UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Dyadic differences in empathy scores are associated with kinematic similarity during conversational question-answer pairs

Trujillo, James P; Dyer, Rebecca M K; Holler, Judith

DOI
10.1080/0163853X.2025.2467605

Publication date
2025

Document Version
Final published version

Published in
Discourse processes

License
CC BY

Link to publication

Citation for published version (APA):
Trujillo, J. P., Dyer, R. M. K., & Holler, J. (2025). Dyadic differences in empathy scores are associated with kinematic similarity during conversational question-answer pairs. *Discourse processes*, *62*(3), 195-213. https://doi.org/10.1080/0163853X.2025.2467605

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Sep 2026

DISCOURSE PROCESSES
2025, VOL. 62, NO. 3, 195–213
https://doi.org/10.1080/0163853X.2025.2467605

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Dyadic differences in empathy scores are associated with kinematic similarity during conversational question–answer pairs

James P. Trujillo[a,b,c*], Rebecca M. K. Dyer[c,d*], and Judith Holler[a,b]

[a]Donders Institute for Brain, Cognition & Behaviour, Radboud University, Nijmegen, The Netherlands; [b]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [c]Institute for Logic, Language & Computation, University of Amsterdam, Amsterdam, The Netherlands; [d]Centre for Human Developmental Science, Cardiff University, Cardiff, UK

ABSTRACT

During conversation, speakers coordinate and synergize their behaviors at multiple levels, and in different ways. The extent to which individuals converge or diverge in their behaviors during interaction may relate to interpersonal differences relevant to social interaction, such as empathy as measured by the empathy quotient (EQ). An association between interpersonal difference in empathy and interpersonal entrainment could help to throw light on how interlocutor characteristics influence interpersonal entrainment. We investigated this possibility in a corpus of unconstrained conversation between dyads. We used dynamic time warping to quantify entrainment between interlocutors of head motion, hand motion, and maximum speech *f0* during question–response sequences. We additionally calculated interlocutor differences in EQ scores. We found that, for both head and hand motion, greater difference in EQ was associated with higher entrainment. Thus, we consider that people who are dissimilar in EQ may need to "ground" their interaction with low-level movement entrainment. There was no significant relationship between *f0* entrainment and EQ score differences.

When engaging in conversation with one another, behaviors, such as the way we move or speak, often converge between interlocutors. This process, called interpersonal entrainment, is thought to facilitate successful communication, due to its association with various positive social outcomes (Bowsher-Murray et al., 2023; Hove & Risen, 2009; Rabinowitch & Meltzoff, 2017; Tarr et al., 2015; Vacharkulksemsuk & Fredrickson, 2012; Zampella et al., 2020). However, there is much variation in the extent to which we entrain our behaviors with our conversational partner. This may depend on the specific communicative context, or individual differences. One possibility is that individuals who are more similar in their social dispositions may find it easier to entrain their behaviors with one another. While there is recent interest in the idea of social dispositions as an influencing factor on social interaction (i.e., the double-empathy problem (Milton, 2012; Milton et al., 2022)), it is currently not known whether or how differences in traits that are relevant for social interaction, such as empathy, impact the extent of behavioral entrainment during conversation.

Recent studies of conversational behavior demonstrate that speakers form dynamical patterns of interaction and behavior coordination with one another. In particular, many studies have investigated entrainment, quantifying the extent to which two interacting individuals exhibit

CONTACT James P. Trujillo j.p.trujillo@uva.nl Institute for Logic, Language & Computation, University of Amsterdam, Science Park 900, Amsterdam 1098 XH, The Netherlands

*These authors contributed equally to this work.

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

similar behaviors, either at the same time or with some lag. In speech, entrainment can refer to linguistic entrainment, such as the use of similar words, syntactic structure, or even converging on a particular topic (Duran et al., 2019), or acoustic entrainment, such as converging on similar prosodic features (Buder & Eriksson, 1997). Similarly, there may be kinematic entrainment, such as moving the hands or head similarly to one's conversational partner (Danner et al., 2021; Louwerse et al., 2012; Trujillo et al., 2023), or gestural entrainment in the form of using similar types of gestures to refer to the same concept (Cienki et al., 2014; Holler & Wilkin, 2011; Holler et al., 2022; Kimbara, 2006; Rasenberg, Dingemanse, et al., 2020). While some earlier studies of kinematic entrainment focused on moment-to-moment behavioral similarity, or so-called synchrony of behavior, we also observe linguistic and kinematic entrainment when comparing sequences of conversational turns, indicating that speakers entrain their movements to one another in fluctuation with the back and forth of dialog. Importantly, entrainment does not necessarily occur across all levels or modalities but. rather, there seems to be a dynamic interplay between different forms of entrainment (Ostrand & Chodroff, 2021; Rasenberg, Özyürek, et al., 2020; Rasenberg et al., 2022; Trujillo et al., 2023). These dynamically emerging patterns of entrainment are thought to facilitate interaction through allowing speakers to converge on shared meaning, resolve ambiguities, elaborate and progress the conversation (Healey et al., 2014; Oben & Brône, 2016) and the emergence of such patterns is generally associated with positive affiliation with one's partner (Hu et al., 2022; Rennung & Göritz, 2016). The term "entrainment" has been measured and defined in different ways, such as on very different timescales (e.g., some studies use entrainment to mean similarity in behavior across an entire interaction; whereas, others refer to similarity of behaviors between adjacent turns; Wynn & Borrie, 2022). Importantly, entrainment occurs along a continuum, whereby entrainment between two individuals may be high, indicating that their behavior is highly similar to one another, or it may be low, indicating that their behavior is highly dissimilar. Low entrainment has also been referred to as "disentrainment," and at least for speech prosody is also associated with positive evaluation of conversation (Pérez et al., 2016).

Interpersonal differences and entrainment

While entrainment is generally seen as a positive aspect of social interaction, individuals may differ substantially in the extent to which they achieve effective coordination (both in terms of entrainment and complementarity). Thus, interpersonal differences appear to play a role. This has primarily been demonstrated by differences between autistic and nonautistic people in entrainment. For example, autistic people, while demonstrating linguistic entrainment during interactions with autistic or nonautistic people in structured tasks (Branigan et al., 2016; Slocombe et al., 2013), demonstrate differences compared with nonautistic people in linguistic entrainment during spontaneous conversation (Bottema-Beutel, 2017). A similar pattern emerges across various modalities of interpersonal entrainment, with conversations involving autistic people demonstrating less entrainment of bodily motion (Georgescu et al., 2020; Noel et al., 2018; Romero et al., 2018) and prosody (Lehnert-LeHouillier et al., 2020; Ochi et al., 2019) compared with conversations involving only nonautistic people. This pattern has also been observed outside of conversation in a tacit communication game, with autistic adults being less likely than nonautistic adults to entrain their nonverbal communicative strategies to that of their partner. Notably, this occurred despite a similar desire and ability to communicate in the task (Wadge et al., 2019).

Importantly, these differences in interpersonal entrainment may not be specific to the autistic population per se. Zampella et al. (2020), as in the studies already described, found that, compared with nonautistic youth, autistic youth demonstrated lower coordination of reciprocal smiling in dyadic conversations with their mothers and with an unfamiliar adult. However, the level of smile coordination, for both autistic and nonautistic participants, was

associated with measures of social ability (Social Responsiveness Scale-2 (SRS-2); Constantino & Gruber, 2012) and dispositional empathy (Interpersonal Reactivity Index (IRI); Davis, 1983), with higher smile coordination occurring in children with higher (parent-reported) empathy. The observed reduction in interpersonal entrainment in autistic people, therefore, may be transdiagnostic—related to characteristics present in but not specific to this population. This is further supported by McNaughton and Redcay (2020), whose review of interpersonal synchrony in autistic participants found that intrapersonal characteristics, such as motor timing, may mechanistically explain the observed entrainment differences in this population. Thus, taking a transdiagnostic approach and investigating the relationship between the characteristics highlighted here, such as empathy (Zampella et al., 2020) and interpersonal entrainment, is necessary to improve our understanding of the mechanisms of interpersonal differences in entrainment.

This is of further relevance, as such a relationship between interpersonal entrainment and empathy has also been described by Fujiwara and Daibo (2022), who found that empathic accuracy (in this case, the ability to classify the thoughts or feelings of an interaction partner from a video recording) is positively correlated with interpersonal entrainment in the form of postural mirroring. Importantly, Zampella et al. (2020) and Fujiwara and Daibo (2022) looked at the empathy of only one individual. Rather than coordination being linked to the magnitude of empathy scores in one individual, it may be that interpersonal coordination is more generally related to the difference in empathy scores between two individuals. This would be more in line with the fact that interpersonal synchrony does not rely on the efforts of one individual or the other but rather the dynamic coupling of two (or more) individuals' behaviors (Schmidt et al., 2012).

Variation in interpersonal entrainment as a two-person problem

Understanding variation in interpersonal entrainment will require us to consider that entrainment arises from the interaction between individuals. Assessing variation in the extent to which two individuals entrain with one another or show behavioral similarity should, therefore, consider the impact of interpersonal differences between entraining interlocutors.

The relationship between interpersonal differences and social behaviors has been described by the double-empathy problem: the idea that people, as a rule, find it more difficult to empathize with those who differ from them than with those who are similar to them (Milton, 2012; Milton et al., 2022). This has been most famously applied to communication breakdown between autistic and nonautistic individuals due to differences in diagnostic status. Such a theoretical reduction in empathy between autistic and nonautistic individuals may negatively impact aspects of social interaction, with Williams et al. (2021) finding that flow, rapport, enthusiasm, and mutual affect were greater in conversations between two autistic adults than in conversations between autistic and nonautistic adults. While the double-empathy problem does not directly refer to empathy traits specifically, the overall framework seems to fit well with previous research suggesting a link between empathy and motor synchrony (Fujiwara & Daibo, 2022; Zampella et al., 2020). At the same time, thinking of interpersonal coordination as a two-person "problem" (Bolis et al., 2018; Schilbach, 2016) and within the framing of interpersonal differences as relating to our ability to connect with one another (Milton, 2012; Milton et al., 2022) suggests that we should consider differences in trait scores, such as empathy, rather than the magnitude of one individual's score.

The impact of such interlocutor differences on interpersonal entrainment in conversations has been investigated. Georgescu et al. (2020) investigated interpersonal entrainment in conversations involving two autistic individuals, involving one autistic and one nonautistic individual, or involving two nonautistic individuals. While these researchers found that conversations containing at least one autistic interlocutor demonstrated less interpersonal entrainment than dyads containing only nonautistic participants, interpersonal entrainment

did not significantly differ between dyads containing participants with matching diagnostic status and those containing participants with mismatched diagnostic status. While an interesting study to systematically compare autistic, nonautistic, and mixed dyads, Georgescu et al. (2020) did not account for interpersonal differences beyond broad neurotype classification. This is particularly important following neurodiversity theories that humans diverge in multiple ways that cannot be captured by diagnostic status alone (Singer, 2016). Furthermore, the double-empathy problem is not only relevant to autistic people but to broader interpersonal differences (Milton et al., 2022). Therefore, it appears a logical next step to investigate whether interpersonal differences in empathy—a transdiagnostic trait observed to relate to interpersonal entrainment (Fujiwara & Daibo, 2022; Zampella et al., 2020)—relates to levels of interpersonal entrainment. Such a relationship could contribute to understanding the mechanisms of interpersonal differences in entrainment. This transdiagnostic approach, focusing more on individual differences, also aligns well with recent work that frames social interaction difficulties (such as in autism) as involving a misattunement between individuals rather than any deficiency in a particular group (Bolis et al., 2018). From this perspective, the most informative way forward is to assess whether we can determine whether there are measures of interpersonal differences that, while collected individually, are predictive of behavioral measures of attunement, or coupling, during conversation.

The present study

The current study aims to better elucidate how variation in interpersonal behavioral entrainment may arise from differences in conversational partners. The idea that partner differences may be associated with interpersonal behavioral (dis)similarity largely builds on the double-empathy problem as a general framework for understanding differences in social interaction dynamics, as well as the findings of a significant relationship between interpersonal entrainment and the strength of autistic traits (Chen et al., 2022; Cheng et al., 2017). More specifically, the present study aimed to determine whether interpersonal behavioral entrainment is associated with differences in empathy quotient (EQ) score between partners engaged in naturalistic, unscripted dyadic conversation.

In the present study, we analyzed 34 video and audio recordings from the Communication in Action (CoAct) corpus (previously analyzed in Nota et al., 2021; Ter Bekke et al., 2024; Trujillo & Holler, 2021) in which nonautistic Dutch speakers engaged in three dyadic 20-minute conversations. We additionally recorded empathy scores based on the EQ questionnaire, and calculated empathy-score differences. From the conversation data, question–response pairs (which form a core coordinative structure and building block of everyday conversational interaction) from each dyad were assessed for interpersonal behavioral entrainment. To capture the multimodal nature of interpersonal behavior, three measures of behavioral entrainment were investigated: head entrainment, hand entrainment, and prosodic matching. We specifically measured what Wynn and Borrie (2022) refer to as *static local proximity* entrainment, calculating the similarity in movement or prosody values between adjacent speaking terms. Movement similarity was used to quantify entrainment for both motor measurements (head and hand), as we made use of dynamic time warping (DTW) to compare kinematic profiles between interlocutors. Utilizing DTW allowed us to compare kinematic profiles between questions and responses, which may differ in length, rather than focusing on small, fixed window sizes inherent in cross-correlation methods. We hypothesized that lower behavioral entrainment would be associated with higher partner EQ differences, which would suggest that behavioral entrainment is lower when conversational partners are able to empathize with one another to a similar extent. Understanding whether differences in empathy affect interaction dynamics during natural conversation will further inform whether interpersonal differences play a role in entrainment and, in particular, whether the patterns expected by the double-empathy problem are relevant for such investigations.

Methods

Data availability and open science

We report all measures and exclusions utilized in the current study. Materials used in this study, including data-analysis scripts and preprocessing scripts, are available via the Open Science Framework (OSF) at the following link: https://osf.io/hp48e/.

Note that the raw data, including audio and video files, cannot be made public due to ethical constraints and related concerns for privacy and anonymity.

The current study was not preregistered, and a priori sample-size estimation could not be performed, as this was a follow-up study using a corpus that had already been collected as part of a larger project (see subsection *Materials*).

Materials

Data for the present study consisted of video, audio, and questionnaire data collected as part of the Communication in Action (CoAct) Corpus (Nota et al., 2021; Ter Bekke et al., 2024; Trujillo & Holler, 2021). The CoACT Corpus consists of 59 dyads (37 female–female, 10 male–male, 12 male–female) of acquainted native Dutch-speakers conversing. For 34 of these dyads (68 participants, 51 female, mean age 23.10 years) question–response and social action coding was completed, and thus these dyads were included in the present study. Each dyad engaged in three 30-minute Dutch conversations: the first a natural conversation, the second a discussion of opinions about predefined topics, and the third a collaborative planning task. Recordings were made in a soundproof room at the Max Planck Institute for Psycholinguistics in the Netherlands. Two cameras filmed the front of each participant, while two more filmed each of their bodies from a 45-degree angle and two others recorded a birds-eye view. Finally, one camera filmed the two participants together. All cameras filmed at 25 frames per second. Each participant's speech was recorded by a directional microphone. All video and audio files were synchronized, leaving a time resolution of 40 ms. Both video and audio files were utilized for the present study. Following the conversation tasks, participants completed two questionnaires: the empathy quotient (EQ; Baron-Cohen & Wheelwright, 2004; Groen et al., 2015) and the brief version of the Fear of Negative Evaluation Scale (Leary, 1983; Watson & Friend, 1969). Responses to the EQ were used for the present study.

The EQ is a 60-question self-report questionnaire, consisting of 40 questions designed to gauge the respondent's empathy skills and 20 filler questions. It is scored out of 80, as participants can achieve up to two points on each of the target questions (Baron-Cohen & Wheelwright, 2004). The Dutch version of this questionnaire has also shown good internal consistency (Cronbach's $\alpha = 0.89$; Groen et al., 2015). It is important to note here that there is some debate about how to classify empathy (Kerem et al., 2001), with researchers identifying different types of empathy. However, it is widely reported that there are at least two types—cognitive and affective (Davis, 1980)—with cognitive empathy referring to the recognition of others' mental states and affective empathy referring to a sympathetic emotional response (Baron-Cohen & Wheelwright, 2004). The EQ does not distinguish between types of empathy.

The present study was conducted within existing ethical approval for a larger project, approved by the ethics committee of the social sciences department of Radboud University, Nijmegen (approval code: ECSW 2018-124).

Question–response pair annotations

Rather than analyzing all turn segments within the corpus, we focused on question–response pairs, as questions are incredibly common in conversation and since questions and their responses are so-called *adjacency pairs* (Schegloff, 2007). Adjacency pairs provide a useful

focus for these analyses, as such focus allows us to zoom in on strongly reciprocal moments of the interactions requiring a high degree of coordination. In other words, the overall behavior of Speaker A will be related to the behavior of Speaker B. This clear question–response format therefore avoids some of the noise typically created by the varying contexts of turn-taking in conversation. As described in Nota et al. (2021) and Trujillo and Holler (2021), questions and responses were manually annotated using the following procedure. First, we made an automatic orthographic transcription of the speech signal using the Bavarian Archive for Speech Signals Webservices (Kisler et al., 2017). Questions were identified and coded in ELAN (5.5; Wittenburg et al., 2006), largely following the coding scheme of Stivers and Enfield (Stivers & Enfield, 2010). In addition to this scheme, more rules were applied on an inductive basis to account for the complexity of the data in the corpus. Specifically, we adopted a holistic approach that took into consideration visual bodily signals, context, phrasing, intonation, and addressee behavior. Nonverbal sounds were excluded (e.g., laughter, sighs). This was done by two human coders, one native speaker of Dutch, and one highly proficient speaker of Dutch. Interrater reliability between the two coders was calculated with raw agreement (Cohen, 1960; Landis & Koch, 1977) on 12% of the total data (four dyads, all tasks). A standard overlap criterion of 60% was used. Reliability between the coders resulted in a raw agreement of 75% for questions and a raw agreement of 73% for responses, indicating substantial agreement. This resulted in a total of 6,778 questions (duration Mdn = 1,114 ms, range = 99–13,145 ms, IQR = 1,138 ms). As not all questions received a response, the present analyses are based on a subset of 4,436 question–response sequences. Note that other levels and types of annotations were made for this corpus (see, e.g., Nota et al., 2021; Trujillo & Holler, 2021), as it is part of a much larger project. However, these other annotations were not used for the present study.

EQ score differences

While all of the participants completed the EQ, in only 19 dyads did both partners complete every question. Following the procedure outlined by Lawrence et al. (2004) for handling missing values within the EQ, scores for questions that were skipped by participants were replaced with the mean score of the group for that question, rounded to the closest integer (Lawrence et al., 2004). See Figure 1 for the distribution of individual EQ values. Following this handling of the missing values, the absolute difference between EQ scores of partners in each of the 34 dyads was calculated and taken as the final measure of EQ score differences.

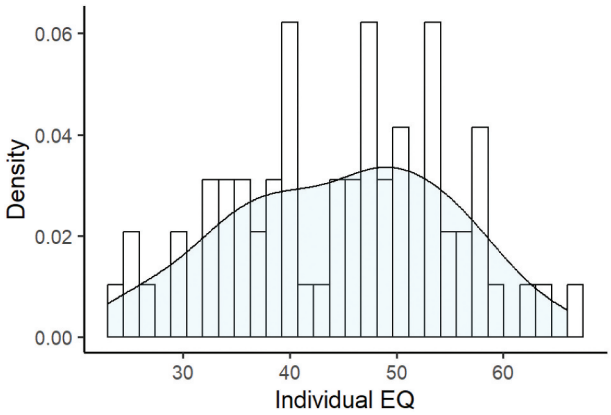

Figure 1. Distribution of individual EQ scores. EQ is given along the *x*-axis, while the density of distributions is given on the *y*-axis. The plot provides both a histogram and a density curve of these values.

Kinematic entrainment

To calculate motor similarity, we used OpenPose (Qiao et al., 2017) motion tracking to extract the two-dimensional motion of the nose, right-hand, and left-hand key points from each participant. Motion data were extracted for each question and response annotation (described above). We used motion of the nose as a measure of head movement.

Entrainment was then calculated based on the similarity between movement produced during the question (by the questioner) and movement during the response (by the responder). Specifically, head and hand asynchrony were calculated using dynamic time warping (DTW). DTW works by using a warping algorithm to nonlinearly align time series while preserving the order of time points in the two time series. After finding a maximal spatial alignment between the two time series, the remaining error, or distance, between the two is taken as a measure of dissimilarity (see Figure 2c). This distance score is normalized for the time-series lengths, ensuring that the distance score is not influenced by differences in time-series length. It is important to note that it is not possible to say whether there is or is not entrainment between two time series from just the distance score. Instead, these scores provide a magnitude of entrainment, where warping two completely unrelated time series will likely result in a very high value (i.e., high dissimilarity) and warping two highly similar time series (e.g., two highly entrained time series) will result in a much lower value. Rather than interpreting the values themselves, it is therefore important to use the values as contrasts between conditions or to test for a linear relationship between distance (henceforth, entrainment) values and some other variable of interest.

The DTW approach differs from other measures of interpersonal entrainment such as cross-recurrence quantification analysis (CRQA; e.g., Fusaroli & Tylén, 2016; Louwerse et al., 2012) in a crucial way. Methods such as CRQA are intended to capture the overall dynamics of recurring patterns of behavior between individuals. For example, the stability and complexity with which two individuals match their behaviors (e.g., both producing a gesture within a short time of one another). Critically, these analyses are focusing on the occurrence of matched behavior events. In contrast, DTW is not quantifying recurring patterns over the course of an interaction—or even capturing the presence of behavior events such as gestures. Instead, we use it to quantify the similarity between two manually annotated events (see also Pouw & Dixon, 2020; Pouw et al., 2021; Trujillo et al., 2023 for similar implementations).

In the case of kinematic asynchrony, DTW therefore calculates the distance between the motion time series during the asking of a question and the motion time series during the uttering of the corresponding response. This distance was taken as a measure of similarity between each other's movements, wherein low values reflect high similarity (i.e., minimal warping to match the time series). DTW was conducted for nose movements, left-hand movements, and right-hand movements between partners for all question–response pairs in each dyad. This was performed in Python using Giorgino's (2009) DTW Python package. The mean similarity of the right and left hand was taken as the final measure of hand-movement similarity. Nose movement was used to calculate head-movement similarity. These similarity scores are then taken as a measure of kinematic entrainment. See Figure 2 for an overview of processing steps.

Prosodic mismatch

To calculate prosodic dissimilarity, mean fundamental frequency (pitch), between 75 and 300 Hz, was first calculated using the Python package Parselmouth PRAAT (Boersma & Weenink, 2007; Jadoul et al., 2018) for each question and response. We then calculated the difference in F0 between the question and corresponding response. This method was inspired by Levitan and Hirschberg (2011) and Edlund et al. (2009), who calculated a correlation coefficient for each dyad based on F0 of adjacent speech turns. We use absolute difference in order to utilize the full distribution of question–response pair data in a mixed-model approach (see below). See Figure 2 for an overview of processing steps.

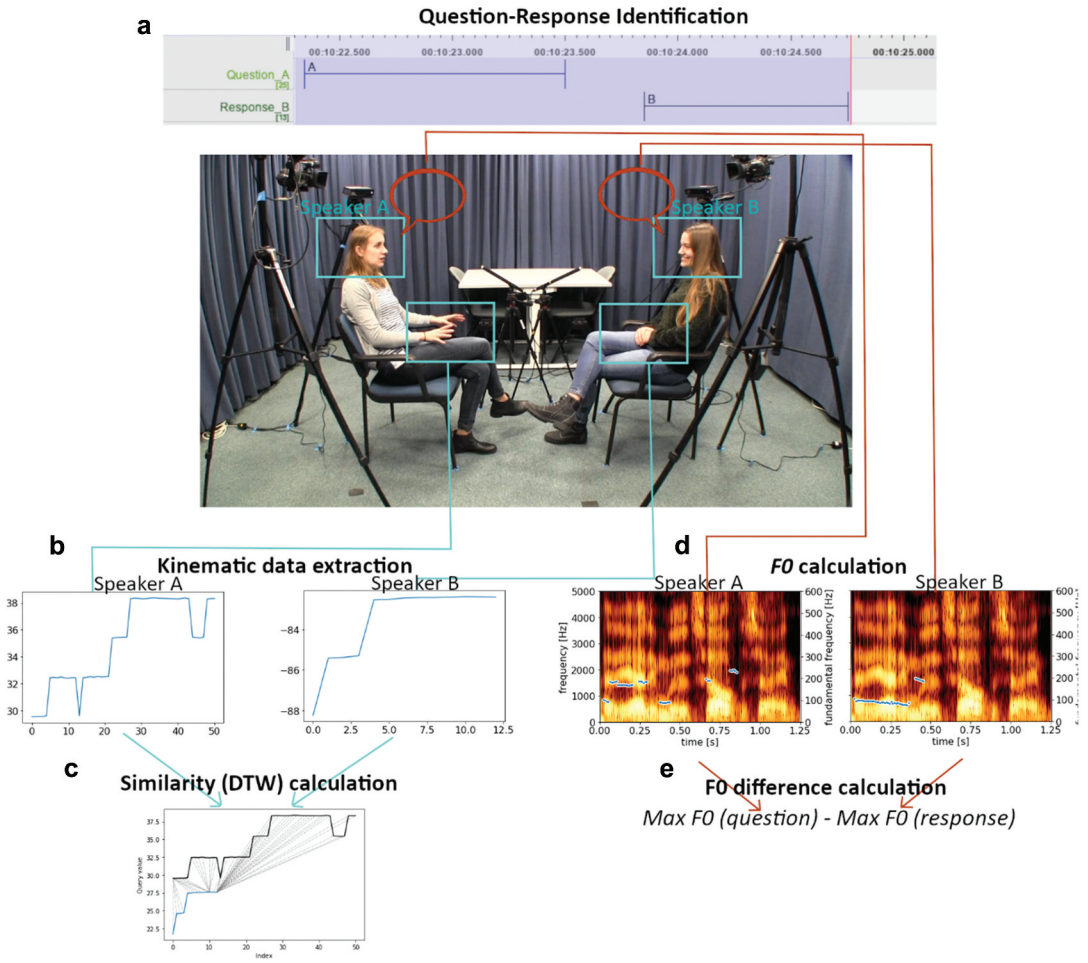

Figure 2. Overview of data and processing steps. In the center, a still frame is presented of the data recording set-up. The procedure involves qualitative, manual annotation (depicted above the still frame), and quantitative extraction of kinematic and acoustic features (depicted below the still frame). As a first step, question–response sequences were identified through manual coding of the data (Panel A). Motion data were extracted from the head and hands, providing displacement time series for each articulator during each question or response segment (Panel B). These motion data were compared using dynamic time warping (DTW) to provide a similarity score (Panel C). Audio recordings were used to extract the pitch contour (Panel D) for each question and response segment. Our prosodic matching score was calculated as the absolute difference in maximum F0 (pitch) between the question and response (Panel E). Participants depicted in this figure gave written, informed consent for their image to be used in academic publications.

Analyses: assessing interpersonal entrainment in question–response pairs

Before assessing whether interpersonal behavioral dissimilarity is related to other factors, we first assessed whether there was any evidence for different-than-chance-level dissimilarity. This was done by comparing the observed question–response dissimilarity against shuffled data, wherein the question–response correspondence was broken. To this end, we shuffled the question–response correspondences within each dyad, so that each question was paired with a response different from that which it had originally corresponded with. By keeping the shuffling within dyads, we also keep any intrinsic level of movement (dis)similarity that may exist between two interlocutors. Our test is thus aimed at determining whether question–response sequences show less dissimilarity (i.e., greater degree of entrainment) than we would expect from any intrinsic movement (dis)similarity occurring

simply by being engaged in dialogue. This surrogate test procedure thus applies the same logic as the standard test for greater-than-chance synchrony (Delaherche et al., 2012), but here applied to dissimilarity. We then built separate models for each of the three dissimilarity variables, with dissimilarity (i.e., head asynchrony, hand asynchrony, prosodic dissimilarity) being the dependent variable and a random intercept for ID, which identified the specific dyad and participant (included a nested term for dyad and participant that led to convergence and singularity issues with model fit. A null model containing only the dependent variable and the random term was then compared, using a likelihood ratio test, against a full model that also contained data set (real versus shuffled) as a fixed effect. A significant model comparison indicates that dissimilarity is different from what would be expected by chance (i.e., in the shuffled data). Holm's correction for multiple comparisons was applied to avoid the increased risk of a Type 1 error (Chen et al., 2017) associated with these three tests. Specifically, following the guidelines of Rubin (2021), Holm's correction was applied to the two kinematic tests (i.e., head asynchrony and hand asynchrony).

Analyses: assessing associations between EQ differences and interpersonal entrainment

To determine whether our measures of interpersonal behavioral dissimilarity were associated with differences in EQ scores, we utilized mixed effects linear regression models, implemented using the lme4 package in R (Bates et al., 2015). We built separate models for each of the three dissimilarity variables, with dissimilarity (i.e., head asynchrony, hand asynchrony, prosodic dissimilarity) as the dependent variable and a random intercept for ID, which identified the specific dyad and participant. A null model containing only the dependent variable and the random term was then compared, using a likelihood ratio test, against a full model that also contained EQ difference as a fixed effect. A significant model comparison indicates that adding the EQ difference explains significantly more variance in the dissimilarity data than is explained by interindividual variability (i.e., the random intercept) alone.

Alpha thresholds were adjusted to account for the increased false positive rate associated with performing three comparisons (Chen et al., 2017). Based on the guidelines of Rubin (2021), we categorize our tests as falling into families: EQ association with prosody and EQ association with kinematics. This follows from the guidelines of Rubin (2021) and is based on our hypotheses that relate to prosody and kinematics separately. We then calculate adjusted p-values for our head and hand kinematic models based on Holm's method for multiple comparisons correction (Chen et al., 2017; Holm, 1979).

Post hoc analysis: assessing associations between entrainment values

Besides behavior entrainment being related to (differences in) individual traits, entrainment in one modality may also be negatively correlated with entrainment in another modality (Trujillo et al., 2023). This would impact the way we must interpret why there may be associations between entrainment and EQ-differences in one modality or articulator and another. We therefore calculated partial correlations between each pair of entrainment variables while excluding the effect of the third entrainment variable. This was done using the partial.r function of R package psych (Revelle, 2020).

Results

Results: significance of interpersonal behavioral entrainment

For our comparison with the pseudo– (i.e., randomly matched) question–response pairs, we find that head asynchrony ($\chi^2(1) = 61.129$, $p < .001$), hand asynchrony ($\chi^2(1) = 8.807$, $p = .003$), and prosodic matching ($\chi^2(1) = 8.190$, $p < .001$) all significantly differed in the real question–response pairs

compared to the pseudo-pairs. This indicates that the dissimilarity that we observe is different from what would be expected in chance circumstances. Specifically, we see higher kinematic asynchrony (i.e., lower similarity) in the real question–response pairs compared to the pseudo-pairs (head asynchrony t-value = 7.84; hand asynchrony t-value = 2.969) but lower F0 difference (i.e., higher prosodic matching) in the real pairs compared to the pseudo-pairs (t-value = 2.862).

These results indicate that all behavioral similarity measures show a difference from the values that would be expected in chance circumstances and, specifically, that kinematic entrainment is lower than would be expected in chance circumstances, while prosodic matching is higher than would be expected in chance circumstances.

Results: associations between EQ differences and behavioral entrainment

For kinematic entrainment, we found that hand movement similarity was significantly positively associated with EQ difference ($\chi^2(1) = 4.295$, $p = .038$, Holm-corrected $p = .044$) such that for each additional point difference in EQ, hand similarity increased by 116.880 ± 50.67 ($t = 2.307$). See Figure 3a for an overview of these results.

Similarly, we found that head movement similarity was similarity positively associated with EQ difference ($\chi^2(1) = 5.264$, $p = .022$, Holm-corrected $p = .044$) such that for each additional point difference in EQ head movement similarity increased by 52.436 ± 25.28 ($t = -2.074$). See Figure 3b for an overview of these results.

We found no evidence for prosodic mismatch being associated with EQ difference ($\chi^2(1) = 0.158$, $p = .691$).

These results indicate that higher kinematic entrainment was associated with higher EQ score differences.

Post hoc results: assessing associations between entrainment values

Our exploratory post hoc analysis showed that only entrainment of the head and hands shows a strong association. Much weaker associations are found between F0 difference and either head or hand kinematic entrainment. See Figure 4 for a visualization of these associations.

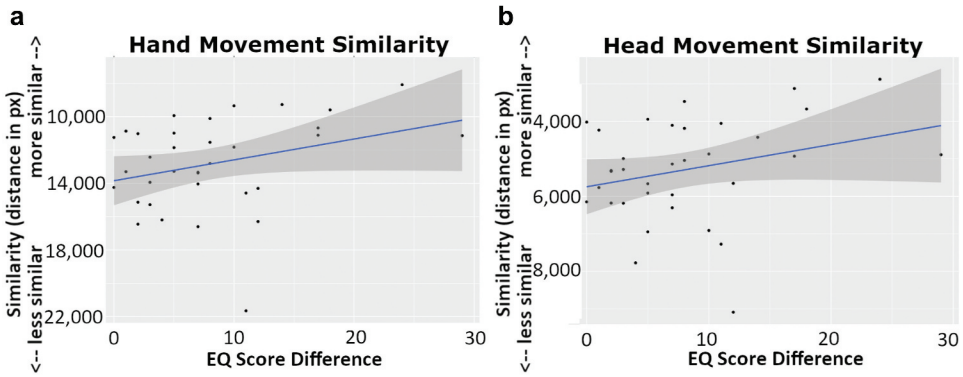

Figure 3. Scatterplots of the per-dyad average movement similarity values plotted against EQ score differences. The left panel (a) shows hand-movement similarity and the right panel (b) shows head-movement similarity. In both panels, similarity (measured as Euclidean distance in pixels) is given on the *y*-axis and EQ score difference (absolute difference) is given on the *x*-axis. A linear fit line is given in blue, with the gray shaded area indicating the 95% confidence band around the fit line.

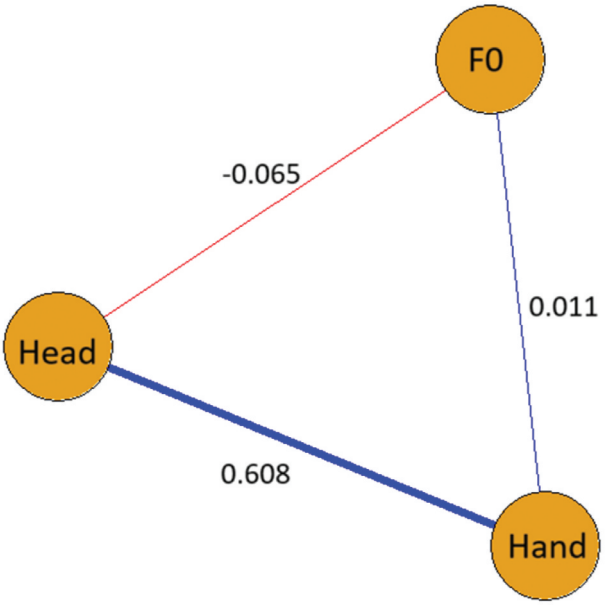

Figure 4. Associations between entrainment variables. Each variable is represented by a node in the graph, with the line (i.e., vertex) between two nodes representing the association (i.e., partial correlation) between the two variables. Stronger absolute associations are indicated by thicker vertices; whereas, direction of association is indicated by color (i.e., blue for positive, red for negative). Correlation coefficients are indicated next to each vertex.

Discussion

Following the double-empathy problem, the present study aimed to investigate whether interpersonal behavioral entrainment during question–response sequences is associated with difference in EQ scores during dyadic conversation. We found that dyads with higher differences in EQ show higher levels of kinematic entrainment of the head and hands.

We first used a pseudo-pair analysis to determine whether behavioral entrainment was significantly different from chance. This was done because different-than-chance entrainment shows that we are measuring something that captures the interactive dynamic of conversation. This first analysis showed that the kinematic and prosodic entrainment observed in our data is significantly different from what would be observed in chance circumstances. However, kinematic entrainment was found to be lower than what would be expected by chance. Although this at first seems counterintuitive, absent or lower-than-chance entrainment has previously been observed, specifically, in relatively unconstrained dialogue (Healey et al., 2014; Howes et al., 2010). Additionally, our analyses focused on question-response pairs, which have an inherent asymmetry between the speakers—that is, speakers are actively responding to an imbalance of knowledge, an ambiguity, or are expressing a particular stance that is likely to elicit a response (e.g., expressing criticism by asking "Do you really think that's a good idea?"). These results are therefore informative on several levels. First, we can be confident that we are observing true coordination between speaker behavior and for one of the core structures prevalent in everyday conversation. Second, our results support the notion that divergence, which may be a form of complementarity, is an important aspect of communicative coordination (Fusaroli et al., 2014; Healey et al., 2014).

The finding of lower-than-chance prosodic mismatch (i.e., high prosodic *matching*) is also interesting in light of work from Syrdal and Kim (2008), who measured F0 range across various interrogative (such as wh- questions), imperative, assertive, and affective speech acts. They found that F0 differed significantly across the various speech acts (Syrdal & Kim, 2008). The fact that we find a tighter coupling of F0 values between speakers than would be expected by chance suggests that,

although questions and responses would be considered separate speech acts (Holtgraves, 2008; Levinson, 2017; Searle, 1969), the sequential organization of these speech acts, in which questions directly elicit a pragmatically fitting response, provides a layer of contingency and, thus, a "binding relationship" between the question and the response.

Interpersonal entrainment as a two-person problem

The main focus of our study was on the association between social predispositions and entrainment. We found associations between kinematic entrainment of both the head and hands with partner differences in EQ score. We initially hypothesized that such an association would be negative, such that higher EQ-score differences would be associated with lower behavioral entrainment. Our findings of higher entrainment associating with higher EQ-score difference is thus in contrast with the pattern expected by the double-empathy problem—that motor similarity would be lower in dyads with greater differences in empathy. Thus, the present study did not support the theory that the patterns expected by the double-empathy problem would extend to motor entrainment, at least in the sense of movement similarity during questions and their responses. When considering why the present study and that by Georgescu et al. (2020) found no evidence for the patterns expected by the double-empathy problem for motor entrainment, it should be noted that the hypotheses of both studies took high entrainment (i.e., low behavioral dissimilarity) to be a positive social outcome, reached with successful social interaction.

Entrainment may be more appropriately viewed as a tool to promote better social outcomes, as entraining behaviors between individuals have been found to be beneficial for various social outcomes (Bowsher-Murray et al., 2023; Galbusera et al., 2019; Hove & Risen, 2009; Rabinowitch & Meltzoff, 2017; Tarr et al., 2015; Vacharkulksemsuk & Fredrickson, 2012). Therefore, entrainment may increase with increasing EQ score differences between conversational partners in order to promote better social outcomes, as a compensatory mechanism for interpersonal differences. When considered within the logic of the double-empathy problem and Williams et al.'s (2021) finding that communication is more successful between people who are more similar to each other, it is possible that communication between those participants who differed more from each other was more difficult initially. Thus, to promote better social outcomes, participants who differed more from their partners may have spontaneously increased entrainment to promote more-positive social outcomes, compensating for their interpersonal differences on the EQ.

This putative compensatory mechanism follows past findings that head-movement entrainment increases with increasing conversation difficulty (Hadley & Ward, 2021). Such a compensatory mechanism leads to the question of why entrainment would not remain at its greatest in all conversations to achieve the best possible social outcomes. While the majority of interpersonal entrainment research has investigated the benefits of entrainment, few studies have investigated the disadvantages of increased entrainment (Galbusera et al., 2019). Galbusera et al. (2019) found that, while motor entrainment increased social affect, it also reduced self-regulation of affect—that is, higher motor entrainment was associated with lower self-reported ability to (prereflectively) regulate or modulate one's own emotions in the moment. Thus, entrainment may be used to compensate for increased conversational difficulty when interpersonal differences are high but reduced when such differences are less prevalent in order to protect intrapersonal processes like self-regulation. Similarly, Wallot et al. (2016) found that lower interpersonal entrainment was associated with better performance on a complex joint-action task (Wallot et al., 2016), again highlighting the importance of behavioral dissimilarity in certain contexts. This mechanism also fits well with the framing of dialogue as interpersonal synergy (Fusaroli et al., 2014), wherein both entrainment and divergence are important aspects of interaction. Of course, future research should also investigate whether the relationship between kinematic entrainment and EQ differences holds in other conversational turn sequences or whether it is specific to question–response sequences.

We found that, contrary to our hypothesis, no significant relationship between EQ score differences and prosodic dissimilarity. Interestingly, the prosodic results did not follow the same pattern as the motor results. One potential explanation is that acoustic entrainment, when visual entrainment is present, may be redundant. Human communication is multimodal, and interference with either the visual or the audio signal can be compensated for by the presence of the other (Alexanderson & Beskow, 2014; Davis et al., 2006; Drijvers & Özyürek, 2017; Drijvers et al., 2018). As motor entrainment was already increasing with EQ score differences, such an increase in prosodic matching may have been redundant, as the full visual signal was available. This theory follows synergy theories of how multimodal forms of entrainment come together. In these framings, entrainment often does not extend to all modalities (Dale et al., 2013; Esteve-Gibert et al., 2022; Fusaroli et al., 2014; Fusaroli & Tylén, 2016; Trujillo et al., 2023). This is also in line with research showing that compensatory exaggeration of one modality, such as speaking louder in noisy environments, does not always go hand-in-hand with exaggeration of other modalities (Garnier et al., 2018; Trujillo et al., 2022, 2021). Given that the current study focused on question–answer sequences, it is also possible that prosodic entrainment and EQ differences are present in other aspects of conversation, or even correspond to a different time scale (e.g., occurring as a general trend over the course of conversation but not specifically according to turn sequences).

Finally, our exploratory post hoc analysis showed no evidence for a strong association, whether positive or negative, between kinematic entrainment and prosodic entrainment. This is contrary to what was found between kinematic and some levels of linguistic entrainment in a previous study on Danish and Norwegian conversation (Trujillo et al., 2023). This is particularly interesting given that F0 entrainment in real pairs was stronger than in pseudo-pairs. This means that while kinematic entrainment is more strongly associated with interpersonal differences in self-reported EQ, it may be relatively independent of the entrainment that we observe in speech prosody. Future analyses will be needed to determine the extent to which and the time scales according to which, prosodic entrainment relates to interpersonal trait differences.

Future directions and practical applications

Future research should assess whether other personality traits, or social predispositions, are also predictive of interpersonal coordination. One possibility is that empathy quotient scores form just one part of a cluster of personality traits that are associated with interpersonal coordination. For example, empathy, extraversion, and agreeableness are positively associated with facial mimicry in simple facial expression viewing paradigms (Perugia et al., 2020; Rymarczyk et al., 2019). An interesting direction for future work would be to quantify the (potential) network of interpersonal trait differences that contribute to the extent of coordination. This avenue of research could be useful both for developing a stronger theoretical understanding of interpersonal coordination, and to inform data-driven–team-management decisions (i.e., predicting who will work best with whom).

It should be noted that the current analyses focused on conversations between speakers who were already acquainted with one another. An interesting future direction would be to test whether the current results generalize to other types of relationships. Additionally, future research should assess whether the associations between EQ and entrainment are apparent very early in a conversation between people who do not yet know each other or if the association is the result of a mutual adaptation that occurred over a longer time period.

While the current study has identified EQ differences as a factor related to the extent of entrainment between two individuals, future studies should additionally investigate whether differences in other traits also impact multimodal entrainment. As the EQ does not distinguish between different types of empathy, future research should utilize a more-refined measure with subscales for cognitive and affective empathy. This is particularly important as the two studies that had found a relationship between interpersonal entrainment and empathy had used different measures of empathy (Fujiwara & Daibo, 2022; Zampella et al., 2020). Future research, investigating the development of familiarity and

(potential) parallel development of motor entrainment, could elucidate how differences in empathy scores might predict the longer-scale development of stable interpersonal motor entrainment patterns among interlocutors who are not already familiar with one another.

While this research was inspired by autistic-person's differences in interpersonal synchrony, our transdiagnostic approach meant we did not investigate this population, specifically. It is hoped that in contributing to an understanding of the mechanisms of interpersonal synchrony, evidence-based advice can be produced to support populations who demonstrate differences in interpersonal synchrony, such as, but not necessarily limited to, autistic people. Such investigations into entrainment mechanisms are prudent because entrainment is associated with various positive social outcomes both in autistic people (Manders et al., 2022) and the wider population (Bowsher-Murray et al., 2023; Hove & Risen, 2009; Rabinowitch & Meltzoff, 2017; Tarr et al., 2015; Vacharkulksemsuk & Fredrickson, 2012). Positive social outcomes are important for various aspects of daily life—including job interviews (Strickland et al., 2013), which have been implicated as one of the potential barriers (Kumazaki et al., 2019) resulting in the underhiring and under-representation of autistic people in the workforce (Taylor & Seltzer, 2011). Understanding the mechanisms of reduced synchrony in this population could facilitate the development of inclusive policies for hiring and for other areas of life wherein autistic people are being systematically disadvantaged by current policies; for example, employers could receive education about reduced entrainment in autistic candidates and be encouraged to consciously consider whether this is influencing their hiring decisions.

In line with our proposal that interpersonal synchrony may work to facilitate positive social interactions, methods such as dance movement therapy, in which dance partners mirror each other, have been found to longitudinally improve social affect and affective engagement (Manders et al., 2022). Thus, group environments in which autistic people experience social exclusion, such as classrooms (Rotheram-Fuller et al., 2010), could utilize such methods to facilitate positive social affect among members. While these applications have been discussed specifically in relation to autistic individuals, differences in entrainment and reduced social outcomes might also be present in other (diagnosed or undiagnosed) neurodivergent groups (Quintero et al., 2019; Wehmeier et al., 2010). By better understanding what underlying interpersonal differences lead to such group differences, these applications can be more appropriately employed.

Finally, it is important to note that, because we focused on question–response sequences, any entrainment occurring between speakers is likely primarily driven by the responder matching the behavior of the questioner. In other words, within the level of dialogue acts (i.e., questions and responses, in this case), entrainment will be largely unidirectional, with the former influencing the latter. However, this does not mean that the entrainment value of a given dyad is based on one particular speaker entraining to the other. As both speakers asked questions and gave responses, the entrainment values essentially reflect entrainment in both directions, in terms of Speaker A entraining to Speaker B and vice versa.

Conclusions

Our study found that greater EQ differences between interlocutors are associated with greater kinematic entrainment during question–response sequences. Entrainment may, therefore, serve to compensate for increased interpersonal differences in empathy. Our findings provide new insights into how interpersonal coordination is associated with interpersonal differences in social traits.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

James P. Trujillo http://orcid.org/0000-0003-4713-376X

References

Alexanderson, S., & Beskow, J. (2014). Animated Lombard speech: Motion capture, facial animation and visual intelligibility of speech produced in adverse conditions. *Computer Speech & Language*, *28*(2), 607–618. https://doi.org/10.1016/j.csl.2013.02.005

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2007). *Praat* (Version 4.5. 25).

Bolis, D., Balsters, J., Wenderoth, N., Becchio, C., & Schilbach, L. (2018). Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. *Psychopathology*, *50*(6), 355–372. https://doi.org/10.1159/000484353

Bottema-Beutel, K. (2017). Glimpses into the blind spot: Social interaction and autism. *Journal of Communication Disorders*, *68*, 24–34. https://doi.org/10.1016/j.jcomdis.2017.06.008

Bowsher-Murray, C., Jones, C. R. G., & von Dem Hagen, E. (2023). Beyond simultaneity: Temporal interdependence of behavior is key to affiliative effects of interpersonal synchrony in children. *Journal of Experimental Child Psychology*, *232*, 105669. https://doi.org/10.1016/j.jecp.2023.105669

Branigan, H. P., Tosi, A., & Gillespie-Smith, K. (2016). Spontaneous lexical alignment in children with an autistic spectrum disorder and their typically developing peers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(11), 1821. https://doi.org/10.1037/xlm0000272

Buder, E. H., & Eriksson, A. (1997). *Prosodic cycles and interpersonal synchrony in American English and Swedish*. https://www.researchgate.net/publication/263965777

Chen, S.-Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease*, *9*(6), 1725. https://doi.org/10.21037/jtd.2017.05.34

Chen, X., Chen, J., Liao, M., & Wang, G. (2022). Early onset of impairments of interpersonal motor synchrony in preschool-aged children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-022-05472-8

Cheng, M., Kato, M., & Tseng, C. (2017). Gender and autistic traits modulate implicit motor synchrony. *PLOS ONE*, *12*(9), e0184083. https://doi.org/10.1371/journal.pone.0184083

Cienki, A., Bietti, L. M., & Kok, K. (2014). Multimodal alignment during collaborative remembering. *Memory Studies*, *7*(3), 354–369. https://doi.org/10.1177/1750698014530624

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale* (2nd ed.). Western Psychological Services.

Dale, R., Fusaroli, R., Håkonsson, D., Døjbak, D., Healey, P., Mønster, D., McGraw, J., Mitkidis, P., & Tylen, K. (2013). Beyond synchrony: Complementarity and asynchrony in joint action. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 35(35).

Danner, S. G., Krivokapić, J., & Byrd, D. (2021). Co-speech movement in conversational turn-taking. *Frontiers in Communication*, *6*, Article 779814. https://doi.org/10.3389/fcomm.2021.779814

Davis, C., Sironic, A., & Kim, J. (2006). Perceptual processing of audiovisual Lombard speech. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology* (pp. 248–252). University of Auckland.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85–104.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

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

Drijvers, L., & Özyürek, A. (2017). Visual context enhanced: The joint contribution of iconic gestures and visible speech to degraded speech comprehension. *Journal of Speech, Language, and Hearing Research*, *60*(1), 212–222. https://doi.org/10.1044/2016_JSLHR-H-16-0101

Drijvers, L., Özyürek, A., & Jensen, O. (2018). Hearing and seeing meaning in noise: Alpha, beta, and gamma oscillations predict gestural enhancement of degraded speech comprehension. *Human Brain Mapping*, *39*(5), 2075–2087. https://doi.org/10.1002/hbm.23987

Duran, N. D., Paxton, A., & Fusaroli, R. (2019). ALIGN: Analyzing linguistic interactions with generalizable techniques—A Python library. *Psychological Methods*, *24*(4), 419. https://doi.org/10.1037/met0000206

Edlund, J., Heldner, M., & Hirschberg, J. (2009). Pause and gap length in face-to-face interaction. In *10th Interspeech 2009 Conference* (pp. 2779–2782). ISCA - International Speech Communication Association.

Esteve-Gibert, N., Lœvenbruck, H., Dohen, M., & d'Imperio, M. (2022). Pre-schoolers use head gestures rather than prosodic cues to highlight important information in speech. *Developmental Science*, *25*(1), e13154. https://doi.org/10.1111/desc.13154

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

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157. https://doi.org/10.1016/j.newideapsych.2013.03.005

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, *40*(1), 145–171. https://doi.org/10.1111/cogs.12251

Galbusera, L., Finn, M. T. M., Tschacher, W., & Kyselo, M. (2019). Interpersonal synchrony feels good but impedes self-regulation of affect. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-50960-0

Garnier, M., Ménard, L., & Alexandre, B. (2018). Hyper-articulation in Lombard speech: An active communicative strategy to enhance visible speech cues? *The Journal of the Acoustical Society of America*, *144*(2), 1059–1074. https://doi.org/10.1121/1.5051321

Georgescu, A. L., Koeroglu, S., Hamilton, A. F. D. C., Vogeley, K., Falter-Wagner, C. M., & Tschacher, W. (2020). Reduced nonverbal interpersonal synchrony in autism spectrum disorder independent of partner diagnosis: a motion energy study. *Molecular Autism*, *11*(1), 11. https://doi.org/10.1186/s13229-019-0305-1

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in *R*: The dtw package. *Journal of Statistical Software*, *31*(7). https://doi.org/10.18637/jss.v031.i07

Groen, Y., Fuermaier, A. B. M., Den Heijer, A. E., Tucha, O., & Althaus, M. (2015). The empathy and systemizing quotient: The psychometric properties of the Dutch version and a review of the cross-cultural stability. *Journal of Autism and Developmental Disorders*, *45*(9), 2848–2864. https://doi.org/10.1007/s10803-015-2448-z

Hadley, L. V., & Ward, J. A. (2021). Synchrony as a measure of conversation difficulty: Movement coherence increases with background noise level and complexity in dyads and triads. *PLOS ONE*, *16*(10), e0258247. https://doi.org/10.1371/journal.pone.0258247

Healey, P. G. T., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLOS ONE*, *9*(6), e98598. https://doi.org/10.1371/journal.pone.0098598

Holler, J., Bavelas, J., Woods, J., Geiger, M., & Simons, L. (2022). Given-New Effects on the Duration of Gestures and of words in face-to-face dialogue. *Discourse processes*, *59*(8), 619–645. https://doi.org/10.1080/0163853X.2022.2107859

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*, 133–153.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Holtgraves, T. (2008). Conversation, speech acts, and memory. *Memory & Cognition*, *36*(2), 361–374.

Hove, M. J., & Risen, J. L. (2009, December 21). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949–960. https://doi.org/10.1521/soco.2009.27.6.949

Howes, C., Healey, P. G. T., & Purver, M. (2010). Tracking lexical and syntactic alignment in conversation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*(32), 2004–2009.

Hu, Y., Cheng, X., Pan, Y., & Hu, Y. (2022). The intrapersonal and interpersonal consequences of interpersonal synchrony. *Acta Psychologica*, *224*, 103513. https://doi.org/10.1016/j.actpsy.2022.103513

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15. https://doi.org/10.1016/j.wocn.2018.07.001

Kerem, E., Fishman, N., & Josselson, R. (2001). The experience of empathy in everyday relationships: Cognitive and affective elements. *Journal of Social & Personal Relationships*, *18*(5), 709–729. https://doi.org/10.1177/0265407501185008

Kimbara, I. (2006). On gestural mimicry. *Gesture*, *6*(1), 39–61.

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347. https://doi.org/10.1016/j.csl.2017.01.005

Kumazaki, H., Muramatsu, T., Yoshikawa, Y., Corbett, B. A., Matsumoto, Y., Higashida, H., Yuhi, T., Ishiguro, H., Mimura, M., & Kikuchi, M. (2019). Job interview training targeting nonverbal communication using an android robot for individuals with autism spectrum disorder. *Autism*, *23*(6), 1586–1595. https://doi.org/10.1177/1362361319827134

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the empathy quotient. *Psychological Medicine*, *34*(5), 911–920. https://doi.org/10.1017/S0033291703001624
Leary, M. R. (1983). A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin*, *9* (3), 371–375. https://doi.org/10.1177/0146167283093007
Lehnert-LeHouillier, H., Terrazas, S., & Sandoval, S. (2020). Prosodic entrainment in conversations of verbal children and teens on the autism spectrum. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00011
Lenth, R. (2019). *Emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.3.01* [R]. https://CRAN.R-project.org/package=emmeans
Levinson, S. C. (2017). *Speech acts* (Y. Huang, Ed.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199697960.013.22
Levitan, R., & Hirschberg, J. B. (2011). *Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions*. https://doi.org/10.7916/D8V12D8F
Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36*(8), 1404–1426. https://doi.org/10.1111/j.1551-6709.2012.01269.x
Manders, E., Koch, S. C., & Fuchs, T. (2022). A case for playful engagement: Synchrony and interaction quality during mirroring in ASD. Conceptual framework and case study. *American Journal of Dance Therapy*, *44*(2), 143–167. https://doi.org/10.1007/s10465-022-09359-0
McNaughton, K. A., & Redcay, E. (2020). Interpersonal synchrony in autism. *Current Psychiatry Reports*, *22*(3), 12. https://doi.org/10.1007/s11920-020-1135-8
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
Milton, D. E. M., Gurbuz, E., & López, B. (2022). The 'double empathy problem': Ten years on. *Autism*, *26*(8), 1901–1903. https://doi.org/10.1177/13623613221129123
Noel, J.-P., De Niear, M. A., Lazzara, N. S., & Wallace, M. T. (2018). Uncoupling between multisensory temporal function and nonverbal turn-taking in autism spectrum disorder. *IEEE Transactions on Cognitive and Developmental Systems*, *10*(4), 973–982. https://doi.org/10.1109/TCDS.2017.2778141
Nota, N., Trujillo, J. P., & Holler, J. (2021). Facial signals and social actions in multimodal face-to-face interaction. *Brain Sciences*, *11*(8), 1017 https://doi.org/10.3390/brainsci11081017
Oben, B., & Brône, G. (2016). Explaining interactive alignment: A multimodal and multifactorial account. *Journal of Pragmatics*, *104*, 32–51. https://doi.org/10.1016/j.pragma.2016.07.002
Ochi, K., Ono, N., Owada, K., Kojima, M., Kuroda, M., Sagayama, S., & Yamasue, H. (2019). Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PLOS ONE*, *14*(12), e0225377. https://doi.org/10.1371/journal.pone.0225377
Ostrand, R., & Chodroff, E. (2021). It's alignment all the way down, but not all the way up: Speakers align on some features but not others within a dialogue. *Journal of Phonetics*, *88*, 101074. https://doi.org/10.1016/j.wocn.2021.101074
Pérez, J. M., Gálvez, R. H., & Gravano, A. (2016). Disentrainment may be a positive thing: A novel measure of unsigned acoustic-prosodic synchrony, and its relation to speaker engagement. In *Interspeech* (pp. 1270–1274). International Speech Communication Association (ISCA).
Perugia, G., Paetzel, M., Castellano, G. (2020). On the role of personality and empathy in human-human, human-agent, and human-robot mimicry, and A. R. Wagner (Ed.), *Social Robotics. ICSR 2020. Lecture Notes in Computer Science* (Vol. 12483, pp. 120–131). Springer. https://doi.org/10.1007/978-3-030-62056-1_11
Pouw, W., de Wit, J., Bögels, S., Rasenberg, M., Milivojevic, B., & Ozyurek, A. (2021, July). Semantically related gestures move alike: Towards a distributional semantics of gesture kinematics. In V. G. Duffy (Ed.), *International Conference on Human-Computer Interaction* (pp. 269–287). Springer International Publishing.
Pouw, W., & Dixon, J. A. (2020). Gesture Networks: Introducing dynamic time warping and network analysis for the kinematic study of gesture ensembles. *Discourse Processes*, *57*(4), 301–319. https://doi.org/10.1080/0163853X.2019.1678967
Qiao, S., Wang, Y., & Li, J. (2017). Real-time human gesture grading based on OpenPose. In Q. Li, L. Wang, M. Zhou, L. Sun, S. Qiu, & H. Liu (Eds.), *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics* (CISP–BMEI) (pp. 1–6). https://doi.org/10.1109/CISP-BMEI.2017.8301910
Quintero, J., Morales, I., Vera, R., Zuluaga, P., & Fernández, A. (2019). The impact of adult ADHD on the quality of life profile. *Journal of Attention Disorders*, *23*(9), 1007–1016. https://doi.org/10.1177/1087054717733046
Rabinowitch, T.-C., & Meltzoff, A. N. (2017). Synchronized movement experience enhances peer cooperation in preschool children. *Journal of Experimental Child Psychology*, *160*, 21–32. https://doi.org/10.1016/j.jecp.2017.03.001
Rasenberg, M., Dingemanse, M., & Özyürek, A. (2020). Lexical and gestural alignment in interaction and the emergence of novel shared symbols. In A. Ravignani & C. Barbieri (Eds.), *The Evolution of Language Conferences* (p. 3). University of Zurich.
Rasenberg, M., Özyürek, A., Bögels, S., & Dingemanse, M. (2022). The primacy of multimodal alignment in converging on shared symbols for novel referents. *Discourse Processes*, *59*(3), 209–236. https://doi.org/10.1080/0163853X.2021.1992235

Rasenberg, M., Özyürek, A., & Dingemanse, M. (2020). Alignment in Multimodal Interaction: An Integrative Framewor. *Cognitive Science*, *44*(11), e12911. https://doi.org/10.1111/cogs.12911
Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony. *Zeitschrift Für Psychologie*, *224* (3), 168–189. https://doi.org/10.1027/2151-2604/a000252
Revelle, W. (2020). *psych: Procedures for personality and psychological research (2.0.12)* [R]. Northwestern University.
Romero, V., Fitzpatrick, P., Roulier, S., Duncan, A., Richardson, M. J., & Schmidt, R. C. (2018). Evidence of embodied social competence during conversation in high functioning children with autism spectrum disorder. *PLOS ONE*, *13* (3), e0193906. https://doi.org/10.1371/journal.pone.0193906
Rotheram-Fuller, E., Kasari, C., Chamberlain, B., & Locke, J. (2010). Social involvement of children with autism spectrum disorders in elementary school classrooms. *Journal of Child Psychology and Psychiatry*, *51*(11), 1227–1234.
Rubin, M. (2021). When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing. *Synthese*, *199*, 10969–11000. https://doi.org/10.1007/s11229-021-03276-4
Rymarczyk, K., Żurawski, Ł., Jankowiak-Siuda, K., & Szatkowska, I. (2019). Empathy in facial mimicry of fear and disgust: Simultaneous EMG-fMRI recordings during observation of static and dynamic facial expressions. *Frontiers in Psychology*, *10*, 701. https://doi.org/10.3389/fpsyg.2019.00701
Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis I*. Cambridge university press.
Schilbach, L. (2016). Towards a second-person neuropsychiatry. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150081.
Schmidt, R. C., Morr, S., Fitzpatrick, P., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior*, *36*, 263–279. https://doi.org/10.1007/s10919-012-0138-5
Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge university press.
Singer, J. (2016). *Neurodiversity: The birth of an idea*. Self-published, Amazon.
Slocombe, K. E., Alvarez, I., Branigan, H. P., Jellema, T., Burnett, H. G., Fischer, A., Li, Y. H., Garrod, S., & Levita, L. (2013). Linguistic alignment in adults with and without Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *43*(6), 1423–1436. https://doi.org/10.1007/s10803-012-1698-2
Stivers, T., & Enfield, N. J. (2010). A coding scheme for question–response sequences in conversation. *Journal of Pragmatics*, *42*(10), 2620–2626. https://doi.org/10.1016/j.pragma.2010.04.002
Strickland, D. C., Coles, C. D., & Southern, L. B. (2013). JobTIPS: A transition to employment program for individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(10), 2472–2483. https://doi.org/10.1007/s10803-013-1800-4
Syrdal, A. K., & Kim, Y.-J. (2008). Dialog Speech Acts and Prosody: Considerations for TTS. In *Speech Prosody. In Proceedings of Speech Prosody* (pp. 661–665). ISCA.
Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, *11*(10), 20150767. https://doi.org/10.1098/rsbl.2015.0767
Taylor, J. L., & Seltzer, M. M. (2011). Employment and post-secondary educational activities for young adults with autism spectrum disorders during the transition to adulthood. *Journal of Autism and Developmental Disorders*, *41*(5), 566–574. https://doi.org/10.1007/s10803-010-1070-3
Ter Bekke, M., Drijvers, L., & Holler, J. (2024). Hand gestures have predictive potential during conversation: An investigation of the timing of gestures in relation to speech. *Cognitive Science*, *48*(1), e13407. https://doi.org/10.1111/cogs.13407
Trujillo, J. P., Dideriksen, C., Tylén, K., Christiansen, M. H., & Fusaroli, R. (2023). The dynamic interplay of kinetic and linguistic coordination in Danish and Norwegian conversation. *Cognitive Science*, *47*(6), e13298. https://doi.org/10.1111/cogs.13298
Trujillo, J. P., & Holler, J. (2021). The kinematics of social action: Visual signals provide cues for what interlocutors do in conversation. *Brain Sciences*, *11*(8), 996. https://doi.org/10.3390/brainsci11080996
Trujillo, J. P., Levinson, S. C., & Holler, J. (2022). A multi-scale investigation of the human communication system's response to visual disruption. *Royal Society Open Science*, *9*(4). https://doi.org/10.1098/rsos.211489
Trujillo, J. P., Özyürek, A., Holler, J., & Drijvers, L. (2021). Speakers exhibit a multimodal Lombard effect in noise. *Scientific Reports*, *11*(1), 16721. https://doi.org/10.1038/s41598-021-95791-0
Vacharkulksemsuk, T., & Fredrickson, B. L. (2012). Strangers in sync: Achieving embodied rapport through shared movements. *Journal of Experimental Social Psychology*, *48*(1), 399–402. https://doi.org/10.1016/j.jesp.2011.07.015
Wadge, H., Brewer, R., Bird, G., Toni, I., & Stolk, A. (2019). Communicative misalignment in autism spectrum disorder. *Cortex*, *115*, 15–26. https://doi.org/10.1016/j.cortex.2019.01.003
Wallot, S., Mitkidis, P., McGraw, J. J., & Roepstorff, A. (2016). Beyond synchrony: Joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PLOS ONE*, *11*(12), e0168306.
Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, *33*(4), 448. https://doi.org/10.1037/h0027806
Wehmeier, P. M., Schacht, A., & Barkley, R. A. (2010). Social and emotional impairment in children and adolescents with ADHD and the impact on quality of life. *Journal of Adolescent Health*, *46*(3), 209–217. https://doi.org/10.1016/j.jadohealth.2009.09.009

Williams, G. L., Wharton, T., & Jagoe, C. (2021). Mutual (mis)understanding: Reframing autistic pragmatic "impairments" using relevance theory. *Frontiers in Psychology*, *12*(616664), 1–21. https://doi.org/10.3389/fpsyg.2021.616664

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN : A Professional Framework for Multimodality Research. In N. Calzolari, K. Choukri, A. Gagemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias, D. (Eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation* (pp. 1556–1559). European Language Resources Association (ELRA). https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_60436

Wynn, C. J., & Borrie, S. A. (2022). Classifying conversational entrainment of speech behavior: An expanded framework and review. *Journal of Phonetics*, *94*, 101173. https://doi.org/10.1016/j.wocn.2022.101173

Zampella, C. J., Bennetto, L., & Herrington, J. D. (2020). Computer vision analysis of reduced interpersonal affect coordination in youth with autism spectrum disorder. *Autism Research: Official Journal of the International Society for Autism Research*, *13*(12), 2133–2142. https://doi.org/10.1002/aur.2334