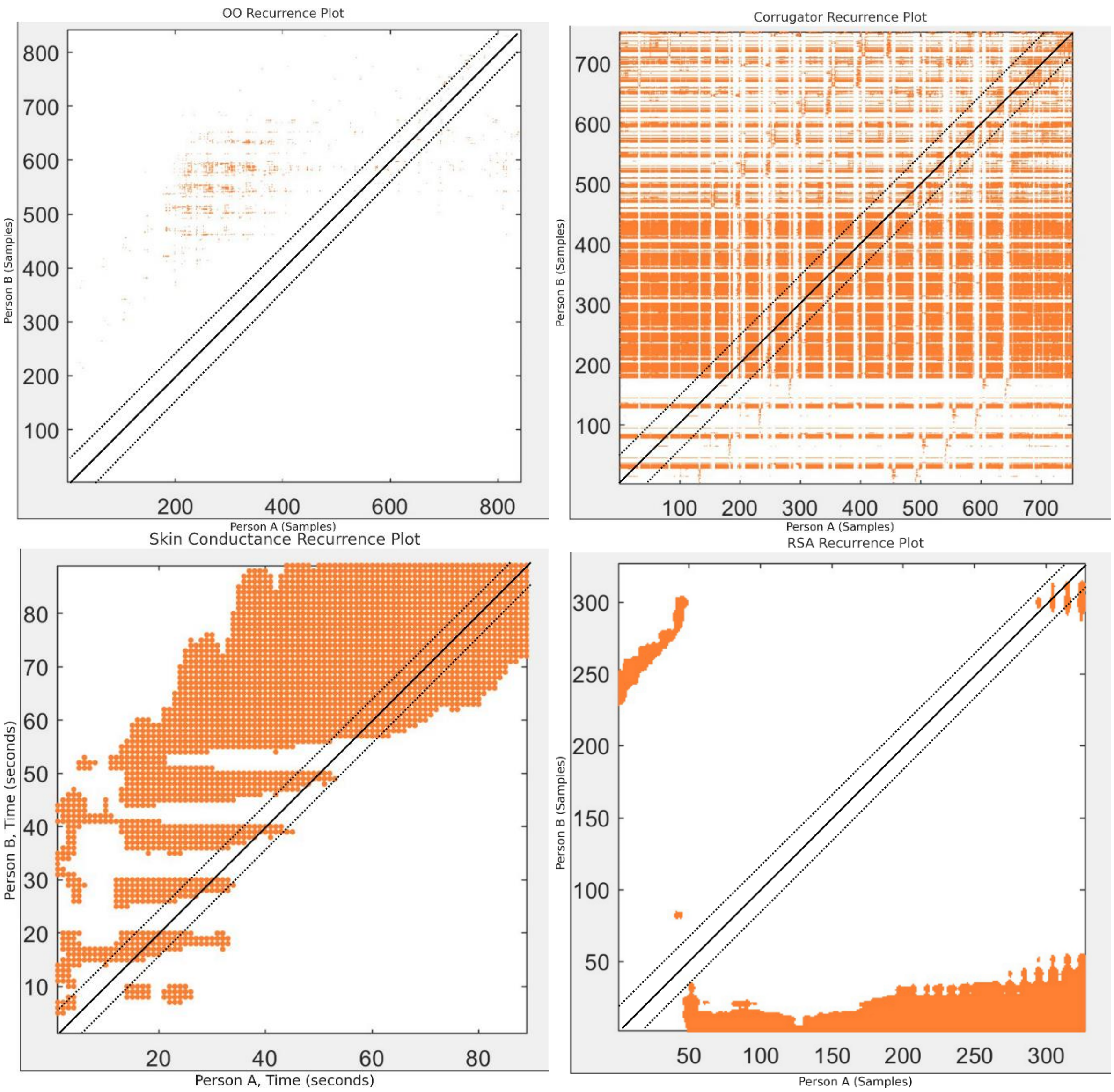
Exemplar Data and CRQA Plots

(3A) Exemplar Data



Note. The four graphs display one dyad's OO, corrugator, skin conductance, and RSA activity over the course of their 2-minute conversation. The RSA plot captures the RSA activity from the 7th to the 114th second, which is the result of RSA calculation from the original data of inter-beat intervals. Data from the beginning and end were lost due to the use of the PolyFilter function (see the RSA calculation program from Abney et al., 2021).

(3B) Corresponding CRQA Plots



Note. The graphs above are recurrence plots corresponding to the data presented in Figure S3A. The x- and y-axis represent each participant's timeline. Every orange dot in the plots, known as a recurrence point, represents a moment when the two participants shared a similar physiological state. The added black line is the main diagonal line of the plot that shows recurrent states at lag0. Orange dots distant from the main diagonal line represent recurrence states at increased time lags (e.g., in this case, for OO and skin conductance we see more orange dots on the top-left side of the main diagonal than on the bottom-right side, suggesting that Person A was a leader on Person B's physiological state). The two dashed lines represent lines that are 5 seconds apart from the main diagonal. The areas between the two dashed lines are the synchrony data we captured for analysis, which indicates synchrony levels within a range of +/- 5-second lags. More orange dots in this area means a higher level of synchrony for this dyad. More explanations of the CRQA calculation can be found in Han et al. (2022).

Supplementary Section 7 –
Results Summary of Communication System Effects on Covariates

Overall Summary

We examined whether the covariate variables were significantly different across the three communication groups at the individual and dyadic levels. Overall, at the individual level, only friendship satisfaction and looking-partner behavior (the latter as expected) were significantly different across the three communication groups. All other covariate variables including talking behavior and relationship closeness were not significantly different across the communication groups. At the dyadic level, no significant difference was found regarding the distributions of personality traits, friendship, total time of talking, and gaze behavior across the three communication conditions.

Summary for Individual Level Analysis

A one-way ANOVA revealed that relational satisfaction (F (2, 147) = 3.80, p = .025, η_p^2 = 0.05, 95% CI [.003, 1.00]) was found to be statistically different across the three communication conditions. Specifically, post hoc analysis with Tukey correction showed that participants in the video conference group reported significantly higher friendship satisfaction (M = 6.35, SD = .59) than those in the in-person communication (M = 5.94, SD = .86, t = 2.75, p = .018). No other statistically significant differences were found (M = 6.15, SD = .77 for the audio conference group).

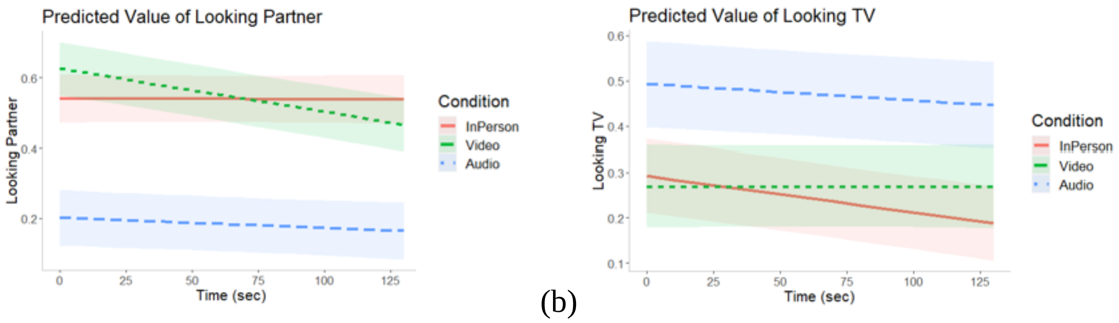
No significant differences were observed in relation to other individual difference variables, including friendship closeness, the Big Five personality traits, and demographic characteristics, across the three communication conditions.

For talking behavior, the results of the multilevel modeling analysis revealed that neither the communication condition, time, nor their interaction had any significant effect on participants' talking behavior.

For participants' gaze behavior, the MLM analysis for both looking-partner and looking-TV behavior revealed significant main effects of communication condition and time, qualified by their interaction effect (see Table S3). In summary, and as expected, the audio conference group displayed a lower frequency of directing their gaze to the laptop screen (where the audio conference tool was launched and through which their partner's voice was delivered), and a higher frequency of looking at the television than the other two groups. In comparison to the in-person group, the video conference group exhibited a higher frequency of looking at their partner at the beginning of their conversation. However, this frequency decreased over time and eventually resulted in a lower frequency of looking at their partner than the in-person group. The in-person group had a consistently high frequency of looking at their partner throughout the conversation, and consequently a low frequency of looking at TV (see Figure S4a and S4b below).

Figure S4

Frequency Distribution of Looking at Conversing Partner (a) and TV (b) as A Function of Communication System and Time



(a) (b)

Table S3

Multilevel Modeling Results of Gaze Behavior

	LookingPartner			**LookingTV**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.63	0.55 – 0.70	**<0.001**	0.27	0.18 – 0.36	**<0.001**
CommunicationSystem [InPerson]	-0.09	-0.19 – 0.02	0.100	0.02	-0.10 – 0.14	0.702
CommunicationSystem [Audio]	-0.42	-0.53 – -0.31	**<0.001**	0.22	0.09 – 0.36	**0.001**
Time	-0.00	-0.00 – -0.00	**<0.001**	-0.00	-0.00 – 0.00	0.993
CommunicationSystem[InPerson] * Time	0.00	0.00 – 0.00	**<0.001**	-0.00	-0.00 – -0.00	**<0.001**
CommunicationSystem [Audio] * Time	0.00	0.00 – 0.00	**<0.001**	-0.00	-0.00 – 0.00	0.100
Random Effects						
σ^2	0.18			0.15		
τ_{00}	0.04_{Dyad}			0.06_{Dyad}		
τ_{11}	$0.05_{\text{Dyad.AorBB}}$			$0.06_{\text{Dyad.AorBB}}$		
ρ_{01}	-0.44_{Dyad}			-0.46_{Dyad}		
ICC	0.18			0.28		
N	72_{Dyad}			72_{Dyad}		
Observations	16941			16939		
Marginal R^2 / Conditional R^2	0.113 / 0.271			0.046 / 0.310		

Note. The audio/video communication, or video conference condition, was the base group in the model.

Summary of Dyadic Level Analysis

For the dyad-difference variables (friendship, personality traits, and gaze and talking behaviors), results of chi-square analysis showed that the distribution of those groups (low-low, high-high, and low-high) was not significantly different across the three communication conditions, for all dyad-difference variables. In other words, the three communication conditions were not different in the distribution of friendship, personality traits, and gaze and talking behavior at the dyadic level.

Supplementary Section 8 – Post Power and Sensitivity Analyses

The complexity of multilevel model analyses makes power calculation very challenging. As Westfall et al. (2014) demonstrated, the commonly used power calculation tool, G*Power, is not appropriate for multilevel models.

In this study, a priori power analysis for sample size estimation is not possible, as the study relied on secondary data analysis. However, a post-hoc power analysis was conducted using G*Power to gain insight into the obtained power and the robustness of the results. Additionally, sensitivity analysis was also conducted to further understand the robustness of the findings.

For Individual Level

Because there is limited research on the effect of communication systems on individuals' physiological levels, we used a relatively small effect size in our post-hoc power analysis, with a partial η^2 of .05. It is important to note that this effect size is much smaller than what has been found in some studies examining similar effects on individuals' physiological activity. For example, Ravaja (2009) found above-medium effect sizes on three facial EMG signals (with partial η^2 values of .18 for zygomatic muscle activity, .39 for orbicularis oculi, and .14 for corrugator activity) when comparing co-located and non-co-located game players during digital game playing.

We chose a conservative effect size of partial η^2 = .05 (equals to effect size Cohen's f = .23), number of measures = 2 (which could be as large as 120 as we had 120 time points measured for each physiological signal), and the minimum sample size = 125, in addition to other parameters (alpha = .05, number of groups = 3). With the most conservative parameter setting, G*Power revealed an achieved power of .98.

For sensitivity analysis, we entered the same values for the parameters above, but with a power of .80. The result revealed an effect size of f = .14. By increasing the number of measurements to 15, the effect size f decreases to .08, which corresponds to a partial η^2 of .019.

For Dyadic Level

Like the individual level analysis, there is limited research on the effect size of communication systems on interpersonal physiological synchrony. Studies examining similar effects have revealed above-medium effect sizes per Cohen's suggestion. For example, Müller and Lindenberger (2011) examined interpersonal synchrony of respiration and heart rate variability among singers in different types of singing conditions including singing with eyes open versus eyes closed, and found an average partial η^2 of .53, with a range of .20 - .83. Moreover, Ham and Tronick (2009) studied mother-infant synchrony in heart rate and RSA as a function of face-to-face versus still face interaction. The results revealed an effect size greater than what Cohen defined as a large effect.

Therefore, for post-hoc power analysis, we chose a small effect size of partial η^2 = .04, corresponding to Cohen's f = .20, with a minimum sample size of 62 dyads and 2 measurements (although the actual and smallest number of measurements in the study was 15, which corresponds to the number of time lags assessed for skin conductance synchrony).

G*Power revealed an achieved power of .79 with these parameters. Increasing the effect size to partial $\eta^2 = .05$ (corresponding to $f = .23$) increases the achieved power to .89. Additionally, increasing the number of measures to 15 resulted in an achieved power of .99.

For sensitivity analysis, we used the same parameters as the above, but set the power to .80. This yielded an effect size f of .20, as expected. We also calculated the effect size when increasing the number of measurements to 15, with the other parameters remaining the same. This resulted in a reduction of the effect size f to .11 (partial $\eta^2 = .012$).

Overall, the post power and sensitivity analyses indicate that the study had adequate power to detect significant differences among the communication conditions, despite the small effect size, sample size, and number of measurements used in the analyses.

The study's R code and data can be found at https://osf.io/7kaxs/?view_only=1f111aa753bb4bad8893b0fce58482d7.

Supplementary Section 9 – Results Summary

Table S4

Multilevel Modeling Results for Dyadic Level Analysis

	REC.OO			REC.CORR			REC.RSA			REC.SC		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.03	0.02 – 0.04	**<0.001**	0.04	-0.02 – 0.10	0.196	0.09	0.05 – 0.14	**<0.001**	0.02	0.01 – 0.03	**0.001**
Lag	-0.01	-0.01 – -0.01	**<0.001**	-0.00	-0.00 – -0.00	**<0.001**	0.00	0.00 – 0.00	**0.002**	-0.00	-0.00 – -0.00	**0.017**
Visual-Nonvisual	0.00	-0.02 – 0.03	0.910	0.03	-0.11 – 0.17	0.654	-0.13	-0.23 – -0.04	**0.006**	0.01	-0.01 – 0.03	0.237
InPerson-Video	0.02	-0.01 – 0.04	0.190	0.09	-0.06 – 0.24	0.230	0.04	-0.08 – 0.16	0.531	-0.02	-0.04 – 0.00	0.084
Lag * (Visual-Nonvisual)	-0.00	-0.00 – 0.00	0.053	0.00	0.00 – 0.00	**0.025**	0.00	-0.00 – 0.00	0.538	0.00	-0.00 – 0.00	0.409
Lag * (InPerson-Video)	-0.01	-0.01 – -0.01	**<0.001**	0.00	-0.00 – 0.00	0.161	0.01	0.01 – 0.01	**<0.001**	0.00	-0.00 – 0.01	0.253
ClosenessGroup [Low-High]				0.20	0.09 – 0.30	**<0.001**						
ClosenessGroup [Low-Low]				0.15	0.06 – 0.24	**0.001**						
(Virtual-Nonvirtual) ClosenessGroup [Low-High]				0.22	-0.01 – 0.44	0.057						
(InPerson-Video) * ClosenessGroup [Low-High]				0.01	-0.24 – 0.27	0.914						
(Virtual-Nonvirtual) * ClosenessGroup [Low-Low]				-0.26	-0.46 – -0.06	**0.010**						
(InPerson-Video) * ClosenessGroup [Low-Low]				-0.10	-0.31 – 0.12	0.371						
AgreeablenessGroup [Low-High]							-0.07	-0.12 – -0.01	**0.012**			
AgreeablenessGroup [Low-Low]							-0.05	-0.11 – 0.00	0.064			
(Virtual-Nonvirtual) * AgreeablenessGroup [Low-High]							0.13	0.02 – 0.23	**0.018**			
(InPerson-Video) * AgreeablenessGroup [Low-High]							-0.03	-0.16 – 0.10	0.658			
(Virtual-Nonvirtual) * AgreeablenessGroup [Low-Low]							0.18	0.07 – 0.30	**0.002**			
(InPerson-Video)* AgreeablenessGroup [Low-Low]							-0.03	-0.17 – 0.10	0.640			
Random Effects												
σ^2	0.00			0.00			0.00			0.00		
τ_{00}	0.00_{Dyad}			0.03_{Dyad}			0.00_{Dyad}			0.00_{Dyad}		
ICC	0.67			0.97			0.92			0.71		
N	69_{Dyad}			67_{Dyad}			69_{Dyad}			62_{Dyad}		
Observations	3519			3417			1794			992		
Marginal R^2 / Conditional R^2	0.066 / 0.687			0.344 / 0.980			0.213 / 0.933			0.051 / 0.724		

Note. REC refers to %REC as the indicator of interpersonal synchrony. OO stands for orbicularis oculi, CORR for corrugator activity, RSA for respiratory sinus arrhythmia, and SC for skin conductance activity. The number of observations in the models varies due to the different frequencies of the data for the four signals, which were explained in Section 2. This led to different numbers of time lags for the calculation of lags for OO, CORR, and RSA, which were set to 5 seconds. For SC, lags up to 5 seconds were originally calculated and tested with the model. Due to its non-significant results and the fact that skin conductance is a slow responding signal, we increased its time lags to 15 seconds for further examination, but again no significant results were found, as reported above.

Table S5

Multilevel Modeling Results for Individual Level Analysis

	OO			CORRUGATOR			RSA			SC		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.38	0.31 – 0.46	**<0.001**	-0.03	-0.04 – -0.01	**<0.001**	-0.20	-0.32 – -0.07	**0.002**	0.80	0.64 – 0.97	**<0.001**
Time	-0.06	-0.07 – -0.06	**<0.001**	0.00	-0.00 – 0.00	0.216	0.05	0.04 – 0.06	**<0.001**	-0.20	-0.21 – -0.19	**<0.001**
Visual-Nonvisual	0.12	-0.05 – 0.29	0.168	0.03	0.00 – 0.07	**0.037**	0.23	-0.04 – 0.50	0.099	-0.29	-0.66 – 0.09	0.131
InPerson-Video	0.11	-0.07 – 0.29	0.232	0.02	-0.01 – 0.05	0.241	0.08	-0.21 – 0.37	0.599	-0.14	-0.54 – 0.25	0.478
Talking	0.05	0.04 – 0.06	**<0.001**	0.01	0.01 – 0.01	**<0.001**	-0.02	-0.03 – -0.01	**0.002**	0.02	0.00 – 0.03	**0.012**
Time * (Visual-Nonvisual)	0.00	-0.01 – 0.02	0.593	-0.00	-0.00 – 0.00	0.822	-0.08	-0.10 – -0.05	**<0.001**	-0.07	-0.10 – -0.05	**<0.001**
Time * (InPerson-Video)	0.00	-0.01 – 0.02	0.738	-0.00	-0.01 – 0.00	0.515	-0.05	-0.08 – -0.02	**<0.001**	0.02	-0.01 – 0.05	0.122
Tonic RSA							-0.25	-0.37 – -0.13	**<0.001**			
Random Effects												
σ^2	0.16			0.01			0.49			0.49		
τ_{00}	0.13 $_{\text{AorB:Dyad}}$			0.01 $_{\text{AorB:Dyad}}$			0.47 $_{\text{AorB:Dyad}}$			0.90 $_{\text{ID\_single}}$		
	0.04 $_{\text{Dyad}}$			0.00 $_{\text{Dyad}}$			0.03 $_{\text{Dyad}}$					
ICC	0.51			0.29			0.51			0.65		
N	2 $_{\text{AorB}}$			2 $_{\text{AorB}}$			2 $_{\text{AorB}}$			123 $_{\text{ID\_single}}$		
	70 $_{\text{Dyad}}$			70 $_{\text{Dyad}}$			70 $_{\text{Dyad}}$					
Observations	16439			16439			14570			14637		
Marginal R^2 / Conditional R^2	0.033 / 0.523			0.017 / 0.307			0.074 / 0.542			0.044 / 0.664		

Note. OO stands for orbicularis oculi, CORRUGATOR for corrugator activity, RSA for respiratory sinus arrhythmia, and SC for skin conductance activity. Data for OO and corrugator activity were multiplied by 1000 in the analysis. The MLM for skin conductance does not include a nested structure for random effects due to convergence issues. Dyad ID was tested for its random effect, with no variance found to be explained by the Dyad ID for skin conductance activity, and therefore was not included in the model.

Table S6

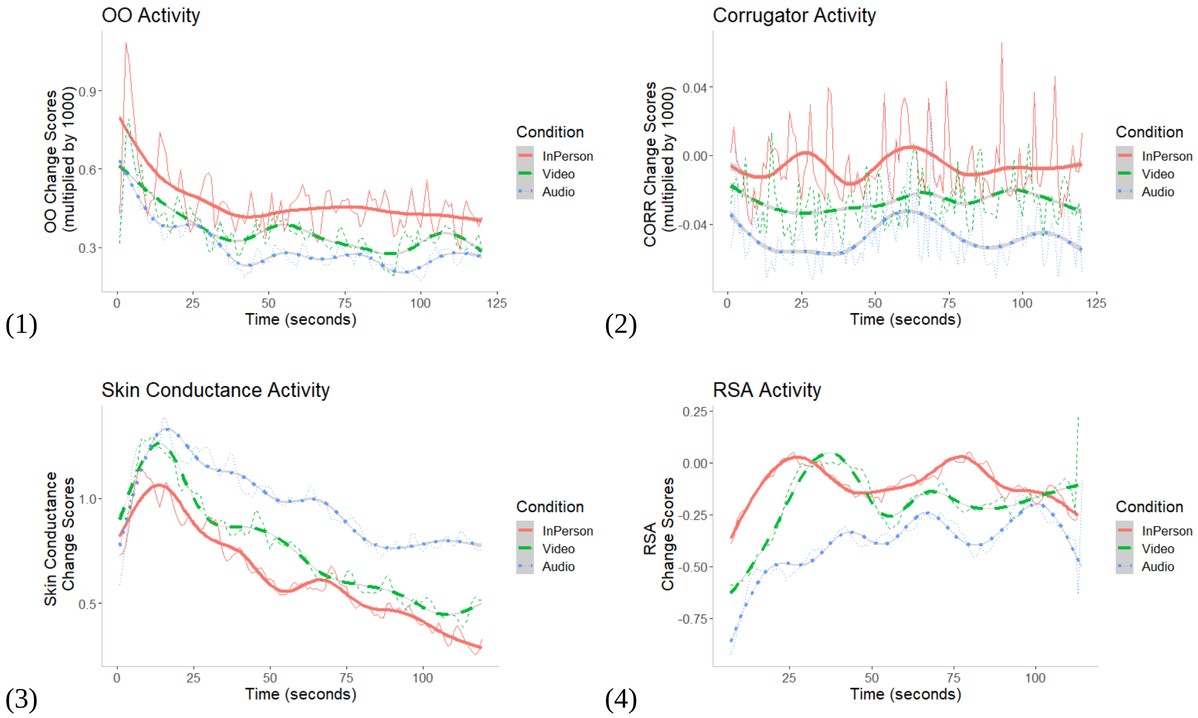
Results Summary

Physiological Signal	**Level of Analysis**	**Visual vs. Non-visual-based Comm. (PC1)**	**In-person vs. Video Comm. (PC2)**
OO	Dyad	---	PC2 × Lag effect (sig. at lag0)
	Individual	---	---
	Summary: Embodied visual information (in-person) better synchronized individuals' positive facial expressions than mediated visual information (video communication) at lag0.		
Corrugator	Dyad	PC1 × Lag effect (trend difference) PC1 × Closeness group effect	---
	Individual	PC1 main effect	---
	Summary: Visual information elicited relatively more (or smaller reduction from baseline in) corrugator activity and helped better maintain its synchrony across longer time lags than the non-visual-based communication.		
Skin Conductance	Dyad	---	---
	Individual	PC1 × Time effect (trend difference)	---
	Summary: Audio-only communication is better at maintaining individuals' excitement over time than visual-based communication.		
RSA	Dyad	PC1 × Agreeableness group effect	PC2 × Lag effect (trend difference)
	Individual	PC1 × Time effect (sig. at the beginning)	PC2 × Time effect (trend difference)
	Summary: Visual information is less mentally challenging when the conversation starts, and embodied visual information (in-person context) allows for a consistently lower level of cognitive effort which can also be more synchronized across time lags than contexts without embodied visual information (video communication). Finally, for dyads made up of people who are very agreeable, audio-only communication had stronger synchronization in cognitive effort than those in visual-based communication.		

Note. "---" refers to no significant effect associated with the tested comparison PC1 or PC2.

Figure S5

Raw Physiological Activities by Communication Systems (Individual Level)



Supplementary Section 10 – Analyses of Speech Content and Turn Taking

We conducted additional analyses on speech content and turn taking to better understand our results. Specifically, for speech content, we analyzed the emotional dimension of speech text to see how it speaks to the current findings from physiology. The reason why the emotional dimension was chosen for analysis is that one of the study's focuses is emotional processing, thus analyzing the emotional dimension of speech content may shed light on the results found in the study, particularly for the results from OO and corrugator data. Specifically, we first transcribed participants' conversations to text, which produced data for 66 dyads out of the total 75. Recordings for the rest seven dyads were either unavailable or had poor audio quality. For speech content, we used the popular sentiment analysis tool, LIWC-22 version (https://www.liwc.app/help/liwc), to extract the affective values of the speech content from each turn. Affective variables offered by LIWC-22 include *affect*, *positive tone*, *negative tone*, *positive emotion*, and *negative emotion* (see data on OSF). Then we built multilevel models with Time*Condition as the fixed effects and Dyad_ID as the random factor and tested their effect on the five emotional indicators, respectively. Results showed no significant differences across the three conditions, either over time or at the aggregated level.

For turn taking, we think it may provide insights into the synchrony results and thus we compared turn taking differences across the communication conditions. Specifically, we tested differences in the total of turns each dyad took, turn difference, and word count difference between two people (which may suggest if there is a dominant role within dyads). Likewise, there were no significant findings from the above analyses (all with $p > .05$, see data on OSF).

The contrast between the findings from physiological data and the non-significant findings from speech content and turn taking confirm the general consensus that interpersonal communication is largely impacted by nonverbal information. For example, interpersonal synchronization and the leader-follower structure can be impacted by patterns of facial expressions and eye contact but may not be necessarily linked to their speech behavior. Further investigations could conduct nonverbal behavioral coding and examine how those nonverbal behaviors speak to the physiological results reported in this study.

References

Abney, D. H., Lewis, G. F., & Bertenthal, B. I. (2021). A method for measuring dynamic respiratory sinus arrhythmia (RSA) in infants and mothers. *Infant Behavior and Development*, *63*, 101569.

Beauchaine, T. P. (2015). Respiratory sinus arrhythmia: a transdiagnostic biomarker of emotion dysregulation and psychopathology. *Current Opinion in Psychology*, *3,* 43–47. https://doi.org/10.1016/j.copsyc.2015.01.017

Eiler, B. A., Kallen, R. W., Harrison, S. J., & Richardson, M. J. (2013). Origins of order in joint activity and social behavior. *Ecological Psychology*, *25*(3), 316-326.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528.

Han, J. (2020). *Physiological and emotional synchrony during television viewing: A nonlinear dynamic systems approach* (Publication No. 28002093) [Doctoral dissertation, Indiana University]. ProQuest Dissertations Publishing.

Han, J., Lang, A., & Amon, M. J. (2022). Can media synchronize our physiological responses? Skin conductance synchrony as a function of message valence, arousal, and emotional change rate. *Communication Monographs*, *89*(1), 47-69.

Ham, J., & Tronick, E. (2009). Relational psychophysiology: Lessons from mother–infant physiology research on dyadically expanded states of consciousness. *Psychotherapy Research*, *19*(6), 619-632.

Huston, T. L., McHale, S. M., & Crouter, A. C. (1986). When the honeymoon's over: Changes in the marriage relationship over the first year. In Gilmour, R. & Duck, S (Eds.). *The emerging field of personal relationships* (pp.109–132). Lawrence Erlbaum Associates.

Kay, B. & Richardson, M. J. (2015). Recurrence quantification analysis conducting using MATLAB code for the APA Advanced Institute of Nonlinear methods for Psychology Science. https://github.com/xkiwilabs.

Müller, V., & Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PloS one*, *6*(9), e24893.

Porges, S. W. (1985). *U.S. Patent No. 4,510,944*. Washington, DC: U.S. Patent and Trademark Office.

Porges, S. W. (1991). Vagal tone: An autonomic mediator of affect. In J. Garber & K. A. Dodge (Eds.), *The development of emotion regulation and dysregulation* (pp. 111–128). Cambridge University Press. https://doi.org/10.1017/CBO9780511663963.007

Porges, S. W. (1995). Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory. *Psychophysiology*, *32*(4), 301-318.

Porges, S. W. (2007). The polyvagal perspective. *Biological psychology*, *74*(2), 116-143.

Ravaja, N. (2009). The psychophysiology of digital gaming: The effect of a non co-located opponent. *Media Psychology*, *12*(3), 268-294.

Read, G. L. (2020). Processing Ambiguous Social Identity: Disclosure of identity and phenotypic prototypicality affect processing and evaluation of persuasive messages. *Journal of Communication*, *70*(5), 693-720. https://doi.org/10.1093/joc/jqaa024

Richardson, M. J., Schmidt, R. C., & Kay, B. A. (2007). Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. *Biological Cybernetics*, *96*(1), 59–78. https://doi.org/10.1007/s00422-006-0104-6

Riniolo, T. C., & Porges, S. W. (2000). Evaluating group distributional characteristics: Why psychophysiologists should be interested in qualitative departures from the normal distribution. *Psychophysiology*, *37*(1), 21-28.

Shockley, K. (2005). Cross recurrence quantification of interpersonal postural activity. In M. A. Riley & G. C. Van Orden (Eds.), Tutorials in contemporary nonlinear methods for the behavioral sciences (pp. 142-177). National science foundation. http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp

Vangelisti, A. L., & Caughlin, J. P. (1997). Revealing family secrets: The influence of topic, function, and relationships. *Journal of Social and Personal Relationships*, *14*(5), 679–705. https://doi.org/10.1177/0265407597145006

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020.