http://www.diva-portal.org

This is the published version of a paper published in *Journal of nonverbal behavior*.

Citation for the original published paper (version of record):

Falk, P., Cañigueral, R., Ward, J A., Hamilton, A F. (2026)
Head Nodding and Hand Coordination Across Dyads in Different Conversational Contexts
Journal of nonverbal behavior
https://doi.org/10.1007/s10919-026-00532-4

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

https://creativecommons.org/licenses/by/4.0/

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-27183

Journal of Nonverbal Behavior
https://doi.org/10.1007/s10919-026-00532-4

ORIGINAL PAPER

Head Nodding and Hand Coordination Across Dyads in Different Conversational Contexts

Patrick Falk[1,2] · **Roser Cañigueral**[2] · **Jamie A Ward**[2,3] · **Antonia F de C Hamilton**[2]

Received: 5 April 2026 / Accepted: 22 September 2026
© The Author(s) 2026

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 found a non-significant trend toward low-frequency coordination during two-way information sharing in the Meal Planning task, but no effects survived correction for multiple comparisons. Finally, nodding varied across contexts, and our exploratory analyses did not provide evidence that it indexes stable individual differences in social traits. 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.

Keywords Nonverbal behavior · Conversational context · Social interaction · Interpersonal coordination · Head nodding · Hand coordination

Introduction

Interpersonal coordination refers to the temporal alignment of two or more individuals while they interact with each other (Hoehl et al., 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 et al., 2020; Ramseyer & Tschacher, 2011; Richardson & Dale, 2005), physiological signals (Feldman et al., 2011; Konvalinka et al., 2011) or brain

Extended author information available on the last page of the article

activity (Hirsch et al., 2018; Stephens et al., 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 et al., 2012) and increasing affiliation (Lakin et al., 2003). From a dynamical and embodied perspective, this tendency to coordinate is not merely incidental but reflects a fundamental basis for social connectedness, whereby individuals become coupled into emergent social units through shared perceptual-motor dynamics (Marsh et al., 2009). Despite these advances, the specific mechanisms underlying interpersonal coordination, and how they operate across multiple modalities and timescales, remain poorly understood (Dale et al., 2013). One influential account proposes that such coordination reflects emergent dynamical processes, in which stable patterns of synchronization arise through the self-organization of coupled behavioural systems rather than solely through internal representations or mirroring mechanisms (Schmidt et al., 2011). Consistent with this view, coordination has been shown to emerge unintentionally from informational coupling in dyadic social interaction, particularly via visual links, rather than from deliberate attempts to synchronize behaviour (Richardson et al., 2005).

However, capturing these dynamics empirically remains challenging. Traditional methods used to record and analyse interpersonal coordination in dyadic interactions often fail to capture the full richness of interaction dynamics. After Chartrand and Bargh (1999) published their work on experimentally manipulating mimicry (i.e., 'the chameleon effect'), researchers started using confederate paradigms to study mimicry (Stel et al., 2011; van Baaren et al., 2004). However, dyads in real-world interactions had seldom been assessed directly. In a study by Ramseyer and Tschacher (2011), based on sessions during psychotherapy, they hypothesized that automated analyses of video might be able to detect mimicry in head and body movements. This automated method tightly controls the variables involved and isolates targeted behavioural or cognitive constructs, which allows a more detailed understanding of patterns of mimicry in real-world conversation. However, standard single-camera frame-differencing approaches do not recover depth or directional kinematics; three-dimensional motion capture can preserve these components when they are relevant to the movement under study (Risko et al., 2016).

Understanding the timing and at which frequencies all these behaviours occur can help us answer how and why we use these signals. Studies have revealed that interpersonal coordination occurs at multiple timescales (Schmidt et al., 2014) and frequencies (Fujiwara & Daibo, 2016). Such multiscale organization is consistent with a dynamical systems perspective, in which coordination emerges from interacting processes operating across nested temporal structures rather than from a single centralized control mechanism (Schmidt et al., 2011). In the study by Schmidt et al. (2014), they investigated how bodily coordination is distributed across different nested timescales in a joke telling task during structured conversation. Dyads performed knock-knock jokes while the researchers recorded their movement using a Kinect camera. To assess the degree and pattern of participants' movement coordination at individual timescales, they used cross-wavelet coherence (CWC) and calculated both the coherence and phase measures for the whole trial, as well as at the minor timescales of the whole joke, the setup of the punch line, the turn-taking exchange, and the utterance. The coherence analysis revealed a greater than chance coordination (high coherence) of the joke teller's (speaker) and joke responder's (listener) movements at all timescales. In addition, the phase measure analysis revealed that the joke teller's movements led those of the joke responder at the longer timescales, which shows that the movements were coordinated in

phase rather than antiphase. These results demonstrate that complex interpersonal coordination is constructed from a set of rhythms associated with nested timescales within a structured social interaction. Fujiwara and Daibo (2016) applied CWC to a combined movement signal derived from fingertip and nose coordinates during six-minute unstructured conversations in which no rhythmic task was prescribed and turn-taking was not controlled. Genuine dyads showed higher average coherence below 4 Hz than pseudo pairs. Within genuine dyads, coherence decreased as frequency increased, with 0.5 Hz marking the reported boundary between higher- and lower-coherence frequency bands. These findings support the use of CWC to detect interpersonal rhythmic similarity in free-flowing conversation.

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 proves consistent across repeated contexts, automated movement measures might contribute, alongside validated measures and appropriately designed samples, to research on individual differences in social behaviour (Heerey, 2015). Such consistency would be a prerequisite, not evidence by itself, for personality inference or clinical assessment. This caution is important because spontaneous gesture rates can predict learning in some contexts (Congdon et al., 2025), whereas dyadic coordination also reflects partner- and context-dependent processes (Fitzpatrick et al., 2017). 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 nonverbal signals during a conversation are centred around the head (e.g., eye-gaze, blinks, facial expressions, and head movements), and listeners' 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 et al., 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). Following Hale et al. (2020), we use fast-nodding dynamics to denote wavelet power above 1.5 Hz and slow nodding to denote components below 1.5 Hz. In behavioural terms, these fast movements resemble the small, brief head-pitch movements that often accompany a conversational backchannel such as "uh-huh," rather than large, emphatic affirmative nods. A frequency of 2.6–6.5 Hz corresponds to approximately 2.6–6.5 complete pitch oscillations, or up-down cycles, per second during a segment containing such oscillatory activity. This does not imply that participants continuously produced this number of conventional nods each second. Because the continuous wavelet analysis did not

segment discrete nod events, it does not estimate the peak-to-peak angular amplitude, duration, or number of cycles within an individual nodding bout. 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 nonverbal 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 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). A long tradition of research into human interpersonal coordination describes patterns of synchrony or mimicry as integral nonverbal behaviours in dyadic social interaction (Bernieri et al., 1991). As noted by Ramseyer and Tschacher (2006), the distinction between 'synchrony' and 'mimicry' is important as these concepts have recently been regarded as two distinct phenomena that have occupied separate research fields. This distinction is central to dynamical accounts of interpersonal coordination, which emphasize that synchrony reflects the temporal organization of coupled behavioural rhythms, whereas mimicry can occur without precise temporal alignment and may arise from different underlying mechanisms (Richardson et al., 2005; Schmidt et al., 2011). However, both perspectives apply the same reasoning and theory. For a simple interaction between two people, we can imagine different mechanisms at play that match with different timing properties. Mimicry occurs when one person performs an action and another performs the action afterwards, which implies a lag between the two actions. The relevant lag depends on the behaviour and analytic scale: it can be subsecond for rapid responses, several seconds for delayed mimicry, and as long as 20–30 s in infant-caregiver vocal coordination (Jaffe et al., 2001). Synchronous behaviour (e.g., rhythmic walking, musical coordination), on the other hand, are those behaviours that are matched in time and occur simultaneously with no time lag (e.g., 0 ms). As such, synchrony and mimicry behaviour can be observed simultaneously during social interactions, and much research has focused on comparing their different causes and effects (Dale et al., 2013; Marsh et al., 2009). In this paper, we will use the term 'mimicry' to refer to behavioural mimicry with a time lag. 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). More broadly, interpersonal coordination, whether synchronous or time-lagged, has been proposed to underpin the emergence of social connection and rapport by coupling individuals into shared behavioural systems (Marsh et al., 2009). Previous work (Hale et al., 2020) identified coordinated slow nodding with a timescale which matches previous reports of mimicry (Hömke et al., 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; Richardson et al., 2007). Eye and head movements are coupled components of gaze shifts: during natural viewing, the head contributes to gaze reorientation, and eye and head latencies can be highly correlated (Einhäuser et al., 2007; Khan et al., 2009). Thus, head-pitch coherence may partly index shared gaze reorientation rather than a discrete nod signal. 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). Cross-wavelet work further shows that gesture-speech synchrony re-stabilizes after delayed auditory feedback, illustrating that the timing of hand movement is dynamically coupled to speech (Pouw & Dixon, 2019). 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 et al., 2004, 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 et al., 2008; Kendon, 1972; Marstaller & Burianová, 2013; McNeill, 1992). Another class of non-co-speech hand movement includes self-grooming and self-touch, often termed self-adaptors (e.g., touching or adjusting the body, clothing, or accessories). In addition to grooming and impression management, conversational self-touch is associated with state anxiety and has been proposed to support self-calming or emotion regulation; because the available evidence is correlational, a causal calming function should be stated cautiously (Daly et al., 1983; Pang et al., 2022). 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 et al., 2013; Pan & Hamilton, 2015; Stel et al., 2010), and a small number of studies have investigated hand-movement coordination in face-to-face interaction. 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 the patient's ratings of affiliation with the therapist. Beyond these video-based studies, de Jonge-Hoekstra (2021, Chapter 5) used motion capture and CWC to examine children's speech, hand, and head coordination during dyadic problem solving. Head and hand coherence was greater at middle (2–0.5 Hz) and slow (0.5–0.25 Hz) than at fast (8–2 Hz) timescales, illustrating the value of a multiscale account.

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 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 s 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 and Analysing Interpersonal Coordination

To understand the changes in movement behaviour across these different conversational contexts, it is important to precisely measure and 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 et al., 2011; Hale et al., 2020; Poppe et al., 2014). 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 et al., 2006). Such dynamical analysis techniques are particularly well suited to capturing the time-dependent and emergent nature of interpersonal coordination, which cannot be fully understood through aggregate or static measures of behaviour (Richardson et al., 2005; Schmidt et al., 2011). Wavelet coherence analysis 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.

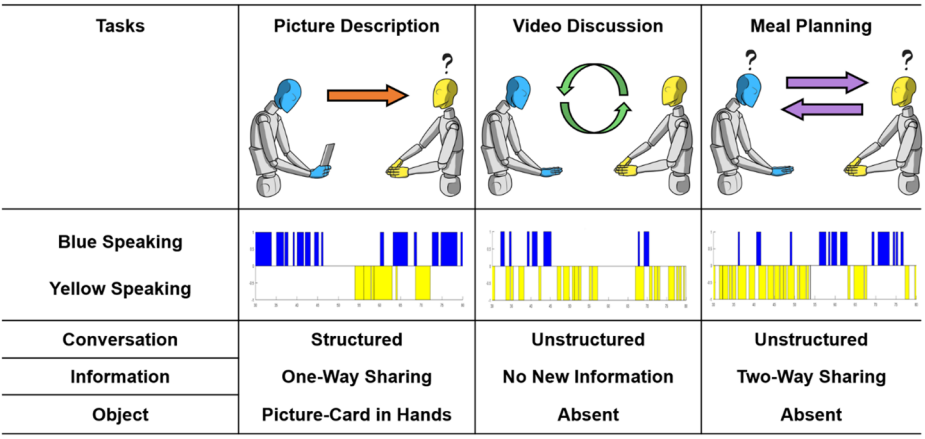

Fig. 1 Conversational tasks. Panel (A) Picture Description; structured one-way information sharing holding an object. Panel (B) Video Discussion; unstructured common ground with no new information sharing. Panel (C) Meal Planning; unstructured two-way information sharing. Graphs (middle) show a sample of the turn-taking structure for each task, highlighting the order and how often blue and yellow participants passed their turns

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 hypothesize 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, a necessary first test is whether it is consistent across contexts; cross-context consistency alone would not establish clinical validity. 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 randomly recruited to arrive at the same time 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

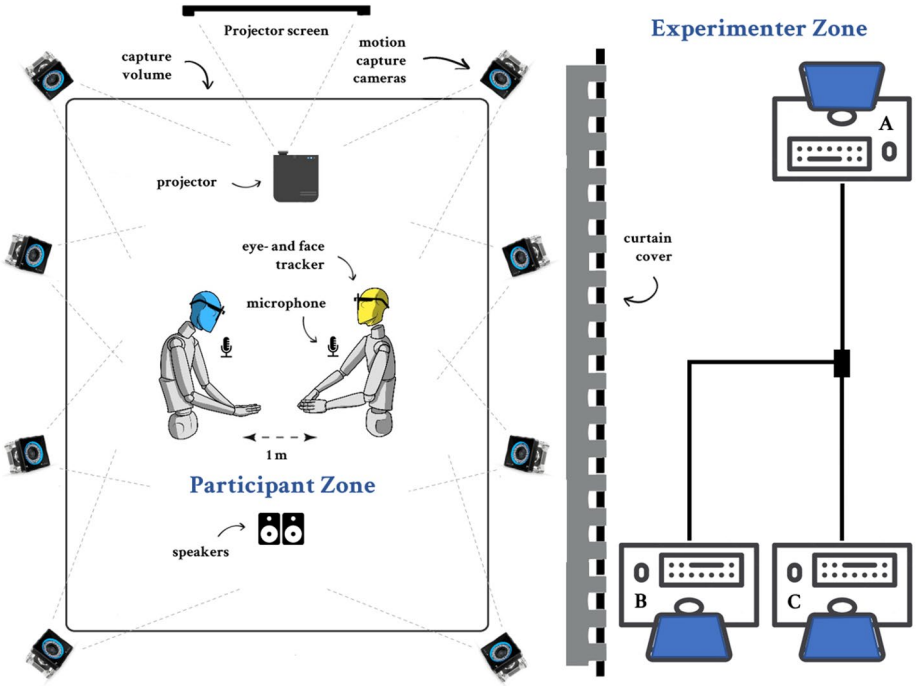

Fig. 2 Experimental Setup and Data Capture. Equipment included motion tracking cameras (4× Optitrack Prime 13, and 4× 13W), a projector, speakers, wearable microphones connected to an audio mixer, eye- and face trackers (Pupil Labs), and a curtain to separate the three computers running the experiment. 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

previous experience with the tasks and were unaware of the purpose of the experiment. 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 Sect. "Analysis of Individual Differences" 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 (Fig. 2). The participants sat on small stools facing each other approximately 1 m 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

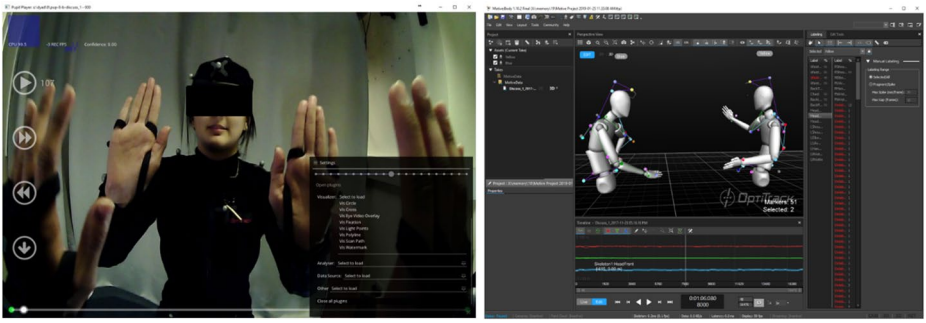

Figure 2 (continued)

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 ran and coordinated the 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 eye tracker, and in addition B captured the Optitrack 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. 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 1 m 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 handclaps in Fig. 3).

First, participants watched a 3 min 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 s of monologue (only leader may speak) followed by 45 s 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

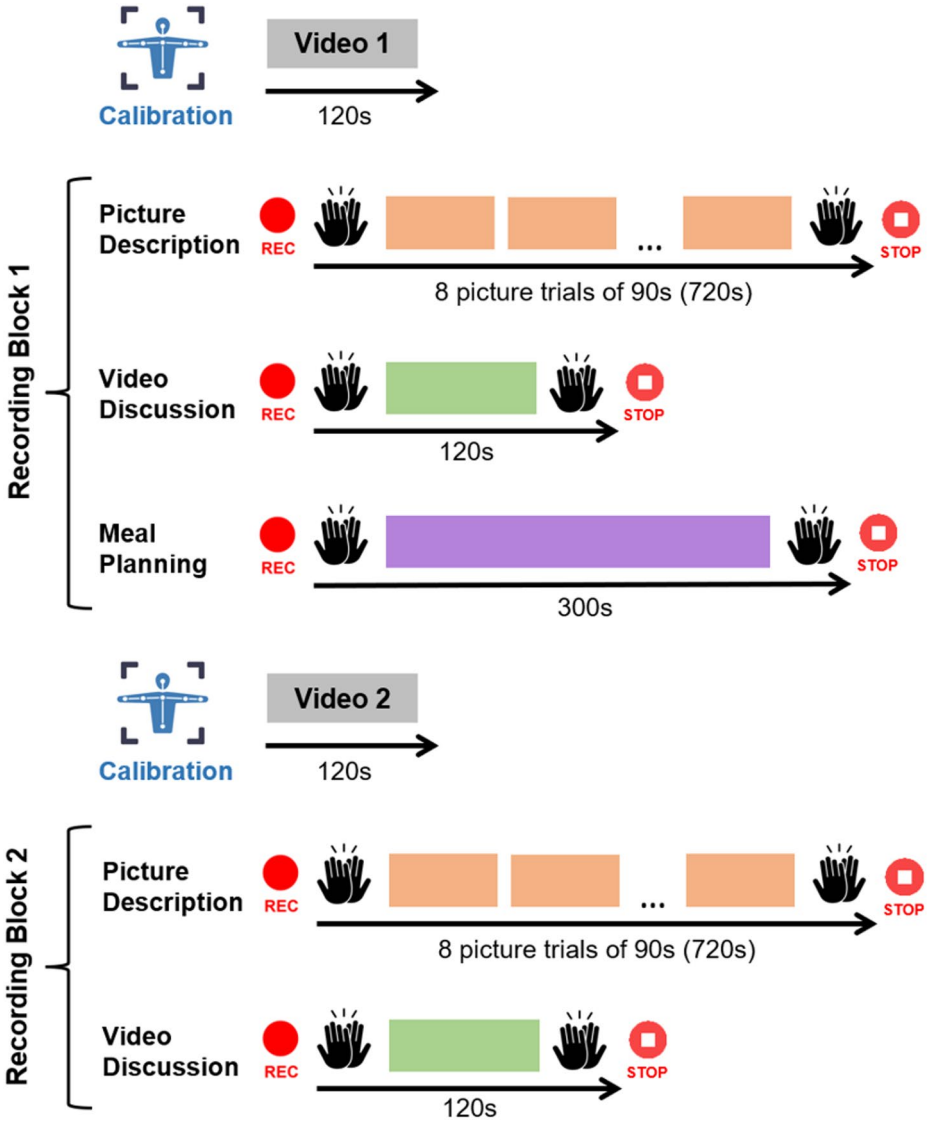

Fig. 3 Experiment timeline. Two calibrations were performed to ensure the eye trackers and motion capture gave high-quality data (T-pose icon) and after each calibration, participants watched a short video together. Data was recorded in two recording-blocks during five task-blocks which occurred in a 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 was completed only once

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.

The task sequence was fixed rather than counterbalanced. There was no specific theoretical rationale for the exact order, and we did not expect serial position itself to influence the outcomes. The order was chosen primarily to separate repeated or similar activities. Nevertheless, because task and serial position are confounded, increasing familiarity, habituation, fatigue, or carryover from earlier tasks could have affected movement. We therefore interpret between-task differences as context-associated rather than uniquely causal effects of task.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 et al., 1994), autistic traits (Baron-Cohen et al., 2001), and the exploratory Experience of Gaze Questionnaire. This measure is being trialled by our group to assess 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"), rated on a five-point scale from "strongly agree" to "strongly disagree". The questionnaire is under development and has no published evidence of reliability or validity; no psychometric validation was conducted in the present sample. Analyses involving this measure are therefore exploratory and should not be interpreted as validating the questionnaire or as testing a well-established gaze-related trait. 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 Fig. 3.

Data Analysis

The present paper focuses only on data from the Optitrack MoCap system. This data was pre-processed in Motive software (supplied by Optitrack) 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 or vertical up-and-down head tilt). We carried out the following pipeline (Fig. 4) using the Matlab toolbox by Grinsted et al. (2004) to identify the wavelet power in the head pitch signals and to calculate cross-wavelet coherence.

For each participant, we took the raw head pitch signal (Fig. 4 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

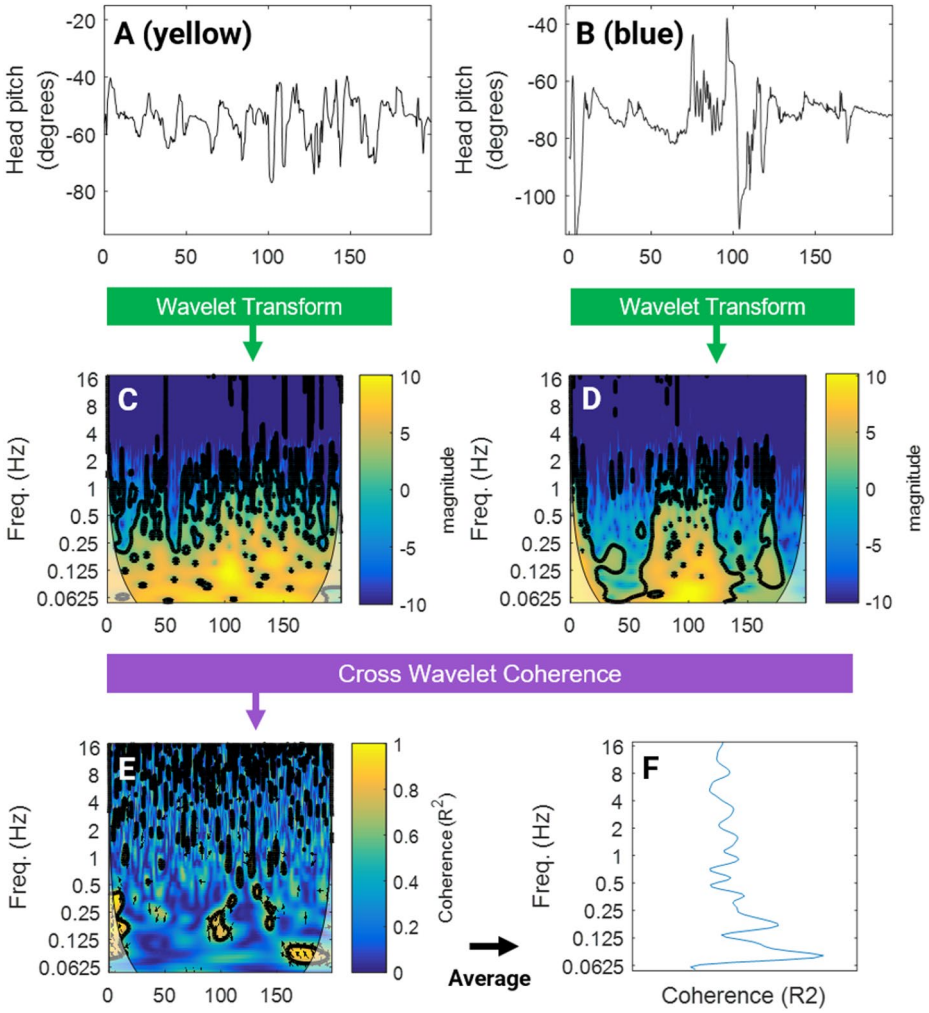

Fig. 4 Cross-wavelet analysis pipeline. For each trial, the head pitch trajectories for both the Yellow and Blue participants (A, B) are subject to a wavelet transform (C, D). Then, the cross-wavelet coherence is calculated between the two participants (E). The magnitude of wavelet power and wavelet coherence is represented by color, where blue is low power, and yellow is high power. The time is represented on the x-axis (200 s) 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)

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 s. In total there were 133 wavelet scales using the Morlet wavelet (periods ranged from 0.1 to 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. We calculated the 'cone of influence' (COI), which marks the region of the time–frequency plot affected by edge effects, and excluded values within this region to ensure reliability of the coherence estimates. 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).

In this study, we analyzed interpersonal coordination across a broad frequency range (0.1–10 Hz) to capture both slow and fast nonverbal behaviours that emerge during naturalistic conversation. While much of the prior literature on interpersonal synchrony focuses on lower frequencies (e.g., <2 Hz), we extended our analysis into higher frequency bands to capture fast, transient behaviours, particularly fast head nodding and beat gestures, which have been linked to real-time listener responses and turn-taking dynamics (Fujiwara & Daibo, 2016; Hale et al., 2020). Our previous work identified fast nodding in the 2.6–6.5 Hz range as a nonverbal backchannel signal that occurs during active listening, motivating the inclusion of these higher frequency bands in the present analysis. To ensure that observed coordination reflects genuine social interaction rather than individual idiosyncrasies or noise, we compared real dyads to carefully matched pseudo-pair baselines. This approach enabled us to detect interaction-specific coherence patterns across the frequency spectrum. Although the full analysis covers 0.1–10 Hz, our interpretation focuses on frequency bands with clear theoretical grounding: slow nodding (0.2–1.1 Hz), fast nodding (2.6–6.5 Hz), and low-frequency hand movements (0.13–1 Hz), each of which has been implicated in distinct social functions in prior literature (e.g., backchanneling, mimicry, or joint attention). The hand band also overlaps the slower movement timescales emphasized in motion-capture/CWC work on dyadic problem solving (de Jonge-Hoekstra, 2021) and complements evidence that gesture-speech synchrony is dynamically stabilized under perturbation (Pouw & Dixon, 2019).

To understand the patterns of head movement present in each task, it is helpful to compare the wavelet coherence values from true pairs, where the two datasets come from the same interaction, to coherence values from pseudo-pairs, where the two datasets come from different interactions. Earlier studies that have used this approach have created pseudo trials by matching datasets from different participants (Bernieri & Rosenthal, 1991; Fujiwara & Daibo, 2016). In this study, we use a task-matched, within-dyad surrogate approach by matching datasets from different trials within the same dyad and task. In the Picture Description Task, each dyad completed 2 blocks of 4 trials, alternating between speaking and listening. For each real trial, we end up with 3 pseudo trial combinations, resulting in 24 pseudo trials per dyad (Table 1). For the video discussion task, this meant matching the discussion of Video 1 to Video 2 from the same participants. For the meal planning task, the 5 min trial was divided in two 150 s chunks which were swapped to create a pseudo-pair. We believe that this approach provides a conservative test with respect to stable individual movement characteristics because the pseudo trials retain the same partners and task while disrupting contemporaneous coupling. It is not intrinsically more rigorous than between-dyad recombination: the two procedures instantiate different null hypotheses, and ideally both would be examined as robustness checks (Moulder et al., 2018; Wiltshire et al., 2019). Consequently, differences between real and pseudo coherence are interpreted as evidence against this specific within-dyad, task-matched null rather than as proof that all alternative non-interaction explanations have been eliminated. This approach to pseudo-pairing we believe also maintain the principles behind cross-wavelet coherence analysis, in such a way as to translate the

Table 1 Example of generating pseudo trials for the picture description task

Block-Trial	True Match		Pseudo Match 1		Pseudo Match 2		Pseudo Match 3	
1-1	1 S	1 L	1 S	3 L	1 S	5 L	1 S	7 L
1-2	2 L	2 S	2 L	4 S	2 L	6 S	2 L	8 S
1-3	3 S	3 L	3 S	1 L	3 S	5 L	3 S	7 L
1-4	4 L	4 S	4 L	2 S	4 L	6 S	4 L	8 S
2-5	5 S	5 L	5 S	1 L	5 S	3 L	5 S	7 L
2-6	6 L	6 S	6 L	2 S	6 L	4 S	6 L	8 S
2-7	7 S	7 L	7 S	1 L	7 S	3 L	7 S	5 L
2-8	8 L	8 S	8 L	2 S	8 L	4 S	8 L	6 S

Note. S = Speaker (Strong Colour); L = Listener (Light Colour).

data into a form that allows the analysis of the degree of similarity between two sets of data – or time-series – and reveals features of movement that are specific to a genuine interaction.

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 Discovery Rate (FDR) of 0.05 (Benjamini & 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

As an exploratory analysis, we used this dataset to test if the tendency to engage in fast or slow nodding shows cross-context stability and relates to self-reported traits. 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 (Fig. 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 (Fig. 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. To examine whether dyadic coordination was related to individual traits during slow nodding, we matched the levels of analysis by aggregating each pair's trait scores. Specifically, we averaged the two

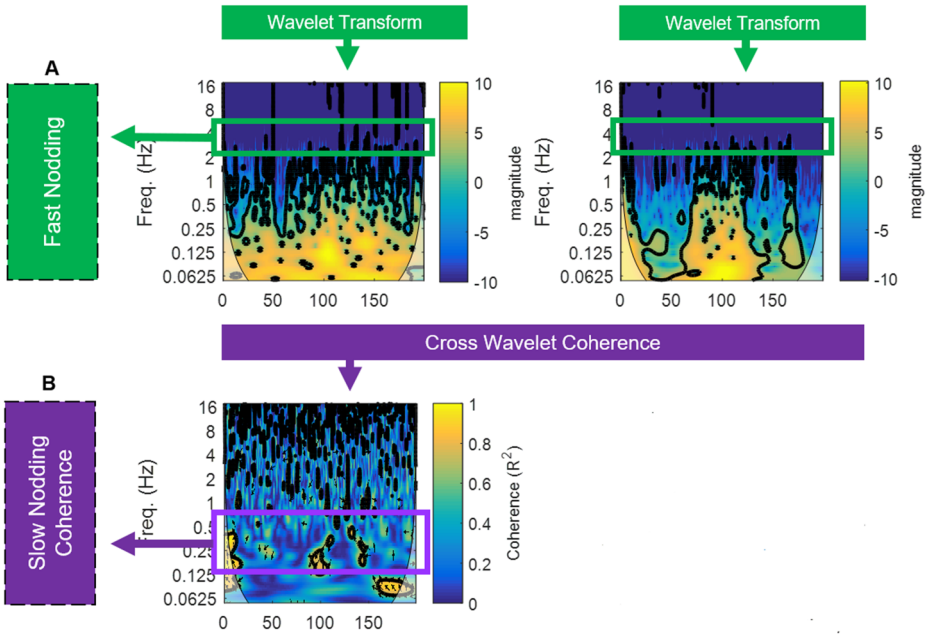

Fig. 5 Fast and slow head nodding scores. The fast nod score was selected from the high frequency band of the individual wavelets (A). The slow nod coherence score was selected from the low frequency band of the cross-wavelet coherence (B)

individuals' scores on each trait questionnaire (LSAS, AQ, TAS-20, Gaze) to create a single dyad-level value per trait. This allowed us to correlate the dyad's slow-nod coherence score (R2) with their mean trait values while maintaining statistical independence and consistency in degrees of freedom. All dyad-level correlations were conducted with N = 21 dyads (df = 19). Because the Experience of Gaze Questionnaire is unvalidated and under development, correlations involving that score were treated as especially exploratory.

Hand Motion Analysis

For the exploratory hand analysis, we focused on vertical (Y-axis; up/down) displacement of each person's right and left hands. The motion-capture system recorded all three spatial axes; selecting Y was an analytic choice rather than a limitation of the system. We selected the Y-axis because it contained the strongest hand-movement component and most clearly reflected vertical beat-like movements. The X- and Z-axis signals were not technically noisier, but they were outside the scope of this exploratory, axis-specific analysis. Accordingly, the results characterize vertical hand-motion coordination only and should not be interpreted as a complete three-dimensional analysis. For each time point, we calculated the absolute level of motion in the y-axis (either up or down) using $y = \sqrt{\left(\frac{dy}{dt}\right)^2}$ where dy is the distance moved in that time window and dt is the duration of the time window (8.3 ms for recordings at 120 Hz). 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 inside 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 hypothesis 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 (Morlet wavelet, w = 6; see Appendix C for w = 8) 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 90 paired-sample t-tests at each frequency (df = 30), comparing coherence values across 31 dyads, 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

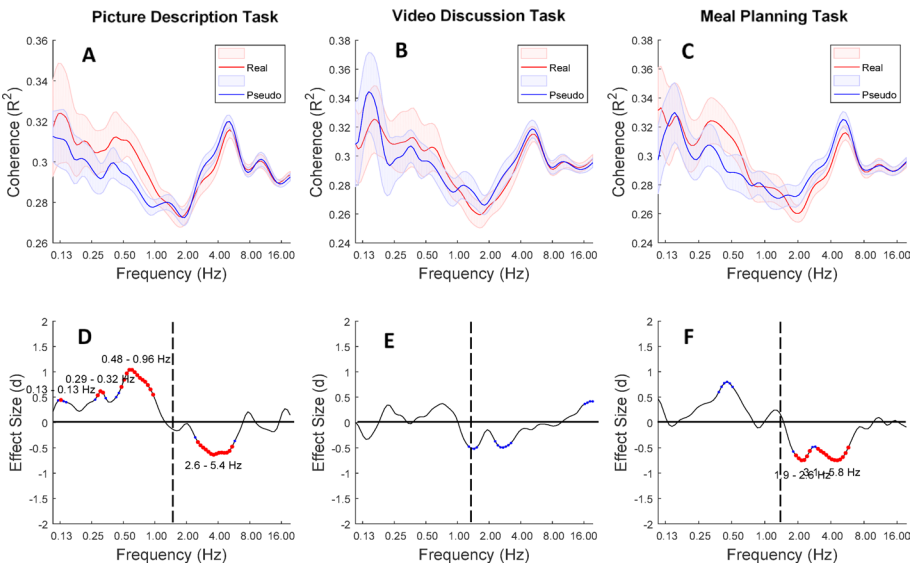

Fig. 6 Head nodding real vs pseudo coherence (A, B, C) and effect sizes (D, E, F) across the full frequency range for the three tasks. Blue dots represent p<.05 (uncorrected), while red dots indicate frequencies that passed FDR correction across 90 frequency bins. Representative frequencies and statistical values are reported in Supplementary Table 1 (Appendix B)

interactions minus the pseudo interactions. The dots indicate frequencies where there is a significant difference of coherence between real and pseudo interactions. FDR correction was applied to control for multiple comparisons. Red dots on the graphs 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 Fig. 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 (Fig. 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 (FDR-corrected $p<.05$ at representative frequencies; see Supplementary Table 1). 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 show each task separately, we also present a cross-task ANOVA in Appendix A.

Individual Differences

We tested if the participants show a reliable pattern of fast or slow nodding by correlating their individual (for fast nods) or dyadic (for slow nods) levels of fast and slow nodding across tasks (Fig. 7).

The results show that there is no reliable positive relationship between fast nodding behavior in any task pair, e.g., $r(29) = .18$, $p = .12$. There was also no reliable positive relationship between slow nodding coherence in any one task paired with any other task. There was a negative correlation between slow-nod coherence in the Meal Planning and Picture Description tasks, r(19) = -.43, uncorrected p = .003; however, it did not survive

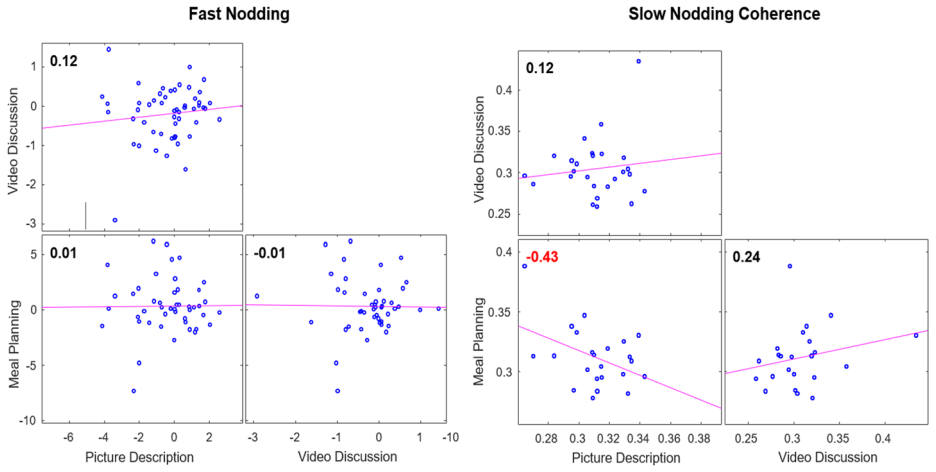

Fig. 7 Correlations of fast and slow nodding power across tasks. The red correlation score mark uncorrected p < .05; none of the displayed correlations survived FDR correction. The axis values for fast nods represent the average power in the 2.6 – 6.5 Hz frequency band. The axis values for slow nods represent the average coherence (R^2) or power in the 0.2 – 1.1 Hz frequency band at the dyad level (one data point for each dyad). Trait scores were averaged within each dyad to match the dyad-level coherence measures

FDR correction, pFDR = .065, and is therefore not interpreted as statistically significant (see Supplementary Table 1). We also tested whether fast or slow nodding behaviour was associated with individual differences in social traits. For fast nods, analyses used participant-level scores (N = 31); for slow nods, analyses used dyad-level coherence scores (N = 21 dyads). We computed correlations between the fast-nod and slow-nod measures and scores on four trait questionnaires: the Liebowitz Social Anxiety Scale (LSAS), the Toronto Alexithymia Scale (TAS-20), the Adult Autism Spectrum Quotient (AQ), and the Experience of Gaze Questionnaire. To assess whether interpersonal coordination was associated with individual traits, we correlated dyad-level slow-nod coherence scores with dyad-level mean trait scores. For each questionnaire, we averaged the two participants' scores to form a single trait value per dyad. No correlations survived FDR correction. Across all tasks and traits, effect sizes were small (all $|r|<.30$), suggesting limited evidence for trait-related variation in nodding behaviour. Post hoc sensitivity analyses indicated that our study was powered to detect medium to large effects ($r \geq .50$, $\alpha = .05$, $1-\beta = .80$).

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 (Fig. 8). We observed an uncorrected trend toward great hand-motion coherence primarily in the Meal Planning task in the low-frequency range (0.13 – 1 Hz). This exploratory pattern could reflect greater social engagement or 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

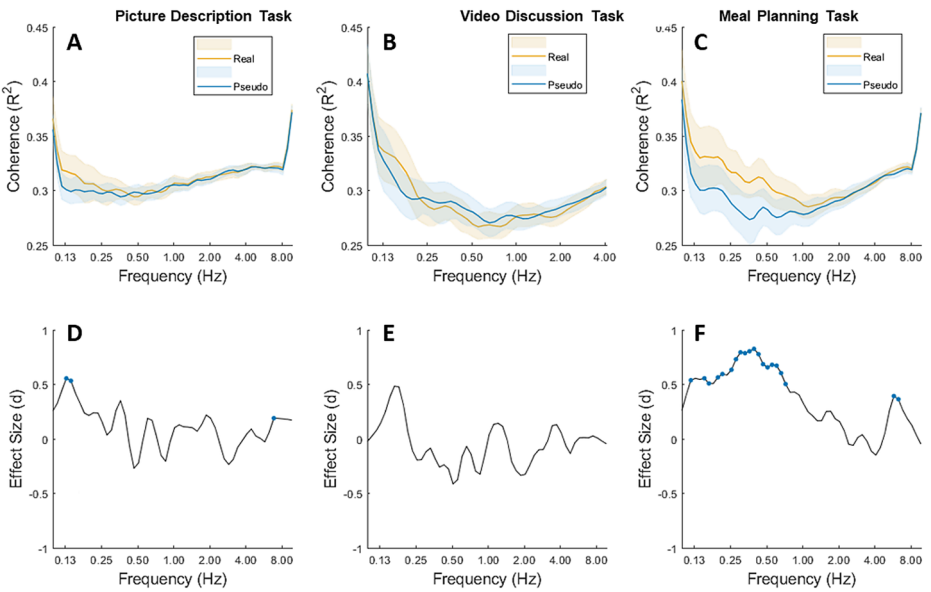

Fig. 8 Hand coordination real vs pseudo coherence (A, B, C) and effect sizes (D, E, F) across the full frequency range for the three tasks. p<.05 uncorrected significance levels shown by blue dots (D, E, F). No frequencies passed FDR correction in the full set, though representative comparisons are included in Supplementary Table 1 (Appendix B)

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, df = 20) 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 (0.2–1.1 Hz 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. The 'Social Glue Hypothesis' states that low frequency slow nodding, or mimicry, is closely related to social bonding and the desire to get on well with others (Lakin et al., 2003), which should create the same motivation to form a social bond across all three conversational contexts. We predicted that the emotional attitudes of bonding should therefore be similar across the different tasks. Our results do not support context-general slow-nod coherence and are more consistent with an account in which the physical gaze target contributes to the observed head movements. This does not make joint attention and affiliation mutually exclusive. A shared target can elicit similar movements mechanically, while the experience of moving together may itself calibrate later interaction or promote prosocial behaviour (Rabinowitch & Meltzoff, 2017). The present design cannot determine whether task-elicited head alignment subsequently altered affiliation or coordination. In the Picture Description Task, participants could alternate gaze between their partner's face and the picture held in one participant's lap, which required vertical 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. This also highlights the design choice of using a picture as a gaze target, because 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 and Ward (2021) has 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, it remains analytically useful to distinguish these actions while recognizing that their mechanisms and consequences may be entangled, because the same movement can have more than one social function. 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 et al., 2008; Richardson et al., 2007). This illustrates the importance of considering the physical and spatial context of actions carefully in any analysis of interpersonal coordination. Because gaze direction is produced jointly by the eyes and head, head motion should not be interpreted independently of ocular behaviour. Natural-exploration and gaze-shift studies show tight eye–head coordination, including correlated response latencies and increasing head contributions to larger gaze changes (Einhäuser et al., 2007; Khan et al., 2009). During storage and retrieval of information, interpersonal coordination and eye-gaze patterns have also been shown to influence the information that is stored (Richardson & Spivey, 2000). For example, people will make eye movements to empty regions of space when retrieving information from memory (Richardson & Kirkham, 2004). Because the success of a linguistic interaction is often dependent on a successful coordination of attention (Clark & Krych, 2004), Richardson and Dale (2005) predicted that the degree of eye movement coordination will reflect the degree to which the listener understood the speaker. Using cross-recurrence analysis, they showed that the more coordinated a listener's eye gaze were with the speaker in a conversation, the better the listener did on a comprehension test. There is little work on specifically head nodding behaviour in relation to memory and learning outcome, and future studies could demonstrate that different frequencies of head nods can be correlated with memory and possibly be used to enhance various teaching strategies.

A limitation of the present study is the gender imbalance in the sample, with the majority of dyads consisting of female-female pairs. As such, the findings primarily reflect coordination patterns in same-gender (female-female) interactions and may not generalize to mixed-gender or male–male dyads. Future studies should aim to include a more balanced sample to systematically examine potential gender effects on nonverbal coordination (Fujiwara et al., 2019; Seibt et al., 2015).

An important methodological consideration concerns the choice of Morlet wavelet order. In this study, we adopted an order of 6, consistent with the default implementation of Grinsted et al. (2004) and our previous analyses of comparable datasets (Hale et al., 2020). This setting prioritizes temporal resolution, which is critical for detecting short, socially meaningful behaviors such as nodding and beat gestures that typically unfold within sub-second timescales (Dikker et al., 2021; Fujiwara & Daibo, 2016). Although higher orders (e.g., 8) improve frequency resolution, they do so at the expense of temporal precision, making order 6 the more appropriate choice for the present research aims (Addison, 2017; Torrence & Compo, 1998). To evaluate the robustness of this parameter choice, we repeated the primary analyses using a Morlet wavelet order of 8, which is also commonly employed in related work (Issartel et al., 2006; Schmidt et al., 2014). The results, reported in the Supplementary Materials, yielded qualitatively comparable outcomes. This convergence indicates that the findings are robust to variations in wavelet order while enhancing transparency and comparability across studies.

A further limitation is the fixed task order. We had no specific theoretical rationale for the exact order and did not predict an effect of serial position; the order was chosen pragmatically to separate repeated or similar activities. However, task context cannot be separated fully from serial-position effects: participants may have become more familiar or comfortable with one another, habituated to the equipment, become fatigued, or carried interactional patterns from earlier tasks into later tasks. Replication should counterbalance tasks where possible or use between-dyad order variants to separate context from order.

Exploring the Fast-Nodding Behaviour

In a secondary analysis, Hale et al. (2020) showed that fast-nodding power arises when a listener makes small, rapid head movements (2.6–6.5 Hz) that do not match the speaker's movements. This frequency range represents approximately 2.6–6.5 up-down pitch cycles per second during an oscillatory segment. The present measure is continuous wavelet power, not an event detector, so it does not provide event-wise peak-to-peak angular amplitude or the number of cycles in a discrete nod bout. This led them to the idea that fast nodding might be a backchannel signal related to listening. However, they were not able to properly test this idea. In the present study, we used motion capture sampled at 120 Hz to examine whether the cross-context pattern of fast-nodding power was consistent with this backchannel account. Fast-nodding power met the FDR-corrected threshold in the Picture Description task (2.6–5.4 Hz) and Meal Planning task (1.9–5.8 Hz), but not in the Video Discussion task. Based on Hale et al. (2020), we hypothesized 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, fast-nodding dynamics were associated with contexts involving the exchange of novel information. In contrast, the Video Discussion Task is about shared recall between members of a dyad with no new information sharing, and the fast-nodding effect did not survive FDR correction here. Thus, one possible explanation for the reduced fast nodding in the Video Discussion task is that, unlike the other tasks, both participants shared full access to the information being discussed. This may have reduced the need for active listener signaling such as fast nodding, which is often associated with feedback during information exchange.

An alternative explanation for the reduced fast nodding observed in the Video Discussion task concerns the role of memory processes in shaping conversational dynamics. Unlike the Picture Description task, which involves continuous visual access to the stimulus, or the Meal Planning task, which draws primarily on stable long-term knowledge such as personal preferences, the Video Discussion task requires participants to jointly reconstruct details of a recently viewed event. In this context, memory representations may be incomplete, uncertain, or differ between interaction partners. Such uncertainty could plausibly alter engagement strategies or reduce the use of rapid listener feedback signals. However, if memory demand or cognitive effort alone were driving reductions in fast nodding, we might have expected comparable or stronger effects in the Meal Planning task, which involves spon-

taneous content generation and coordination. This pattern was not observed. Instead, fast nodding was selectively reduced in the Video Discussion task, where both participants had equal access to the information under discussion and no novel information was being introduced. We therefore interpret the reduction in fast nodding primarily as reflecting reduced informational asymmetry and a diminished need for listener backchanneling, rather than memory load alone. Nonetheless, these factors are not mutually exclusive, and memory-related uncertainty may interact with informational structure to shape nonverbal signaling during conversation. With these results we can now more reliably say that the fast-nodding pattern is a socially meaningful signal, and it would be interesting in the future to test if fast nodding behaviour can provide a marker of successful information transfer in a conversation, that might be linked to later learning. Furthermore, although the fast-nod score was calculated at the individual level, its occurrence is embedded in the partner's speaking and movement timing and may therefore be a relational feature of the dyad. The absence of cross-task stability cannot distinguish context-specific individual behaviour from partner-contingent dyadic dynamics. Round-robin designs and turn-level models are needed to separate actor, partner, relationship, and context effects.

Individual Differences in Nodding Behaviour

If nonverbal behaviour can be robustly measured and linked to individual differences in personality or socio-cognitive 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 (Fig. 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 dyadic slow nodding scores 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 exploratory Experience of Gaze Questionnaire. Evidence that spontaneous gesture production can predict learning in children illustrates that individual movement differences may be meaningful, but their interpretation can be outcome- and context-specific (Congdon et al., 2025).

These results did not support our hypothesis, as 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, nor did it correlate with any questionnaire measures. 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). The present results are more consistent with strong contextual and partner dependence and do not justify using these fast- or slow-nodding measures, in the current form, as indices of an individual's social skills or for clinical assessment. This distinction matters because

automated interactive-behaviour measures are being explored in research on social-interaction difficulties, including autism, but clinical validity requires independent validation, adequate classification designs, and evidence that extends beyond a single dyad or context (Georgescu et al., 2019). The Experience of Gaze measure also has no established psychometric properties, so the null correlations involving it should not be treated as evidence about a validated gaze-related trait.

Interpreting questionnaire associations is challenging when the aim is to isolate each person's contribution within a social interaction. This is difficult because data from naturalistic interactions are not independent but depend on the way people adapt their behaviours over time. Whereas social interactions have characteristics in addition to those of the individuals involved, the interaction might nonetheless be influenced by individual differences. For example, in the case of autistic traits in dyadic interactions, studies have demonstrated that the dissimilarity, as much as the similarity, between subjects is crucial to the outcome of social interactions (Bolis et al., 2017, 2021; Milton, 2012). Future studies employing round-robin design (e.g., Bolis et al., 2021; Milton, 2012) could more effectively test how trait (mis)alignment influences social interaction outcomes.

We believe that studying multimodal aspects of social interaction is an essential next step for this research area, and even though we only focus on head and hand movements in this study, our multimodal data capture setup enables us to study these interactions in detail. However, a common problem with interpersonal dyadic studies (Fujiwara & Daibo, 2016; Paxton & Dale, 2013), motion capture studies (Kijima et al., 2017), or even corpus studies on conversation behaviour (Aubrey et al., 2013; Carletta, 2007), is that the setup is very obtrusive and requires specialized equipment, expertise, and dedicated recording space. Thus, it is very costly and time consuming, which often limits the sample size from which data can be collected. Regardless of having hundreds of hours of motion capture data to work with, the relatively small sample size (i.e., 62) contributes to the risk that this measure is not powerful enough to detect an effect based on a specific experimental manipulation. This problem can be more serious, especially when looking at exploratory data and measures of individual differences. In light of this, scalable video-based methodologies should not be underestimated. Frame-differencing or motion-energy methods quantify pixel-wise change across successive images (Ramseyer & Tschacher, 2010; Schmidt et al., 2012), whereas OpenPose estimates two-dimensional body keypoints in each frame and does not rely on motion-energy differences (Cao et al., 2017). Romero et al. (2017) demonstrated comparability between motion-capture and frame-differencing measures for some applications, and Fujiwara and Yokomitsu (2021) validated OpenPose-based tracking for face-to-face interpersonal coordination. With synchronized, calibrated views from multiple cameras, two-dimensional pose estimates can also be triangulated into three-dimensional kinematics using pipelines such as Pose2Sim (Pagnon et al., 2022). These developments create important possibilities for scaling research on nonverbal coordination in naturalistic or resource-limited settings.

Coordination of Hand Movements

Previous work has examined interpersonal hand-movement coordination using video coding (Holler & Wilkin, 2011; Ramseyer & Tschacher, 2016) and, in child dyads, motion capture with CWC across multiple timescales (de Jonge-Hoekstra, 2021, Chapter 5). Together, these

studies show that hand coordination can be detected with both observational and kinematic methods. 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 pattern is compatible with, but does not establish, an influence of spontaneous conversation and novel-information sharing on hand-motion coordination. 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 (a design limitation that unfortunately stems from the choice of prioritizing our main hypothesis of the slow nodding coherence), 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 is that the hand analysis characterizes only vertical movement, selected because the Y-axis contained the strongest hand-movement component and most clearly reflected vertical beat-like movements. The X and Z axes may contain meaningful information but were outside the scope of this exploratory analysis. We also cannot discriminate gesture types (e.g., iconic, beat, or self-adaptor movements). Future work should preregister axis-specific hypotheses and compare individual axes, three-dimensional vector magnitude, and gesture-segmented analyses. Although no hand-movement comparison survived FDR correction, the uncorrected Meal Planning pattern warrants targeted replication with gesture segmentation and adequate power.

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, while our analysis primarily focused on coherence values averaged across trials and dyads, this approach was chosen to generate stable, interpretable estimates of coordination within each conversational context, consistent with prior work using wavelet coherence in dyadic interaction (e.g., Fujiwara & Daibo, 2016; Hale et al., 2020). Another methodological consideration is the construction of pseudo-pairs. Rather than randomly recombining participants from different dyads, we generated pseudo-pairs by matching data from different trials within the same dyad and within the same task. This approach preserved individual movement characteristics and task structure while disrupting real-time temporal coupling, offering a task-matched null for assessing interaction-specific coherence. This baseline is not intrinsically superior to between-dyad recombination: different surrogate constructions test different null hypotheses, and results can depend on that choice (Moulder et al., 2018; Wiltshire et al., 2019). Future studies should report both within-dyad and between-dyad surrogate schemes as robustness checks. The present conclusions are therefore limited to the within-dyad, task-matched pseudo-pair comparison used here.

We complemented this approach with a 2×3 repeated-measures ANOVA (Real vs. Pseudo × Task) to examine broader interaction patterns (Appendix A). However, we acknowledge that future studies using larger and more balanced samples could benefit from trial-level

or mixed-effects analyses that better capture within-subject variability. Designs such as round-robin pairings (Salazar Kämpf et al., 2017) may also help disentangle individual contributions to interactional dynamics across contexts. In addition, there are also statistical considerations related to power. Although the overall pattern of results was consistent across tasks and analysis types, we acknowledge that the study may have been underpowered to detect small effects. Post hoc sensitivity analyses indicated that our individual-level correlational analyses (N = 31) were powered to detect effects of $r \geq .50$, while dyadic-level correlations (N = 21) and paired *t*-tests were powered to detect effects of $r \geq .61$ and $d \geq .64$, respectively. For the 2×3 repeated-measures ANOVA, the detectable effect size was partial $\eta^2 \geq .35$. These thresholds suggest that while the study was sufficiently powered to detect medium-to-large effects, smaller effects may have gone undetected. We encourage future work with larger, more diverse samples to extend these findings and more fully explore individual-level variation. Nonetheless, the use of conservative corrections (FDR), control comparisons (pseudo pairs), and consistent task-level findings support the robustness of the observed effects.

Second, the complex multimodal data-collection setup in our lab requires substantial equipment, which reduces the naturalness of the conversations. Future studies could use pose estimation such as OpenPose, which extracts two-dimensional body keypoints rather than motion energy (Cao et al., 2017), or wearable motion sensors (Sun et al., 2023; Ward & Pinti, 2019) to track social coordination less obtrusively outside the lab. Where calibrated, synchronized multi-camera video is available, workflows such as Pose2Sim can reconstruct three-dimensional kinematics from multiview pose estimates (Pagnon et al., 2022). 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 et al., 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 backchannel 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. Moreover, in everyday conversations, we sometimes nod when the other person speaks with hand gestures, which involves a lot of information exchange. While our analyses focused on head–head and hand–hand coordination, future studies could explore cross-modal synchrony, such as whether listeners' head nods align with their partners' hand gestures. Investigating such head–hand coupling may reveal additional layers of nonverbal alignment, particularly in spontaneous and gesture-rich conversation. 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 et al., 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 et al., 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-nodding dynamics were associated with contexts involving novel information and may function as listener feedback, while slow head nods may be coordinated to direct gaze to a shared object. We also report that neither behaviour appeared to reflect a stable personality trait; instead, both varied with conversational context. For hand movements, we observed only uncorrected trends toward low-frequency coordination 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.

Appendix A

Coherence Across Tasks and Real/Pseudo Interactions

Our measures of head motion coherence show different patterns in our three different tasks, which leads to the question of whether coherence is greater in one task than another. An interaction analysis in an ANOVA might seem the ideal way to analyse this, but as the baseline levels of coherence in the pseudo-datasets are different, it is not so easy to interpret the results of such an ANOVA. For this reason, we do not include the ANOVA in the main text of the paper, but we report the findings here for completeness. We performed a 2 (Real vs. Pseudo) × 3 (Tasks) repeated measures ANOVA on head motion coherence values, with 31 dyads (df = 1,30 and 2,60 respectively). Results are shown separately for each main effect together with the interaction effect (Fig. 9)

Graphs A, B and C show the mean and standard error of coherence (R^2) of each effect. 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 all the conditions, we calculated the effect size as an indication of the % of variance between them. Graphs D, E and F show the effect sizes (partial eta-squared, ηp^2). The plots show that there is a substantial main effect of task (Fig. 9 B and E) with more coherence in the video and picture tasks than in the Meal Planning task, especially at low frequencies. There is also a marginal main effect of Real interaction ($p = .047$ uncorrected, FDR-corrected $p = .058$; see Supplementary Table 1) (Fig. 9A and D) with more coherence for real interaction only in the 0.25 – 0.35 Hz range. There is an effect for fast-nodding around 2 – 4 Hz with less fast-nodding for real interactions. Finally, there were no significant interaction effects (Fig. 9 C and F).

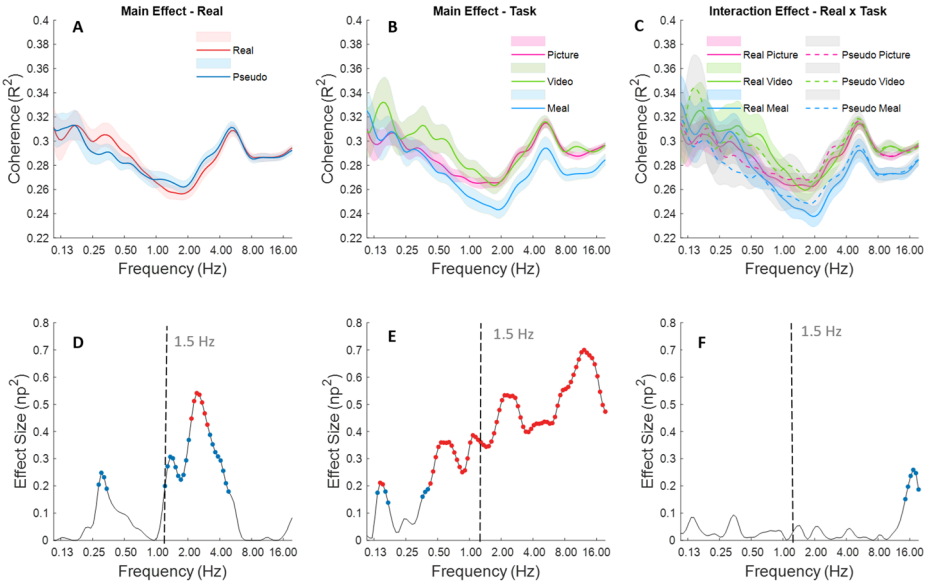

Fig. 9 ANOVA cross-wavelet coherence. Graphs A, B, and C show the mean and standard error of coherence (R^2) of each effect. Graphs D, E, and F show the effect sizes (partial eta-squared, ηp^2). The dotted line indicates frequencies where there is a significant difference of coherence. Red dots represent points on the frequency range that pass a $p<0.05$ FDR significance threshold; blue dots represent significant differences that did not pass

Appendix B

Supplementary Table 1. Comprehensive Statistical Output

This table summarizes key statistical tests, including representative frequency-level results for the wavelet coherence analyses. Because FDR correction is applied across 90 frequency bins, exact corrected p-values for each bin are not reported. Instead, we report uncorrected p-values and effect sizes for representative frequencies that passed FDR correction.

Test Type	Comparison	df	Test Statistic	p (uncorrected)	p (FDR-corrected)	Effect Size
Paired t-test	Real vs Pseudo (Head, Picture) – Representative Freq: 0.45 Hz	t(30)			< .05 (FDR)	d ≈ 0.43 (est.)
Paired t-test	Real vs Pseudo (Head, Meal) – Representative Freq: 0.30 Hz	t(30)			< .05 (FDR)	d ≈ 0.39 (est.)
Correlation	Slow Nod Coherence (Picture vs Meal)	r(19)	-0.43	0.003	0.065	r = −0.43
Correlation	Fast Nod × LSAS (Picture)	r(29)	0.19	0.317	0.54	r = 0.19
Paired t-test	Real vs Pseudo (Hand, Picture) – Representative Freq: 0.25 Hz	t(20)			< .05 (FDR)	d ≈ 0.51 (est.)

Test Type	Comparison	df	Test Statistic	p (uncorrected)	p (FDR-corrected)	Effect Size
Paired t-test	Real vs Pseudo (Hand, Meal) – Representative Freq: 0.33 Hz	t(20)			< .05 (FDR)	d ≈ 0.45 (est.)
2×3 ANOVA	Real × Task (Head Coherence)	F(2,60)		0.047	0.058	η^2 = 0.08 (est.)

Appendix C

Robustness Check: Wavelet Order Comparison (Order 6 vs. Order 8)

In our primary analysis, we used a Morlet wavelet order of 6, consistent with the default setting in the widely used toolbox by Grinsted et al. (2004) and with our previous work (Hale et al., 2020), which analyzed a similar dataset using identical parameters. A wavelet order of 6 offers improved temporal resolution, which we prioritized due to our focus on short, socially meaningful behaviors such as fast nodding and beat gestures that unfold on sub-second timescales (e.g., Dikker et al., 2021; Fujiwara & Daibo, 2016). While a higher wavelet order (e.g., 8) improves frequency resolution, it does so at the cost of temporal precision, potentially obscuring transient coordination signals. This trade-off has been well-documented in applied wavelet literature, which recommends tuning the order based on the nature of the signal and task demands (Addison, 2017; Torrence & Compo, 1998). To promote transparency, we ran our core analyses using a Morlet wavelet of order 8 as well. The results are shown below (Fig. 10). Importantly, all key patterns of coherence and statis-

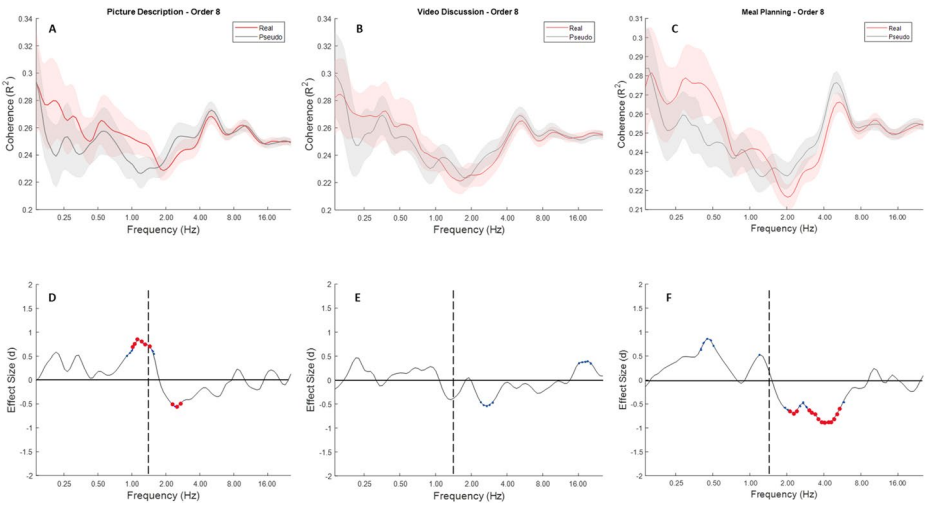

Fig. 10 Wavelet Order 8. Head nodding real vs pseudo coherence (A, B, C) and effect sizes (D, E, F) across the full frequency range for the three tasks. Blue dots represent p<.05 (uncorrected), while red dots indicate frequencies that passed FDR correction across 90 frequency bins

tical outcomes remain qualitatively consistent with those obtained using order 6, supporting the robustness of our findings and facilitating comparability across studies using different parameter conventions.

Author Contribution Conceptualization: P.F., R.C., J.W., A.H. Data curation & investigation: P.F., R.C., J.W. Formal analysis: P.F., R.C., J.W., A.H. Manuscript – original draft: P.F. Manuscript – review and editing: P.F., R.C., J.W., A.H. Project supervision: J.W., A.H. Funding acquisition: P.F., A.H. All authors approved the final submission of the manuscript.

Funding Leverhulme Trust

Data Availability All data supporting the findings of this study are available within the paper and its Appendix. The raw data are available upon request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aburumman, N., Gillies, M., Ward, J. A., & Hamilton, AFdeC. (2022). Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions. *International Journal of Human-Computer Studies, 164*, Article 102819.

Addison, P. S. (2017). *The illustrated wavelet transform handbook: Introductory theory and applications in science, engineering, medicine and finance* (2nd ed.). CRC Press.

Allwood, J., & Cerrato, L. (2003). A study of gestural feedback expressions. In Paggio et al. (Eds.), *Proceedings of the First Nordic Symposium on Multimodal Communication*, Copenhagen.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge University Press.

Aubrey, A. J., Marshall, D., Rosin, P. L., Vandeventer, J., Cunningham, D. W., & Wallraven, C. (2013). Cardiff conversation database (CCDb): A database of natural dyadic conversations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Portland, Oregon, USA, 277–82.

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

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

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. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of Nonverbal Behavior* (pp. 401–432). Cambridge University Press.

Bolis, D., Balsters, J., Wenderoth, N., Becchio, C., & Schilbach, L. (2017). Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. *Psychopathology, 50*(6), 355–372. https://doi.org/10.1159/000484353

Bolis, D., Lahnakoski, J. M., Seidel, D., Tamm, J., & Schilbach, L. (2021). Interpersonal similarity of autistic traits predicts friendship quality. *Social Cognitive and Affective Neuroscience, 16*(1–2), 222–231. https://doi.org/10.1093/scan/nsaa147

Bosker, H. R., & Peeters, D. (2021). Beat gestures influence which speech sounds you hear. *Proceedings of the Royal Society b: Biological Sciences*. https://doi.org/10.1098/rspb.2020.2419

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.

Carletta, J. (2007). Unleashing the killer corpus: Experiences in creating the multi-everything AMI meeting corpus. *Language Resources and Evaluation, 41*, 181–190. https://doi.org/10.1007/s10579-007-9040-x

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.

Chovil, N. (1991). Discourse oriented facial displays in conversation. *Research on Language and Social Interaction, 25*, 163–194. https://doi.org/10.1080/08351819109389361

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*(1), 62–81. https://doi.org/10.1016/j.jml.2003.08.004

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

Congdon, E. L., Novack, M. A., & Wakefield, E. M. (2025). Exploring individual differences: A case for measuring children's spontaneous gesture production as a predictor of learning from gesture instruction. *Topics in Cognitive Science, 17*, 569–585. https://doi.org/10.1111/tops.12722

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

Dale, R., Fusaroli, R., Duran, N. D., & Richardson, D. C. (2013). The self-organization of human interaction. *Psychology of Learning and Motivation, 59*, 43–95. https://doi.org/10.1016/B978-0-12-407187-2.00002-2

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

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., & Poeppel, D. (2021). Brain-to-brain synchrony tracks real-world dynamic group interactions. *Nature Neuroscience, 24*, 1–10. https://doi.org/10.1038/s41593-021-00850-6

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. https://doi.org/10.1145/3462244.3479962.

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

Einhäuser, W., Schumann, F., Bardins, S., Bartl, K., Böning, G., Schneider, E., & König, P. (2007). Human eye–head co-ordination in natural exploration. *Network: Computation in Neural Systems, 18*(3), 267–297. https://doi.org/10.1080/09548980701671094

Emery, N. J. (2000). The eyes have it: The neuroethology, function, and evolution of social gaze. *Neuroscience and Biobehavioral RevieWs, 24*, 581–604. https://doi.org/10.1016/S0149-7634(00)00025-7

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. https://doi.org/10.1109/ISWC.2011.31

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

Fitzpatrick, P., Romero, V., Amaral, J. L., Duncan, A., Barnard, H., Richardson, M. J., & Schmidt, R. C. (2017). Evaluating the importance of social motor synchronization and motor skill for understanding autism. *Autism Research, 10*(10), 1687–1699. https://doi.org/10.1002/aur.1808

Fujiwara, K., & Daibo, I. (2016). Evaluating interpersonal synchrony: Wavelet transform toward an unstructured conversation. *Frontiers in Psychology, 7*, Article 516. https://doi.org/10.3389/fpsyg.2016.00516

Fujiwara, K., Kimura, Y., & Daibo, I. (2019). Gender differences in interpersonal synchrony. *European Journal of Social Psychology, 49*(5), 1042–1054. https://doi.org/10.1002/ejsp.2587

Fujiwara, K., & Yokomitsu, K. (2021). Validation of OpenPose-based motion tracking for interpersonal synchrony. *Behavior Research Methods, 53*, 2700–2711. https://doi.org/10.3758/s13428-021-01612-7

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

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*, Article 132. https://doi.org/10.3389/frobt.2019.00132

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. https://doi.org/10.5194/npg-11-561-2004

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

Hadley, L. V., Naylor, G., & Hamilton, AFdeC. (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

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), Article e0258247. https://doi.org/10.1371/journal.pone.0258247

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 Nonverbal Behavior, 44*, 63–83.

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

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. https://doi.org/10.1177/0963721415570731

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

Hoehl, S., Fairhurst, M., & Schirmer, A. (2020). Interactional synchrony: Signals, mechanisms and benefits. *Social Cognitive and Affective Neuroscience, 16*(1–2), 5–18. https://doi.org/10.1098/scan/nsaa024

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

Hömke, P., Holler, J., & Levinson, S. (2018). Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLoS ONE, 13*, Article e0208030. https://doi.org/10.1371/journal.pone.0208030

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.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development, 66*(2), 1–132.

de Jonge-Hoekstra, L. (2021). How hand movements and speech tip the balance in cognitive development: A story about children, complexity, coordination, and affordances [Doctoral dissertation, University of Groningen]. https://doi.org/10.33612/diss.172252039

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

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

Kendon, A. (1972). Some relationships between body motion and speech: An analysis of an example. In A. W. Siegman & B. Pope (Eds.), Studies in dyadic communication (pp. 177–210). Pergamon Press.https://doi.org/10.3758/s13428-013-0398-y

Kendon, A. (2002). Some uses of head shake. *Gesture, 2*(2), 147–182. https://doi.org/10.1075/gest.2.2.03ken

Khan, A. Z., Blohm, G., McPeek, R. M., & Lefèvre, P. (2009). Differential influence of attention on gaze and head movements. *Journal of Neurophysiology, 101*(1), 198–206. https://doi.org/10.1152/jn.90815.2008

Kijima, A., Shima, H., Okumura, M., Yamamoto, Y., & Richardson, M. J. (2017). Effects of agent-environment symmetry on the coordination dynamics of triadic jumping. *Frontiers in Psychology, 8*, Article 3. https://doi.org/10.3389/fpsyg.2017.00003

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., 2003*(48), 16–32.

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

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.

Liebowitz, M. R. (1987). Social phobia. In T. A. Ban., P. Pichot., & W. Pöldinger (Eds.), *Modern Problems of Pharmacopsychiatry*: *Vol. 22. Anxiety* (pp. 141–173). https://doi.org/10.1159/000414022

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

Loehr, D. (2007). Aspects of rhythm in gesture and speech. *Gesture, 7*(2), 179–214. https://doi.org/10.1075/gest.7.2.04loe

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science, 1*(2), 320–339. https://doi.org/10.1111/j.1756-8765.2009.01022.x

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

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Milton, D. E. M. (2012). On the ontological status of autism: The 'double-empathy problem.' *Disability & Society, 27*(6), 883–887. https://doi.org/10.1080/09687599.2012.710008

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

Pagnon, D., Domalain, M., & Reveret, L. (2022). Pose2Sim: An open-source Python package for multiview markerless kinematics. *Journal of Open Source Software, 7*(77), Article 4362. https://doi.org/10.21105/joss.04362

Pang, H. T., Canarslan, F., & Chu, M. (2022). Individual differences in conversational self-touch frequency correlate with state anxiety. *Journal of Nonverbal Behavior, 46*, 299–319. https://doi.org/10.1007/s10919-022-00402-9

Paxton, A., & Dale, R. (2013). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavioural Research Methods, 45*(2), 329–343.

Pan, X., de Hamilton, A. F., & C. (2015). Automatic imitation in a rich social context with virtual characters. *Frontiers of Psychology, 6*, 790. https://doi.org/10.3389/fpsyg.2015.00790

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.*

Poppe, R., van der Zee, S., Heylen, D. K. J., & Taylor, P. J. (2014). AMAB: Automated measurement and analysis of body motion. Behavior Research Methods, 46(3), 625–633. https://doi.org/10.3758/s13428-013-0398-y

Pouw, W., & Dixon, J. A. (2019). Entrainment and modulation of gesture–speech synchrony under delayed auditory feedback. *Cognitive Science, 43*, Article e12721. https://doi.org/10.1111/cogs.12721

Rabinowitch, T. C., & Meltzoff, A. N. (2017). Joint rhythmic movement increases 4-year-old children's prosocial sharing and fairness toward peers. *Frontiers in Psychology, 8*, Article 1050. https://doi.org/10.3389/fpsyg.2017.01050

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.

Ramseyer, F., & Tschacher, W. (2006). Synchrony: A core concept for a constructivist approach to psychotherapy. *Constructivism in the Human Sciences, 11*(1–2), 150–171.

Ramseyer, F., & Tschacher, W. (2010). Nonverbal synchrony or random coincidence? How to tell the difference. In A. Esposito, N. Campbell, C. Vogel, A. Hussain, & A. Nijholt (Eds.), *Development of multimodal interfaces: Active listening and synchrony* (Vol. Vol. 5967, pp. 182–196). Springer. https://doi.org/10.1007/978-3-642-12397-9_15

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. https://doi.org/10.1037/a0023419

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

Reece, A., Cooney, G., Bull, P., Chung, C., Dawson, B., Fitzpatrick, C., Glazer, T., Knox, D., Liebscher, A., & Marin, S. (2023). The CANDOR corpus: Insights from a large multimodal dataset of naturalistic conversation. *Science Advances*. https://doi.org/10.1126/sciadv.adf3197

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*(6), 1045–1060. https://doi.org/10.1207/s15516709cog0000_29

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological Science, 18*(5), 407–413. https://doi.org/10.1111/j.1467-9280.2007.01914.x

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and infants. *Journal of Experimental Psychology: General, 133*(1), 46–62. https://doi.org/10.1037/0096-3445.133.1.46

Richardson, D. C., & Spivey, M. J. (2000). Representation, space, and Hollywood squares: Looking at things that aren't there anymore. *Cognition, 76*(3), 269–295. https://doi.org/10.1016/S0010-0277(00)00084-6

Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 62–79. https://doi.org/10.1037/0096-1523.31.1.62

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. Current Directions in Psychological Science, 25(1), 70–74. https://doi.org/10.1016/B978-0-08-015867-9.50013-7.

Roberts, S. (2011). *DipDap.* [BBC]. Ragdoll Productions.

Romero, V., Fitzpatrick, P., Roulier, S., Schmidt, R. C., & Richardson, M. J. (2017). Using frame-differencing methods to measure interpersonal synchrony. *Behavior Research Methods, 49*, 1–14. https://doi.org/10.3758/s13428-016-0799-y

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. https://doi.org/10.1177/0956797617727121

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second person neuroscience. *Behavioral and Brain Sciences, 36*(4), 393–462. https://doi.org/10.1017/S0140525X12000660

Schmidt, R. C., Fitzpatrick, P., Caron, R., & Mergeche, J. (2011). Understanding social motor coordination. *Human Movement Science, 30*(5), 834–845. https://doi.org/10.1016/j.humov.2010.05.014

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

Schmidt, R. C., Nie, L., Franco, A., & Richardson, M. J. (2014). Bodily synchronization underlying joke telling. *Frontiers in Human Neuroscience, 8*, Article 633. https://doi.org/10.3389/fnhum.2014.00633

Seibt, B., Mühlberger, A., Likowski, K., & Weyers, P. (2015). Facial mimicry in its social setting. *Frontiers in Psychology, 6*, 1125. https://doi.org/10.3389/fpsyg.2015.01125

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

Stel, M., van Baaren, R. B., & Vonk, R. (2011). Effects of mimicking: Acting prosocially by being emotionally moved. *European Journal of Social Psychology, 41*(1), 33–43. https://doi.org/10.1002/ejsp.722

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

Sun, Y., Greaves, D., Orgs, G., Hamilton, A. F. de C; Day, 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

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society, 79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079%3c0061:APGTWA%3e2.0.CO;2

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology, 5*, Article 1323. https://doi.org/10.3389/fpsyg.2014.01323

van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science, 15*(1), 71–74. https://doi.org/10.1111/j.0963-7214.2004.01501012.x

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

Ward, J. A., & Pinti, P. (2019). Wearables and the brain. *IEEE Pervasive Computing, 18*(1), 94–100. https://doi.org/10.1109/MPRV.2019.2898536

Wiltshire, T. J., Steffensen, S. V., & Fiore, S. M. (2019). Multiscale movement coordination dynamics in collaborative team problem solving. *Applied Ergonomics, 79*, 143–151. https://doi.org/10.1016/j.apergo.2018.07.007

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Patrick Falk[1,2] · Roser Cañigueral[2] · Jamie A Ward[2,3] · Antonia F de C Hamilton[2]

✉ Patrick Falk
patrick.falk@his.se

Roser Cañigueral
roser.canigueral.15@ucl.ac.uk

Jamie A Ward
J.Ward@gold.ac.uk

Antonia F de C Hamilton
a.hamilton@ucl.ac.uk

[1] School of Bioscience, University of Skövde, Skövde, Sweden

[2] University College London, London, United Kingdom

[3] Goldsmiths University of London, London, United Kingdom