scientific reports

Check for updates

OPEN

(Mis)align: a simple dynamic framework for modeling interpersonal coordination

Grace Qiyuan Miao [1]✉, Rick Dale [1] & Alexia Galati [2]

As people coordinate in daily interactions, they engage in different patterns of behavior to achieve successful outcomes. This includes both synchrony—the temporal coordination of the same behaviors at the same time—and complementarity—the coordination of the same or different behaviors that may occur at different relative times. Using computational methods, we develop a simple framework to describe the interpersonal dynamics of behavioral synchrony and complementarity over time, and explore their task-dependence. A key feature of this framework is the inclusion of a task context that mediates interactions, and consists of active, inactive, and inhibitory constraints on communication. Initial simulation results show that these task constraints can be a robust predictor of simulated agents' behaviors over time. We also show that the framework can reproduce some general patterns observed in human interaction data. We describe preliminary theoretical implications from these results, and relate them to broader proposals of synergistic self-organization in communication.

People coordinate with each other to perform joint tasks in a wide variety of circumstances. Such coordination involves individuals performing synchronized and complementary behaviors to achieve successful outcomes. Researchers of interpersonal coordination have used a variety of terms to describe patterns associated with behavioral synchrony: alignment[1], entrainment[2], mimicry[3], accommodation[4], and contagion[5], among other terms (see[6] for a systematic review). These terms capture phenomena where individuals exhibit the same behavior at about the same time. Beyond this interpersonal synchrony, researchers have identified other coordination patterns that allow for more complementarity in how individuals organize their behavior over time. Terms like synergy, coupling, misalignment, or complementarity highlight that individuals can coordinate without exhibiting the identical behaviors[7–10]. Indeed, these two patterns can be accommodated in many everyday collaborative situations. For example, when searching for a missing item in a room, two people could both look at the bed (i.e. synchronized joint attention), or one could look at the bed while the other at the table (i.e. complementary visual search). These behavioral patterns are intuitive and pervasive. How do people balance synchronized and complementary behaviors during collaboration, and what are the underlying cognitive processes that make them possible? These questions pertain to theories of social interaction and have implications for how we optimize collaboration and task performance[11–13].

One possibility is that synchrony and complementarity are supported by cognitive control processes. Consider two interlocutors engaging in friendly turn-taking, where one speaks and the other listens (i.e. complementarity). Inhibition in this case (e.g., when one partner suppresses what they want to say until the other finishes) could be important for minimizing the interference of information that would not advance the joint task goals. Indeed, there is evidence that, in situations involving multiple perspectives and potential ambiguity, language users need to actively inhibit their own perspective to process language appropriately[14]. These cognitive functions, central to interpersonal synchrony, could also be used to identify leader–follower roles in human communication[15], where the leader initiates more actions and the follower might inhibit actions to allow smooth turn-taking and consequently successful collaboration.

One way to investigate these questions about cognitive processes is through computational simulation[16,17]. Computational models represent a simplification of a real-world system, allowing researchers to test experimental treatments or what-if scenarios that are practically infeasible in real-life settings. Moreover, simulated data complement empirical data collected in laboratories, granting researchers greater power to draw theoretical implications ([16,18–20] for discussions of how modeling supports theory development).

[1]Department of Communication, University of California, Los Angeles, CA, USA, , . [2]Department of Psychological Science, University of North Carolina at Charlotte, Charlotte, USA. ✉email: q.miao@ucla.edu

nature portfolio

One open question is *how* humans determine when to synchronize with or complement each other, particularly in tasks that permit both options. There is ongoing debate about the prevalence of alignment in conversation[21], its reliance on priming as a mechanism[1,22], the default status of egocentric processes during alignment[23–27], and its benefits to joint performance. Some studies show negative effects of excessive alignment on joint performance[28,29], others highlight its importance[30,31–33], and yet others suggest that alignment is highly dynamic and complex, so its benefits are dependent on context[28,34]. Although computational models have been developed for joint action, most are related to motor control[35], and few account for the emergence of partners' behavior in conversation (see[36] for review). Proposals of computational models of coordination in conversation are often schematic and have not yet been implemented and tested (e.g.,[37,38]).

In the present work, we develop a simplified framework to describe dynamics of synchrony and complementarity in dyads and explore their task-dependence. The framework is deliberately simple, using coupled linear equations that could implement behavioral dynamics akin to coupled oscillators[37] yet accounts for additional interactional features. This permits potentially integrating hypotheses about cognitive and social constraints. In the version presented here, we show that this modeling framework could offer insight into the emergence of synchrony and complementarity, especially by including parameters related to the task and context of interaction. In the next section, we summarize our modeling approach, and compare it to prior computational strategies. Following this, we give a formal specification of the model, consisting of two agents coupled by an interactive context. We then simulate thousands of "conversations" by testing all possible parameters of the model, and discuss their theoretical implications.

Modeling approach

Recently, computational investigation of human communication and related behavior has often used machine learning and related methods[40]. The power of these methods derives from their ability to take in huge amounts of data and train a model from scratch. These methods produce models that can sometimes generate surprisingly human-like behavior[41]. Despite these strengths, the inner workings of many-layered deep learning neural networks can be difficult to interpret[42].

Another approach to computational investigation is to use models that are derived from "first principles." These are models in which modeling choices are theoretically inspired so that parameters can be interpreted transparently in their development and resulting behavior. Starting from first principles, with theoretically-motivated equations, means that there are relatively fewer parameters plugged in directly to the model, and can be interpreted relatively transparently (see also[39] for a recent model on turn-taking).

Interpersonal coordination is a domain in which computational models often draw from first principles in this way, allowing for greater transparency and interpretability (see[43] for review). For example, some models that follow this approach specify dynamical systems equations, with parameters carefully chosen to represent particular aspects of interpersonal coordination. Many such models have been developed to account for joint action in terms of motor control, in tasks such as finger movements[44–46] or using hand movements to shepherd target objects[35] or to avoid collision[47].

Other models of interpersonal coordination transparently capture dynamics in activities that extend beyond motor coordination, including conversation. These models typically focus on activities where agents aim to achieve synchrony or convergence. One prominent approach uses the theory of coupled oscillators as a mathematical foundation, defining a system that executes periodic behavior[48]. A key characteristic of the coupled oscillator model is that the system will always achieve synchronization eventually, whether or not the oscillators are identical[48]. Models using this approach have been developed for various contexts, including psychotherapy[49], improvisation or musical coordination[50,51], and romantic couples discussing health behaviors[52]. In these types of interpersonal activities, the goal often involves synchronization. For instance, music ensembles aim to play notes at synchronized beats, while romantic couples or therapist-client dyads strive to reach consensus by the end of their conversations. The nature of these activities—centered on convergence—makes coupled-oscillator models suitable for abstracting formalisms of behavior.

However, human interaction does not depend solely on synchrony[47,53]. There is a need to create models representing more diverse and generalizable patterns of human interactions, which encompass both synchrony and complementarity in various contexts. One example of this effort is a simple dynamic system of coupled agents programmed to "take turns" using parameters that govern their dynamics[54]. In another model exploring marital relationships, researchers implemented "emotion" states and parameters that dictate how married couples influence each other's emotions[55]. Yet, these models often remain abstract, with behavior not directly represented (e.g., Buder[54] interprets a coupled dynamical system *as* interaction).

With the model we present here, we pursue a theoretical framework in which we specify some basic properties of agents' interaction and then examine their coupled dynamics. Importantly, our approach facilitates the integration of the agents' "task" in a manner that is flexibly defined, enabling it to account for the emergence of synchrony and complementarity in different contexts. We identify a set of components comprising an interaction, including the task context, the behaviors of two agents, and how they influence each other. The parameters of these components can be specified manually and simulated under various conditions, resulting in a comprehensive exploration of how and why two simulated agents become aligned or complementary according to specific principles. Importantly, this computational investigation can inform how the task itself may govern the emergence of these behavioral processes. This framework contrasts with prior computational efforts that have focused on model development in a way that centers on the task. In prior coupled-oscillator models, for example, there typically exists a conception of two or more systems that are directly coupled in an interaction, shaping their mutual dynamics as a function of time. We don't discount this coupling as a key feature of interaction, but our model centers the task as a potential source of constraint on coupled dynamics. A theoretical implication of

this focus on the task is that interaction partners may not need to retain all details of an interaction to support their joint performance. Instead, the task environment can influence the agents' dynamics without the agents incurring high cognitive demands.

Beyond devising a computational framework to explore alignment and complementarity, a second motivation of this approach is to bridge the model with human data from real tasks eventually. This integration with human data is facilitated by having a model whose inner-workings are relatively easy to understand and modify. The model can be employed to predict how various interactive tasks with human participants yield distinct signatures of alignment and complementarity. After presenting our modeling framework, we illustrate two ways in which the model can capture data collected in experimental human tasks. In our concluding discussion, we also address how bridging the model to human data can be explored in future model development.

Modeling framework

Our initial approach is to draw from discrete dynamical systems models that permit direct specification of parameters and behaviors[56]. We model a single behavior for two interacting partners, and represent these behaviors as a two-dimensional vector that is updated iteratively over time (one dimension per simulated agent). We take their behavior as numeric descriptions, scalars, which could describe many quantitatively measured behaviors in real interaction (e.g., eye movements, body motion, etc.). Specifically, the scalar number could represent the attentiveness of an agent at a specific time point, their rate of speaking, or their arm movements. The model is deliberately abstract in this regard, allowing for its application across a wide variety of operationalizations in human tasks.

Taking B to be this vector $(b_{\text{Person 1}}, b_{\text{Person 2}})$ at time t, we have:

$$B(t) = C \cdot I \cdot B(t-1) + U(-0.5, 0.5) - \alpha B(t-1).$$

The behavior of two people $B(t)$ is a function of a given context C and power of influence I. We take this C parameter to be a "context matrix", because it transforms two-person behaviors based on a prior time step ($t-1$), with added noise (U). In this iteration, I is set as constant ($I = 1$), which represents 100% receptivity of two people towards all signals. U is set in range ($-0.5, 0.5$), which represents a uniform source of noise per turn, from -0.5 to 0.5. The final subtractive term is a decay term with scalar α, generally a number between 0 and 1, which we specify as 0.1. This decay term proportionally reduces the behavioral signals by 10% to ensure that the model does not saturate and veer towards positive or negative infinity.

The task context C is numerically described as a 2×2 matrix whose values can be 0, 1, or -1, representing inactive, active, or inhibitory constraints on communication. For example, during a presentation, the speaker is actively speaking, while the audience is inactive in speaking. If an audience member has a question in mind but inhibits the act of asking questions immediately to maintain the presentation flow, they are practicing inhibition. This C matrix specifies how much individuals influence each other, how much they are autocorrelated (influence themselves), and can vary across *specific* tasks that individuals face. The 2×2 matrix consisting of elements S_1, O_1, O_2 and S_2 can be interpreted in the following way:

$C = (S_1$ how much *Person 1* influences self, O_1 how much *Person 2* influences *Person 1*; O_2 how much *Person 1* influences *Person 2,* S_2 how much *Person 2* influences self $)$.

Figure 1 further illustrates the interpersonal influences constructed in C. A consequence of our modeling approach is that the configuration of values in the context matrix C represents task constraints and reflects different interaction types. For example, if C is "null" with $C = (0, 0; 0, 0)$, the model generates agent behaviors that are random with no meaningful structure or interaction, and $C = (1, 1; 1, 1)$ represents a perfectly synchronizing interaction, such as when two people are singing a song in unison. When $C = (1, 0; 1, 0)$ or $(0, 1; 0, 1)$, there is a role asymmetry, with one person actively influencing the other's (and their own) behavior. This would be relevant to educational scenarios, such as lectures, where one partner primarily drives the interaction. We can also implement *inhibitory* constraints, such as $(1, 0; 1, -1)$, which could again represent a lecture scenario, but this time, the
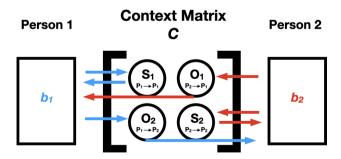

Figure 1. How context matrix C mediates interactions between two agents (blue, red are Person 1 and Person 2) and their resulting behaviors (b_1 and b_2, for Person 1 and 2 respectively). S_1 represents how much *Person 1* influences self and S_2 how much *Person 2* influences self. O_1 represents how much *Person 2* influences *Person 1* and O_2 how much *Person 1* influences *Person 2*. Arrows represent the direction of influence, and the color of the arrows corresponds to the person initiating the action (blue for Person 1-initiated action and red for Person 2-initiated action).

listener has a question in mind and inhibits from asking the question so that the speaker can maintain the communication flow (see Fig. 2.iii for example runs of this simulation; see also Fig. 4 for a zoomed-in illustration).

Simulation methods

We implemented a simulation under this framework to explore the impact of *C* on the behavior of the two coupled agents. We aimed to investigate *all* potential variations of the *C* matrix. Exploring all possible combinations of the parameters of *C* will allow us to isolate what governs the emergence of synchrony or complementarity in this model.

Independent variable: *C*

Our model of a two-agent (Person 1 and Person 2) system permitted us to codify dyadic interactions using context matrix *C*, which represents four possible combinations of influence: (i) Person 1 to Person 1, (ii) Person 2 to Person 1, (iii) Person 1 to Person 2, and (iv) Person 2 to Person 2. Recall that each of these four directions of influence has three possible states—active, inactive, and inhibitory—represented by 1, 0, and −1. With these numeric codes, we are able to characterize different tasks or contexts in human interaction using the context matrix *C*.

Dependent variable: *r*

To account for the relational nature of behaviors, Pearson's *r*, or correlation coefficient, is often used as the dependent variable when investigating interpersonal synchrony[59]. In addition to being one of the most common metrics used to describe interpersonal synchrony, *r* also allows clear and familiar interpretation of the correlation between two signals. If *r* is strongly positive, it would indicate two agents that are synchronizing in whatever behavior they exhibited in the "conversation"; if *r* is strongly negative, it would suggest a kind of complementarity—doing the opposing behavior of the other (akin to taking turns in speech, or taking on separate complementary roles in an interaction). Since *r* is a statistic describing how two agents correlate with each other in interactions, the meaning of "complementarity" could have multiple interpretations in different contexts. For example, it could represent the "division of labor" in collaboration, where agents may engage in distinct, concurrent behaviors throughout the task. Complementarity could also signify "turn-taking" in conversations, where agents engage in distinct behaviors simultaneously (such as one being silent while the other talks), which alternate between agents over time.
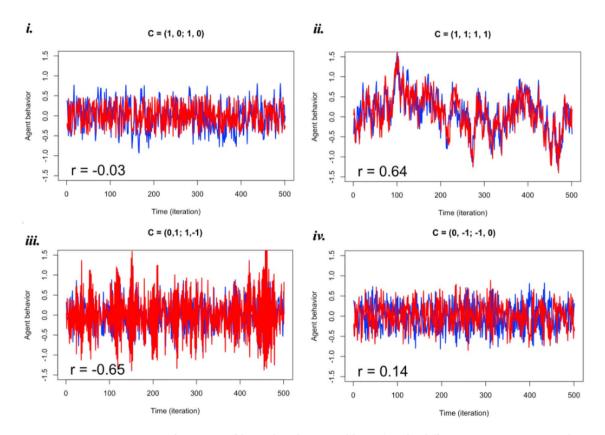

Figure 2. Time series of two agents (blue, red are $b_{Person\,1}$ and $b_{Person\,2}$) under different context matrices. i. When Person 1 is active but Person 2 is inactive, they have no interdependence. ii. Total coordination between two agents (such as singing a song together) produces close synchrony. iii. In a leading/following scenario (such as teaching), blue precedes red and correlates in their simulated behaviors. (Fig. 4 shows a zoomed in segment of these time series). iv. With inhibitory parameters, a weaker correlation (calculated using 0-lag cross-correlation) is exhibited by the simulated dyad, and their behavior fluctuates between states.

In the simulations discussed below, we approach interpersonal synchrony and complementarity by directly modeling the correlation coefficient (r) between two agents. We use the simulated correlation coefficient (r) as the dependent variable, we then perform regression analysis to further evaluate and interpret this computational quantification of interpersonal synchrony. Finally, we apply the model to simulate human data from prior studies[32,60] using two different measures of interpersonal coordination, cross-correlation and turn-taking lag. We show that our model's C matrix produces distinct trends that resemble the patterns of synchrony and complementarity documented in these studies. By directly modeling the statistics that other human studies use to quantify synchrony, we bridge computational simulation and empirical quantification as methodologies for exploring behavioral dynamics.

Simulation procedure

Using the model presented above, we obtained simulation data with the following steps:

Step 1 We obtained all possible combinations of C. With 4 parameters and 3 possible values (-1, 0, 1) for each, there are $3^4 = 81$ combinations of the C matrix. These combinations represent all possible interactive scenarios involving the two agents. We generated all such permutations of C in rounds of our simulation.

Step 2 We obtained simulation data and calculated the dependent variable r. We ran 100 simulations under random starting conditions for each of the 81 possible C matrices, resulting in 8100 simulations. For each simulation, one of the agents is chosen to "start the conversation," and displays its behavior (element of B) to the other; the model then operates with its matrix C, modifies agent behaviors, and the next agent's turn commences. This is done for 500 turns for each simulation, and we record each sequence of behavior separately. These behaviors represent two time series, and after each such simulated "conversation," the correlation coefficient r between these behavioral sequences is recorded. As noted above, this coefficient r represents relative synchrony or complementarity and is utilized as the dependent variable for later analysis.

Step 3 We then conducted a regression analysis, using C to predict r. This regression model takes values in C to predict the resulting r across the 8100 simulated interactions. Specifically, we performed multicategorical analysis and built various regression models that include different subsets of C in order to isolate the strongest contributors to determining r. Multicategorical analysis is commonly used to unpack all aspects/categories of a variable with multiple features, and regression models are used to understand how each aspect of the independent variable explains/predicts the dependent variable. In this case, components of C include S_1, O_1, O_2, and S_2. These regressions are simple linear models that account for the simulation and help us understand how different context matrices C, or different combinations of active, inactive, and inhibitory states of communication influence the relative dynamics between the two agents. The regression analysis helps us understand what drives this theoretical model, offering answers to some key questions: Is it possible for a simple regression model to capture complex interpersonal dynamics from the task matrix C? How many parameters and what kinds do we need to make this prediction? For example, is it possible that the self-related parts of C—how agents influence themselves—are sufficient to predict r across simulated interactions?

All code and data necessary to reproduce Steps 1 to 3 are found at the following GitHub repository: https://github.com/miaoqy0729/sim-syn-sims.

Results

Trends in *C* for predicting *r*

As noted, we performed 100 simulations for each of the 81 possible configurations of the task context C, involving active, inactive, or inhibitory states for each of the 4 possible directional influences within and across interacting partners. Illustrations of the output of model simulations are shown in Fig. 2, including when agents have no or weak interdependence (Fig. 2.i and Fig. 2.iv), when two agents perform optimal synchrony and yield high interdependence (Fig. 2.ii), and when one agent leads the interaction and the dyad yields complementary interdependence (Fig. 2.iii; see also Fig. 4).

The model exhibited a range of behaviors. The distribution of r across the 8,100 simulated "interactions" is shown in Fig. 3. Sometimes agents show perfect synchrony, with $r \sim 1$; agents can also exhibit extreme complementarity, $r \sim -1$. Since we were interested in how the context matrix may determine these behaviors, we inspected what characteristics of C seemed to accompany extreme values in r.

We considered the negative correlations in the simulated results, which represent complementarity—complementary actions performed at the same time when engaging in an interpersonal activity. For example, Fig. 4 illustrates a leading/following scenario where Person 1's behavior (blue) precedes Person 2's behavior (red), while being highly complementary to each other ($r = -0.65$).

Across the 8100 simulations, we examined when models exhibited a strong negative correlation in agent behavior akin to complementarity. We defined the threshold as $r < -0.25$, as this captures the left side of the observed distribution of r outside of its mean (see Fig. 3). In these cases, the task matrix (C) contained at least one -1 parameter 99.8% of the time. To test for statistical significance of this observation, we conducted a chi-square test on uniform probability, yielding a significant result ($\chi^2 = 1317.00$, $p < 0.00001$). We also conducted an additional test using probabilities of the corresponding positive correlations in the simulations as baseline (using a $r > 0.25$ threshold). The observed rate of negative task parameters was 83.3%, which is significantly different from that in the negative-r threshold ($\chi^2 = 259.98$, $p < 0.00001$). This suggests that in order for there to be a robust complementary pattern (low, negative r), the combination of constraints in C must include at least one negative

Histogram of Simulated r

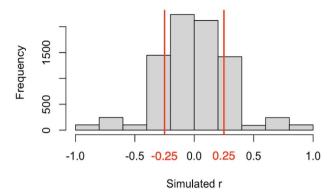**Figure 3.** Frequency distribution of correlation coefficient (r) obtained from the 8100 simulations. Red lines mark the thresholds ($r < -0.25$ for complementarity and $r > 0.25$ for synchrony) for statistical tests.
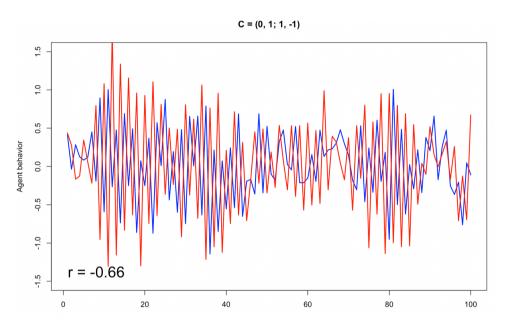

Figure 4. Zoomed in time series of Fig. 2.iii. In a leading/following scenario (such as teaching), blue precedes red and correlates in their simulated behaviors.

value almost all the time. It is important to note that the structural features of the model may create certain consequences. For example, constraints on some models might result in patterns like "turn-taking" simply by virtue of their underlying mathematical formulation, such as in predator–prey population levels modeled with difference equations[56]. A similar constraint may be present in our formulation as well. Nevertheless, our result supports the broader hypothesis, discussed in the introduction, that inhibition is a critical aspect of cognitive processing central to successful coordination. In fact, there are almost no conditions of C that can yield complementarity dynamics without inducing inhibitory constraints. To delve into the subtleties of this result, we performed further analyses using multiple regression.

Predicting r from C with multiple regression

The inhibitory effect we documented suggests that the model's formulation produces constraints that can predictably result in specific values of r. In the most pronounced case we observed, extreme negative correlations appeared to require inhibition. Consequently, we sought to understand how specific values of C contribute to specific behavioral patterns, such as which values predicted synchrony and which predicted complementarity. To answer these questions, we performed a multicategorical analysis on the simulated data described earlier.

Specifically, all four parameters of the C matrix (S_1, O_1, O_2, S_2) were dummy coded with value 0 as the reference group, resulting in 8 parameters (S_1p, S_1n, O_1p, O_1n, O_2p, O_2n, S_2p, S_2n) representing each
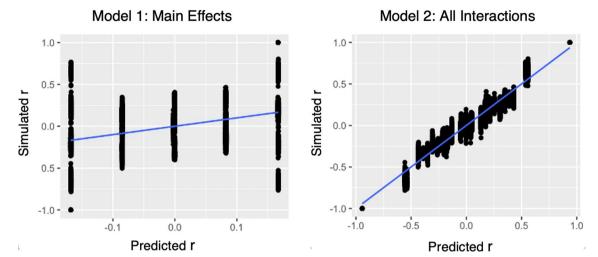
Figure 5. Predicted results of Model 1 (main effects model) and Model 2 (all interactions model) (x-axis) and *r* value between two agents (y-axis). Model 1 is not able to make fine-grained predictions, while Model 2 predicts *r* almost perfectly. Further context for Model 1's failure to capture the trends in *r* comes from the AIC and BIC values in Table 1: they indicate that all models with interactions are more efficient and should be selected over the main effects model. In particular, the AIC and BIC both suggest that Model 2 is a more efficient model in terms of complexity and data fit.

Model	Number of parameters	R^2	Adjusted-R^2	AIC	BIC
(1) Main effects	8	0.101	0.1001	2851.535	2921.532
(2) Interactions	48	0.9451	0.9449	− 19,743.45	− 19,505.47
(3) Overall	56	0.9451	0.9449	− 19,743.45	− 19,505.47
(4) Initiator-focused	28	0.9084	0.9081	− 15,606.81	− 15,396.82
(5) Self-influence	28	0.9405	0.9403	− 19,102.45	− 18,892.46

Table 1. Information of various regression models. Model (1) includes all main effect terms (S_1p, S_1n, O_1p, O_1n, O_2p, O_2n, S_2p, S_2n) as regressors. Model (2) includes all the interaction terms (for example, S_1p * (O_1p + O_2p + O_2p + O_1n + O_2n + S_2n) contains all interaction terms for S_1p) for each of the 8 main effect terms. Model (3) includes all of the regressors in Model (1) and Model (2). Model (4) focuses on Person 1 as the initiator, and include regressors on how Person 1 self-influences and how Person 1 influences Person 2. Model (5) includes Person 1's self-influence and Person 2's self-influence. In other words, how agents auto-correlate with themselves, or how Person 1's past behavior predicts Person 1's future behavior.

C matrix ("p" and "n" stand for "positive" and "negative"). With the positive and negative values specified for each of the four parameters, regression analysis can reveal whether the valence of three different states of communication plays impacts the communication flow. In addition, to better explore the components that account for correlations of dyadic behavior in the simulated data, linear interactions across all possible dummy-coded parts of the model are calculated.

We focus on three regression models containing different predictors that account for the dyadic correlation: (1) a model that includes the main effects of the 8 dummy coded parameters in the C matrix (S_1p, S_1n, O_1p, O_1n, O_2p, O_2n, S_2p, S_2n), (2) a model that includes the possible interactions among different parameters (8*6 = 48 interactions), and (3) an overall model consisting of all predictors (8 + 48 = 56 predictors).

To evaluate the amount of variability in the simulated data the regression model could explain, we considered the R^2 values of the three regression models. The main effects model had an R^2 value of 0.101, accounting for 10% of variability in the dependent variable *r*. The interaction model had an R^2 value of 0.945, thus accounting for 94.5% of the variation in *r*. The overall model with 56 predictors had an identical R^2 value as the interaction model by itself. These R^2 values suggest that when predicting synchrony between two agents, optimal results could be achieved solely by the interaction parameters. As illustrated in Fig. 5, the regression model with only main effect terms failed to yield fine-grained predictions, whereas the model with interaction terms predicted *r* almost perfectly. In other words, the key features of how the task constrains two agents *relative to each other* can alone predict the agents' coordination (*r*) almost perfectly.

We further examined whether an R^2 value greater than 0.9 can be obtained using subsets of the predictors of the overall model, instead of all 56. We found that more than 90% of variability could be accounted for with fewer than half of the 56 predictors. For example, the R^2 of the model (4) that accounts for how Person 1 influences

themselves (S_1p, S_1n) and Person 2 (O_2p, O_2n) is 0.91 (see Table 1). And the R^2 of the model (5) that accounts for how Person 1 influences themselves (S_1p, S_1n) and how Person 2 influences themselves (S_2p, S_2n) is 0.94. This means more than 90% of the variability in the dyadic system could be estimated by focusing on a specific part of the dyadic system—either focusing on the initiator of the time series or on the agents' auto-correlation with themselves. We explain further our rationale for selecting the predictors of models (4) and (5) below, and evaluate their efficiency.

To determine the optimal model that accounts for an adequate amount of variability while using a minimal number of predictors[57,58], we calculated the two popular model selection criteria—Akaike information criterion (AIC) and Bayesian information criterion (BIC) —for each regression model. In general, the *lower* AIC and BIC for a given model, the more effective the model is at balancing complexity and variance captured of the observed data—namely, a more efficient model.

The lowest AIC and BIC values came from the models with interactions only or with the full model. These two models had the same information score (AIC = − 19,743.45, BIC = − 19,505.47, see Table 1). Because the inter-action-only model has fewer terms compared to the overall model, we can regard it as more efficient. The least effective model was the one with main effect terms on their own, which had the highest score on AIC and BIC.

In addition to testing models based on their statistical properties (namely, main effects or interactions), we used subsets of the 56 predictors to devise models to answer specific questions of interests: How much variability in the overall distribution can we account for by focusing on the influences of the initiator? Or by focusing on the self-influence within each agent?

To answer the first question, we focused on Person 1 as the action initiator in Model (4), which involves S_1 (how much *Person 1* influences self) and O_2 (how much *Person 1* influences *Person 2*). The parameters in Model (4) include s_1p, s_1n, o_2p, o_2n and all the interaction terms involved. To answer the second question about the role of self-influence, we considered model (5) focusing on S_1 (how much *Person 1* influences self) and S_2 (how much *Person 2* influences self). The parameters in Model (5) include s_1p, s_1n, s_2n, s_2n and all their relevant interaction terms. With an equal number of parameters, the self-influence model (see Table 1 for model 5) has noticeably better AIC and BIC results than the initiator-focused model (see Table 1 for model 4).

Despite being a highly simplified exploration of these effects, the framework offers a foundation for under-standing how behavioral dynamics between partners may be accounted for with task constraints. These regression results suggest that relatively efficient information from a task (the task matrix C) can be used to predict how interlocutors in the simulation will synchronize under that task.

Comparison to human tasks

So far, we have demonstrated that our model can capture essential features of how the task can mediate behav-ior among agents, and that the task alone can predict the aggregate dynamics exhibited by the agents. In this section, we show how this framework could bridge to properties of human experimental data, providing two specific illustrations.

In the first illustration, we use cross-correlation in a manner related to prior interpersonal dynamics[32,59], and show that the model's C matrix can moderate this signal in interesting ways. Results indicate that while the C matrix generates the common observation of a peaked cross-correlation function, this function is modulated depending on the task setup of C.

In the second illustration, we show that turn-taking dynamics can also be captured with this model, with C again playing a modulatory role. As Stivers and colleagues[60] have shown, turn transitions can be quite rapid, across different cultures. Together, these illustrations show how our simple framework can produce nuanced patterns resembling those observed in human tasks. They also reveal instances where the simulation *diverges* from what would be expected or possible in the human scenarios.

Cross-correlation function

In the first case, we examined several example C matrices that demonstrate various patterns including synchrony, anti-synchrony, and leading and following, all of which have been observed in human dynamic data[38,47,53]. When neither agent is coupled to the other (O_1, $O_2 = 0$), the cross-correlation function appears flat (Fig. 6, green). However, a specific C matrix that induces leading and following, in that the cross-correlation function is shifted, further emphasizes the potential role of inhibition. In situations where one agent follows the other ($O_1 = 1$) and the task "inhibits" their own autocorrelation ($S_1 = − 1$), the result is a shifted cross-correlation function towards one of the agents leading (Fig. 6, purple).

In examining this concordance with human data, we discovered that some C matrices produce implausible cross-correlation functions. Rather than being a weakness, this can be seen as a strength of the model: it suggests that this modeling framework could be used to evaluate whether certain C matrices are cognitively plausible at all. For example, if both agents inhibit their own behavior while trying to mimic the other, their mutual "attrac-tor" becomes unstable, because the agents are simultaneously seeking "mimicry" (the same behavior) and the C matrix restricts this possibility. The result is an erratic cross-correlation function (Fig. 6, right). We revisit this issue in the conclusion, arguing that the modeling framework need not produce coherent human-like responses in all cases; when it does not, it may be suggestive of unstable task contexts.

We also examined the behavioral time series from both agents in $\mathbf{B}(t)$ and assessed when they "take turns." We did this by thresholding the state variables of these systems, taking $b_1(t) > \mu_1$ and assessing the closest point in time that $b_2(t)$ showed an "on" state. We recorded these "lags" between the times when the two agents became more activated, drawing an analogy to the human data in Stivers et al.[60], who showed rapid transitions dur-ing polar (yes–no) questions. Humans tend to respond so quickly that the highest maximum is near 0. The researchers also found some subtle cultural influences that altered this turn-taking lag (see their Fig. 1). When
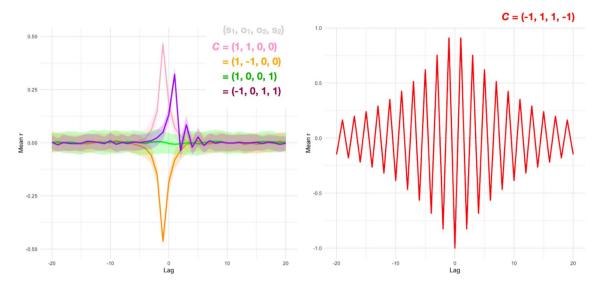
Figure 6. Mapping human task dynamics with simulation data. Left: Average cross-correlation functions for different C's (ribbon width = SD). C parameters that promote other-influence yields positive correlations, but this profile is shifted when one agent "inhibits" its own states but follows another. Right: Some C configurations are unstable, such as two systems avoiding their own behavior while seeking to "mimic"; the system cannot stabilize because it at once seeks to converge and avoid similar behaviors.

Figure 7. Mapping gaps in human conversational turn-taking patterns from Stivers et al.[60] with simulated data. Different colored lines represent various C matrix inputs and the generated patterns over different time lags (from t_{-20} to t_{20}).

we performed a similar analysis with our simulated agent data, we found a qualitative resemblance to the lags described in Stivers et al.[60]. Moreover, these lags do seem to be influenced by C, showing that subtle impacts of the task setup may slightly alter the structure of turns. Additionally, and somewhat speculatively, these subtle impacts of C may hint at methods for conceptualizing cross-cultural differences. In this context, culturally specific patterns of turn-taking can be regarded as subtle changes to how participants in interactions perceive the C that underlies their interaction (see discussion of subtle cross-cultural variations in[60]. This lag result is shown in Fig. 7 according to the same conditions shown with cross-correlation, revealing the slight modulation of turn-taking patterns by the task matrix.

Discussion and conclusion

Dyadic collaboration has been widely studied empirically, but computational models in this domain have focused on accounting for motor coordination, such as joint actions[35,44–47], or testing specific cognitive mechanisms, such as action prediction and signaling[39]. We introduced a simple modeling framework, relevant to both joint action and conversation, that simulates the emergence of the dynamics of partners' behavior across varied collaborative task contexts. The framework allows us to transparently couple task parameters and behavioral dynamics: It allows us to implement and test that changes in the task constraints, through the context matrix C, lead to different stable patterns of relative behavior, expressed with r and other metrics of interpersonal synchrony (e.g.,

cross-correlation function). Results from various matrices can, in combination, simulate dyadic interaction from the onset of interaction to emerging synchrony and complementarity.

We evaluated variation in the context matrix C to compare how different inducements during simulated communication (active, inactive, or inhibitive) would lead to different levels of alignment between task partners. Our analysis suggests that in order for complementarity to emerge in the simulation (such as in turn-taking scenarios), at least one of the two agents needs to be actively inhibited by the task. This inhibition is linked to self-restraint from action.

Although this result might be an inherent outcome of the model's specification, it provides an essential insight revealing that our framework is capable of probing the fundamental structural ingredients of interactions. It is important to recognize, however, that this initial finding allows only limited interpretation because our simple model lacks the complex and subtle distinctions between different kinds of inhibition, such as the exogenous inhibition introduced by tasks, and the endogenous inhibition involved in the internal cognitive process.

Despite this limitation, our discovery regarding the significance of inhibitory parameters in producing clear complementary dynamics is consistent with other research findings that highlight the role of executive processes in communication and collaboration[14,61]. For example, there is evidence that individual differences in inhibitory control significantly predicted people's ability to inhibit their own perspective and consider discrepancies in shared information during language processing[14]. Conceptually, these findings are consistent with wider theoretical proposals of synergistic self-organization, a concept used to explain the behavior of complex dynamic systems. Under this framework, cognition and behavior in a multi-agent environment exhibit distinct and partly stable patterns of coordination emerging from constraints across people and the task[8,56,62]. We would argue that a perspective rooted in synergy elegantly frames the initial modeling endeavors reported here.

In this general and simple framework, we have included only a few essential components of human interactions, such as agent behaviors and context. This simplicity provides flexibility for future iterations of the model that may focus on various aspects of human interactions, such as different potential loci for where inhibition resides (whether cognitive control, or task control). By mapping the human data collected in Stivers et al.[60], we demonstrated how this simple framework can be applied to specific scenarios like turn-taking patterns of polar question answering in daily conversations. Future iterations of the model can increase in complexity to investigate the nuances in human interactions that are abstracted away in this first illustration, so that model and theory may serve together as a framework for exploring key characteristics of interpersonal processes through a computational lens.

Though promising as a preliminary framework, our model has various limitations in addition to the simplified nature described above. For example, we currently represent behavior as scalar signals instead of specific behavioral metrics. This is a common limitation shared among computational models that aim to represent interpersonal interactions[35,54]. To move beyond these initial steps, we identify three possible future directions to explore how computational endeavors can achieve increased internal and external validity in computationally characterizing human social interaction and cooperation.

Information receptivity

The speaker's power of influence and the listener's level of information receptivity are described by the I parameter in the model. The current model assumes $I = 1$ in order to exhaustively explore the effects of context matrix C. However, in real life, it seems very unlikely that individuals attend to every meaning, and few if any individuals can exert maximum power of influence on another person's behavior and thoughts. This limitation is not restricted to the model we derived. In fact, there is a lack of modeling endeavors looking into the dynamics of influence during social interaction and cooperation.

The I parameter in our model can be adapted to capture the dynamics of influence between interlocutors. To achieve this, we plan to build on the theoretical construct of communicative potential (CP)—an attribute that describes the level of receptivity of the information recipient, or the potential that a certain piece of information will be communicated[63]. For example, we may transform I from a scalar to a function with various parameters that account for factors involving information receptivity, such as the listener's level of interest, complexity of the topic, and psycho-physiological states for speaker and speaker at a specific time point.

Modality of behavior

The current model uses a generated number to represent one behavior, such as eye gaze or speaking. These generated numbers are extracted from a uniform distribution with random deviations, which represents a single modality of behavior in real life. However, the interpersonal processes in real life are multimodal in nature[64]. When we interact with people, multiple modalities of perceptual signals take place concurrently—including vocal, facial, gestural signals—along with neurocognitive activities. Different behaviors can have complementary roles—there could be complementarity in one dimension (e.g., eye movements) but still high alignment in language use. This limitation is not restricted to the model we derived. In fact, most dynamical models focus on one aspect of interaction, such as conversational turn-taking[54] and motor control in spatial movements[35]. However, dynamical models don't typically synthesize these modalities in a single model to capture the multimodal nature of human interaction.

We plan to update the B parameter in our model from a two-dimensional vector with one dimension per simulated agent with one behavior (such as eye movement), to a three-dimensional vector with two or more behaviors per agent (such as both eye movement and conversation). This three-dimensional vector will be able to describe multiple behavioral components concurrently to mimic real-life interactions.

Behavioral consistency over time

The current model includes the power of influence, context of the interaction, and behavior represented by a generated number as the initial parameters, then regurgitates the behavior at previous time points to generate behavior of the current time point. In other words, the interactional process that happens at an early time point (e.g., $t = 1$) is identical to that at a later time point (e.g., $t = 100$). While permitting statistical testing of critical cognitive processes in interaction, such as inhibition, this stable interactive dynamic is not an accurate representation of interpersonal processes in real life. Interactions have complex structures, and are not given to one single stream of behavioral patterning[65].

One possible solution is the segmentation of interaction—transforming a continuous interaction into segments with distinct features. For example, a recent paper by Microsoft Semantic Machines shows how dialogues can be segmented and represented as flow charts[66]. This relates to statistical methods in machine learning, such as the automatic distillation of structure (ADIOS) algorithm[67], and other discourse analytic methods, such as dialogic resonance[68]. Other work in cognition and communication suggests that distilling structural patterns of interaction is important for understanding complex combinations of emergent patterns that include both synchrony and complementarity. For example, Fusaroli and Tylén[69] found that people are most effective in an interactive problem when they form more complex patterns of communication, referred to as interactive "synergy" (see also[32]), and not merely synchrony. Our model may be able to capture some trends in conversation, such as emerging complementarity, but because their dynamics are sustained in the same way across simulated conversations, they cannot easily model abrupt changes in behavior. Integrating these ideas into the model would diversify the possible patterns emerging in dyadic interaction. For example, an interaction could show distinct leading and following patterns as conversation topics shift, or as goals shift in the interaction. Such sequential structuring of interaction dynamics would allow greater ecological validity in our computational investigations of interpersonal dynamics.

The initial results and versatile extensions of this simple computational framework could further accommodate the variability and complexity of social interactions in human cognition and behavior. Interpersonal communication is a complicated multiscale process involving coordination in conversational turn-taking[70], multimodal behavior, including gaze and gestures[64], and even correspondence in neurocognitive activity[71,72]. Unpacking this multiscale nature of communication requires computational models that are highly flexible, extendable, and interpretable. Devising such a model for dyadic interactions can potentially further inspire the development of computational models for teams and role emergence in groups[73,74]. In conjunction with computational modeling, bridging to the efforts of other methodologies and disciplines that investigate social interactions—including conversation analysis, experimental social psychology, and social neuroscience—could help to address open queries about the multiscale nature of human communication.

Data availability

Code used in this study is available on GitHub at https://github.com/miaoqy0729/sim-syn-sims.

References

1. Pickering, M. J. & Garrod, S. Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* **27**(2), 169-190 (2004).
2. Brennan, S. E. & Clark, H. H. Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol. Learn. Mem. Cogn.* **22**, 1482–1493 (1996).
3. Chartrand, T. L. & Bargh, J. A. The chameleon effect: The perception–behavior link and social interaction. *J. Pers. Soc. Psychol.* **76**, 893–910 (1999).
4. Giles, H., Edwards, A. L. & Walther, J. B. Communication accommodation theory: Past accomplishments, current trends, and future prospects. *Lang. Sci.* **99**, 101571 (2023).
5. Hatfield, E., Cacioppo, J. T. & Rapson, R. L. Emotional contagion. *Curr. Dir. Psychol. Sci.* **2**, 96–100 (1993).
6. Paxton, A. E. *Coordination: Theoretical, Methodological, and Experimental Perspectives*. Dissertation, University of California, Merced (2015).
7. Fusaroli, R., Rączaszek-Leonardi, J. & Tylén, K. Dialog as interpersonal synergy. *New Ideas Psychol.* **32**, 147–157 (2014).
8. Riley, M. A., Richardson, M. J., Shockley, K. & Ramenzoni, V. C. Interpersonal synergies. *Front. Psychol.* **2**, 38 (2011).
9. Shockley, K., Santana, M.-V. & Fowler, C. A. Mutual interpersonal postural constraints are involved in cooperative conversation. *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 326–332 (2003).
10. Shockley, K., Richardson, D. C. & Dale, R. Conversation and coordinative structures. *Top. Cogn. Sci.* **1**, 305–319 (2009).
11. Galati, A., Alviar, C., Dale, R., & Coco, M.I. Task goals modulate the alignment of eye-movements and linguistic strategies. In *Talk given at the 31st Annual Meeting of the Society for Text and Discourse, Virtual Conference* (2021).
12. Gorman, J. C., Amazeen, P. G., Crites, M. J. & Gipson, C. L. Deviations from mirroring in interpersonal multifrequency coordination when visual information is occluded. *Exp. Brain Res.* **235**, 1209–1221 (2017).
13. Hasson, U. & Frith, C. D. Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150366 (2016).
14. Brown-Schmidt, S. The role of executive function in perspective taking during online language comprehension. *Psychon. Bull. Rev.* **16**, 893–900 (2009).
15. Takamizawa, K. & Kawasaki, M. Transfer entropy for synchronized behavior estimation of interpersonal relationships in human communication: Identifying leaders or followers. *Sci. Rep.* **9**(1), 10960 (2019).
16. Hilbert, H. *et al.* Computational communication science: A methodological catalyzer for a maturing discipline. *Int. J. Commun.* **13**, 3912–3934 (2019).
17. McClelland, J. L. The place of modeling in cognitive science. *Top. Cogn. Sci.* **1**, 11–38 (2009).
18. Guest, O. & Martin, A. E. How computational modeling can force theory building in psychological science. *Perspect. Psychol. Sci.* **16**(4), 789–802 (2021).
19. Álvarez-Gálvez, J. Computational simulation methods. In *The International Encyclopedia of Communication Research Methods* 1–11 (2017).

11

20. Smaldino, P. E. Models are stupid, and we need more of them. In *Computational Social Psychology* 311–331 (2017).
21. Healey, P. G., Purver, M. & Howes, C. Divergence in dialogue. *PLoS ONE* **9**, e98598 (2014).
22. Pardo, J. S. On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* **119**, 2382–2393 (2006).
23. Keysar, B., Barr, D. J. & Horton, W. S. The egocentric basis of language use. *Curr. Dir. Psychol. Sci.* **7**, 46–49 (1998).
24. Brennan, S. E., Galati, A. & Kuhlen, A. K. Two minds, one dialog. In *Psychology of Learning and Motivation* 301–344 (2010).
25. Barr, D. J. Perspective taking and its impostors in language use: Four patterns of deception. In *The Oxford Handbook of Language and Social Psychology* (2014).
26. Dale, R. *et al.* Interacting timescales in perspective-taking. *Front. Psychol.* **9**, 1278 (2018).
27. Mozuraitis, M., Stevenson, S. & Heller, D. Modeling reference production as the probabilistic combination of multiple perspectives. *Cogn. Sci.* **42**, 974–1008 (2018).
28. Fusaroli, R. *et al.* Coming to terms. *Psychol. Sci.* **23**, 931–939 (2012).
29. Coco, M. I., Dale, R. & Keller, F. Performance in a collaborative search task: The role of feedback and alignment. *Top. Cogn. Sci.* **10**, 55–79 (2017).
30. Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B. & Zelinsky, G. J. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition* **106**, 1465–1477 (2008).
31. Richardson, D. C. & Dale, R. Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cogn. Sci.* **29**, 1045–1060 (2005).
32. Paxton, A. & Dale, R. Argument disrupts interpersonal synchrony. *Q. J. Exp. Psychol.* **66**, 2092–2102 (2013).
33. Reitter, D. & Moore, J. D. Alignment and task success in spoken dialogue. *J. Mem. Lang.* **76**, 29–46 (2014).
34. Riordan, M. A., Kreuz, R. J. & Olney, A. M. Alignment is a function of conversational dynamics. *J. Lang. Soc. Psychol.* **33**, 465–481 (2013).
35. Nalepka, P. *et al.* Human social motor solutions for human–machine interaction in dynamical task contexts. *Proc. Natl. Acad. Sci.* **116**, 1437–1446 (2019).
36. Dale, R., Fusaroli, R., Duran, N. D. & Richardson, D. C. The self-organization of human interaction. In *Psychology of Learning and Motivation* 43–95 (2013).
37. Wilson, M. & Wilson, T. P. An oscillator model of the timing of turn-taking. *Psychon. Bull. Rev.* **12**, 957–968 (2005).
38. Pickering, M. J. & Garrod, S. Forward models and their implications for production, comprehension, and dialogue. *Behav. Brain Sci.* **36**, 377–392 (2013).
39. Donnarumma, F., Dindo, H., Iodice, P. & Pezzulo, G. You cannot speak and listen at the same time: A probabilistic model of turn-taking. *Biol. Cybern.* **111**, 165–183 (2017).
40. Luo, B., Lau, R. Y., Li, C. & Si, Y. W. A critical review of state-of-the-art chatbot designs and applications. *WIREs Data Min. Knowl. Discov.* **12**(1), e1434 (2021).
41. Chen, C. *et al.* Gunrock: Building a human-like social bot by leveraging large scale real user data. In *2nd Proceedings of Alexa Prize* (Alexa Prize, 2018).
42. Zhang, Y., Tino, P., Leonardis, A. & Tang, K. A survey on neural network interpretability. *IEEE Trans. Emerg. Top. Comput. Intell.* **5**, 726–742 (2021).
43. Delaherche, E. *et al.* Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Trans. Affect. Comput.* **3**, 349–365 (2012).
44. Haken, H., Kelso, J. A. & Bunz, H. A theoretical model of phase transitions in human hand movements. *Biol. Cybern.* **51**, 347–356 (1985).
45. Noy, L., Dekel, E. & Alon, U. The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proc. Natl. Acad. Sci.* **108**, 20947–20952 (2011).
46. Kostrubiec, V., Dumas, G., Zanone, P.-G. & Kelso, J. A. The virtual teacher (VT) paradigm: Learning new patterns of interpersonal coordination using the Human Dynamic Clamp. *PLoS ONE* **10**, e0142029 (2015).
47. Richardson, M. J. *et al.* Self-organized complementary Joint Action: Behavioral Dynamics of an interpersonal collision-avoidance task. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 665–679 (2015).
48. Strogatz, S. H. & Stewart, I. Coupled oscillators and biological synchronization. *Sci. Am.* **269**, 102–109 (1993).
49. Wiltshire, T. J., Philipsen, J. S., Trasmundi, S. B., Jensen, T. W. & Steffensen, S. V. Interpersonal coordination dynamics in psychotherapy: A systematic review. *Cogn. Ther. Res.* **44**, 752–773 (2020).
50. Dahan, A., Noy, L., Hart, Y., Mayo, A. & Alon, U. Exit from synchrony in joint improvised motion. *PLoS ONE* **11**, e0160747 (2016).
51. Sabharwal, S. R. *et al.* Husync—A model and system for the measure of synchronization in small groups: A case study on musical joint action. *IEEE Access* **10**, 92357–92372 (2022).
52. Butler, E. A. & Barnard, K. J. Quantifying interpersonal dynamics for studying socio-emotional processes and adverse health behaviors. *Psychosom. Med.* **81**, 749–758 (2019).
53. Wallot, S., Mitkidis, P., McGraw, J. J. & Roepstorff, A. Beyond Synchrony: Joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PLoS ONE* **11**, e0168306 (2016).
54. Buder, E. H. A nonlinear dynamic model of Social Interaction. *Commun. Res.* **18**, 174–198 (1991).
55. Gottman, J. M., Murray, J. D., Swanson, C. C., Tyson, R. & Swanson, K. R. *The mathematics of Marriage: Dynamic Nonlinear Models* (MIT Press, 2002).
56. Richardson, M. J., Dale, R. & Marsh, K. L. Complex dynamical systems in social and personality psychology. In *Handbook of Research Methods in Social and Personality Psychology* 253–282 (2014).
57. Wagenmakers, E.-J. & Farrell, S. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–196 (2004).
58. Myung, I. J., Forster, M. R. & Browne, M. W. Model selection [Special issue]. *J. Math. Psychol.* **44**, 1–2 (2000).
59. Ramseyer, F. & Tschacher, W. Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome. *J. Consult. Clin. Psychol.* **79**, 284–295 (2011).
60. Stivers, T. *et al.* Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci.* **106**, 10587–10592 (2009).
61. Qureshi, A. W., Monk, R. L., Samson, D. & Apperly, I. A. Does interference between self and other perspectives in theory of mind tasks reflect a common underlying process? Evidence from individual differences in theory of mind and inhibitory control. *Psychon. Bull. Rev.* **27**, 178–190 (2019).
62. Richardson, M. J. *et al.* Modeling embedded interpersonal and multiagent coordination. In *Proceedings of the 1st International Conference on Complex Information Systems* (2016).
63. Steen, F. F. & Chakraborty, S. Exploring the possible: A unifying cognitive and evolutionary approach to art. *Front. Psychol.* **12**, 787789 (2022).
64. Rasenberg, M., Pouw, W., Özyürek, A. & Dingemanse, M. The multimodal nature of communicative efficiency in social interaction. *Sci. Rep.* **12**(1), 19111 (2022).
65. Schegloff, E. A. *Sequence Organization in Interaction* (Cambridge University Press, 2007).
66. Semantic Machines *et al.* Task-oriented dialogue as dataflow synthesis. In *Proceedings of Transactions of the Association for Computational Linguistics* **8**, (TACL, 2020).
67. Solan, Z., Horn, D., Ruppin, E. & Edelman, S. Unsupervised learning of natural languages. *Proc. Natl. Acad. Sci.* **102**, 11629–11634 (2005).

68. Du Bois, J. W., Hobson, R. P. & Hobson, J. A. Dialogic resonance and intersubjective engagement in autism. *Cognit. Linguist.* **25**, 411–441 (2014).
69. Fusaroli, R. & Tylén, K. Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cogn. Sci.* **40**, 145–171 (2015).
70. Stivers, T. & Sidnell, J. Introduction: Multimodal interaction. *Semiotica* **2005**, 1–20 (2005).
71. Kuhlen, A. K., Allefeld, C. & Haynes, J.-D. Content-specific coordination of listeners' to Speakers' EEG during communication. *Front. Hum. Neurosci.* **6**, 266 (2012).
72. Lieberman, M. D. Seeing minds, matter, and meaning: The CEEing model of pre-reflective subjective construal. *Psychol. Rev.* **129**, 830–872 (2022).
73. Kazi, S. *et al.* Team physiological dynamics: A critical review. *Hum. Factors J. Hum. Factors Ergon. Soc.* **63**, 32–65 (2019).
74. Goldstone, R. L., Andrade-Lotero, E. J., Hawkins, R. D. & Roberts, M. E. The emergence of specialized roles within groups. *Top. Cognit. Sci.* https://doi.org/10.1111/tops.12644. *Advance online publication (2023).*

Acknowledgements

Author contributions

A.G., G.M. and R.D. planned the research. G.M. and R.D. designed and programmed the simulation. G.M. programmed the regression analysis. G.M., A.G and R.D. wrote and revised the paper repeatedly.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.Q.M.

Reprints and permissions information is available at www.nature.com/reprints.

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