Head Nodding and Hand Coordination Across Dyads in Different Conversational Contexts

Patrick Falk (patrick.falk@ki.se)
UCL
Roser Cañigueral
UCL
Jamie A Ward
Goldsmiths University of London
Antonia F de C Hamilton
UCL

Research Article

Keywords: Non-verbal behaviour, conversational context, social interaction, interpersonal coordination, head nodding, hand coordination

Posted Date: November 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3526068/v1

Additional Declarations: No competing interests reported.

Abstract

This paper aims to explore what different patterns of head nodding and hand movement coordination mean in conversation by recording and analysing interpersonal coordination as it naturally occurs in social interactions. Understanding the timing and at which frequencies such movement behaviours occur can help us answer how and why we use these signals. Here we use high-resolution motion capture to examine three different types of two-person conversation involving different types of information-sharing, in order to explore the potential meaning and coordination of head nodding and hand motion signals. We also test if the tendency to engage in fast or slow nodding behaviour is a fixed personality trait that differs between individuals.

Our results show coordinated slow nodding only in a picture-description task, which implies that this behaviour is not a universal signal of affiliation but is context driven. We also find robust fast nodding behaviour in the two contexts where novel information is exchanged. For hand movement, we find hints of low frequency coordination during one-way information sharing, but found no consistent signalling during information recall. Finally, we show that nodding is consistently driven by context but is not a useful measure of individual differences in social skills. We interpret these results in terms of theories of nonverbal communication and consider how these methods will help advance automated analyses of human conversation behaviours.

Main text

Interpersonal coordination refers to the temporal alignment of two or more individuals while they interact with each other (Hoehl, Fairhurst, & Schirmer, 2020). Growing interest in the dynamics of real-world social interactions (Redcay & Schilbach, 2019; Schilbach et al., 2013) has shown that interpersonal coordination is present across various domains, such as bodily movements (Chartrand & Bargh, 1999; Hale, Ward, Buccheri, Oliver, & Hamilton, 2020; Ramseyer & Tschacher, 2011), physiological signals (Feldman, Magori-Cohen, Galili, Singer, & Louzoun, 2011; Konvalinka et al., 2011) or brain activity (Hirsch, Noah, Zhang, Dravida, & Ono, 2018; Stephens, Silbert, & Hasson, 2010). Across all these domains it is widely agreed that interpersonal coordination has positive effects on social interactions (Hoehl et al., 2020), by facilitating communication (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012) and increasing affiliation (Lakin, Jefferis, Cheng, & Chartrand, 2003). However, the specific patterns of interpersonal coordination remain poorly understood. A reason for this could be that traditional methods to record and analyse interpersonal coordination in dyadic social interactions fail to capture the full richness of interaction dynamics.

Here, we examine the interpersonal coordination of head nods and hand movements, using advanced methods (high-resolution automated motion capture and wavelet coherence analysis) and in three different conversational contexts. Tracking conversation behaviour across different contexts will allow us to test hypotheses about why people engage in particular patterns of nodding or hand movements, and thus to interpret what these actions might mean. In addition, we can test if the behaviour of individual participants is consistent from one context to another, that is, do some people always engage in a lot of nodding regardless of context while others rarely nod? If individual behaviour is consistent, this would support the development of automated methods that could discriminate personality (Heerey, 2015) or even diagnose psychiatric disorders from social behaviour patterns (Georgescu et al., 2019). Thus, this paper aims to explore what nodding and hand movements mean in conversation and how this type of data could be used for future research. We first review current knowledge about head nodding and hand movement patterns in conversation and detail our experimental manipulations.

Head Nodding Behaviour

Many non-verbal signals during a conversation are centred around the head (e.g., eye-gaze, blinks, facial expressions, and head movements), and listener's attention is typically drawn to the speaker's head and face during conversation (Argyle & Cook, 1976). Head nodding is regarded as a distinct social signal that is particularly sensitive to conversational demands and can convey several different meanings (Poggi, Errico, & Vincze, 2010), from signalling attention and understanding (Hadar et al., 1983; Kendon, 2002), to requests for information and passing turns (Duncan, 1972). Recent work from our lab has developed an automated method which can identify and quantify two distinct types of nods – fast nods and coordinated slow nods (Hale et al., 2020). From that work, we define fast nods as vertical head movements that are faster than 1.5 Hz, and slow nods as below 1.5 Hz. By examining how fast-nods and slow-nods are used across different conversational contexts, we aim to understand the meaning of nodding as a social signalling behaviour.

We consider three potential meanings of a head nod: backchannelling, mimicry and joint attention. First, backchannelling is the information flow in a conversation where the listener signals 'back' to the speaker. Verbal backchannels include linguistic vocalizations such as 'uh-uh', whereas non-verbal backchannels include facial expressions and head movements like nodding. For example, a listener may nod their head to show that they are listening, or even to indicate that one is agreeing with what the speaker is saying (Allwood & Cerrato, 2003; Duncan, 1972). Previous research (Hale et al. 2020) showed that a fast nods are produced mainly when participants are listening and receiving new information, which suggests this might be a backchannel. The present study will test if this is true across different information-exchange contexts.

Second, nodding could be a type of mimicry. Mimicry arises when one person copies the gestures, actions, or postures of another (Chartrand & Bargh, 1999). This mimicry typically occurs spontaneously during interactions and is believed to act as a 'social glue' to facilitate bonding and affiliation between people (Lakin et al., 2003). Previous work (Hale et al, 2020) identified coordinated slow nodding with a time-scale which matches previous reports of mimicry (Hömke, Holler, & Levinson, 2018). If this behaviour is a form of mimicry linked to social affiliation, we would expect the behaviour to be present across many different conversation contexts, regardless of the topic of conversation.

Third, nodding could represent a type of joint attention, which arises when two people gaze at the same object at the same time, typically with one leading the gaze and the other following (Emery, 2000). In raw motion capture data, this gaze following pattern might look like a nodding action if both people are looking downwards to an object in their hands, which was the case in Hale et al (2020). That is, it is possible that the 'nodding coordination' in the previous study arose primarily because both participants were jointly attending towards an object held in the hands. If this interpretation is correct, then conversations in a different context without the picture should not show coordinated slow nodding behaviour.

To summarise, we have described two types of head nodding behaviour – fast nods and coordinated slow nods – and we consider three different social meanings that could be applied to these behaviours – backchannelling, joint attention and mimicry. Changing context of a conversation provides a way to distinguish the social meanings of the different nodding behaviours. Here, we create three different conversation tasks, which allow us to manipulate information sharing and joint attention targets to distinguish between these different interpretations of nodding behaviour. Before detailing these different tasks, we will describe the hand movement behaviours which are the second focus of the present paper.

Hand Movement in Conversation

During conversation, co-speech hand movements are tightly linked to speech at the temporal and semantic level (Kita & Özyürek, 2003; Loehr, 2007). For instance, beat gestures are rapid movements used as temporal cues to emphasise relevant information (McNeill, 1992), whereas iconic gestures have high semantic content and are used to describe an object or action to disambiguate complex sentences (Kelly, Kravitz, & Hopkins, 2004; Kelly, Ward, Creigh, & Bartolotti, 2007). In fact, several studies show that co-speech hand gestures facilitate attention capture, affect speech comprehension, and improve learning and memory in both speakers and listeners (Cook, Mitchell, & Goldin-Meadow, 2008; Kendon, 1972; Marstaller & Burianová, 2013; McNeill, 1992). Another type of (non-co-speech) hand movements are self-grooming gestures. These are actions used to clean or maintain parts of the body (e.g. fixing the hair) in order to give a positive impression to others and increase affiliation (Daly, Hogg, Sacks, Smith, & Zimring, 1983). Despite the critical role of hand gestures in conversation and social interactions, little is known about their dynamics at the interpersonal level.

Single-participant studies have shown that individuals coordinate or imitate hand actions from a video-clip or virtual characters (Genschow, Florack, & Wänke, 2013; Pan & Hamilton, 2015, Stel et al., 2010), but to our knowledge only two previous studies have investigated hand gesture coordination in face-to-face conversation. Holler and Wilkin (2011) used a referential communication task (Clark & Wilkes-Gibbs, 1986) where dyads were given two equal sets of cards depicting figure-like stimuli, and were instructed to discuss them with the aim of placing the cards on a table in the same order. They found that participants imitated each other's co-speech gestures during the conversation, and that such imitation played an important role in establishing mutual shared understanding. In another study, Ramseyer and Tschacher (2016) investigated the presence of hand movement imitation during conversation in the context of a natural psychotherapy session. In a single-case analysis, they found that patient and therapist imitated each other's hand movements, and that the levels of interpersonal coordination over the sessions were positively associated with patient's ratings of affiliation with the therapist.

Although these studies provide evidence of interpersonal coordination of hand gestures, they rely on slow but precise coding of video recordings by trained observers. Here, we aimed to determine if hand movement dynamics can be captured with high resolution motion capture and interpreted using the same automated framework that we used for head nods. This is an exploratory analysis, which will test if there is interpersonal coordination of hand movements that can be detected with motion capture and if this coordination varies across conversational contexts.

Changing Conversational Contexts

In the study of both head nodding and hand movements, it is clear that examining behavior in a single context is not enough to interpret the social meaning of a behavior or to provide a general analysis. Thus, the present study placed participants in three different conversational contexts. First, we used a picture description task which has previously been valuable in our lab and elsewhere (Chartrand & Bargh, 1999; Hale et al., 2020; van Baaren et al., 2009). Here, one participant holds a picture of a complex scene and must describe it to their partner, who listens and then can ask questions about this picture. Each trial lasts only 90 seconds and is divided equally into monologue and dialogue phases. This task is highly structured, with one person in the role of the 'leader' who holds the picture and who speaks most of the time. The presence of the picture also provides a clear target for joint attention.

The second 'video recall' task was selected to create a conversation with common ground (i.e., shared knowledge) that engages memory but did not involve the exchange of any new information. At two points during the data collection session, participants watched a 3 min wordless children's animation together. Later, they were asked to recall the animation in detail, working together to describe as much as they could. This tended to be a slow unstructured conversation where both participants discuss events which they are familiar with.

The third 'meal planning' task was developed by Chovil (1991) and Tschacher et al. (2014) as a way to encourage a fun and relaxed conversation between strangers. Participants were asked to spend 5 min planning a meal using ingredients they both dislike. This conversation topic induces some general exchange of information about food preferences together with joint planning of the meal. The exchanges tend to be short and dynamic with laughter and overlaps in speaking.

Figure 1 provides an illustration of these three conversation tasks and a sample of the turn-taking behavior in each one. Panel A illustrates the picture description task where one participant (here blue) speaks for the majority of the time, providing information about the picture to their partner. Here, the picture itself provides a joint attention target. Panel B illustrates the video discussion task, where participants recall the short movie (i.e., share 'common ground') but do not exchange any new information. Panel C illustrates the meal planning task where both participants share information and often speak quickly with overlaps.

Measuring Interpersonal Coordination

To understand the changes in movement behaviour across these different conversational contexts, it is important to precisely measure and appropriately analyse the behaviour of our participants. Traditional video-coding methods have high accuracy but are very time-consuming and hard to quantify (Holler & Wilkin, 2011). Video-based analyses can quantify motion energy (Ramseyer & Tschacher, 2011, 2016), but their resolution is limited because they quantify pixel changes on a flat image. Motion capture technologies provide high resolution recordings of bodily movement in a 3D space (Feese, Arnrich, Tröster, Meyer, & Jonas, 2011; Hale et al., 2020; Poppe, Zee, Heylen, & Taylor, 2013). The present study uses this method to record head and hand position at a high sampling rate (120 Hz) while two participants interact face-to-face.

To analyse the data, we use wavelet coherence analyses (Fujiwara & Daibo, 2016; Issartel, Marin, Gaillot, Bardainne, & Cadopi, 2006). This provides a measure of interpersonal correlation for each *frequency* component and *time-point* in the interaction. Information on the frequency domain has been useful in distinguishing different types of nodding behaviour. For instance, recent studies in our lab using wavelet coherence analysis (Hale et al., 2020) have identified fast and slow nods as distinct behaviours which are visible in a wavelet analysis. The present study extends this to different contexts to test how context changes nodding behaviour.

The Present Study

The present study combined a high-resolution motion capture system with wavelet coherence analysis to investigate head and hand motion patterns of dyads as they engaged in three conversational tasks with varying degrees of structure and common ground. The aim of the study was to address three major questions.

Question 1: What do head-nodding signals mean? We hypothesis that, if fast-nods are a backchannel that signals 'information received', they should be found in the contexts where the participants exchange novel information (picture description and meal planning tasks) but not in the video discussion task. If coherence in slow nodding reflects affiliation it should be found across all contexts, but if it reflects joint attention it should be found only in the picture description task when an object (the picture) is available to look at.

Question 2: Are individual levels of head nodding correlated across contexts? If head-nodding is a robust individual signature with the potential to be a clinical marker, it should be consistent across contexts. For example, an individual who nods a lot in the picture description task should also nod in the video discussion task and this might correlate with personality measures. By testing for this pattern, we can explore the potential of nodding measures as a way to quantify individual differences in social behaviour.

Question 3: What are the patterns of interpersonal coordination of hand movement across contexts? This question is more exploratory, as there is little prior data on hand coordination, so we considered two aspects. First, can the wavelet coherence methods we used for nodding detect any robust pattern of hand movement coordination, and if so, what frequencies are seen? And second, does hand movement coordination change across contexts? Given the absence of previous studies on this topic, we did not make any specific predictions for the patterns found in each conversational task. However, we hypothesised that, if interpersonal coordination of hand gestures facilitates communication, dyads would generally show more interpersonal coordination of hand gestures when the task was unstructured and there was no common ground.

Methods

Participants

62 participants were recruited from the UCL Psychology Subject Pool and the ICN Subject Database. All were fluent in English and the mean age was 24 years. All participants were recruited and tested in pairs (31 dyads); they were all strangers prior to the study and no participant was included in more than one pair. Six pairs were mixed gender, and 25 pairs were female-female. The participants did not have any previous experience with the tasks and were unaware of the purpose of the experiment. Ethical approval was given by the UCL Research Ethics committee, and all participants gave their written informed consent. A monetary reimbursement was offered for participating in the study at a rate of £7.50/hour.

For head motion capture data, all participants were included in the final sample. For hand motion capture data, 10 dyads were excluded because one or both members of the dyad had poor quality hand motion data (see Section 2.5 on Data Analysis for details on data processing). Thus, the final valid sample for hand data consisted of 42 participants (39 females, 3 males) assigned to 21 pairs (3 pairs were mixed gender and 18 pairs were female-female).

Experimental Setup

The testing room was divided in two spaces separated by a curtain – the participant space and the experimenter space (Figure 2). In the participant space, the participants sat on small stools facing each other and approximately 1m apart. The participant space was equipped with a motion capture system (OptiTrack, NaturalPoint Inc., v.1.10), which consisted of eight cameras recording at a frequency of 120 Hz. This system tracks head and body movements by detecting the position of reflective markers which were placed on an upper body suit and cap (n=25 markers per person) worn by the participants. Participants were equipped with lapel microphones to record their voices, as well as wearable eye-trackers (Pupil Core, Pupil Labs GmbH., Germany) to record their eye movements (See Dobre et al., 2021 for some of this data). To one side of the participant space, a projector screen played video stimuli, a speaker played audio instructions and a webcam recorded the room.

On the other side of the curtain, the experimenter space was equipped with three computers (A, B and C) that run and coordinated the whole experiment. Computer A was the client computer and communicated with computers B and C, which acted as servers. B and C each captured data from the eyetracker, and in addition B captured the Optitrak data and captured audio and video data. This setup allowed us to handle the large amount of data recorded by the various sources in a synchronised manner, by generating precise, machine-specific timestamps for each recording. We also inserted audio-visual synchronisation gestures (i.e. 3 hand claps) into our experimental protocol to allow for post-hoc synchronisation of data streams if needed.

Procedure

Participants arrived at the lab and were shown all the equipment and informed of the procedures before they signed the informed consent. They were asked to remove eye-makeup, bulky clothes and jewelry and were randomly assigned roles as ‘Blue’ or ‘Yellow’. The roles enabled our data labelling but had no impact on the tasks performed. Participants put on the motion capture suits, eye-trackers, and microphones and sat down 1m from each other to begin the study. The experimental session began with calibration of the motion capture and eyetracker systems and then a synchronization event where participants were asked to clap their hands three times with each other. This synchronization event was repeated as needed (See claps in Figure 3).

First, participants watched a 3 minutes animated video together (DipDap) (in preparation for the Video Discussion task). Then they completed 8 trials of the Picture Description task, which was adapted from earlier behavioral studies (Chartrand & Bargh, 1999). This task involves one-way information sharing, as one participant (leader) holds a picture of a complex social scene and is asked to describe it to their partner. The conversation is highly structured, with 45 seconds of monologue (only leader may speak) followed by 45 seconds of dialogue (listener may ask questions about the picture), and participants took turns in the role of leader. Audible cues signaled the start and end of each trial, as well as the transition from monologue to dialogue. Next, they completed a Video Discussion task, where they were instructed to recall the short video which they had watched together earlier. The cartoon (DipDap, Roberts 2011) is a simple animation with no words where a character encounters a variety of interesting objects that may transform into different things. In recalling the details of the video, participants had an unstructured conversation to describe the video, but did not need to exchange any novel information.

Next, they completed the Meal Planning task, based on Chovil (1991), and recently adapted by Tschacher et al. (2014). In this task the participants have five minutes to come up with a menu together consisting of an appetizer, main course, and dessert. However, they can only use ingredients that they both dislike, which introduces a fun cooperative element to the conversation. Like the Video Discussion task, this is an unstructured conversation, but with two-way information sharing or joint planning. Participants completed a single five-minute trial of this task. At this point, a second calibration was performed and then participants watched a second short animation of DipDap. Then they completed a further 8 trials of the Picture Description task and 1 trial of the Video Discussion task. The Meal Planning task was not repeated because participants by the end of the first block were already familiar with each other's meal preferences.

After finishing all the tasks, the participants removed their equipment and were seated at separate desks, where they completed four questionnaires measuring social anxiety traits (Liebowitz, 1987), alexithymia traits (Bagby, Parker, & Taylor, 1994), autistic traits (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), and a novel questionnaire on Experience of Gaze. This is being trialed by our group to measure the participants subjective experience of eye contact. It consists of 20 statements (e.g., "I need to think about whether or not to make eye-contact"), with a forced choice on a five-point scale between "strongly agree" to "strongly disagree". Finally, participants were debriefed about the real purpose of the study and were paid for their participation. A summary of the experimental procedure is shown in Figure 3.

Data Analysis

The present paper focuses only on data from the Optitrak mocap system. This data was preprocessed in Motive software (supplied by Optitrak) to match the markers to a skeleton model and linearly interpolate over any time points with missing markers. In a minority of trials where Motive was unable to reliably track data, the trial was excluded. This applied to 28/248 picture description trials; 5/62 video discussion trials; and 3/31 meal planning trials.

For head motion analysis, we follow Hale et al. (2020) and focus solely on the head pitch, or nod, data (i.e., degrees of rotation in the y-plane). We carried out the following pipeline (Figure 4) using the Matlab toolbox by Grinsted, Moore, and Jevrejeva (2004) to identify the wavelet power in the head pitch signals and to calculate cross-wavelet coherence. For each participant, we take the raw head pitch signal (Fig4 A and B) and calculated the wavelet transform for each trial to get the time-frequency representation of each time-series (C, D). We used default parameters (Morlet wavelet, w = 6), and the head signal for each participant was decomposed in 54 wavelets that ranged between frequencies of 0.1 Hz and 10 Hz, with a wavelength difference of 0.185 seconds. In total there were 133 wavelet scales using the Morlet wavelet (periods ranged from 0.1 – 19 Hz on the sampled data). Next, we calculated the cross-wavelet coherence between each of the two wavelet transforms (E), which gives a measure of the time-frequency coordination between their movements. To ensure that the analysis was free from edge effects, we calculated the 'cone of influence' (COI) and zeroed any data outside it. We also applied cone-of-influence zeroing around the monologue-to-dialogue transition in the Picture Description task, this helped to minimize the influence of stimuli outside the dyad. We discarded data that was outside the 0.1 – 19 Hz range. In the final step, we averaged the cross-wavelet coherence (R^2) over the time-course of each trial to obtain a measure of the frequency of coherence without regard to the specific time within the trial at which it occurred (F).

To understand the patterns of head movement present in each task, it is helpful to compare the wavelet coherence values from true pairs to coherence values from pseudo-pairs (Bernieri & Rosenthal, 1991; Fujiwara & Diabo, 2016; Hale et al., 2020). We created pseudo-pairs by shuffling data *within* dyad and *within* task. For the video discussion task, this meant matching the discussion of Video 1 to Video 2 from the same participants. For the picture description task, this meant shuffling trials within the same participants. For the meal planning task, the 5 min trial was divided in two 150sec chunks which were swapped to create a pseudo-pair. This gives us a strong test where the pseudo trials have the same general movement characteristics as our real trials, and any differences in the coherence levels between them must be due to a genuine live social interaction and will not be attributed to any individual differences between them.

We carried out wavelet analysis on the pseudo-data using the same pipeline as in the real trials. This gave us a set of coherence values for each real and pseudo trial of each dyad. Separately for the real dataset and the pseudo dataset, we averaged the coherence values across all trials for all dyads. We then calculated a coherence difference for each dyad, representing the average coherence in real interactions minus the average coherence in pseudo interactions, and performed t-tests on the coherence differences at each frequency (90 tests, one for each wavelet scale). To correct for multiple comparisons, we used a False Detection Rate (FDR) of 0.05 (Benjamin & Hochberg, 1995). By comparing real and pseudo trials in this way, we can see if interpersonal coordination that occurs in real conversations is different from the same people just speaking.

Analysis of Individual Differences. We aimed to test if the tendency to engage in fast or slow nodding behaviour is a fixed personality trait that differs between people and is consistent across tasks. For each participant, we calculated a 'fast nod score' as that person's mean power in the 2.6–6.5 Hz range across the complete dataset (Figure 5A). To avoid 'double dipping' in our analysis, this frequency range was chosen directly from the findings of Hale et al. (2020). The 'fast nod score' measures how much a person engages in fast nodding behaviour. Then, we correlated the fast-nod scores across tasks for each participant, to test if some participants consistently show high levels of fast-nodding while others rarely nod in any tasks.

We also calculated a 'slow nod coherence score' from the mean dyadic coherence level in the 0.2–1.1 Hz range (Figure 5B) – using frequencies reported by Hale et al. (2020). This gives a measure of how much each dyad engages in coordinated slow nodding. Again, this is a single value measure but this time represented as an R^2 coherence score (between 0 and 1) revealing how much each dyad engages in mutual slow nods. We correlated this slow-nod coherence across dyads to test if some dyads consistently engage more in coordinated slow nodding. Finally, we calculated if fast nod scores or slow nod coherence scores were related to the subjective reports at the end of the experiment. For this analysis, we used the same dyad slow-nod coherence score for each of the two individuals in the dyad.

Hand Motion Analysis. For the analysis of hand motions, we focus on the Y-axis (up/down) movements recorded from each person's right- and left-hand. We decided to focus on the Y axis because the video recordings from the testing sessions revealed that participants mostly moved the hands along this axis, and inclusion of the X and Z position would increase noise in the dataset. For each time point, we calculated the absolute level of motion in the y-axis (either up or down) using $y = \sqrt{(dy/dt)^2}$ where dy is the distance moved in that time window and dt is the duration of the time window (8.3 msec for recordings at 120Hz). We then computed wavelet coherence for each combination of right- and left-hand pairs across partners (i.e., right-right hands, left-left hands, right-left hands, left-right hands). As with the head analysis, these coherence values were then averaged across the trial time-course, with values outside the cone-of-influence excluded. We then averaged coherence values across trials of the same task, and across the four combinations of hand pairs, because we did not have a hypotheses for specific hand pairs. Overall, this resulted in a mean

coherence value for each wavelet, dyad, and task. For the pseudo interactions, we used the same approach as we used with the head motion analysis, shuffling data *within* dyad and *within* task. Following the same pipeline used for real coherence values, we calculated a mean coherence value for each wavelet, pseudo-dyad, and task. Finally, we compared real and pseudo coherence levels for each task by performing *t*-tests for each wavelet component, and corrected for multiple comparisons with an FDR of 0.05.

Results

Head Nodding Across Contexts

We use wavelet coherence to quantify nodding behaviour in each of the three conversation tasks, comparing data in real interactions to pseudo interactions to identify when fast and slow nodding occurs. Figure 6- A, B and C show the mean and standard error of coherence (R^2) for real (red) and pseudo (blue) interactions. High coherence means a high degree of coordination, as it indicates that two people are moving with the same frequency. To assess the difference in coherence between real and pseudo interactions, we performed t-tests (90 tests) at each frequency and calculated the effect size. Figure 6- D, E and F show the effect sizes (Cohen's d) calculated from the average coherence in real interactions minus the pseudo interactions. The dots indicate frequencies where there is a significant difference of coherence between real and pseudo interactions. Red dots represent points on the frequency range that pass a $p<0.05$ FDR significance threshold, while blue represent significant differences that did not pass this threshold.

From the graphs in Figure 6 we can observe two distinct patterns of coherence across the range of frequencies. These patterns are divided into two frequency ranges, above and below 1.5 Hz, as indicated by the dashed vertical line (Figure 6D, E, F). In the low frequency range (<1.5 Hz) results show greater coherence in the real compared to the pseudo interactions for the Picture Description Task. However, this pattern was not observed in the Video Discussion and the Meal Planning Tasks. In the high frequency range, there is less coherence in real compared to pseudo interactions in the Picture Description and Meal

Planning Tasks but this did not reach significant FDR corrected thresholds for the Video Discussion Task. In addition to these plots which analyse each task separately, we also present a cross-task ANOVA analysis in Appendix A.

Individual Differences

We tested if the participants show a reliable pattern of fast or slow nodding by correlating their individual (or dyadic) levels of fast and slow nodding across tasks (Figure 7). The results show that there is no reliable positive relationship between fast nodding behavior in any one task paired with any other task. There was also no reliable positive relationship between slow nodding coherence in any one task paired with any other task. There was a significant negative correlation, r=-0.43, p=.003, in slow nod coherence in the Meal Planning and Picture Description Task, but it did not pass FDR correction. In addition, we also tested if the tendency to nod is related to any of the personality traits measured in the questionnaires by performing correlations of the measures with relevant frequency bands of the wavelet data (high and low frequency nods) for each task separately. The questionnaires included the Liebowitz Social Anxiety Scale (LSAS), the Toronto Alexithymia Scale (TAS), the Adult Autism Spectrum Quotient (AQ), and the Experience of Gaze Questionnaire. The results show no correlations between the nodding measures and the questionnaire scores that passed an FDR correction.

Hand Coordination Across Contexts

Our analysis of hand movement matches the nodding analysis reported above, using wavelet coherence patterns to explore hand motion in real interactions and pseudo-data (Figure 8). We found a trend towards hand motion coordination primarily for the Meal Planning task in the low frequency range (0.13 – 1 Hz). The greater coherence across a wide range of frequencies in this task could reflect greater social engagement and more beat gestures. A small effect was seen in the Picture Description task at 0.13 Hz and in the Meal Planning task at 8 Hz, but 8 Hz effects are unlikely to be psychologically relevant. It is important to note that our sample size here was smaller than for the nodding analysis (n=21) and none of the results reached our FDR corrected threshold, so they must be considered marginal.

Discussion

This study tracked how head nodding and hand movement behaviours change across different conversational contexts to understand the social meaning of these behaviours. We aimed to discover what head nodding signals mean, if they are robust indicators of individual differences and if hand movements coordination can be quantified in the same way. Results showed that head nodding patterns differed between different contexts in line with our hypothesis. However, patterns of head nodding were not consistent in individuals across contexts, and patterns of hand movement were not easy to recover. We will discuss these results in relation to current studies of human social coordination. Coherent Slow Nodding Behaviour We find that participants show coherent patterns of slow nodding (0.2-1.1Hz range) during the picture description task but not during the meal planning or video discussion task. The results for the picture description task replicate the findings from Hale et al. (2020) using a higher resolution motion capture system. However, it is now clear that this pattern of behaviour does not generalise across contexts. This argues against the idea that slow head nods could be a form of social mimicry that facilitates bonding and affiliation between people (Lakin et al., 2003). Because if that were true, the coherence of slow nods should have been similar across the different conversational contexts due to the equal motivation to form social bonds during conversation. Instead, these results support the idea that slow head nods are a product of gaze following, which arises in the specific context of the Picture Description Task because here participants can alternate gaze between their partners face and the picture (held in one participant's lap) which requires up-down head movements. In contrast, the video discussion and meal planning tasks do not have a distinct gaze target. A prediction arising from these findings is that if participants were in a context where a shared gaze target was located beside them, rather than in one person's hands, then we would instead see coherence of 'head shaking' as they turn their heads towards the target. The question of how gaze following relates to other types of mimicry and social coordination remains open. One possibility would be to consider gaze-following to be a subset of a more general rubric of 'interpersonal coordination' or interaction. For example, Hadley & Ward (2021) have found an increase of low frequency overall head movement in triadic interactions when two people listened to a third (i.e., joint attention). Indeed, some studies which score mimicry behaviour based on video may not distinguish between gaze following and mimicry (Salazar-Kämpf et al., 2017). However, we suggest that it can be useful to make this distinction, because the two actions could have different social meanings. Gaze following is specific to the target of gaze (if an object is located on the left of A and on the right of B, then gaze following implies that A looks leftwards and B looks rightwards), whereas mimicry might be defined according to body-centred coordinates (I mimic a right-hand action with my right hand) (Liepelt, von Cramon, & Brass, 2008). This illustrates the importance of considering the physical and spatial context of actions carefully in any analysis of interpersonal coordination. Exploring the Fast-Nodding Behaviour Fast nodding arises when a listener makes small rapid head movements (2.6-6.5Hz) that do not match the movements of the speaker (Hale et al., 2020). Here, we find fast-nodding behaviour is present in the Picture Description task and the Meal Planning task but does not meet FDR correction in the Video Discussion task. Based on Hale et al. (2020), we suggested that fast nodding might be a backchannel related to listening and receiving information, and predicted that it should be present more often in the contexts where new information is exchanged. This prediction is supported in the current data. The Picture Description Task is a one-way information sharing context where the speaker is sharing new information to the listener about the picture. Similarly, the Meal Planning Task is a two-way information sharing context in which both participants are unaware of the other's meal preferences, while also having to share their own preferences. In both tasks, the exchange of new information seems to be linked to the presence of fast-nodding. In contrast, the Video Discussion Task is about shared recall between members of a dyad with no new information sharing, and fast nodding was not present here. It would be interesting in future to test if fast-nodding behaviour can provide a marker of successful information transfer in a conversation, and might be linked to later learning. Individual Differences in Nodding Behaviour If nonverbal behaviour can be robustly measured and linked to individual differences in personality or sociocognitive processes, this would be valuable for both research and clinical applications. Here, we took advantage of our data collection to test if the tendency to engage in fast or slow nodding behaviour is a fixed personality trait that differs between individuals. First, we tested if participants show a reliable pattern of fast or slow nodding by correlating these across tasks (Figure 7). That is, if a participant nods a lot in the Meal Planning Task, does that person also nod a lot in the Video Discussion and Picture Description Tasks? Second, we tested if the tendency to nod is related to any of the personality traits measured in questionnaires by correlating individual scores on fast nodding and coherent slow nodding with the questionnaire measures. If reliable individual differences in nodding behaviour could be identified, this would motivate us to test in future studies if the tendency to nod reflects broader social skills. The four questionnaires included the Liebowitz Social Anxiety Scale (LSAS), the Toronto Alexithymia Scale (TAS), the Adult Autism Spectrum Quotient (AQ), and the Experience of Gaze Questionnaire. In general, we did not find any evidence for reliable individual differences in nodding behaviour. Fast nodding behaviour in one task did not correlate with fast nodding in another task, nor did it correlate with any questionnaire measures. Slow nodding coherence in one task did not correlate with the same measure in a different task, and nor did it correlate with any questionnaire measures. This means we can reject the idea that head nodding is linked to fixed personality traits or provides a stable individual difference. Our study is limited in that each person only appears in one dyad, so we are not able to quantify each person's behaviour independent of their interaction partner, as done by Salazar-Kämpf et al. (2017). However, at present there is no strong reason to use fast or slow nodding behaviour as a measure of an individual's social skills or as a clinical assessment. This is relevant because studies are attempting to use automated analyses of interactive behaviour to identify and even diagnose disorders of social interaction such as autism (Georgescu et al., 2019). Coordination of Hand movements To our knowledge, only two previous studies have investigated interpersonal coordination of hand gestures during conversation (Holler & Wilkin, 2011; Ramseyer & Tschacher, 2016). These studies found that pairs of participants coordinate hand gestures during conversation, based on video recordings. Here, we use high-resolution motion capture and wavelet coherence analysis to determine if there are robust patterns of coordination in hand movement which could be detected with simple automated methods. Comparison of real versus pseudo interactions revealed that dyads showed weak coordination of low-frequency hand gestures (0.13 Hz to 1 Hz) during the Meal Planning task, although this effect was absent for the Picture Description and Video Discussion tasks. This implies that it is the combination of spontaneous forms of conversation plus the sharing of novel information that incentivises the coordination of hand gestures. The Meal Planning task, which was more dynamic with more overlaps, may have promoted more use of communicative gestures (McNeill, 1992) or beat gestures (Bosker & Peeters, 2021). In contrast, one participant's hands were occupied with the picture during the Picture Description task, while the Video Discussion task involves less information exchange. Note, however, that these results did not pass the correction for multiple comparisons, which limits the extent of our interpretations. A further limitation of our analysis is that we focus only on vertical hand movements, because these were the largest and clearest in our data, and we cannot discriminate between different types of gestures (iconic vs. beat gestures). It is possible that hand movements are much more complex and multidimensional than head movements, so a simple wavelet analysis is unable to capture the richness of the data and a more detailed video coding approach would be needed to understand hand motion coordination. However, it is promising that even with a small sample size, our automated analysis of the Meal Planning task was able to show some evidence of coordination of hand movements, and future work could examine the types of gesture involved and what they mean in more detail. Future Directions The present paper provides evidence that the social coordination of head nods and hand gestures changes in different conversational contexts, and opens up a large number of possible future directions. First, more detailed study will be needed to resolve some limitations of the present work. For example, studies using a round-robin design (Salazar-Kämpf et al., 2017), could provide more robust analyses of individual differences in nodding behaviour and what they might mean. Second, the complex multimodal data collection setup in our lab requires a lot of equipment, which reduces the naturalness of the conversations. Future studies could use video tracking in conjunction with machine learning (i.e., OpenPose) (Cao, Simon, Wei, & Sheikh, 2017) or wearable motion sensors (Ward & Pinti, 2019; Sun, Greaves, Orgs, Day, Hamilton & Ward, 2023) to track social coordination in a less obtrusive fashion and in novel contexts outside of the lab. The insights gained from motion capture studies of human social interaction can also contribute to the challenge of building realistic virtual humans who are able to interact in meaningful ways (Aburumman, Gillies, Ward, & Hamilton, 2022). Now that methods for tracking the coordination of nodding and head movements are becoming more established, it will also be possible to expand our understanding of how these signals relate to other cognitive processes. For example, we suggest here that fast-nodding signals are a back channel related to the exchange of information. It would be interesting to test if fast-nodding is related to successful learning. It could also be interesting to test if the coherence of slow head movements is related to joint attention in other contexts and with other potential gaze target locations. Finally, integrating the study of behaviour across modalities remains a major challenge for researchers in this area. The present paper examines head and hand movements, while another paper based on the same data (Dobre, Gillies, Falk, Ward, Hamilton, & Pan, 2021) examines gaze and speech. Analyses which can integrate these diverse signals will be very valuable in gaining a more rounded understanding of the richness of human social interaction. Such analyses will likely require a detailed consideration of the social meaning behind different behaviours (Hadley, Naylor, & Hamilton, 2022) and also integration with verbal behaviours (Reece et al., 2023). Rapid advances in these areas provide a lot of promise for future studies of nonverbal communication.

Conclusions

In the present paper, we report on patterns of fast nodding, slow nodding and hand movement coordination in three different types of conversation between pairs of strangers. Our data suggest that fast head nods are a signal of having received new information, while slow head nods may be coordinated to direct gaze to a shared object. We also suggest that neither of these behaviours are linked to stable personality traits, but rather they differ strongly with the type of conversation. We also find weak evidence for slow coordination of hand movements in some contexts. Overall, these results advance our understanding of how nonverbal coordination works, how it can be measured, and how these measures could be used to answer a wide range of questions in the domain of human social interaction.

Declarations

The authors declare that they have no conflict of interest.

Acknowledgements

The research was supported by the Leverhulme Trust.

References

1. Aburumman, N., Gillies, M., Ward, J. A., & de Hamilton, A. F. C (2022). Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies*, *164*, 102819.
2. Allwood, J., & Cerrato, L. (2003). A study of gestural feedback expressions. In.
3. Paggio, et al. (Eds.). *Proceedings of the First Nordic Symposium on Multimodal*.
4. Communication, Copenhagen.
5. Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge University Press.
6. Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale: I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1.
7. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/A:1005653411471.
8. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B Methodological*, *57*(1), 289–300. doi.org/10.1111/j.2517-6161.1995.tb02031.x.
9. Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman, & B. Rimé (Eds.), *Studies In Emotion & Social Interaction. Fundamentals of Nonverbal Behavior* (pp. 401–432). Cambridge University Press.
10. Bosker, H. R., & Peeters, D. (2021). Beat gestures influence which speech sounds you hear. *Proc Roy Soc B*. 10.1098/rspb.2020.2419.
11. Cao, Z., Simon, T., Wei, S., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7291–7299.
12. Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. doi.org/10.1037//0022-35er14.76.6.893.

13. Chovil, N. (1991). Discourse oriented facial displays in conversation. *Research on Language and Social Interaction*, *25*, 163–194. doi.org/10.1080/08351819109389361.
14. Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7.
15. Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. https://doi.org/10.1016/j.cognition.2007.04.010.
16. Daly, J. A., Hogg, E., Sacks, D., Smith, M., & Zimring, L. (1983). Sex and relationship affect social self-grooming. *Journal of Nonverbal Behavior*, *7*(3), 183–189. https://doi.org/10.1007/BF00986949.
17. Dobre, G. C., Gillies, M., Falk, P., Ward, J., Hamilton, A., & Pan, X. (2021). Direct gaze triggers higher frequency of gaze change: An automatic analysis of dyads in unstructured conversation. In *Proceedings in International Conference on Multimodal Interaction (ICMI '21)*, Montreal, Canada, 735–739. 10.1145/3462244.3479962.
18. Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*(2), 283–292. https://doi.org/10.1037/h0033031.
19. Emery, N. J. (2000). The eyes have it: The neuroethology, function, and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*, 581–604. doi.org/10.1016/S0149-7634(00)00025-7.
20. Feese, S., Arnrich, B., Tröster, G., Meyer, B., & Jonas, K. (2011). Detecting posture mirroring in social interactions with wearable sensors. *Proceedings of the 15th International Symposium on Wearable Computers*, *USA*, *5959582*, 119–120. 10.1109/ISWC.2011.31.
21. Fujiwara, K., & Daibo, I. (2016). Evaluating interpersonal synchrony: Wavelet transform toward an unstructured conversation. *Frontiers in Psychology*, *7*, 516. doi.org/10.3389/fpsyg.2016.00516.
22. Genschow, O., Florack, A., & Wänke, M. (2013). The power of movement: Evidence for context-independent movement imitation. *Journal of Experimental Psychology: General*, *142*(3), 763–773. https://doi.org/10.1037/a0029795.
23. Georgescu, A. L., Koehler, J. C., Weiske, J., Vogeley, K., Koutsouleris, N., & Falter-Wagner, C. (2019). Machine Learning to Study Social Interaction Difficulties in ASD. *Frontiers in Robotics and AI*, *6*, 132. doi.org/10.3389/frobt.2019.00132.
24. Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*, 561–566. doi.org/10.5194/npg-11-561-2004.
25. Hadar, U., Steiner, T. J., Grant, E. C., & Rose, F. C. (1983). Kinematics of head movements accompanying speech during conversation. *Human Movement Science*, *2*(1–2), 35–46. https://doi.org/10.1016/0167-9457(83)90004-0.
26. Hadley, L. V., Naylor, G., & de Hamilton, A. F. C (2022). A review of theories and methods in the science of face-to-face social interaction. *Nature Reviews Psychology*, *1*, 42–54. https://doi.org/10.1038/s44159-021-00008-w.
27. Hadley, L. V., & Ward, J. A. (2021). Synchrony as a measure of conversation difficulty: Movement coherence increases with background noise level and complexity in dyads and triads. *PLOS ONE*,

16(10), e0258247. https://doi.org/10.1371/journal.pone.0258247.

28. Hale, J., Ward, J. A., Buccheri, F., Oliver, D., & de Hamilton, A. F. C (2020). Are you on my wavelength? Interpersonal coordination in naturalistic conversations. *Journal of Non-Verbal Behaviour*, *44*, 63–83.
29. Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in cognitive sciences*, *16*(2), 114–121. https://doi.org/10.1016/j.tics.2011.12.007.
30. Heerey, E. A. (2015). Decoding the dyad: Challenges in the study of individual differences in social behavior. *Current Directions in Psychological Science*, *24*(4), 285–291. doi.org/10.1177/0963721415570731.
31. Hirsch, J., Adam Noah, J., Zhang, X., Dravida, S., & Ono, Y. (2018). A cross-brain neural mechanism for human-to-human verbal communication. *Social cognitive and affective neuroscience*, *13*(9), 907–920. https://doi.org/10.1093/scan/nsy070.
32. Hoehl, S., Fairhurst, M., & Schirmer, A. (2020). Interactional synchrony: Signals, mechanisms and benefits. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 5–18. 10.1098/scan/nsaa024.
33. Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, *35*(2), 133–153. https://doi.org/10.1007/s10919-011-0105-6.
34. Hömke, P., Holler, J., & Levinson, S. (2018). Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLOS ONE*, *13*, e0208030. 10.1371/journal.pone.0208030.
35. Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant behavior & development*, *34*(4), 569–577. https://doi.org/10.1016/j.infbeh.2011.06.008.
36. Issartel, J., Marin, L., Gaillot, P., Bardainne, T., & Cadopi, M. (2006). A practical guide to time-frequency analysis in the study of human motor behaviour: The contribution of the wavelet transform. *Journal of Motor Behaviour*, *38*(2), 139–159. doi.org/10.3200/JMBR.38.2.139-159.
37. Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, *89*(1), 253–260. https://doi.org/10.1016/S0093-934X(03)00335-3.
38. Kelly, S. D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An intentional stance modulates the integration of gesture and speech during comprehension. *Brain and Language*, *101*(3), 222–233. https://doi.org/10.1016/j.bandl.2006.07.008.
39. Kendon, A. (2002). Some uses of head shake. *Gesture*, *2*(2), 147–182. https://doi.org/10.1075/gest.2.2.03ken.
40. Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic co-ordination of speech and gesture reveal? Evidence of an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.
41. Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E. M., Wallot, S., Van Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-

walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8514–8519. https://doi.org/10.1073/pnas.1016955108.

42. 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 Behaviour*, *27*(3), 145–162. doi.org/10.1037/e413812005-152.
43. Liebowitz, M. R. (1987). Ban., P. Pichot.,W. Pöldinger (Ed.), *Modern Problems of Pharmacopsychiatry*: *Vol. 22. Anxiety* (pp. 141–173). 10.1159/000414022.
44. Liepelt, R., Cramon, D. Y. V., & Brass, M. (2008). What is matched in direct matching? Intention attribution modulates motor priming. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 578–591. https://doi.org/10.1037/0096-1523.34.3.578.
45. Loehr, D. (2007). Aspects of rhythm in gesture and speech. *Gesture*, *7*(2), 179–214. https://doi.org/10.1075/gest.7.2.04loe.
46. Marstaller, L., & Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychonomic Bulletin & Review*, *20*(3), 496–500. https://doi.org/10.3758/s13423-012-0365-0.
47. McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
48. Pan, X., & de Hamilton, A. F. C (2015). Automatic imitation in a rich social context with virtual characters. *Frontiers of Psychology*, *6*, 790. 10.3389/fpsyg.2015.00790.
49. Poggi, I., D'Errico, F., & Vincze, L. (2010). Types of nods. The polysemy of a social signal. *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, Valetta, Malta, 596.*
50. Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: Coordinated body-movements reflects relationship quality and outcome. *Journal of Consulting and Clinical Psychology*, *79*(3), 284–295. 10.1037/a0023419.
51. Ramseyer, F., & Tschacher, W. (2016). Movement coordination in psychotherapy: Synchrony of hand movements is associated with session outcome. A single-case study. *Nonlinear Dynamics Psychology and Life Sciences*, *20*(2), 145–166.
52. Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature reviews Neuroscience*, *20*(8), 495–505. https://doi.org/10.1038/s41583-019-0179-4.
53. Reece, et al. (2023). The CANDOR corpus: Insights from a large multimodal dataset of naturalistic conversation. *Science Advances*, *9*, 10.1126/sciadv.adf3197.
54. Roberts, S. (2011). *DipDap.* [BBC]. Ragdoll Productions.
55. Salazar-Kämpf, M., Liebermann, H., Kerschreiter, R., Krause, S., Nestler, S., & Schmukle, C. (2017). Disentangling the sources of mimicry: Social relations analyses of the link between mimicry and liking. *Psychological Science*, *29*(1), 131–138. 10.1177/0956797617727121.
56. Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second person neuroscience. *Behavioural and Brain Sciences*, *36*(4), 393–462.

10.1017/S0140525X12000660.
57. Stel, M., van Baaren, R. B., Blascovich, J., van Dijk, E., McCall, C., Pollmann, M. M., van Leeuwen, M. L., Mastop, J., & Vonk, R. (2010). Effects of a priori liking on the elicitation of mimicry. *Experimental psychology*, *57*(6), 412–418. https://doi.org/10.1027/1618-3169/a000050.
58. Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14425–14430. https://doi.org/10.1073/pnas.1008662107.
59. Sun, Y., Greaves, D., Orgs, G., Hamilton, A. F., de Day, C., S., & Ward, J. A. (2023). Using Wearable Sensors to Measure Interpersonal Synchrony in Actors and Audience Members During a Live Theatre Performance. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *7*(1), 27. ISSN 2474–9567.
60. Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*, 1323. 10.3389/fpsyg.2014.01323.
61. van Baaren, R. B., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2381–2389. https://doi.org/10.1098/rstb.2009.0057.
62. Ward, J. A., & Pinti, P. (2019). Wearables and the Brain. In *IEEE Pervasive Computing*, vol. 18, no. 1, pp. 94–100, Jan.-March 2019, 10.1109/MPRV.2019.2898536.

Figures

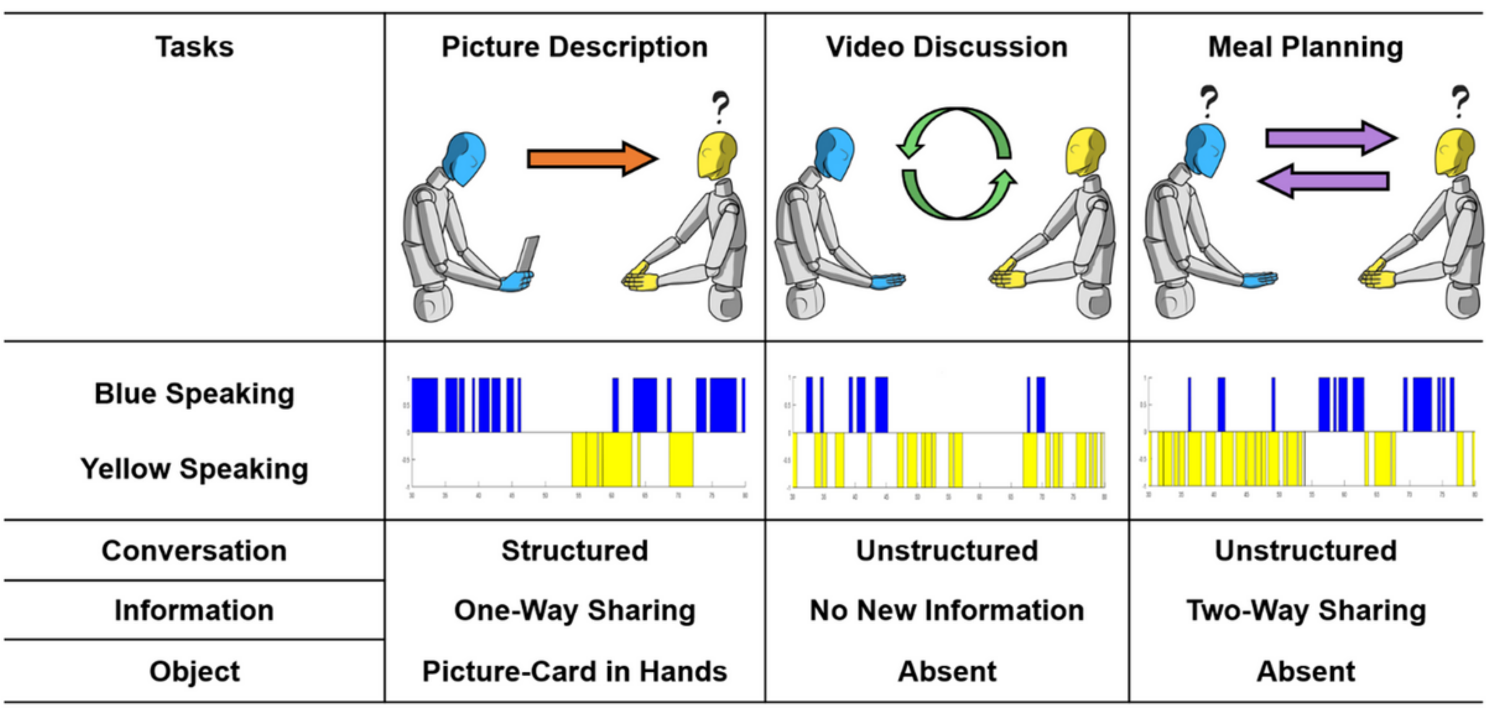

Figure 1

Conversational tasks. Panel (A) Picture Description; structured one-way information sharing holding an object. Panel (B) Video Discussion; unstructured common ground with no new information sharing. Panel (C) Meal Planning; unstructured two-way information sharing. Graphs (middle) show a sample of the turn-taking structure for each task in this experiment, highlighting the order and how often blue and yellow participants passed their turns.

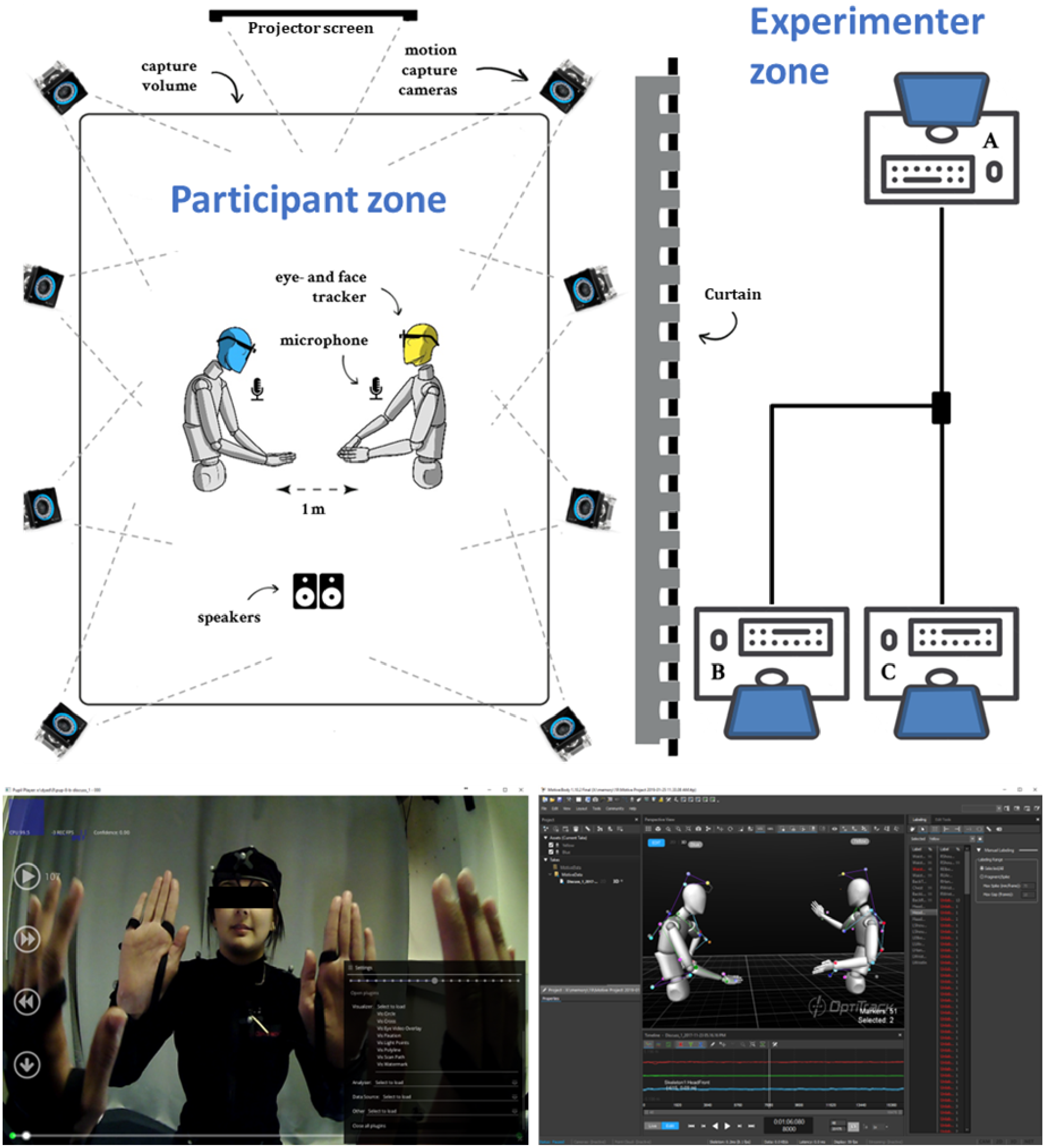

Figure 2

Experimental Setup and Data Capture. Equipment included motion tracking cameras (4x Optitrack Prime 13, and 4x 13W), a projector, speakers, wearable microphones connected to an audio mixer, eye- and face trackers (Pupil Labs), and a curtain to separate the three computers running the experiment. Computer A acted as the client computer, that communicated with the two computers B and C acting as servers running the recording software; *Bottom left*: Pupil Player output; *Bottom right:* Optitrack Motive output. Audacity was used to record the verbal components of the interaction.

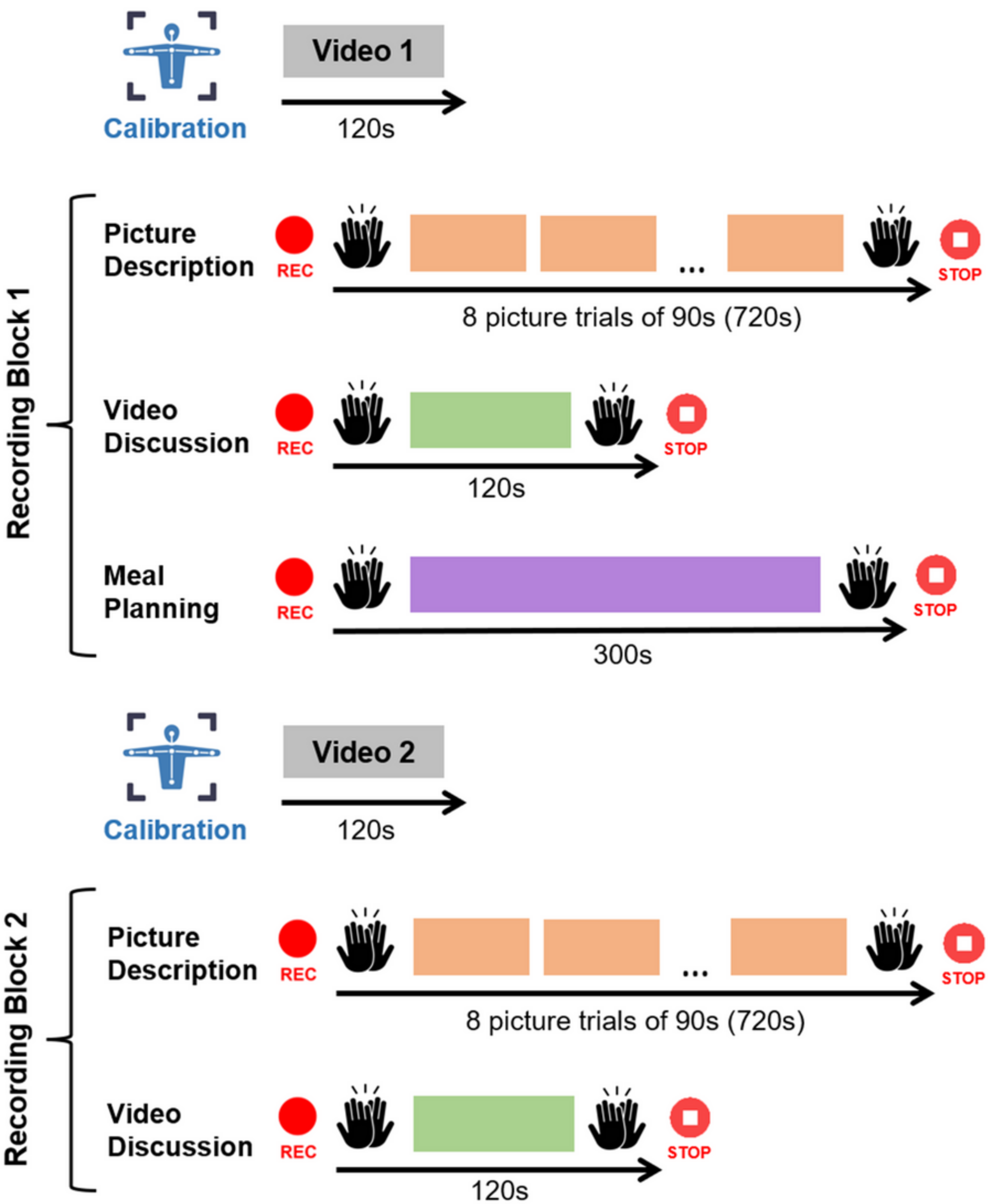

Figure 3

Experimental session timeline. Two calibrations were performed to ensure the eyetrackers and motion capture gave high-quality data (T-pose icon) and after each calibration, participants watched a short video together. Data was recorded in two recording-blocks during five task-blocks which occurred in a fixed order as shown in the figure. Each task-block began and ended with a synchronization event (handclap icon). The Picture Description and Video Discussion tasks were completed twice, whereas the Meal Planning Task completed only once.

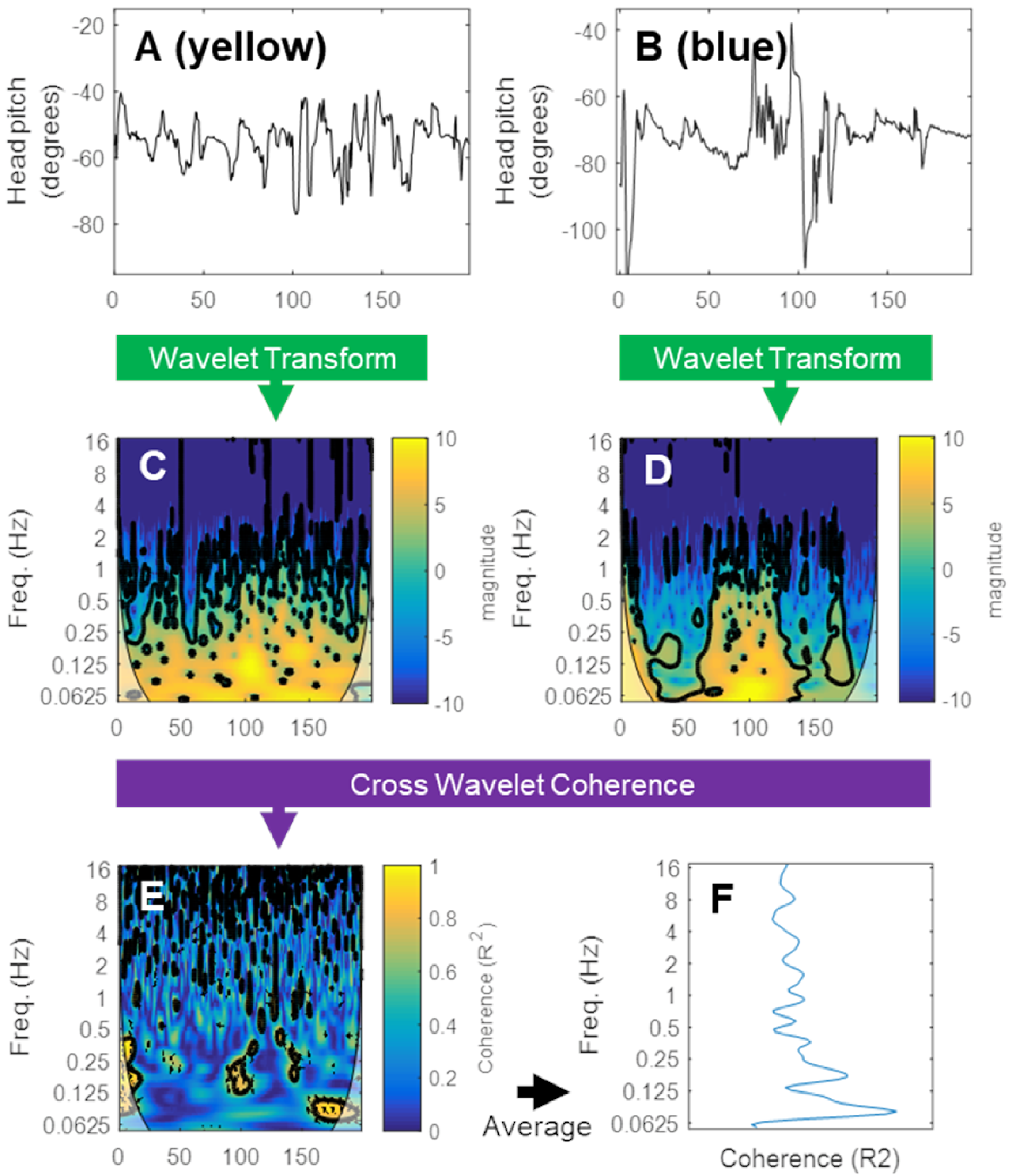

Figure 4

Cross-wavelet analysis pipeline. For each trial, the head pitch trajectories for both the Yellow and Blue participants (A, B) are subject to a wavelet transform (C, D). Then, the cross-wavelet coherence is calculated between the two participants (E). The magnitude of wavelet power and wavelet coherence is represented by color, where blue is low power, and yellow is high power. The time is represented on the x-axis (200s) and each frequency on the y-axis. The coherence value (R^2) is then averaged over time and over all trials to obtain the overall frequency of coherence in head pitch between the two participants (F).

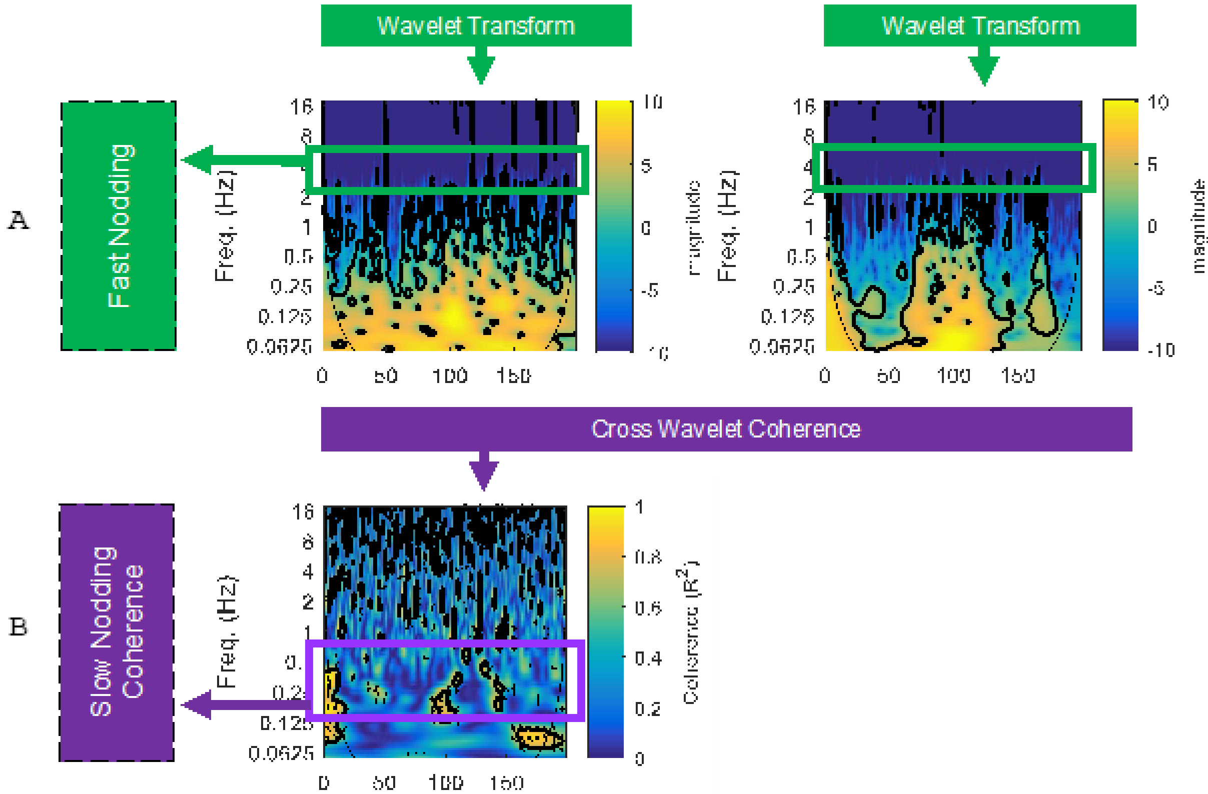

Figure 5

Fast and slow head nodding scores. The fast nod score was selected from the high frequency band of the individual wavelets (A). The slow nod coherence score was selected from the low frequency band of the cross-wavelet coherence (B).

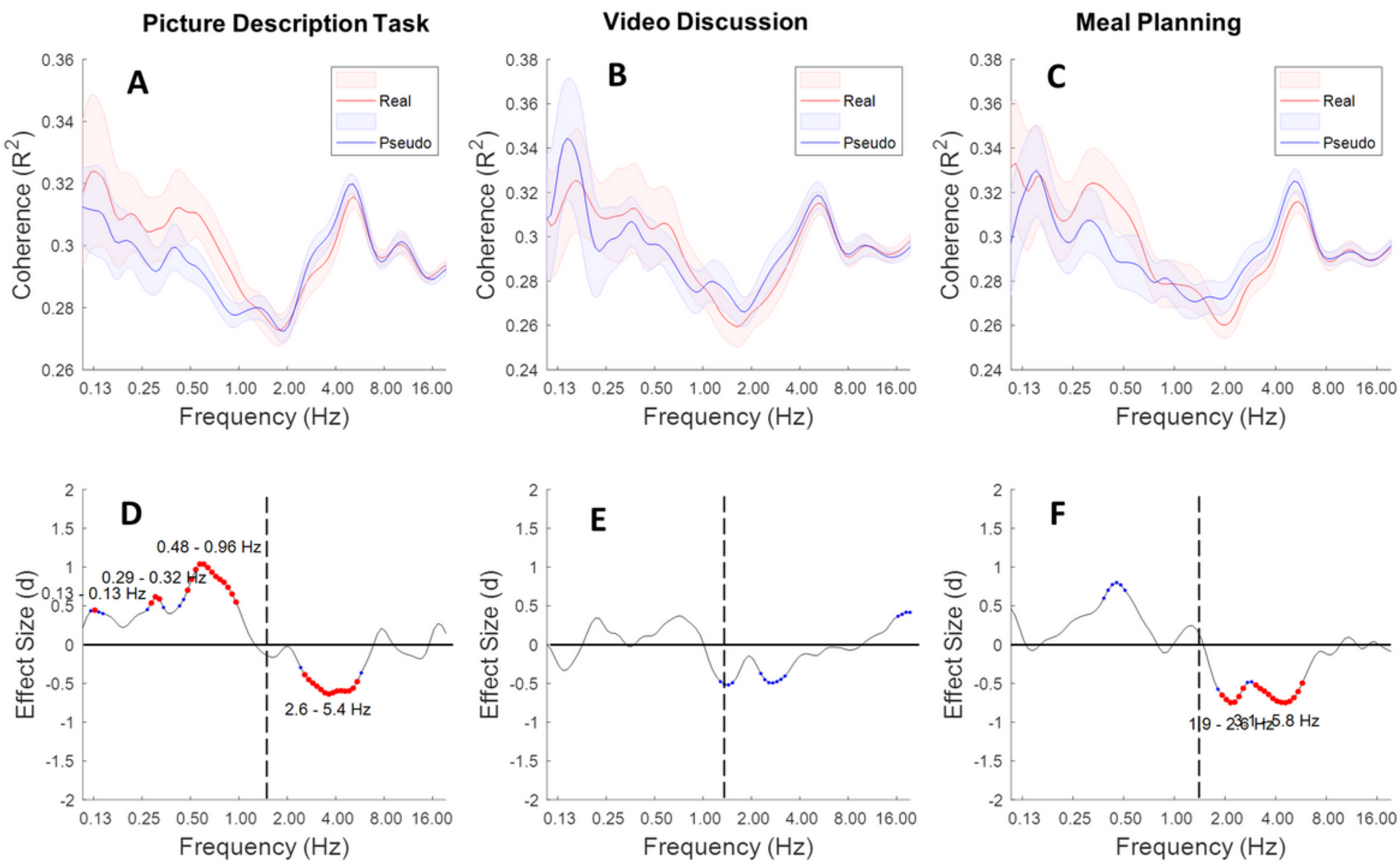

Figure 6

Head nodding real vs pseudo coherence (A,B,C) and effect sizes (D,E,F) across the full frequency range for the three tasks. p<0.05 significance levels shown by blue dots, with FDR adjusted significance highlighted in red (D,E,F).

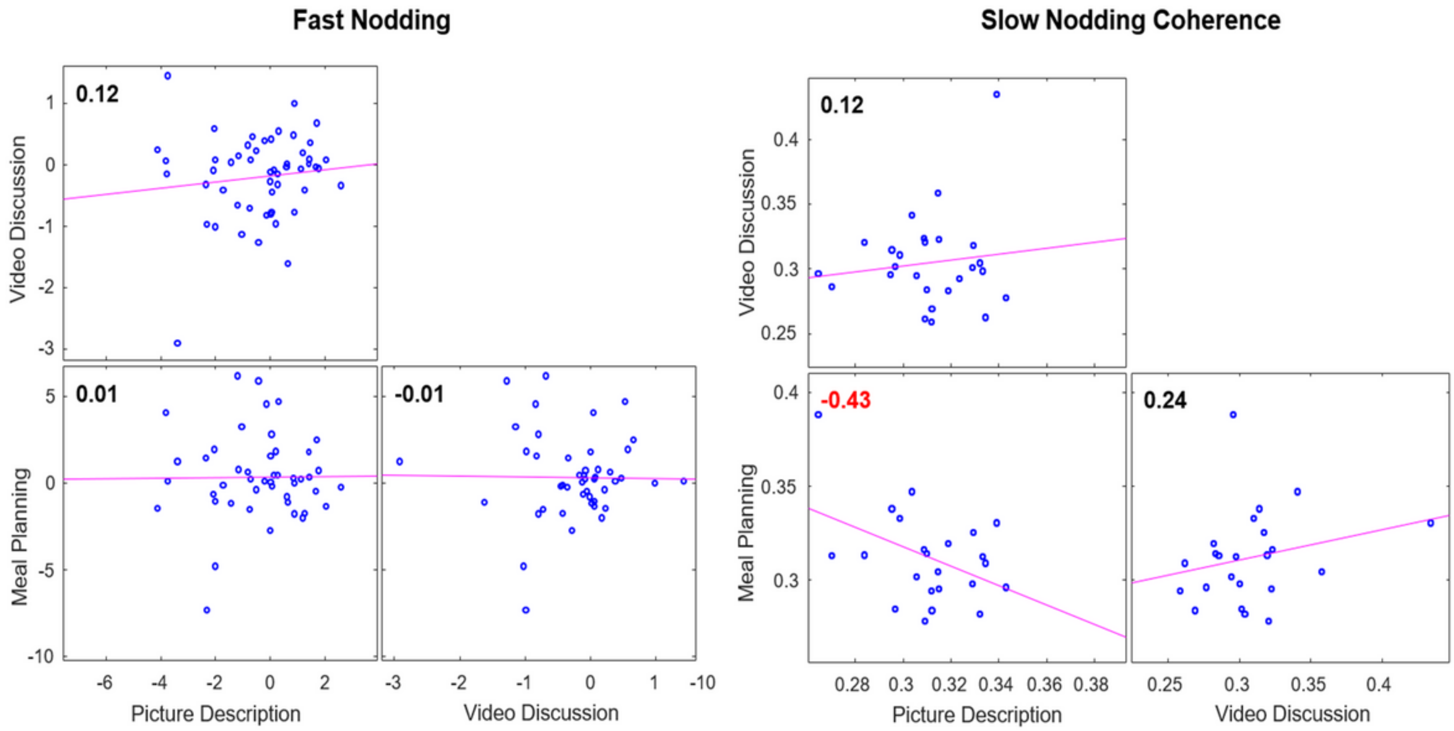

Figure 7

Robustness of individual differences in head nodding behaviour. Red correlation scores indicate if the correlation is significantly ($p<0.05$) different from zero. The axis values for fast nods are the average power in the 2.6 – 6.5 Hz frequency band.

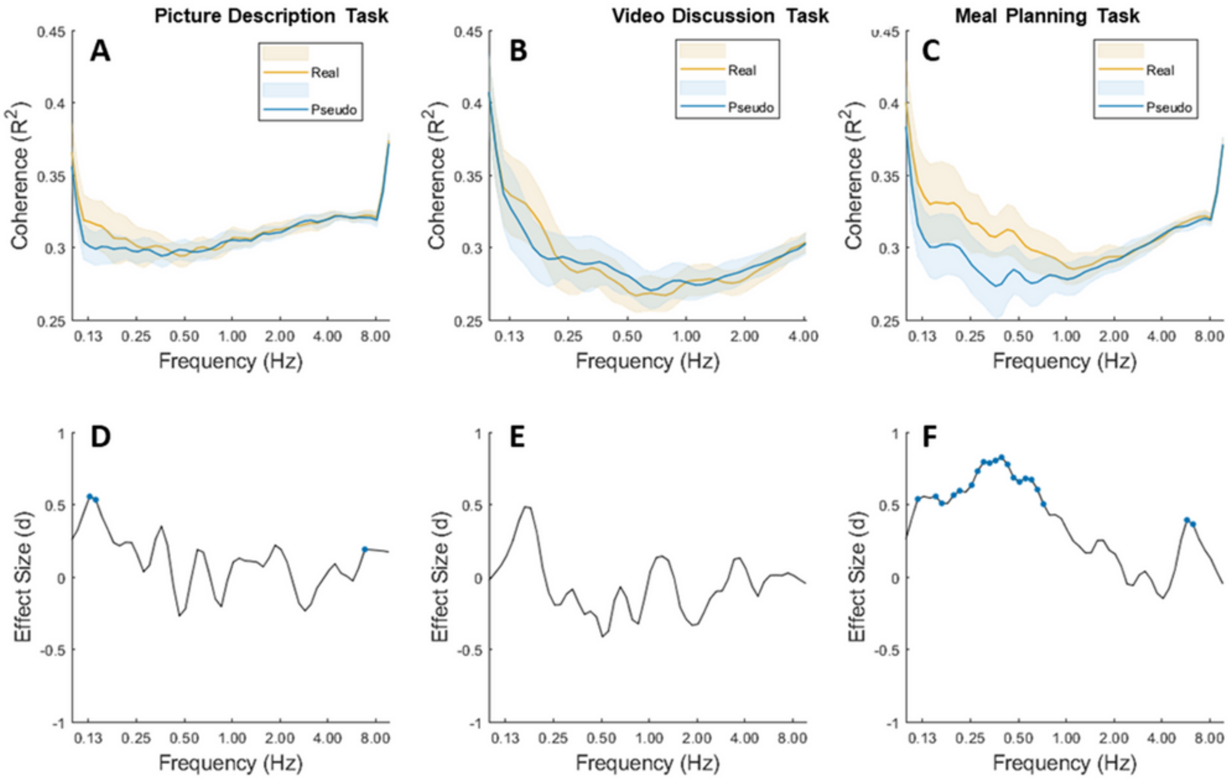

Figure 8

Hand coordination real vs pseudo coherence (A,B,C) and effect sizes (D,E,F) across the full frequency range for the three tasks. *p*<0.05 uncorrected significance levels shown by blue dots (D,E,F).

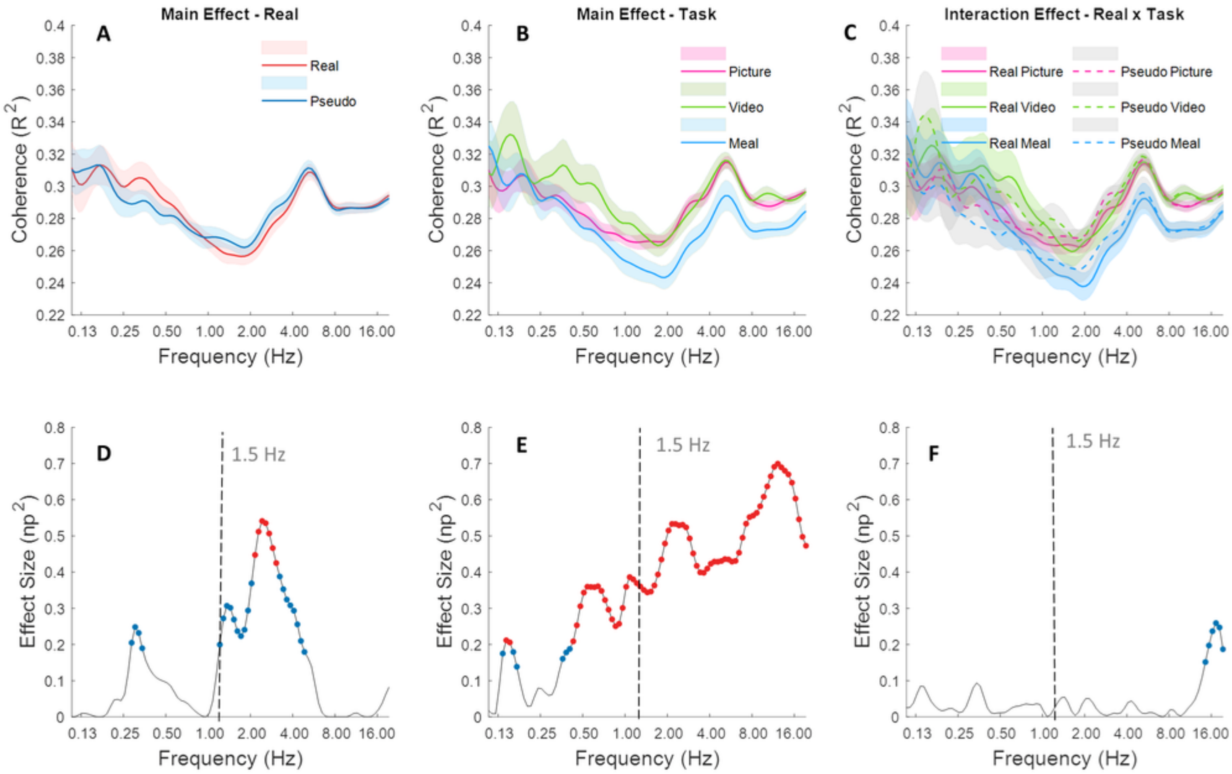

Figure 9

ANOVA cross-wavelet coherence. Graphs A, B, and C show the mean and standard error of coherence (R^2) of each effect. Graphs D, E, and F show the effect sizes (partial eta-squared, ηp^2). The dotted line indicates frequencies where there is a significant difference of coherence. Red dots represent points on the frequency range that pass a *p*<0.05 FDR significance threshold, whereas blue dots represent significant differences that did not pass this threshold.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AppendixA.docx