ORIGINAL PAPER

Nonverbal Synchrony in Technology-Mediated Interviews: A Cross-Cultural Study

Ken Fujiwara[1] [iD] · **Christopher D. Otmar[2]** · **Norah E. Dunbar[2]** · **Mohemmad Hansia[2]**

Abstract

Technology-mediated communication has changed the way we interact. Since the onset of the COVID-19 pandemic in March 2020, this trend became even more pronounced. Media interviews are no exception. Yet, studies on nonverbal behaviors, especially nonverbal synchrony during such mediated settings, have been scarce. To fill the research gap, this study investigated synchronized patterns between interview hosts' and guests' facial emotional displays and upper body movement during mediated interviews recorded in the countries in Western (mainly the US, with the addition of the UK) and Eastern cultures (Japan). The interviews were categorized into information- or entertainment-driven interviews, depending on the social attributes of the guest. The time series of the valence in facial displays and upper body movement was automatedly measured using FaceReader and Motion Energy Analysis software, respectively, which was analyzed in terms of simultaneous movements, a primary component of synchrony. As predicted, facial synchrony was more prevalent in information-driven interviews, supporting the motivational and strategic account of synchrony. In addition, female-hosted interviews had a higher degree of synchrony, especially in information-driven interviews. Similar patterns were seen in movement synchrony, although not significant. This study is the first evidence of synchrony in technology-mediated interviews in which a host and a guest appear on split-screen to inform or entertain audiences. However, no cultural differences in synchrony were observed. Situational demands in front of the interactants and the goal-driven nature of communication seemed to play a more prominent role than cultural differences in nonverbal synchrony.

Keywords Nonverbal synchrony · Facial expressions · Technology-mediated interview · Thin slice · Cross-Culture · Dynamic time warping

✉ Ken Fujiwara
psykf@ccu.edu.tw

1 Department of Psychology, National Chung Cheng University, No.168, Sec. 1, University Rd., Minhsiung 621301, Taiwan, ROC

2 Department of Communication, University of California, Santa Barbara, Santa Barbara, USA

Introduction

Nonverbal Synchrony in Technology-Mediated Interviews: A Cross-Cultural Study

Interpersonal communication is fundamental in our social life. Although face-to-face interaction has played an essential role, technology-mediated communication changed the way we interact. Since the onset of the COVID-19 pandemic in March 2020, this trend became even more pronounced. To prevent the spread of the virus, communication through computer- mediated channels have become a central component of everyday life (Choi & Choung, 2021; Shufford et al., 2021). Universities (Watermeyer et al., 2021) and businesses (Karl et al., 2021) have quickly adapted to the ever-evolving public health crisis, transitioning many traditionally face-to-face encounters to mediated contexts. Media interviews are no exception. Talk shows are one of the most ubiquitous and controversial types of television programming worldwide (Tolson, 2001). These broadcasts cover many events and topics, from news shows that interview experts and political figures about current events, to entertainment shows that interview actors and authors about their latest projects. In most cases, guests are interviewed by talk show hosts in the studio, which is recorded. However, during the COVID-19 era, many talk shows have started using a mediated environment through videoconferencing software and subsequently publishing the interviews online.

In this trend, studies on nonverbal behaviors, especially interactional or nonverbal synchrony during such mediated interviews, are still scarce. The mediated format created a unique opportunity to analyze synchrony for two reasons. First, the mediated interviews are real phenomena that took place outside the laboratory, yet are standardized in camera angles. Interactions that occur outside the laboratory ensure ecological validity, but the video footage tends to be inconsistent, which could make synchrony analysis unreliable. In this regard, since the mediated interviews often have just one camera each, they only capture faces and not the entire body, and guests and hosts face forward to speak to the camera on a split-screen, which even enables them to monitor the interview partner (and themselves).[1] Thus, this study can examine whether the synchrony phenomenon is still evident outside the laboratory. Second, the availability of these shows on the internet rather than on local or cable television allows for a cross-cultural examination that would ordinarily be difficult to conduct. The COVID-19 has spread throughout the world, which encourages TV programs in various countries to adopt the mediated format in a similar (standardized) way. To our best knowledge, no studies seem to have directly investigated cross-cultural differences of synchrony using the same analysis method. Synchrony itself should be a robust and ubiquitous communication pattern (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016), but it seems not conclusive which factors drive/ prevent synchrony during the interview and whether it differs across cultures.

---

[1] Although there is no way to confirm that each media company provided the host and the guest with a monitor to see the other person, it is highly likely because they responded to one another's gestures naturally, as if they could see them. In fact, it would be strange if a media company (such as CNN or ABC) did not provide their host with the opportunity to see the guest, which makes the conversation less smooth (not knowing when the appropriate time to speak is nor when the other person wants to jump in with a comment). Overall, because these are professional media companies that recorded the interviews, it is more likely that they could see each other. It would be the rare exception if they could not see.

Interpersonal Coordination and Synchrony

Bernieri and Rosenthal (1991) defined interpersonal coordination as "the degree to which the behaviors in an interaction are non-random, patterned, or synchronized in both timing and form" (p. 403). Although scholars still debate the underlying reason people tend to synch up during interactions, many believe there is an evolutionary basis for these behaviors (Lakin et al., 2003). Indeed, coordinated communication patterns during an interaction are broadly observed in several nonverbal behaviors such as head nods/shakes, hand gestures, and postural sway (Dunbar et al., 2020), and they can also help promote affiliative social dynamics and rapport among the speakers (Bernieri et al., 1994; Fujiwara et al., 2020; Tickle-Degnen & Rosenthal, 1990). The "social glue" function of coordination has been confirmed in several different meta-analyses (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). Interactants can also strategically adapt, or accommodate, their communication behaviors by converging or diverging from their conversational partner to accomplish their goals (Knoblich et al., 2011).

Nonverbal synchrony (or termed *interactional* synchrony) is one of the major facets of interpersonal coordination (Bernieri & Rosenthal, 1991). Most of the previous works consistently emphasize that synchrony involves a temporal component, rhythmic coordination of behavioral patterns (Bernieri, 1988; Burgoon et al., 1995). Delaherche et al. (2012) define synchrony as "the dynamic and reciprocal adaptation of the temporal structure of behaviors between interactive partners" (p. 351). As a neighbor of the synchrony phenomenon, behavioral matching and mimicry is the similarity in body postures between individuals (LaFrance & Broadbent, 1976). Both are considered to be the unintentional tendency to imitate someone else's behavior at a particular moment in time (for a review, see Chartrand & Lakin, 2013, 2013; Lakin, 2013). Mimicry analysis examines whether the same or similar behavior is observed at a given point in time while behaviors can differ because synchrony analysis focuses on the time process of nonverbal behavior (Chartrand & Lakin, 2013, 2013; Fujiwara & Daibo, 2022). However, it is not always possible to distinguish between synchrony and matching/mimicry based solely on whether the behavior is the same or different because synchrony studies sometimes require participant pairs to engage in the same repetitive task such as stepping (Miles et al., 2010) and arm curl (Miles et al., 2011). Even in such cases, it should be noted that the focus of synchronization studies is still on temporal (timing and rhythmic) convergence.

Similar to various forms of full-body synchrony (Dunbar et al., 2014; Fujiwara et al., 2021), facial expressions can also be synchronized (Hess & Bourgeois, 2010; Riehle et al., 2017). It should be noted that *emotional mimicry* (Hess & Fisher, 2014) is quite relevant because as is in synchrony research, it emphasizes temporal proximity between behaviors. In terms of emotional mimicry, matching emotional expressions are time-locked and occur shortly after each other, usually within a second. However, this study will use the term "facial synchrony" or synchrony in facial displays, not "emotional mimicry" because this study does not investigate the co-occurrence of categorized (facial) behaviors such as smiling and frowning, which is typically assumed in mimicry studies (Chartrand & Lakin, 2013, 2013; Lakin, 2013), nor specific emotions. Instead, we will perform a time series analysis of negative-positive affective valence in facial displays (see details in Method). Thus, although it could depend on the definition of "matching nonverbal expressions of emotion" (Hess & Fisher, 2014, p. 46), it would be better not to use the term mimicry to avoid unnecessary confusion.

Automated Coding of Nonverbal Signals

Research on nonverbal synchrony has advanced with the development of technology. The use of video images forms the basis of a rigorous approach for synchrony, however, in early methods, manual coding has been the gold standard (see Fujiwara et al., 2021). Human observers carefully watch video recordings of interactions and performed the temporal coding of specific actions to evaluate movement changes in the form of initiations and terminations of body part movements or vocal activity to judge whether temporal co-occurrence of actions was present (Schmidt et al., 2012). For example, Newtson and colleagues (Newtson et al., 1977, 1987) placed a transparency over a still frame on a video screen and located 15 different body parts at 1.0 or 0.5 s intervals, and tallied the number of changes per frame toestablish a time series. Although a groundbreaking method at the time, this technique was very laborious and not widely adopted by other researchers. Schmidt et al. (2012) argue these types of coding methods, in addition to being difficult to employ, provide a rather coarse grain view of synchrony because they are limited in the number of behaviors that can be measured.

An alternative that has been used is a rating approach in which third-party raters watching a video make gestalt judgments about synchrony at regular intervals (see Bernieri et al., 1994, for an example). In Bernieri (1988), a group of untrained judges rated the interactional synchrony from the perspective of simultaneous movement, tempo similarity, and general smoothness. The rating approach could save the coder's time and effort, but it would be still laborious in the frequent human judgments to be made. To overcome the labor-intensive nature of manual coding, researchers have had to resort to focusing on the segmented clip (Murphy & Hall, 2021). A previous study demonstrated nonverbal behaviors in a "thin-sliced" segment could represent nonverbal behavior throughout the interaction (Murphy et al., 2015). Predictive validity is also confirmed (Murphy et al., 2019), showing that thin-slice coding or rating is also correlated with outcomes of interest reported after the interaction (Ambady & Rosenthal, 1992).

Another solution is using automated coding techniques enabled by advances in computer modeling and vision analysis (see Delaherche et al., 2012, for an introduction). Regarding facial emotional expressions, FaceReader (Noldus) is one of the best-regarded automated systems for facial expression analysis, whose validity was confirmed (see Lewinski et al., 2014). This study used the FaceReader software because it recognizes several specific properties in facial images, including discrete emotional expressions (e.g., happy, sad, angry, fear, disgust, and surprise) as well as an integrated measure (i.e., valence). Affectiva Affdex (iMotions) could be another emotion software, which offers comparable performance to the facial EMG technique (Kulke et al., 2020). Not only commercial software but also a variety of open software is offered to the public by researchers. OpenFace (Baltrušaitis et al., 2016), for example, analyzed the activation of facial Action Units (AUs), which are publicly available for free. Such technologies cover bodily movement (Motion Energy Analysis [MEA], Ramseyer & Tschacher, 2011; Ramseyer, 2020) as well as human posture (OpenPose, Cao, et al., 2017; Fujiwara & Daibo, 2022). Although scholars studying nonverbal synchrony have struggled with laborious manual coding for decades (see Fujiwara et al., 2021), now they have benefitted greatly from advances in automated coding techniques, which are comparable to manual observation (Fujiwara et al., 2021). The great advantage of the automated coding technique is not only its cost efficiency but also its high reliability.

If the same software was used on the same video footage, everyone will always get the same result. It is a powerful tool to find out what nonverbal behaviors occurred during the interaction.

In this study, the valence of the hosts' and guests' facial emotional displays during mediated interviews was measured by FaceReader software. In a laboratory setting, the electrical activity of facial muscles can be recorded, which is submitted to a particular form of time series analysis (Riehle et al., 2017). However, provided our interest in mediated interviews, we instead employed a video analysis technique to acquire time series data of facial displays, which is examined as a primary focus of the hypothetical testing in this study because, in interviews in which the host and guest are filmed from the chest up, facial displays and their synchrony are considered to play a prominent role. Moreover, we explored movement synchrony in our analyses to investigate whether the results are specific to facial synchrony or whether they are applicable to another type of nonverbal signal (i.e., upper body movements). Bodily movements were quantified using MEA.

Time Series Analysis for Synchrony

Those tools offer time series data of the targeted nonverbal cue(s) from video images. Once the time series is obtained from two speakers, they are analyzed in terms of synchrony. Synchrony is computed as the convergence of timing and rhythm, not as the similarity of "behavior" that is categorized and counted/rated via manual coding. Now, different analysis methods are proposed for each component of synchrony. For the convergence of timing, cross- correlation is one of the most common approaches (Ramseyer & Tschacher, 2011; Schoenherr et al., 2019). Cross-correlation is a simple extension of Pearson's correlation to time series data, which can include several time lags (e.g., Schoenherr et al., 2019). Cross-recurrence quantification analysis (CRQA; Coco & Dale, 2014) is a non-linear method that extracts co- visitation patterns between two systems (i.e., how one time series revisits a state that another time series has visited), which can also capture synchronous patterns occurred in varying lags throughout the conversation. Dynamic time warping (Berndt & Clifford, 1994) is also a non-linear method, which calculates a distance between two time series using the "warping" sequences in the time dimension. The distance score is considered as an inverted metric of synchrony (i.e., similarity) (Fujiwara et al., 2022; Van Der Zee et al., 2021). Although each method holds different assumptions as well as parameters to be determined prior to the analysis, in general, they capture the convergence of timing in two speakers' movements.

As for the convergence of rhythm, another property of synchrony, spectrum analysis is a promising option. Spectrum analysis is an analysis technique for time series signals in the frequency domain, which deconstructs complex time series into rhythmic components. This technique calculates a spectral power, the magnitude at each component frequency. Moreover, a cross-spectrum analysis can provide a coherence measure if there are two time series. Coherence, which ranges on a scale from 0 to 1, is a metric of similarity between the two time series at each frequency component. A coherence of 1 reflects a perfect rhythmic match between the two movements, whereas 0 reflects no match. As an option of spectrum analysis, the cross-wavelet transform has been considered a powerful approach for unstructured interactions because it does not require constant properties (i.e., stationarity) within time series (Fujiwara & Daibo, 2022; Fujiwara et al., 2020, 2021).

In this study, we used dynamic time warping due to the feature of the segmented video images (see also Method) because the time series obtained seemed too short to capture the

rhythmic components. Also, cross-correlation and CRQA were excluded from consideration because they have several parameters to be determined prior to the analysis. It was not easy to find the optimal parameters since there was no clear assumption regarding the time series signal in this study.

Context-based Nature of Synchrony

Communication behaviors, in general, shift depending on the social and relational context and communication goal of the conversation (Berger & Palomares, 2011), as is the case for nonverbal synchrony. For example, the greater extent of synchrony has been observed in well-formed relationships such as child-parent (Bernieri et al., 1988), teacher-student (Bernieri, 1988), and friends (Fujiwara et al., 2020). On the other hand, individuals do not try to synchronize their movement with those who look dishonest (Brambilla et al., 2016) or were lazy and late to the experiment (Miles et al., 2010). However, somewhat surprisingly, previous studies that used group membership, an important factor in social relationships, provided mixed evidence. For facial expressions, the expresser's group membership had less or no impact on mimicked expressions (Bourgeois & Hess, 2008). Miles et al. (2011) used arbitrary group membership (i.e., minimal groups) and asked their participants to perform a repetitive arm curl task together with a member of the same or a different minimal group, which showed that synchrony was most pronounced when they interacted with a member of a different minimal group. This implies that group membership itself may not play a significant role in synchrony. Rather, the communication goal for diminishing minor interpersonal differences toward the out-group member results in the greater extent of synchrony that should serve as a means to reduce social distance.

The goal-based account for synchrony has been confirmed by previous studies. Individuals who have prosocial goals show greater synchrony (Lumsden et al., 2012). More specifically, having an affiliation goal increases coordinated movement (Lakin et al., 2003). On the contrary, competitive goals associated with a particular social context hamper synchrony. In Weyers et al. (2009), the participants primed for competition exhibited less facial synchrony. The same is true for the full-body synchrony in a debate (Bernieri et al., 1994) and an argument (Paxton & Dale, 2013). It is then expected that different types of interviews could entail different goals, then increase or decrease synchrony.

In this study, we utilized two different types of underlying interview goals: to inform an audience and to entertain an audience. An example of the former from the database created in this study (see Method), is Judy Woodruff, a PBS Newshour anchor interviewing California Governor, Gavin Newsom about the COVID-19 outbreak. The primary purpose of this conversation was to inform the audience about what factors are involved in the decision to "reopen" the state. In such an interview, the credibility of the source (i.e., the interview guest) must be essential because the information provided during the interview has social value. Thus, the guest is supposed to be knowledgeable and/or responsible professionals so that trustworthy (Hovland & Weiss, 1951), such as politicians, medical doctors, and scientists. On the other hand, an exemplar entertainment interview stated as entertainment-driven interview later, is Jimmy Fallon interviewing Dolly Parton about a recent holiday album she recorded. The underlying goal of this interview was to promote a product and share stories between the host and the guest (as evidenced through the content of the conversation). Indeed, there could be exceptions, such as non-professional entertainers providing important information (e.g., a comedian describes how they recovered from the COVID-19 and recommend daily infection control measures). However, the expertise

attributed to the source, compared to a story of one's personal experience, seems to be more influential on how the interview host behaves.

When disseminating information to an audience, credibility is a key attribute. Given that the media bias effect revealed that how the interviewer behaves had a significant impact on viewers' impressions of the interviewee (Babad & Peer, 2010; Tikochinski & Babad, 2022), matching a conversational partner's professional demeanor may be a promising way for the host to ensure the guest's credibility. In addition, if the host is speaking with a guest expert for the first time, they would try more to build rapport to encourage ease of conversation with the conversational partner. Thus, the affiliative goal could be more pronounced in information-driven interviews rather than entertainment-driven interviews, which is supposed to drive synchrony. Given this, the hypothesis regarding the impact of the type of interview on synchrony is as follows:

H1. Greater synchrony will occur in information-driven interviews compared to entertainment-driven interviews.

Meanwhile, as for the valence of facial displays, a different prediction can be made. In entertainment-driven interviews with a non-professional guest, having a positive and interpersonally warm interaction should be required. It is more likely that the facial expressions

of hosts and guests become positive there, whereas it is noted that the preexisting lighthearted nature in entertainment-driven interviews will not necessarily promote simultaneous facial displays (i.e., synchrony). As such, the hypothesis regarding the impact of the type of interview on the valence of facial displays is as follows:

H2. In entertainment-driven interviews compared to information-driven interviews, the valence of facial displays is more positive.

Gender Effect of Synchrony

Previous studies have found that synchrony in social situations is associated with the speaker's sociality (Brambilla et al., 2016; Lumsden et al., 2012). Taking this view, it seems that females should show a greater degree of synchrony since previous studies have demonstrated that females have more social motives compared to their male counterparts (e.g., Costa et al., 2001; Feingold, 1994). Indeed, Fujiwara et al. (2019) showed that females exhibited greater synchrony in their face-to-face unstructured conversation. Although same-gender dyads with no particular roles for each of the communicators in conversation were investigated in the previous study, this study focuses on the interview host. Tickle-Degnen (2006) suggests that interpersonal coordination during a social interaction demonstrates respect toward the other person and can help build rapport (i.e., the coordination-rapport hypothesis). Thus, interview hosts will likely play a primary role to facilitate the conversation and try to coordinate one's facial displays to build rapport with their conversational partners. The hypothesis regarding the gender effect on synchrony is as follows:

H3a. Female-hosted interviews will have a higher degree of synchrony than male-hosted interviews.

While H3a hypothesizes the main effect of the host's gender, an interaction effect with the type of interview can also be expected because of female's greater likelihood to respond to tense situations in a more affiliative manner (Bikmen et al., 2022; Hall & Halberstadt, 1986; Taylor et al., 2000). Given the information-driven interviews entail expertise and hence tension, female hosts will be more motivated to coordinate their facial displays in order to overcome such a situation. Thus, the interaction term of the host's gender and the type of interview is hypothesized as follows:

H3b. Female-hosted interviews will have a higher degree of synchrony especially in information-driven interviews.

Cultural Differences in Synchrony

To date, no studies seem to have investigated cross-cultural differences in synchrony using the same analysis method. Synchrony is believed to be a robust and ubiquitous communication pattern (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016), and there seems an evolutionary basis for these behaviors (Lakin et al., 2003). In other words, there is no strong rationale to predict cultural differences in synchrony. However, although synchrony itself is a robust phenomenon, it seems not conclusive whether or not factors that drive/prevent synchrony during the interview differ across cultures. In terms of this, cross- cultural comparison studies have a certain significance in the research field.

Besides, a number of cultural differences have been identified in facial expressions. For two reasons, the Japanese sample seems to be a good target for this study. First, the manner of facial expression is vastly different between Americans and Japanese, yet, facial synchrony is still present (Tamura & Kameda, 2006). Japanese facial expressions tend to be restricted in general (Matsumoto, 1990; Matsumoto et al., 1998), and especially, negative emotions such as anger, sadness, fear, and disgust, which are seldom expressed in public situations (Inamine & Endo, 2009). The mediated interview of this study is a public activity such that the difference in display rules between Japan and the Western countries (e.g., the US, the UK) will be more salient.

Second, more importantly for this study, Japanese facial expressions are sensitive to situational influences (Matsumoto & Ekman, 1989). For instance, Japanese people mask their true emotion in public displaying an intentional smile if a situation encourages social harmony (Ekman, 1972). In addition, as E. T. Hall (1976) described as high context culture, the relational history and common backgrounds between speakers are usually incorporated in their communication, which implies that their facial expressions will be adjusted according to the relational and situational requirements. Therefore, we reasoned Japanese interviews would be a good contrasting sample for this study. As a research question, it is explored whether or not the difference in culture (the US and UK, Japan) has a moderating effect on synchrony.

RQ. Does the culture of the interview context have a moderating effect on synchrony?

Method

Sample Size Calculation

No previous studies exist on synchrony in mediated interviews, so there is no effect size to which we can directly refer. Thus, as for H1, we followed Miles et al. (2011) who demonstrated the situational effect of the different minimal group membership to facilitate synchrony, with the effect size η_p^2 of 0.21. Regarding the gender effect (H3a, H3b), Fujiwara et al. (2019) provided the effect size η_p^2 of 0.15 for the main effect of gender. Given these effect sizes, we performed the power analysis using the "pwr.anova.test" function of the *pwr* package in R to calculate the necessary sample size to examine the main effect of the type of interview and host's gender. More specifically, the parameter was set as $k = 2$,

power = 0.80, sig.level = 0.05, and f = .516, .420, respectively [2]. The results suggested that each cell should secure 24 samples at least. Since the cross-cultural analysis is exploratory, we just targeted 100 samples (about 25 samples in each 2 by 2 cell) in each culture.

Target Interview Videos

We collected 178 mediated interviews (116 from the US, 2 from the UK, 60 from Japan) mainly from YouTube between April 2020 to March 2021.[3] Each interview was between 1 to 69 minutes long (Median = 9 minutes and 50 seconds) and had only one interview host and one interview guest on the screen. Additionally, the interview must have been conducted using video conferencing technology in which each person was on the split-screen looking at the camera for most of the interview.

Research assistants who were blind to the study's hypotheses coded each video for the type of interview (i.e., informative or entertainment) depending on the social attributes of the guest. For example, the interview was coded as an information-driven interview when the guest was an expert or professional such as a government official, medical doctor, or scientist.

Conversely, if the guest was an artist, athlete, comedian, etc., the interview was coded as an entertainment-driven interview. All the videos were coded as either information-driven or entertainment-driven interviews. The US and UK videos were classified by American RAs, and Japanese videos were categorized by a Japanese RA. Since only one Japanese RA could be employed, one of the authors (Japanese) independently classified the data in advance, and the results were compared with the classification of Japanese RA, which was 100% matched. The gender of each interactant was also coded.

Collecting Time Series

A synchronized time series of facial valence and bodily movement between the guest and host allowed us to measure facial and movement synchrony, respectively. However, in collected video clips, they were not always split into a single screen (e.g., appearing on the screen one by one). Thus, the "thin slice" technique was employed for the split-screen, which can represent the entire interaction even with a short, segmented clip (Murphy & Hall, 2021; Murphy et al., 2015). Since more clips can be included in the analysis, the length of the sliced segment was set to 20 seconds. At first, we extracted all the start and end points of the split-screen time during the interview from all the videos. Then, from these, one was selected for each interview based on a random number from among the chunks that were longer than 20 seconds. Within the chunk, the analyzed portion was selected using another random number.

Valence of Facial Displays

The facial displays of the host and guest were cropped from the segmented clip, respectively, because the FaceReader software only analyzes a single person at a time. In this

---

[2] $f = \sqrt{\eta_p^2/(1 - \eta_p^2)}$

[3] In the video database, only one Japanese host-guest pair appeared twice. The data obtained from them were averaged and used in subsequent analyses.

study, as an integrated measure, the valence time series was obtained. As for the missing values (mainly due to not facing forward towards the camera), we applied spline interpolation using the "na.spline" function of the *zoo* package in R. It is noted that among the videos we collected, there were several in which the FaceReader software could not analyze valence mainly because the filmed face was too small. Thus, as a result, 151 clips (94 from the US, 2 from the UK, 55 from Japan) were used in the subsequent facial synchrony analysis.

Movement

Using the segmented clips, the host's and guest's (mainly upper) body movements were quantified using MEA software. Split-screen showing host and guest each was covered as a region of interest (ROI), in which the software calculates the change in greyscale pixels between consecutive video frames. The raw value for MEA was used in the subsequent analysis.[4] For movement synchrony, 159 clips (101 from the US, 2 from the UK, 56 from Japan) were used since several clips cannot be analyzed because the camera and/or background were moved.

Synchrony Analysis

To compute facial synchrony for H1, H3a, and H3b, we performed the dynamic time warping using the "dtw" function of the *dtw* package in R. The default parameters of the function were employed, and no locality constraints were added. The normalized distance score, the inverse to the amount of synchrony, was obtained for each host-guest dyad. As for the valence score for H2, the average score during the interaction was calculated for the host and guest, respectively. Then, they are further averaged to represent the valence score of the pair. Regarding movement synchrony, the dynamic time warping was also performed for the movement time series in the same manner.

For the synchrony measures obtained, as is in previous studies (e.g., Fujiwara et al., 2022), we first confirmed that it was not the product of chance. More specifically, we created artificial interactions using data shuffling within a time series, which is known as surrogate data (Moulder et al., 2018). That is considered as a time series equivalent of a randomization/permutation test since all of the time-dependent properties in the original series are destroyed. Still, even after the shuffling, the time-independent information (e.g., mean, variance) representing the entire series remains the same. The rationale of this technique is to determine a baseline synchrony level so that researchers can investigate the level of synchrony in genuine dyadic interactions. In this study, each data point of time series in each pair (H, G) was randomly shuffled to create a new surrogate time series (H^s, G^s). The distance score between H^s, G^s was computed via dynamic time warping to compare that in the genuine interaction (i.e., H, G). Before testing the hypotheses, we ensure that the distance of the genuine interaction is significantly smaller than that of the surrogate data, confirming that the host and guest in the interview exhibited synchrony beyond chance.

---

[4] Nine of the videos analyzed (4 from the US, 5 from Japan) had a different size of split-screen for the host and guest; i.e., the size of the guest was larger than the host. Although the raw value for MEA could depend on the screen size, it was confirmed that the results remained the same even after removing the nine interviews. Thus, this study reports the results of the data that includes them.
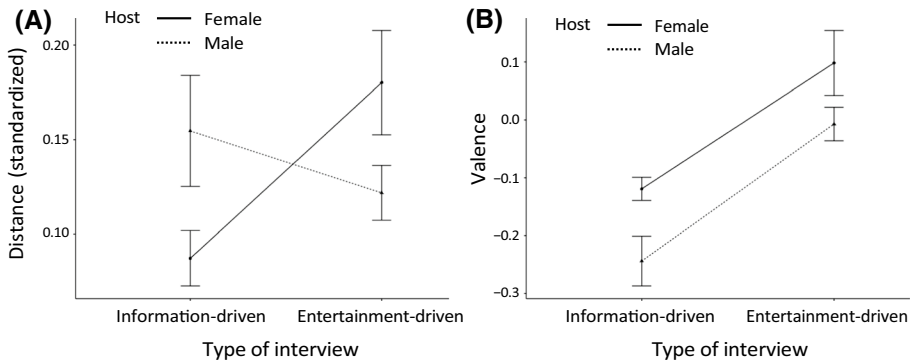
Fig. 1 A Two-way interaction of the type of interview and host's gender on the distance. **B** Main effect of the type of interview on the facial valence

Results

To begin, since we have taken 20-second segments from interviews of different time lengths, we examined whether the overall interview time was related to the DVs (i.e., facial synchrony, valence, and movement synchrony). The results showed that the time length of the interview was not significantly correlated with facial synchrony ($r = -.070$, $p = .379$), valence ($r = -.047$, $p = .556$), and movement synchrony ($r = -.058$, $p = .466$).

To test the existence of synchrony, the distance measure was compared between the genuine and surrogate data. The result of the paired t-test showed that the distance of facial valence was significantly smaller in the genuine data ($M = 0.132$, $SD = 0.129$) rather than the surrogate data ($M = 0.140$, $SD = 0.111$), $t(150) = 2.27$, $p = .025$, $d = 0.185$. The same was true for the bodily movements; the distance measure was significantly smaller in the genuine data ($M = 660.973$, $SD = 694.350$) compared to that in the surrogate data ($M = 756.356$, $SD = 728.436$), $t(158) = 7.27$, $p < .001$, $d = 0.577$.

Hypothesis Testing for Facial Synchrony

Since it was confirmed that the host and guest in the interview exhibited synchrony beyond chance, the distance score was further submitted to a 2 (interview type: information- driven, entertainment-driven) by 2 (host's gender: female, male) by 2 (culture: the US and the UK, Japan) between-subjects 3-way ANOVA (Table 1). Regarding H1, the results yielded a significant main effect of interview type in which greater synchrony occurred in information- driven interviews compared to entertainment-driven interviews (Table 2). In addition, as for H3a and H3b, not the main effect of host's gender, but the 2-way interaction of interview type and host's gender was also significant. The simple effect analysis revealed that female hosts showed greater synchrony (smaller distance) in the information-driven interview, compared to the entertainment-driven, $F(1, 143) = 8.62$, $p = 0.004$, $\eta^2 = 0.055$, $\eta_p^2 = 0.057$ (Fig. 1A). Conversely, male hosts did not show the significant difference between both types of interview, $F(1, 143) = 0.02$, $p = 0.899$, $\eta^2 = 0.0001$, $\eta_p^2 = 0.0001$. In information-driven interviews, female hosts exhibited greater level of synchrony, which was significant, $F(1, 143) = 4.57$, $p = 0.034$, $\eta^2 = 0.029$, $\eta_p^2 = 0.031$. In the entertaining interview, there was no significant effect of the host's gender, $F(1, 143) = 1.34$, $p = 0.250$, $\eta^2 = 0.009$,

Table 1 Facial Synchrony and Valence in Mediated Interviews

	Host	Information-driven			Entertainment-driven		
		N	M	SD	n	M	SD
Western interview							
Dist	Female	27	0.091	0.093	12	0.212	0.211
	Male	25	0.168	0.172	32	0.107	0.083
Val	Female	27	− 0.093	0.139	12	0.079	0.213
	Male	25	− 0.152	0.166	32	0.034	0.296
JP interview							
Dist	Female	7	0.071	0.043	18	0.159	0.095
	Male	5	0.089	0.056	25	0.141	0.136
Val	Female	7	− 0.220	0.125	18	0.077	0.183
	Male	5	− 0.155	0.081	25	0.002	0.175

Note. Dist. = Distance (inverted measure of facial synchrony), Val. = Valence of facial expressions.

Table 2 ANOVA results of the facial synchrony

	df	Sum Sq	Mean Sq	F	p	η^2	η_p^2
Type	1	0.063	0.063	4.000	.048	.026	.027
Gender	1	0.001	0.001	0.084	.772	.001	.001
Culture	1	0.022	0.022	1.400	.239	.009	.010
Type × Gender	1	0.075	0.075	4.745	.031	.030	.032
Type × Culture	1	0.010	0.010	0.637	.426	.004	.004
Gender × Culture	1	0.001	0.001	0.083	.774	.001	.001
Type × Gender × Culture	1	0.034	0.034	2.140	.146	.014	.015
Residuals	143	2.260	0.016				

Note. Type: Type of interview, Gender: Host's gender.

$\eta_p^2 = 0.009$. As for RQ, the 3-way interaction was not significant, and any interaction terms including culture were not significant. Thus, culture did not moderate the effects of synchrony.

Regarding the valence score ($M = −0.034$, $SD = 0.218$), it was firstly found that it was not significantly correlated to the distance measure ($r = .127$, $p = .122$). Then, to examine H2, the valence score was submitted to the same 3-way ANOVA. Only the main effect of interview type was significant (Table 3), which showed that the hosts and guests showed more positive displays in entertainment-driven interviews (Figure 1B).

Table 3 ANOVA results of the facial valence

	df	Sum Sq	Mean Sq	F	p	η^2	η_p^2
Type	1	1.044	1.044	25.708	<.001	.149	.152
Gender	1	0.020	0.020	0.498	.482	.003	.004
Culture	1	0.043	0.043	1.054	.306	.006	.007
Type×Gender	1	0.025	0.025	0.614	.435	.004	.004
Type×Culture	1	0.015	0.015	0.366	.546	.002	.003
Gender×Culture	1	0.014	0.014	0.341	.560	.002	.002
Type×Gender×Culture	1	0.037	0.037	0.915	.340	.005	.006
Residuals	143	5.809	0.041				

Note. Type: Type of interview, Gender: Host's gender.

Table 4 Movement Synchrony in Mediated Interviews

	Host	Information-driven			Entertainment-driven		
		n	M	SD	n	M	SD
Western interviewDist	Female	27	577.440	944.583	15	386.421	246.732
	Male	26	436.052	463.304	35	1031.539	882.632
JP interview Dist	Female	7	440.440	351.998	19	554.834	366.056
	Male	5	581.455	789.408	25	789.365	507.624

Note. Dist. = Distance (inverted measure of movement synchrony)

Exploratory Analyses

Movement Synchrony

First, correlation analysis revealed that the movement synchrony was not significantly correlated to facial synchrony ($r = -.001$, $p = .990$).[5] Then, similar to facial synchrony, the distance score of bodily movement was submitted to a 2 (interview type: information-driven, entertainment-driven) by 2 (host's gender: female, male) by 2 (culture: the US and the UK, Japan) between-subjects 3-way ANOVA (Table 4). The results showed that no effect was significant, whereas the main effect of host's gender and the 2-way interaction of interview type and host's gender was relatively close to the level of significance (Table 5).

Adjustment of Multiple Entries of the Host

No same host–guest pair was included in the data analyzed, however, some hosts, with a different guest, appeared multiple times. For example, 6 interviews by Laura Ingraham were included, followed by Jimmy Fallon and Andy Katz (5 times), which could bias the results. Thus, to adjust the multiple entries, for each host, we averaged the distance measures for facial and movement synchrony and the valence, respectively. Since a cell of smaller size resulted in the newly created dataset ($N=92$), cross-cultural comparisons were

---

[5] The correlation analysis was performed using the data for facial synchrony ($N = 151$).

Table 5 ANOVA results of the movement synchrony

	df	Sum Sq	Mean Sq	F	p	η^2	η_p^2
Type	1	0.871	0.871	1.942	.167	.025	.013
Gender	1	1.275	1.275	2.843	.094	.0003	.019
Culture	1	0.007	0.007	0.016	.900	.010	.0001
Type×Gender	1	1.278	1.278	2.849	.094	.034	.019
Type×Culture	1	0.011	0.011	0.025	.875	.005	.0002
Gender×Culture	1	0.027	0.027	0.060	.806	.0004	.0004
Type×Gender×Culture	1	0.792	0.792	1.766	.186	.015	.012
Residuals	151	67.736	0.449				

Note. For ease of visibility of Sum Sq. and Mean Sq., the original movement value was divided by 1000. Type: Type of interview, Gender: Host's gender.

not conducted. Instead, a 2 (interview type: information-driven, entertainment-driven) by 2 (host's gender: female, male) between-subjects 2-way ANOVA was performed to each distance and valence measure. The results were almost the same before the adjustment; the main effect of interview type was significant on facial synchrony, $F(1, 88)=4.03$, $p=0.048$, $\eta^2=0.042$, $\eta_p^2 = 0.044$. and valence, $F(1, 88)=25.32$, $p<0.001$, $\eta^2=0.221$, $\eta_p^2 = 0.224$, which supports H1 and H2, respectively.

Although the interaction effect on facial synchrony (H3b) did not reach the level of significance presumably due to the low powered analysis ($F(1, 88) = 3.12$, $p = .081$, $\eta^2 = .033$, $\eta^2p = .034$), the simple effect for female hosts was still significant; they showed smaller distance in information-driven interviews compared to entertainment-driven interviews ($F(1, 88) = 6.42$, $p = .013$, $\eta^2 = .067$, $\eta^2p = .068$). In addition, for movement synchrony, the main effect of interview type became significant ($F(1, 88) = 5.16$, $p = .026$, $\eta^2 = .051$, $\eta^2p = .055$), in which the direction was the same as facial synchrony (H1, greater synchrony in information-driven interviews). The interaction effect was also significant ($F(1, 88) = 4.72$, $p = .047$, $\eta^2 = .005$, $\eta^2p = .055$), which showed that male hosts showed smaller distance in information-driven interviews compared to entertainment-driven interviews ($F(1, 88) = 11.08$, $p = .001$, $\eta^2 = .110$, $\eta^2p = .112$)). The detailed results were reported in Supplementary Results.

Discussion

Nonverbal synchrony is one of the fundamental communication patterns, which has been found in a wide variety of face-to-face interactions. The current study is the first evidence of synchrony in technology-mediated interviews in which a host and a guest appear on split-screen to inform or entertain audiences. Indeed, the synchrony phenomenon was captured as it occurred outside the laboratory because it was advantageous to use a mediated format that ensured a fixed-angle camera and front-facing subjects. Throughout the interview, both interactants coordinated their facial displays and beyond the chance level, which varied according to the communication goals set for specific interviews. More specifically, greater synchrony occurred in information-driven interviews compared to entertainment-driven interviews, which supports H1.

As H2 was supported, hosts and guests displayed a particular type of facial expressions such as negative-valence displays in information-driven interviews, whereas positive displays in entertainment-driven interviews. It implicates that each type of interview has its own goal and the interactants followed it during the show. In information-driven interviews, the host and guest would be oriented to inform the audience with credibility. In this regard, positive display or smiling does not always enhance credibility (Reed et al., 2018). Also, since positive displays could not fit the situational requirement, especially when dealing with serious content such as the COVID-19 pandemic, the valence of the facial display was lower in the information-driven interview. Alternately, in entertainment-driven interviews, having a positive and interpersonally warm interaction should be required. Our data successfully captured these patterns of the facial display. Also, it is noteworthy that there was no significant gender difference in the valence measure, which suggests females' greater social skills for emotional expressions within a general population (Fisher & LaFrance, 2015; LaFrance et al., 2003) cannot be generalized to a particular population such as professional communicators in this study. Female hosts, who might be supposed to display friendly expressions, did not just smile but behaved in a professional manner, which canceled out the generally assumed difference in the valence of facial displays. That also indicates that the hosts and guests were behaving in accordance with the goal of the interview.

However, more importantly, displaying specific types of facial expressions does not mean they were congruent in timing. Indeed, the results showed that the correlation between the distance and valence measures was not significant. In information-driven interviews, credibility is a key attribute, and matching a conversational partner's professional demeanor could be a promising tool to ensure credibility when disseminating information to an audience. Therefore, the affiliative goal could be more pronounced in information-driven interviews, which leads to a greater extent of synchrony. Conversely, the preexisting light-hearted nature of entertainment-driven interviews, shown in the positive-valence facial displays, did not drive interactants to further simultaneous facial displays. The results of this study seem to be explained by the contextual demands and the accompanying increased motivation, which is congenial to "planned" coordination (Knoblich et al., 2011). Note, however, that the motivational (Miles et al., 2010, 2011) and strategic account (Dunbar et al., 2020) emphasize that synchrony itself is not the intended goal nor the intentional behavior. Rather, it is believed that synchrony is chosen as one of the promising tools in various options to achieve relational or communication goals.

There could be the possibility of an alternative explanation. For example, Van Der Zee et al. (2021) suggested that synchrony is an automated process during interpersonal interaction such that it can be salient when people were under cognitive load. Van Der Zee and colleagues demonstrated that the deceivers showed greater synchrony when the interview was cognitively resource-demanding. Although this study does not deal with deception, a similar propensity in terms of cognitive load may be assumed because social interaction itself makes the conversant cognitively busy (e.g., Gilbert & Osborne, 1989). That is, the greater degree of synchrony might derive from the high cognitive load of the informational interview. However, if only the cognitive load causes synchrony, the interaction effect with the host's gender cannot be accounted for. As H3b was supported, female-hosted interviews had a higher degree of synchrony, especially in information-driven interviews. This should be because of females' greater interpersonal sensitivity (Hall et al., 2016) and greater likelihood to respond to tense situations in a more affiliative manner (Bikmen et al., 2022; Hall & Halberstadt, 1986; Taylor et al., 2000). The information-driven interviews entail expertise and hence tension, female hosts would be more motivated to coordinate their facial

displays. Given that females' greater social motivation leads to increased synchrony, the motivational account seems to explain the interaction effect better. Still, it may be worthwhile to consider which of the two accounts (cognitive load, motivational account) are more explanatory, although it is possible that both are not mutually exclusive.

Another contribution of this study was to examine the interaction effect of culture on synchrony by taking data from the countries in Western culture (mainly the US, with the addition of the UK) and Japan. Although the manner of facial expression is vastly different between Americans and Japanese (Matsumoto, 1990; Matsumoto & Ekman, 1989; Matsumoto et al., 1998), facial synchrony has been also confirmed in Japan (Tamura & Kameda, 2006). Therefore, as RQ, it seemed worth directly considering whether the difference in culture has a moderating effect on synchrony. As an evolutionary basis is assumed for coordinated behaviors (e.g., Lakin et al., 2003) and compatible patterns have been confirmed cross-culturally (Fujiwara et al., 2020, 2021), previous studies did not propose a strong rationale to predict cultural differences in synchrony. Indeed, in the current study, the results showed that the interaction effects, including the factor of culture, were not significant. The direct comparisons further deepened that understanding. Synchrony should be considered as a fundamental communication pattern, and situational demands in front of the interactants play a more prominent role than cultural differences. However, there was a difference in the sample sizes obtained, with Japan's being smaller than that of the US and UK. Thus, it is noted that further investigation should be needed to determine if there are cultural differences in synchrony.

As for movement synchrony, bodily movements were synchronized more than by chance, even under the limitation of the mediated interview that only the upper body of the host and guest was primarily filmed. It confirms that synchrony is a fundamental pattern in our communication, which was also evident even in humang-avatar interaction (Fujiwara et al., 2022). Still, the contextual effect of interview type and host's gender did not have a clear impact on movement synchrony, which suggests that situational demands could be particularly reflected in nonverbal signals that are prominent in the given circumstances (e.g., facial displays in mediated interviews). It is also interesting and somewhat surprising that the two types of synchrony measures were not significantly correlated, whereas their resulting patterns are similar to each other. That is, nonverbal synchrony occurred more in the information-driven interviews, but which signals were synchronized depended on the pair, and not all signals were synchronized at the same time. This implies that there might be an optimal level of synchrony where the host and guest felt comfortable. Indeed, research has already established that more synchrony is not always beneficial (Butler, 2011, 2015). According to Communication Accommodation Theory (Bernhold & Giles, 2020; Giles et al., 1991), overaccommodation where the interlocutors show a greater extent of coordination than expected is believed to be as worse as underaccommodation where they failed to show coordination. For some hosts and guests, showing facial *and* movement synchrony during the interview might go beyond the level of comfort. Yet, things are not so simple because the correlation between the two types of synchrony was not even negative. It means that they did not give priority to synchrony in either facial expressions or movements and ignored the other. Some (not all) hosts and guests might avoid overaccommodation, which made the correlation less clear. This study could not pursue the exact mechanism of the non-significant correlation, but it would be an interesting question as to which signals are to be synchronized in each dyad and whether there is an intentional or strategic choice involved.

Limitations and Future Directions

One of the main drawbacks of this study stems from one of its strengths—these were previously recorded media interviews. This provided an excellent representation of naturally occurring mediated conversations, but we could not manipulate the conditions of the conversations. For instance, the overall length of the interviews differed considerably though the thin-sliced time segment analyzed was the same. The overall length of the interview was not significantly correlated to the other DVs (i.e., facial synchrony, valence, and movement synchrony), however, whether or not rehearsals and careful preparation are required may vary depending on the length of the interview. The lack of control in prior preparation and pre-existing relationships between the host and guest may be considered a trade-off for the high ecological validity of this study. Yet, one particularly noteworthy finding in our study was that the goal-driven nature of communication played an important role in nonverbal synchrony, and this should be interpreted collectively with previous findings (e.g., Bernieri et al., 1994; Paxton & Dale, 2013). We believe this study will help social scientists and communication researchers derive a deeper understanding of social interactions and provide additional exigency to study how nonverbal synchrony differs based on the underlying goal of the conversation.

This study could not include the verbal content or the topic of each interview in the analysis, which may be mentioned as another limitation. The content of the interview, as well as the interview type (i.e., information- and entertainment-driven), could be related to how the host and guest behaved. Indeed, many information-driven interviews discussed things related to the COVID-19 pandemic, but few entertainment-driven interviews did so. We should not rule out the possibility that the unprecedented outbreak made information-driven interviews more serious, resulting in a further increase in synchrony. With the aid of the current development in computer science, automated coding tools for nonverbal behavior are more accessible (e.g., Baltrušaitis et al., 2016; Fujiwara & Yokomitsu, 2021; Lewinski et al., 2014; Ramseyer, 2020; Ramseyer & Tschacher, 2011). Further investigation with larger samples will bring us more in- depth findings.

In the future, it should be ascertained what impression the audiences had of the interviews, in which the host and guest exhibited different degrees of synchrony, which was not examined in the current study. While interactants who showed synchrony were perceived as "mutually interested" by third parties who observe their interaction (Fujiwara et al., 2019), there seems much remains unknown in the mediated interview. Especially after the pandemic, interviews (mainly in information-driven interviews) could include reminders to viewers about infection control measures, and some of which would be more effective. If the greater synchrony between the host and the expert guest leads to more credibility for the interview program, the audiences may follow the reminder. As many studies have shown that interactants benefit from synchrony, it will be beneficial to examine the effects of synchronization on third parties.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

References

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*(2), 256–274. https://doi.org/10.1037/0033-2909.111.2.256

Babad, E., & Peer, E. (2010). Media bias in interviewers' nonverbal behavior: Potential remedies, attitude similarity and meta-analysis. *Journal of Nonverbal Behavior, 34*(1), 57–78. https://doi.org/10.1007/s10919-009-0078-x

Baltrušaitis, T., Robinson, P., & Morency, L.-P. (2016). Openface: An open source facial behavior analysis toolkit. *IEEE Winter Conference on Applications of Computer Vision (WACV), 2016*, 1–10.

Berger, C. R., & Palomares, N. (2011). Knowledge structures and social interaction. In M. L. Knapp & J. A. Daly (Eds.), *The sage handbook of interpersonal communication* (4th ed., pp. 169–200). Sage.

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in timeseries. In Proceedings of the Workshop on Knowledge Discovery in Database, 359–370.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior, 12*(2), 120–138. https://doi.org/10.1007/BF00986930

Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin, 20*(3), 303–311. https://doi.org/10.1177/0146167294203008

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother–infant interactions. *Journal of Personality and Social Psychology, 54*(2), 243–253. https://doi.org/10.1037/0022-3514.54.2.243

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rime (Eds.), *Fundamentals of nonverbal behavior: Studies in emotion & social interaction* (pp. 401–432). Cambridge University Press.

Bernhold, Q. S., & Giles, H. (2020). Vocal accommodation and mimicry. *Journal of Nonverbal Behavior*, *44*(1), 41–62. https://doi.org/10.1007/s10919-019-00317-y

Bikmen, N., Koneczny, A., & Caballero, K. (2022). Duchenne smiles of white American college students in same-race and interracial interactions. *Journal of Nonverbal Behavior, Advance online publication.* https://doi.org/10.1007/s10919-021-00393-z

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology, 77*(3), 343–352. https://doi.org/10.1016/j.biopsycho.2007.11.008

Brambilla, M., Sacchi, S., Menegatti, M., & Moscatelli, S. (2016). Honesty and dishonesty don't move together: Trait content information influences behavioral synchrony. *Journal of Nonverbal Behavior, 40*(3), 171–186. https://doi.org/10.1007/s10919-016-0229-9

Burgoon, Judee K., Stern, Lesa A., & Dillman, Leesa. (1995). *Interpersonal adaptation: Dyadic interaction patterns*. Cambridge University Press. https://doi.org/10.1017/CBO9780511720314

Butler, E. A. (2011). Temporal interpersonal emotion systems: The "TIES" that form relationships. *Personality and Social Psychology Review, 15*(4), 367–393. https://doi.org/10.1177/1088868311411164

Butler, E. A. (2015). Interpersonal affect dynamics: It takes two (and time) to tango. *Emotion Review, 7*(4), 336–341. https://doi.org/10.1177/1754073915590622

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7291–7299.

Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology, 64*(1), 285–308. https://doi.org/10.1146/annurev-psych-113011-143754

Choi, M., & Choung, H. (2021). Mediated communication matters during the COVID-19 pandemic: The use of interpersonal and masspersonal media and psychological well- being. *Journal of Social and Personal Relationships, 38*(8), 2397–2418. https://doi.org/10.1177/02654075211029378

Coco, M. I., & Dale, Rick. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An R package. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.00510

Costa,J., P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322–331. https://doi.org/10.1037/0022-3514.81.2.322

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing, 3*(3), 349–365. https://doi.org/10.1109/T-AFFC.2012.12

Dunbar, N. E., Giles, H., Bernhold, Q. S., Adams, A., Giles, M., Zamanzadeh, N., Gangi, K., Coveleski, S., & Fujiwara, K. (2020). Strategic synchrony and rhythmic similarity in lies about ingroup affiliation. *Journal of Nonverbal Behavior, 44*(1), 153–172. https://doi.org/10.1007/s10919-019-00321-2

Dunbar, N. E., Jensen, M. L., Tower, D. C., & Burgoon, J. K. (2014). Synchronization of nonverbal behaviors in detecting mediated and non-mediated deception. *Journal of Nonverbal Behavior, 38*(3), 355–376. https://doi.org/10.1007/s10919-014-0179-z

Ekman, P. (1972). Universal and cultural differences in facial expressions of emotion. *Nebraska Symposium of Motivation, 19*, 207–283.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*(3), 429–456. https://doi.org/10.1037/0033-2909.116.3.429

Fischer, A., & LaFrance, M. (2015). What drives the smile and the tear: Why women are more emotionally expressive than men. *Emotion Review, 7*(1), 22–29. https://doi.org/10.1177/1754073914544406

Fujiwara, K., Bernhold, Q. S., Dunbar, N. E., Otmar, C. D., & Hansia, M. (2021). Comparing manual and automated coding methods of nonverbal synchrony. *Communication Methods and Measures, 15*(2), 103–120. https://doi.org/10.1080/19312458.2020.1846695

Fujiwara, K., & Daibo, I. (2022). Empathic accuracy and interpersonal coordination: Behavior matching can enhance accuracy but interactional synchrony may not. *The Journal of Social Psychology, 162*(1), 71–88. https://doi.org/10.1080/00224545.2021.1983509

Fujiwara, K., Hoegen, R., Gratch, J., & Dunbar, N. E. (2022). Synchrony facilitates altruistic decision making for non-human avatars. *Computers in Human Behavior, 128*, 107079. https://doi.org/10.1016/j.chb.2021.107079

Fujiwara, K., Kimura, M., & Daibo, I. (2019). Gender differences in synchrony: Females in sync during unstructured dyadic conversation. *European Journal of Social Psychology, 49*(5), 1042–1054. https://doi.org/10.1002/ejsp.2587

Fujiwara, K., Kimura, M., & Daibo, I. (2020). Rhythmic features of movement synchrony for bonding individuals in dyadic interaction. *Journal of Nonverbal Behavior, 44*(1), 173–193. https://doi.org/10.1007/s10919-019-00315-0

Fujiwara, K., & Yokomitsu, K. (2021). Video-based tracking approach for nonverbal synchrony: A comparison of Motion Energy Analysis and OpenPose. *Behavior Research Methods, 53*(6), 2700–2711. https://doi.org/10.3758/s13428-021-01612-7

Gilbert, D. T., & Osborne, R. E. (1989). Thinking backward: Some curable and incurable consequences of cognitive busyness. *Journal of Personality and Social Psychology, 57*(6), 940–949. https://doi.org/10.1037/0022-3514.57.6.940

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *The contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge University Press.

Hall, J. A., Gunnery, S. D., & Horgan, T. G. (2016). Gender differences in interpersonal accuracy. In J. A. Hall, M. Schmid Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 309–327). Cambridge University Press.

Hall, E. T. (1976). *Beyond culture*. Anchor Press.

Hall, J. A., & Halberstadt, A. G. (1986). Smiling and gazing. In J. S. Hyde & M. C. Inn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 136–185). John Hopkins University Press.

Hess, U., & Bourgeois, P. (2010). You smile–I smile: Emotion expression in social interaction. *Biological Psychology, 84*(3), 514–520. https://doi.org/10.1016/j.biopsycho.2009.11.001

Hess, U., & Fisher, A. (2014). Emotional mimicry: Why and when we mimic emotions. *Social and Personality Psychology Compass, 8*(2), 45–57. https://doi.org/10.1111/spc3.12083

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly, 15*(4), 635–650. https://doi.org/10.1086/266350

Inamine, M., & Endo, M. (2009). The effects of context and gender on the facial expressions of emotions: A study of japanese college students. *The Japanese Journal of Research on Emotions, 17*(2), 134–142. https://doi.org/10.4092/jsre.17.134

Karl, K. A., Peluchette, J. V., & Aghakhani, N. (2021). *Virtual work meetings during the COVID-19 pandemic: The good, bad, and ugly*. Advance online publication. https://doi.org/10.1177/10464964211015286

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 59–101). Elsevier Academic Press.

Kulke, L., Feyerabend, D., & Schacht, A. (2020). A comparison of the affectiva imotions facial expression analysis software with EMG for identifying facial expressions of emotion. Frontiers in Psychology, 11, Article 329. https://doi.org/10.3389/fpsyg.2020.00329

LaFrance, M., & Broadbent, M. (1976). Group rapport: Posture sharing as a nonverbal indicator. *Group & Organization Management, 1*(3), 328–333. https://doi.org/10.1177/105960117600100307

LaFrance, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin, 129*(2), 305–334. https://doi.org/10.1037/0033-2909.129.2.305

Lakin, Jessica L. (2013). 18 Behavioral mimicry and interpersonal synchrony. In Judith A. Hall & Mark L. Knapp (Eds.), *Nonverbal Communication:* (pp. 539–576). DE GRUYTER. https://doi.org/10.1515/9783110238150.539

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior, 27*(3), 145–162. https://doi.org/10.1023/A:1025389814290

Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and facs aus in facereader. *Journal of Neuroscience, Psychology, and Economics, 7*(4), 227–236. https://doi.org/10.1037/npe0000028

Lumsden,J., Miles, L. K., Richardson, C. A., M. J. amd Smith, & Macrae, C. N. (2012). Who syncs? Social motives and interpersonal coordination. Journal of Experimental Social Psychology, 48(3), 746–751. https://doi.org/10.1016/j.jesp.2011.12.007

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion, 14*(3), 195–214. https://doi.org/10.1111/1467-839X.00010

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13*(2), 143–157. https://doi.org/10.1007/BF00992959

Matsumoto, David, Takeuchi, Sachiko, Andayani, Sari, Kouznetsova, Natalia, & Krupp, Deborah. (2002). The Contribution of Individualism vs. Collectivism to Cross-national Differences in Display Rules. *Asian Journal of Social Psychology, 1*(2), 147–165. https://doi.org/10.1111/1467-839X.00010

Miles, L. K., Griffighs, J. L., Richardson, M. J., & Macrae, C. N. (2010). Too late to coordinate: Contextual influences on behavioral synchrony. *European Journal of Social Psychology, 40*(1), 52–60. https://doi.org/10.1002/ejsp.721

Miles, L. K., Lumsden, J., Richardson, M. J., & Macrae, C. N. (2011). Do birds of a feather move together? group membership and behavioral synchrony. *Experimental Brain Research, 211*, 495–503. https://doi.org/10.1007/s00221-011-2641-z

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology, 72*, 13–20. https://doi.org/10.1016/j.jesp.2017.03.009

Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychological Methods, 23*(4), 757–773. https://doi.org/10.1037/met0000172

Murphy, Nora A., & Hall, Judith A. (2021). Capturing behavior in small doses: A review of comparative research in evaluating thin slices for behavioral measurement. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.667326

Murphy, N. A., Hall, J. A., Mast, M. S., Ruben, M. A., Frauendorfer, D., Blanch-Hartigan, D., Roter, D. L., & Nguyen, L. (2015). Reliability and validity of nonverbal thin slices in social interactions. *Personality and Social Psychology Bulletin, 41*(2), 199–213. https://doi.org/10.1177/0146167214559902

Murphy, N. A., Hall, J. A., Ruben, M. A., Frauendorfer, D., Schmid Mast, M., Johnson, K. E., & Nguyen, L. (2019). Predictive validity of thin-slice nonverbal behavior from socialinteractions. *Personality and Social Psychology Bulletin, 45*(7), 983–993. https://doi.org/10.1177/0146167218802834

Newtson, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology, 35*(12), 847–862. https://doi.org/10.1037/0022-3514.35.12.847

Newtson, D., Hairfield, J., Bloomingdale, J., & Cutino, S. (1987). The structure of action and interaction. *Social Cognition, 5*(3), 191–237. https://doi.org/10.1521/soco.1987.5.3.191

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology, 66*(11), 2092–2102. https://doi.org/10.1080/17470218.2013.853089

Ramseyer, F. T. (2020). Motion energy analysis (mea): A primer on the assessment of motion from video. *Journal of Counseling Psychology, 67*(4), 536–549. https://doi.org/10.1037/cou0000407

Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: Coordinated body-movement reflects relationship quality and outcome. *Journal of Consulting and Clinical Psychology, 79*(3), 284–295. https://doi.org/10.1037/a0023419

Reed, Lawrence Ian, Stratton, Rachel, & Rambeas, Jessica D. (2018). Face value and cheap talk: How smiles can increase or decrease the credibility of our words. *Evolutionary Psychology, 16*(4), 147470491881440. https://doi.org/10.1177/1474704918814400

Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A meta-analysis. *Zeitschrift für Psychologie, 224*(3), 168–189. https://doi.org/10.1027/2151-2604/a000252

Riehle, M., Kempkensteffen, J., & Lincoln, T. M. (2017). Quantifying facial expression synchrony in face-to-face dyadic interactions: Temporal dynamics of simultaneously recorded facial EMG signals. *Journal of Nonverbal Behavior, 41*(2), 85–102. https://doi.org/10.1007/s10919-016-0246-8

Schmidt, R. C., Morr, S., Fitzpatrick, P., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior, 36*(4), 263–279. https://doi.org/10.1007/s10919-012-0138-5

Schoenherr, D., Paulick, J., Worrack, S., Strauss, B. M., Rubel, J. A., Schwartz, B., Deisenhofer, A. K., Lutz, W., Stangier, U., & Altmann, U. (2019). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods, 51*(1), 361–383. https://doi.org/10.3758/s13428-018-1139-z

Shufford, K. N., Hall, D. L., Randall, A. K., Braunstein, B. M., O'Brien, M. M., & Mickelson, K. D. (2021). Connected while apart: Associations between social distancing, computer-mediated communication frequency, and positive affect during the early phases of COVID-19. *Journal of Social and Personal Relationships, 38*(10), 2906–2920. https://doi.org/10.1177/02654075211041316

Tamura, R., & Kameda, T. (2006). Are facial expressions contagious in the Japanese? *Shinrigaku Kenkyu: The Japanese Journal of Psychology, 77*(4), 377–382. https://doi.org/10.4992/jjpsy.77.377

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or- flight. *Psychological Review, 107*(3), 411–429. https://doi.org/10.1037/0033-295X.107.3.411

Tickle-Degnen, L. (2006). Nonverbal behavior and its functions in the ecosystem of rapport. In V. Manusov & M. L. Patterson (Eds.), *The sage handbook of nonverbal communication* (pp. 381–399). Sage.

Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 1*(4), 285–293. https://doi.org/10.1207/s15327965pli0104_1

Tikochinski, R., & Babad, E. (2022). Perceived epistemic authority (source credibility) of a TV interviewer moderates the media bias effect caused by his nonverbal behavior. *Journal of Nonverbal Behavior, 46*(2), 215–229. https://doi.org/10.1007/s10919-022-00397-3

Tolson, A. (2001). *Television talk shows: Discourse, performance, spectacle*. Lawrence Erlbaum Associates.

Van Der Zee, S., Taylor, P., Wong, R., Dixon, J., & Menacere, T. (2021). A liar and a copycat: Nonverbal coordination increases with lie difficulty. *Royal Society Open Science, 8*, 200839. https://doi.org/10.1098/rsos.200839

Vicaria, I. M., & Dickens, L. (2016). Meta-analyses of the intra- and interpersonal outcomes of interpersonal coordination. *Journal of Nonverbal Behavior, 40*(4), 335–361. https://doi.org/10.1007/s10919-016-0238-8

Watermeyer, R., Crick, T., Knight, C., & Goodall, J. (2021). COVID-19 and digital disruption in UK universities: Afflictions and affordances of emergency online migration. *Higher Education, 81*(3), 623–641. https://doi.org/10.1007/s10734-020-00561-y

Weyers, P., Mühlberger, A., Kund, A., Hess, U., & Pauli, P. (2009). Modulation of facial reactions to avatar emotional faces by nonconscious competition priming. *Psychophysiology, 46*(2), 328–335. https://doi.org/10.1111/j.1469-8986.2008.00771.x