Capturing the Dynamics of Trust and Team Processes in Human-Human-Agent Teams via Multidimensional Neural Recurrence Analyses

LUCCA ELOY*, University of Colorado Boulder, USA
CARA SPENCER, University of Colorado Boulder, USA
EMILY DOHERTY, University of Colorado Boulder, USA
LEANNE HIRSHFIELD, University of Colorado Boulder, USA

As collaborative technologies evolve from supportive tools to interactive teammates, there is a growing need to understand how trust and team processes develop in human-agent teams. To contribute effectively, these systems must be able to support human teammates in a task without disrupting the delicate interpersonal states and team processes that govern successful collaboration. In order to break down the complexity of monitoring multiple actors in human-agent collaborations, there is a need to identify interpretable, generalizable measures that can monitor the emergence of interpersonal and team-level processes that underlie effective teaming. We address this gap by using functional Near-Infrared Spectroscopy to concurrently measure brain activity of two individuals in a human-human-agent team during a complex, ecologically valid collaborative task, with a goal of identifying quantitative markers of cognition- and affect-based trust alongside team processes of coordination, strategy formulation, and affect management. Two multidimensional extensions of recurrence quantification analysis, a nonlinear method based in dynamical systems theory, are presented to quantify interpersonal coupling and team-level regularity as reflected in the hemodynamics of three cortical regions across multiple time-scales. Mixed-effects regressions reveal that neural recurrence between individuals uniquely reflects changes in self-reported trust, while team-level neural regularity inversely predicts self-reported team processes. Additionally, we show that recurrence metrics capture temporal dynamics of affect-based trust consistent with existing theory, showcasing the interpretability and specificity of these metrics for disentangling complex team states and processes. This paper presents a novel, interpretable, and computationally efficient model-free method capable of differentiating between latent trust and team processes a complex, naturalistic task setting. We discuss the potential applications of this technique for continuous monitoring of team states, providing clear targets for the future development of adaptive human-agent teaming systems.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: Human-agent teaming, Recurrence quantification analysis, Neuroergonomics

ACM Reference Format:
Lucca Eloy, Cara Spencer, Emily Doherty, and Leanne Hirshfield. 2023. Capturing the Dynamics of Trust and Team Processes in Human-Human-Agent Teams via Multidimensional Neural Recurrence Analyses. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 122 (April 2023), 23 pages. https://doi.org/10.1145/3579598

Authors' addresses: Lucca Eloy, lucca.eloy@colorado.edu, University of Colorado Boulder, , Boulder, Colorado, USA, ; Cara Spencer, cara.spencer@colorado.edu, University of Colorado Boulder, , Boulder, Colorado, USA, ; Emily Doherty, emily.doherty@colorado.edu, University of Colorado Boulder, , Boulder, Colorado, USA, ; Leanne Hirshfield, leanne.hirshfield@colorado.edu, University of Colorado Boulder, , Boulder, Colorado, USA, .

2573-0142/2023/4-ART122
https://doi.org/10.1145/3579598

1 INTRODUCTION

It need not be restated that computer-supported collaborations are all but necessary in the modern world— be it the workplace, classroom, lab, or other dynamic environments. These technologies are shifting roles from simply aiding human-human collaboration to joining as uniquely skilled members of human-agent teams (HATs). The evolution of HATs in military, business, medical, and educational domains is requiring intelligent systems to adapt from facilitators into full-fledged teammates [12, 66]. However, human teams are complex, evolving dynamical systems that must coordinate expectations, share information, and maintain interpersonal relations [23, 29, 53]. Collaboration allows a team to outperform the sum of its parts on a given task, but a team may also be subject to process loss by which the degradation of cooperative processes results in underperformance [42]. Thus, a goal of computer-supported work is to mitigate process loss by monitoring and supporting effective processes such as coordination, affect co-regulation, and trust. This mitigation and support work best when unobtrusive to the humans' teaming processes. Successful HAT systems will leverage the computational power and sensory capabilities of computer agents to monitor and adapt to emerging team dynamics, encouraging human operators to effectively coordinate and maintain cohesion with each other while offloading tasks onto the agent when appropriate to increase team efficiency and task performance.

Few empirical studies have attempted to map these theoretical processes onto ecologically valid human-agent collaborations with more than one human [27]. Studies restricted to only one human team member fail to generalize to the complexities of real-world group settings in which HAT systems are most needed. Due to the difficulty of objectively measuring team processes in naturalistic settings, most efforts are also hindered by rigid task settings that reduce participants' potential decisions and strategies to approach the task. Moreover, machine learning and neural network approaches, though powerful classification and prediction tools, are often difficult to interpret. As such, these studies are largely unable to draw clear conclusions from their results, highlighting the need for more robust, ecologically valid, and generalizable studies. Ideally, intelligent HAT systems will provide insight into the emergence and behavior of latent states over time. To accomplish this, we must explore interpretable computational techniques grounded in existing theory that capture the complexities of multiparty interaction.

We address this challenge by using functional Near-Infrared Spectroscopy (fNIRS) to concurrently measure brain activity of two individuals in a human-human-agent team during a complex, open-ended collaborative task. We explore methods to measure latent trust and team states as an important step toward integrated HAT systems that support cooperative work by monitoring and adapting to shifting dynamics. A novel method of quantifying neurophysiological coupling between individuals and regularity within a team is presented to monitor interpersonal trust and collaborative team processes. Specifically, multi-dimensional recurrence quantification analysis (MdRQA) and multidimensional cross-recurrence quantification analysis (MdCRQA) applied to neurophysiological sensor data are shown to hold great potential for the continuous and unintrusive extraction of team states. These techniques are computationally tractable, readily interpretable, and impose no assumptions or training requirements on the data.

The ability to separate and individually measure interpersonal trust alongside team states is crucial to adaptive systems, allowing these collaborative technologies to monitor and directly respond to lapses in any individual processes via supportive interventions. Figure 1 outlines how these methods can function in a robust adaptive human-human-agent collaboration system: during a HAT collaboration, neurophysiological sensor recordings extract multidimensional time series from each participant, which are continuously fed into interpersonal (MdCRQA) and team-level (MdRQA) analyses, subsequently translating the data into markers of emergent states like trust, coordination, and co-regulation. Information about these states is then sent to the assistive agent, which can appropriately modify its behavior or shed tasks onto either participant to increase or otherwise modulate trust and collaborative team processes.

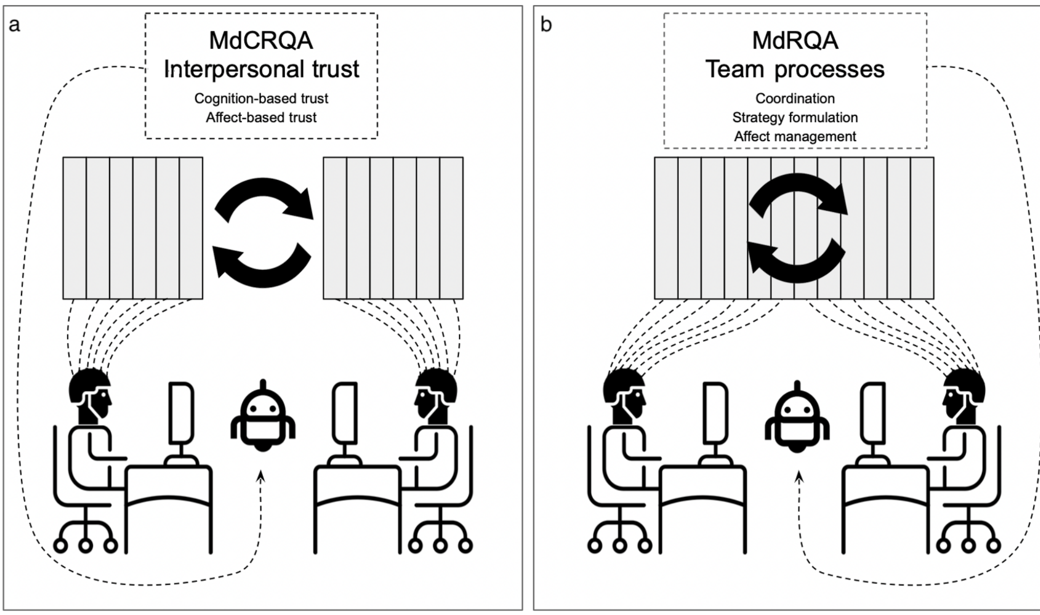

Fig. 1. Relevant states are simultaneously extracted from neurophysiological sensor data via (a) multidimensional cross-recurrence quantification analysis, identifying interpersonal trust, and (b) multidimensional recurrence quantification analysis, identifying team processes.

The overarching goal of this research is to break down the complexity of monitoring multiple actors in a human-agent collaboration. There is a clear need to identify single measures that can monitor the emergence of interpersonal and system-level processes that underly effective collaboration. To address this goal, we designed a study to test our ability to disentangle the interpersonal and team-level processes using recurrence metrics based solely on fNIRS data. This study makes three main contributions to the CSCW domain:

(1) We identify metrics that can independently monitor changes in interpersonal trust and team-level collaborative processes in an ecologically valid context.
(2) We show how these methods uncover dynamics consistent with existing theory, allowing us to draw conclusions about the temporal development of trust and team states during collaboration.
(3) Our results suggest that the diagnosticity of recurrence methods is robust to noise and changes in task context, addressing a major gap in existing human-agent teaming research.

We believe this is one of the first efforts to successfully pick these apart fine processes in a complex, naturalistic task setting— all from one sensor suite with minimal preprocessing and model-free analyses.

1.1 Background and Related Work

In this section, we summarize literature at the intersections of social/teaming science, human-computer interaction, complex systems theory, and neuroergonomics to provide a foundation for the present study. We first introduce trust and team dynamics to identify target variables that underly successful teaming. Then, we discuss the ecological validity of neurophysiological sensors for measuring these variables. Finally, we present theory behind recurrence quantification analysis as a tool for modeling human-agent teams as complex systems.

1.1.1 Trust and team processes. Trust is a particularly hot topic in the human-agent teaming domain (see [27, 35, 46, 48, 64] for reviews). It is a driving force of decision making in both human-human and human-agent teams [16, 39, 46, 60, 83], and has been proposed as a key factor in computer-supported cooperative work [40, 75]. Thimbleby, et al. [75] argue that the success of computer-supported work and everyday interactions hinges on appropriate, or calibrated, trust. Trust, generally defined as a willingness to accept vulnerability based on positive expectations of others [68], is properly calibrated when the trustor's expectations appropriately match the capabilities of the trustee. For example, trust calibration drives individuals to offload tasks onto a teammate when the teammate is capable of properly handling the task (increased trust), and to take over tasks other teammates are not fit to handle (decreased trust). This promotes team performance by ensuring each individual is capable of their task. This formulation of trust extends to human-human teams, human-agent teams, and hybrid teams of multiple humans and/or multiple agents [27, 46, 48, 63].

Much of the literature on trust in teams follows a seminal framework proposed by McAllister [56], which characterizes two main dimensions of trust especially relevant to interpersonal collaborations: cognition-based trust and affect-based trust. Cognition-based trust, or cognitive trust, is grounded in one's perceived ability or competency. In team collaborations, cognition-based trust is found to promote cooperation among individuals, stemming from the belief that the trustee will be successful in a given task [39, 41, 54]. Studies on the effects of cognitive trust in teams found that competency increased cognition-based trust, which in turn increased individuals' effort and, subsequently, team performance. Affect-based, or affective trust, on the other hand, is fostered from the perceived motives behind the trustee's actions toward the trustor. Affective trust is increased by displays of personal concern and assistance, especially when performed outside of an individual's role and without expectation of reward [32, 56]. Affect-based trust is slower to form than cognition-based trust, but is more robust to conflict and time [83]. Affective trust encourages individuals contribute to a team and increases communication during collaboration [13, 70]. In a review of trust in virtual teams, Hacker et al. [32] found that cognition- and affect-based trust together improve knowledge sharing and promote coordination and cohesion, highlighting the role of both types of trust in interpersonal teaming dynamics. Notably, they also identify that much literature is hindered by viewing trust and team processes as static, rather than dynamic constructs.

We ground the present work in a dynamic, cyclical teams framework of hybrid human-human-agent teams [10], wherein individuals are influenced by current events, past events, and developing interpersonal and team-level processes [28, 44, 50]. These team processes, denoted as either transition processes, action processes, or interpersonal processes, determine the success of cooperative teams [50, 52, 53]. Transition processes, such as strategy formulation, are characterized by reflection on past occurrences or preparation for an upcoming task; action processes, such as coordination, describe the actions of team members working toward a goal; and interpersonal processes, such as affect management, consist of social and emotional behaviors that foster relationships between individuals. As teammates interact in a collaborative setting, cognitive and affective trust (interpersonal states) develop as system-level and interpersonal processes emerge, forming a dynamic feedback loop where current states (e.g., levels of trust and/or coordination) govern subsequent interaction [6, 19]. Several reviews identify these states as important to performance outcomes, explicitly mentioning trust as a key factor [6, 28, 36].

In a recent study of trust in human-human-agent teams [21], we found that lowering the agent's level of communication reduced participants' reported trust in their agent and human teammates alike, as well our measure of team processes (adapted from [53]). We proposed that human-human-agent teams do exhibit similar characteristics as human-only teams in complex tasks, as a lapse in communication from one member disrupted individuals' overall sense of team cohesion. A similar study of team cognition in hybrid human-agent triads [71] concluded that human-human-agent teams exhibit similar qualities of iterative mental model formation as human teams. They also distinguish that while communication improved team cognition regardless of composition, goal-oriented communication was significantly more beneficial in human-agent teams. These findings inform the present study's aim to extend the investigation of team dynamics traditionally constrained to human-human teams into the domain of human-agent teams. Indeed, teams of any size and composition can be modeled as complex dynamical systems, with varying states of interpersonal trust and team-level processes emerging over time. Beyond the existing literature, there is a need for empirical methods to consistently measure latent variables like trust. This reflects a main gap in the effort to interface intelligent systems with complex human interactions. Theory and initial findings in [10] detail the effects of agent transparency and reliability on trust, providing a cyclical framework to describe how interpersonal and team-level processes develop in a human-agent interaction. Expanding on this existing groundwork, we link patterns of brain activity to changes in interpersonal trust and team processes while accounting for the effects of agent behavior.

1.1.2 Neuroergonomics and neurophysiological sensors. One of the main hurdles of monitoring trust and team states during collaborations is the inherent reliance on periodic surveys and behavioral measures of latent constructs. While this is a necessary characteristic of experimental paradigms, further applied development of supportive human-agent teaming systems in realistic settings requires continuous, nondisruptive monitoring. A promising yet challenging answer to this is the application of modern neurophysiological sensors to continuously predict levels of trust and other latent variables. The ability to monitor activity across several functionally unique regions via a single sensor suite combined with the brain's role in governing human cognition, affect, and behavior makes a captivating case for the use of neurophysiological measures in investigating individual and team-level internal states.

While fMRI remains the standard for precise functional measurements of brain activity, it is not feasible for real-time measurement in a realistic teaming setting due to the cost and physical restrictions of the scanner. Less invasive technologies have become an attractive alternative as their lightweight sensors, usually consisting of a head cap with probes, allow studies in a much wider variety of environments. In particular, functional near-infrared spectroscopy (fNIRS) holds great potential for measuring social, cognitive, and affective states in ecologically valid scenarios. In a recent review, Curtin & Ayaz identify fNIRS as one of the most promising technologies for applications of continuous brain monitoring, highlighting the technology's ecological advantages and suitability for applied neuroergonomics [17]. fNIRS uses infrared light to measure hemodynamic responses that occur in the brain following neuronal activation— specifically, Δoxyhemoglobin (HbO) and Δdeoxyhemoglobin (HbR)— and thus boasts higher spatial resolution than electroencephalography (EEG) and temporal resolution comparable to fMRI. Furthermore, fNIRS devices are easily configurable, do not require any gel or special preparation, and can function wirelessly allowing for virtually no movement restriction while in use.

Since fNIRS remains in its nascent stage when compared EEG and fMRI, few studies have used the device to measure complex social constructs, such as trust and suspicion, in the brain. Thus, we turn to the body of literature on fMRI as a roadmap for identifying regions of interest (ROIs) to measure

[8, 20, 24, 45]. Brain regions including the frontopolar area (FPA), medial prefrontal cortex (MPFC), dorsolateral prefrontal cortex (DLPFC), and the bilateral temporoparietal junction (TPJ) have often been linked to trust, decision making, and interpersonal reasoning [1, 24, 25, 34, 49, 61, 65, 69, 82]. Dorsal MPFC (DMPFC) in particular has been linked to reasoning and making social judgements about others (vs. oneself) during decision-making [58], a process crucial to evaluating one's intent when deciding to trust. DLPFC has often been implicated in executive processes, such as task supervision and affect regulation, which underlie teaming processes [11, 57]. Taken together, these prior studies identify three key regions of interest and encourage the viability of real-time team monitoring via fNIRS.

Recent studies have attempted to classify trust in automated agents via neurophysiological signals [2, 31, 81]. However, many such efforts use rigid task environments and complex machine learning techniques that are poorly interpretable and prone to overfitting, thus failing to generalize to ecologically valid scenarios. In our most recent study [21], we found DMPFC, DLPFC, and FPA sensitive to changes in trust after manipulating agent transparency and reliability in teams of two humans and one agent in a naturalistic task setting. Through their findings and limitations, these studies encourage the potential of neurophysiological sensors for collaborative technologies and provide a clear direction to improve upon previous efforts.

1.1.3 Recurrence quantification analysis: teams as dynamical systems. Extending the groundwork laid by the studies above, we take a neuroergonomics approach to model teaming dynamics as reflected by patterns of brain activity. However, applying these rich data streams to real teaming scenarios requires sound methods to appropriately model and capture the complex team-level dynamics at play. Following the framework of dynamic team collaborations mentioned in section 1.1.1, we draw from dynamical systems theory to quantify interpersonal and team processes during human interactions [30]. Complex systems like teams exhibit nonlinear shifts in states that can be captured best with nonlinear analyses [18, 51]. Such patterns are typical to cooperative work, where teams exhibit qualitatively distinct behaviors, for example, as they move from discussing possible strategies, to executing a select strategy, to evaluating its effects. The inherent messiness of human behavior often challenges the assumptions required of traditional linear techniques, which are often limited to measuring dyadic synchrony [3, 26, 55, 62]. In line with this perspective, we utilize an analytical method from nonlinear dynamical systems theory called recurrence quantification analysis (RQA) to quantify team-level and interpersonal patterns which emerge during the collaboration [84].

Specifically, RQA captures points during which states are revisited, giving insight to the temporal and structural organization of a dynamical system (such as a human-agent team). Though traditionally used to quantify recurrent patterns within one signal (simple RQA) [9, 47, 76, 84] or alignment between two signals (cross-RQA, or CRQA) [14, 33, 67], newer methods of recurrence quantification can examine multidimensional data to measure coordination and co-evolution of multi-person systems during interactions [14, 18, 73]. For example, recent work has quantified team-level recurrence across physiological and behavioral signals, including eye tracking, speech rate, galvanic skin response, and body movement, during collaborative problem solving and found it predictive of team processes and performance [3, 22, 77, 78].

Many prior applications of recurrence analysis in the CSCW domain have used eye-tracking as a primary signal to capture constructs like shared attention, shared understanding, and interpersonal synchrony [3, 26, 37, 38, 78]. Using a pair programming task, [38] found higher gaze cross-recurrence consistent with increased shared understanding within a team. Demonstrating the interpretability of recurrence analyses coupled with eye tracking data, CRQA was able to quantify instances of attention coupling during the task. Similarly, [78] found that gaze coupling linked to shared knowledge during a collaborative problem solving task. Using multidimensional RQA (MdRQA), they also showed team-level regularity to be predictive of successful negotiation and coordination. Indeed, studies of interpersonal CRQA are able to link recurrence to increased synchrony and coordination, while system-level patterns captured by MdRQA are often sensitive to changes in team performance and collaborative processes [3, 22, 78]. Nevertheless, MdRQA and CRQA are both uniquely tuned to capture different dynamics of human interaction.

Though research applying recurrence quantification to brain data is limited, prior work taking a complex systems approach to fMRI analysis has shown that univariate RQA performs similarly to traditional statistical analyses in identifying patterns of activation (namely, general linear models and independent component analysis) without requiring any a priori assumptions or model specifications beyond a single fixed parameter [9, 47]. Thus, we focus our current effort on MdRQA and a multidimensional extension of CRQA, examining their individual specificity for measuring trust and team processes from neurophysiological activity.

1.2 Current Approach and Hypotheses

Two multidimensional extensions to RQA exist— multidimensional-RQA (MdRQA) and multidimensional cross-RQA (MdCRQA). MdRQA quantifies the level to which multiple signals exhibit collective regularity, or repeated states with a system, though the individual signals may not be in alignment with each other. For example, one repeated state in a system of two individuals' brain signals might include high activation in one person's DLPFC & FPA and simultaneously low activation in the other person's DLPFC & FPA (potentially signaling differences in task engagement), whereas another state might include low DLPFC & FPA activity in both individuals. This is valuable from a neuroergonomics perspective, as concurrent activity in different functional brain regions provides a much richer picture than any one region alone.

The second multidimensional modification of RQA, multidimensional cross-recurrence quantification analysis (MdCRQA), quantifies the co-occurrence of states between two multidimensional systems. This differs from MdRQA which examines states within a single multidimensional system; Figure 1 depicts how these methods examine a team's data differently. Using the above example, MdCRQA might quantify how often both individuals displayed simultaneously high DLPFC & FPA activation across all time lags. **We hypothesize that [H1] applying MdCRQA to fNIRS data will (a) predict cognitive and affective trust but (b) not team processes, while [H2] MdRQA will (a) predict team processes but (b) not interpersonal trust**. In testing these hypotheses, we are investigating whether the difference between these two multidimensional recurrence analyses allows us to tap into distinct, high-dimensional latent states.

Recurrence quantification methods are well suited to fNIRS data as they require no assumptions of linearity or stationarity and are thus robust to noise. Additionally, the ability to quantify regularity across multiple signals at once allows for multiple brain regions of interest (ROIs) and signals (HbO & HbR) to be examined concurrently as they co-evolve without losing information. Furthermore, by restricting comparisons to values that occur only within a specific time of each other, RQA can easily be constrained to quantify recurrence across short, medium, and long time scales. For example, at a 5 second lag each value is compared only to other values occurring within 5 seconds before or after it. These different scales allow us to comprehensively assess the evolution of different team and interpersonal dynamics over time. In other words, we will be able to distinguish states linked to short-term alignment from those that emerge from longer-term alignment. Given previous theory of how cognitive and affective trust form differently over time, we predict that recurrence across different time scales will promote the emergence of different states. **Specifically, we hypothesize that [H3] affect-based trust will stem from longer-term recurrence than cognition-based trust, and [H4] team-level processes will strengthen with longer-term**

recurrence. By piecing apart these temporal dynamics, we aim to uniquely identify markers of interpersonal and team-level states as targets for implementing continuous, real-time monitoring of computer-supported teaming systems.

We test our hypotheses on data obtained from teams of two humans and one agent collaborating virtually on an open-ended (i.e., no 'optimal' solution or strategy is easily attainable) geospatial mapping task that requires knowledge sharing and agreement between both human and agent teammates. The task context simulates a realistic collaboration scenario where individuals are able to approach the problem and coordinate with each other however they choose, introducing a deal of uncertainty. Results will thus indicate that recurrence methods may be readily generalizable across changes in team and task contexts, a major hurdle to building adaptive systems that many machine learning techniques fail to overcome. Furthermore, RQA boasts higher interpretability and lower computational demand than machine learning and neural network approaches traditionally used in HAT research. Results of these analyses can thus be used to draw greater theoretical conclusions of how high-level interpersonal and team dynamics are reflected in the self-organization and interdependencies between brain signals. Therefore, investigation of our hypotheses will reveal whether neural recurrence analysis is a viable metric for both the technical and theoretical advancement of adaptive systems in complex, real-world environments.

2 METHODS

2.1 Experiment

The dataset used in this paper was collected as part of an ongoing study [10, 21], detailed in this section.

2.1.1 Participants. Participants were 30 students (average age 21, 53% male) from a large public university in the United States. Students completed the task in teams of 2 participants and 1 agent, for a total of 15 teams and data collection sessions. Participants were compensated $15/hr for their time and additionally with a variable cash bonus based on task score. Recruitment and experimental procedures were approved by the university's Institutional Review Boards. Participants completed informed consent forms upon arriving at the lab.

2.1.2 Task and testbed environment. The experimental task was hosted in the Computer-Human Allocation of Resources Testbed (CHART), which allows collaboration between a human-agent team through a crime mapping task. Teams of participants are tasked with using past spatiotemporal crime data made publicly available via the Denver Police Department and allocate a limited number of crime prevention resources throughout the city based on this historical crime data. The collective goal of each human-agent team is to 'prevent' as many crimes as possible through resource allocation.

The CHART interface consists of two displays: (right) an interactive crime map unique to each user that allows for selection of past data and category(s) of crimes (e.g., traffic incidents, assault, theft) to overlay on a map of Denver, and (left) a shared map which reflects where each participant places their resources (represented as pins). Figure 2 shows a screenshot of both CHART displays. The shared map reflects real-time changes as participants place their resources, which is done by adding, moving, and deleting pins on the shared map. A sidebar on the right screen also reports the team's current cash bonus, as well as each individual's contribution, based on their performance on prior rounds.

Participants are able to place two different types of resources, differentiated by pin color, which correspond to differing levels of crimes. 'Surveillance/monitoring' resources, represented by green pins, are placed to prevent non-violent crimes. 'First response' resources, represented by red pins,

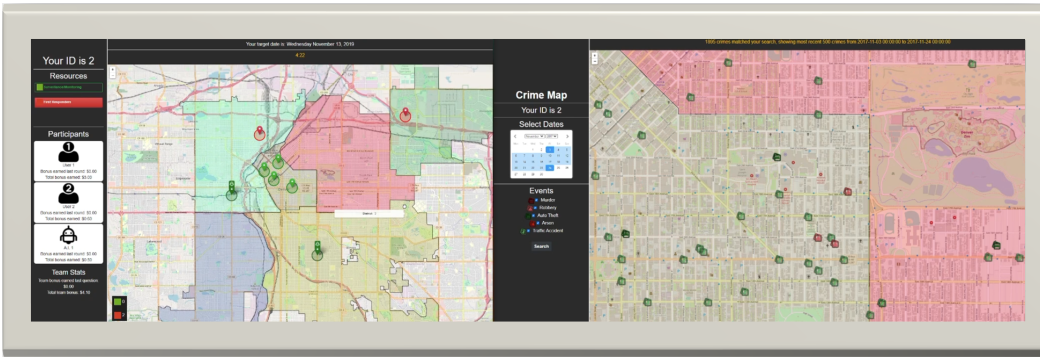

Fig. 2. Screenshot of CHART interface. The right side monitor has each participant's individual map for searching crime data. The left side monitor is a shared map where participants place, move, or delete pins in real time.

address violent crimes. Each participant is given access to the historical data corresponding to different categories of crime and must share knowledge with the other participant to perform well. Participants have an allotment of 6 green pins and 6 red pins (shared between the two) to place during each round. The rest of the pins are placed by the 'AI' agent (detailed below) and can be moved by either participant.

2.1.3 AI agent and manipulations. CHART allows for the use of a pre-programmed (wizard-of-oz-style) 'AI' agent. In the current study, the agent places 6 green pins and 6 red pins in alternating order during each round. Participants are informed that the agent is another team member and its placement of pins contribute to their cumulative cash bonus. When the agent places a pin, a text box message announces its placement and, depending on the experimental (transparency) condition, provides a justification. Agent-placed pins can be moved by the participants, but not deleted. Both reliability (high/low performance) and transparency (high/low communication) of the agent teammate were manipulated in the current study. See [21] for more details on CHART and the agent manipulations.

2.1.4 Experimental procedure. Following arrival to the lab, participants completed informed consent forms and were then separated into different rooms with a desktop computer on which they completed the collaborative task. They first completed a pre-survey and read a slideshow presentation of instructions on how to complete the task. Following review of the instructions, participants independently completed a CHART training session consisting of two shortened 2-minute rounds, during which the agent places only 2 pins, in order to learn how to use the CHART interface.

After the training sessions, the experimenter placed and calibrated a NIRx NIRSport2 head cap on each participant. Following sensor fitting, participants communicated over a Zoom call on their desktop computers. They were directed to click 'I'm Ready' on the testbed screen when they were ready to begin, and they were left alone in their respective rooms. Following each round, the individual's and team's total bonus earned was presented on the screen as well as a short post-round survey completed in Qualtrics. Following this short survey, the next round began when each participant selected "I'm Ready." This procedure was repeated for a total of 8 rounds, after which a final post-round survey was completed and the experimenters removed the sensors and paid participants for their time plus bonus earned.

2.2 Measures

2.2.1 Surveys. Self-report surveys, administered after each experimental round, were selected as dependent measures of emergent individual and team states, including affect, trust, mental demand, and team processes. Due to the length of the study, we chose only the items with the highest factor loadings to be able to assess constructs of interest while minimizing time spent on surveys. Behavioral measures were extracted after data collection by a script using CHART's comprehensive event logs. Timestamps were used to sync neurophysiological time series data with experimental rounds and events. For more detail, see [10] and [21].

Cognition-based and affect-based trust were measured with the highest factor loading items from the respective surveys by McAllister [56].

We assessed team processes under the framework proposed by Marks [50] using the top factor loading items from each category of Mathieu's team process survey for brevity. To cover the three distinct types of processes, we selected the most relevant construct under each category: strategy formulation (transition processes), coordination (action processes), and affect management (interpersonal processes) [53]. We also include an aggregate measure of team processes calculated as the mean across these items.

2.2.2 Neurophysiological recording. Data were recorded on 38 participants using 2 NirSport 16x16 machines (NIRx, Berlin, Germany) with a sampling rate of 5.0863 Hz. We selected optode locations to maximize coverage of relevant cortical regions identified from previous literature in section 1.1.2 (FPA, DMPFC, and DLPFC) using fOLD Toolbox [59] for spatial registration of fNIRS channels onto the anatomical regions of interest. Figure 3 depicts this montage of NIRS source-detector pairs.

2.3 Data Analysis

Each participant's time series data were extracted as follows: band-pass filtered fNIRS data (between 0.01 – 0.2 Hz) from each round was averaged across channels in each ROI to obtain time series for HbO and HbR of FPA, DPLFC, and DMPFC, for a total of six measures per participant. Data were z-scored within each ROI/Hb signal across all participants to prevent differences in magnitude from skewing Euclidean distance computation.

2.3.1 Recurrence quantification analysis methodology. To perform multidimensional cross-recurrence quantification analysis (MdCRQA), we included HbO and HbR together to capture the neural dynamics of each individual more fully. To account for high-dimensional dynamics, state-space reconstruction was performed by first estimating the appropriate delay and embedding parameters using average mutual information and false nearest neighbors techniques as outlined in [79, 80]. Parameters were estimated for each dyad per round (16 dyads x 8 rounds), and the medians were chosen as the final delay and embedding parameters to be applied to the entire dataset. The radius parameter was estimated in a similar fashion using a grid search and performing MdCRQA on all dyad/round units to obtain a median recurrence rate between 4.95% and 5.05%, following standard procedures suggested in [80, 84]. Final parameters were: embedding dimension = 2, delay = 6, and radius = 0.532.

MdCRQA was performed on all team/round units with the final parameter set held constant to allow for comparison between observations. For each round, the 6-dimensional time series of participant 1 and participant 2 were compared across all lags to construct a recurrence matrix. Recurrence rates were then calculated as the percentage of recurrent points across three different lags: 5 seconds, 90 seconds, and 180 seconds in both directions. At the 5 second lag, for example, each value is compared pairwise with each other value occurring within 5 seconds before or after it (a total 10-second window). Thus, 180-second lags represent the entire 6-minute block.

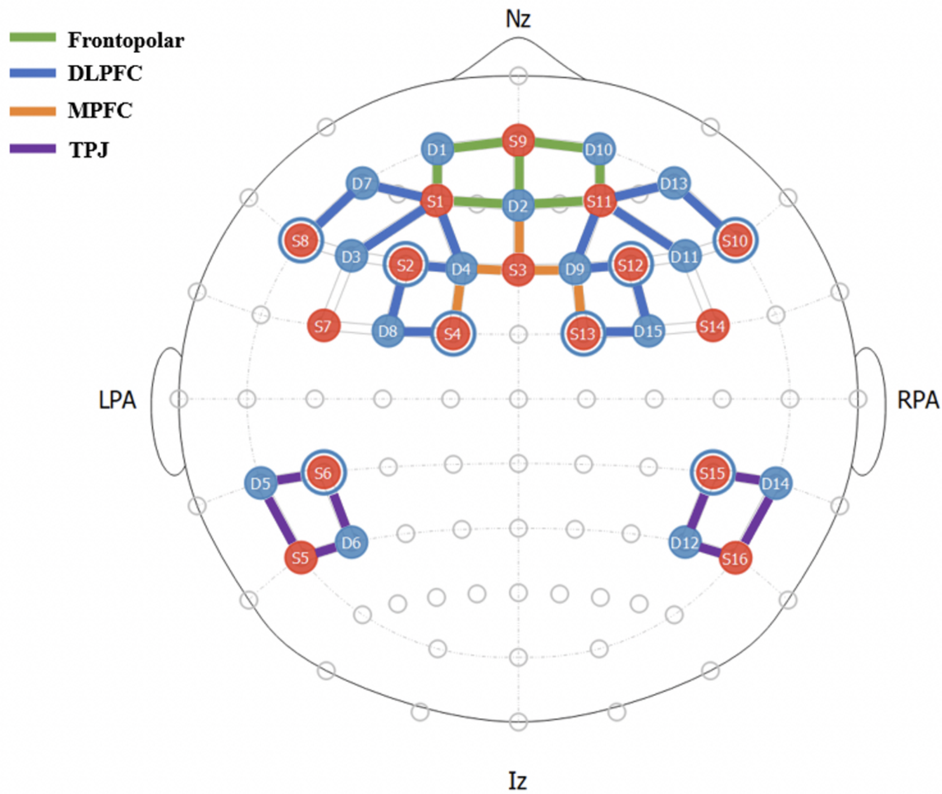

Fig. 3. fNIRS Montage and ROI Mappings. Each channel consists of a source-detector (S#-D#) pair, represented by colored lines.

Multidimensional recurrence quantification analysis (MdRQA) analyses follow the same procedure, including state-space reconstruction and parameter estimation, except that participant time series are merged to create a single data frame. Only participants' HbO signal was used in MdRQA to combat high dimensionality since including both HbO and HbR would result in 12 dimensions (state-space reconstruction further increases dimensionality), which can complicate intuitions behind Euclidean distance computation and skew the radius parameter. Final parameters for MdRQA were: embedding = 2, delay = 6, and radius = 0.407.

3 RESULTS

To investigate the main effects and predictive power of recurrence, we regress each trust and team process measure on MdCRQA and MdRQA recurrence rates at 5, 90, and 180 second windows. We fit linear mixed-effects models using the R package lme4 [5] to best capture our experimental design, accounting for repeated measures (multiple rounds) and expected differences between teams as a random effect. Each model included recurrence rate, transparency condition, reliability condition, and round number as fixed effects, with each team as a random effect. The interaction between recurrence rate and round number was also included in each model to account for the expected drift in the fNIRS signal due to physiological effects of time [15, 72], which may affect recurrence metrics. Additionally, observations with a total recurrence rate greater than 2.5 standard deviations away from the mean were excluded from analyses (6 observations discarded for a total n = 114) to account for instances of poor signal quality (due to improper optode contact with the skull) that result in inflated recurrence rates.

Final models took the following form: $response = recurrence\ rate + transparency + reliability + round + round * recurrence\ rate + (1|team)$. We chose to keep this formula the same across all models to allow for comparison of effects and inference on the predictive potential of recurrence metrics for interpersonal and team processes. Likewise, each response measure was z-scored across all participants to obtain easily interpretable effect sizes (B) in terms of standard deviations.

3.1 Multidimensional Cross-Recurrence Quantification Analysis (MdCRQA)

First, we examined interpersonal neural coupling by running MdCRQA on brain data from both participants of each dyad. Regressing each dependent measure on these predictors (with separate models for each RR lag) obtained the results shown in Table 1 and Table 2.

Table 1. Effect sizes (B) and p-values of mixed-effects models regressing trust on MdCRQA recurrence rate (RR). Individual models were fit for each recurrence rate window length. For brevity, effect sizes and p values for covariates on the right side are summarized across lags. Bold values indicate $p \leq 0.05$.

Predictor	**RR 5s**		**RR 90s**		**RR 180s**		Transparency		Reliability		Round #	
Metric	B	p	B	p	B	p	B	p	B	p	B	p
Cognitive trust	0.10	**0.02**	0.18	$<$ **0.01**	0.18	**0.01**	-0.25 — -0.21	$>$ 0.07	0.10 — 0.12	$>$ 0.30	0.19 — 0.21	$<$ **0.01**
Affective trust	0.06	0.09	0.11	**0.03**	0.12	**0.02**	-0.22 — -0.18	**0.04** — 0.09	0.05 — 0.07	$>$ 0.40	0.19 — 0.21	$<$ **0.01**

Mixed-effects regression results in Table 1 indicate the following main effects of multi-dimensional cross-recurrence rate on trust measures. Cognitive trust is significantly predicted by recurrence rate at 5s (B = 0.10, p = 0.02), 90s (B = 0.18, p < 0.01), and 180 second (B = 0.18, p = 0.01) windows, supporting H1a. Recurrence rates at 90s (B = 0.11, p = 0.03) and 180s (B = 0.12, p = 0.02), but not 5s (B = 0.06, p = 0.09), had a significant effect on affect-based trust. The effect grows slightly in magnitude as the window increases. This result is in line with H3, following previous theory on trust development and suggesting that affect-based trust is indeed captured by interpersonal recurrence at larger time scales.

Regarding covariates, low agent transparency linked to a decrease in affective trust (B = -0.22, p = 0.04) significantly in the 180s lag model, and the trend (non-significant) is observed for all models. The significant effects of round number show that reported trust increased with subsequent rounds, reflecting the continued development of trust as teammates collaborate for longer periods of time. An interaction between 90s recurrence rate and round number was observed (B = -0.02, p = 0.04), signifying that the effect of recurrence rate on cognition-based trust decreased with each subsequent round. This may be due to the increase in trust over time, and the inability of 90s recurrence rate to account for this added variance.

In partial support of H1b, MdCRQA present significant effects on only one individual team process: strategy formulation decreased with an increase in 5s recurrence rate (B = -0.07, p = 0.04) and 180s recurrence rate (B = -0.13, p = 0.03). Our aggregate team processes measure was not significantly linked to MdCRQA recurrence rate at any window length. The low agent transparency condition corresponded to a significant decrease in all measures, consistent with prior findings [21].

Table 2. Effect sizes (B) and p-values of mixed-effects models regressing team process measures on MdCRQA recurrence rate (RR). Individual models were fit for each recurrence rate window length. For brevity, effect sizes and p values for covariates on the right side are summarized across lags. Bold values indicate $p \leq 0.05$.

Predictor	**RR 5s**		**RR 90s**		**RR 180s**		Transparency		Reliability		Round #	
Metric	B	p	B	p	B	p	B	p	B	p	B	p
Coordination	0.01	0.83	0.00	1.00	0.00	0.94	-0.39 — -0.38	< **0.01**	0.09 — 0.10	> 0.30	0.04 — 0.06	> 0.20
Strategy formulation	-0.07	**0.04**	-0.09	0.09	-0.13	**0.03**	-0.38 — -0.37	< **0.01**	0.07 — 0.09	> 0.40	-0.04 — 0.00	> 0.40
Affect management	-0.07	0.08	-0.10	0.08	-0.12	0.06	-0.29 — -0.27	**0.02**	0.05 — 0.07	> 0.50	-0.05 — -0.01	> 0.40
Team process (avg)	-0.04	0.19	-0.05	0.33	-0.08	0.12	-0.39 — -0.40	< **0.01**	0.07 — 0.08	> 0.40	0.02 — 0.05	> 0.30

While no other covariates displayed significant effects, the interaction between recurrence rate and round was significantly positive for strategy formulation in the 180s lag model ($B = 0.02$, $p = 0.04$). Similar to the interaction effect on cognitive trust, this indicates that the effect of recurrence rate became more positive (decreasing in magnitude) as round number increased. Lastly, we note that affect management marginally decreases with recurrence rate at 180s (the marginal trend is observed for all lags). This effect, though opposite in direction, could be due to its classification as an interpersonal affective process, similar to affect-based trust.

3.2 Multidimensional Recurrence Quantification Analysis (MdRQA)

The same modeling procedure was performed to regress trust and team process measures on MdRQA recurrence rate. To allow for interpretable comparison between results, models included the same set of covariates (transparency & reliability conditions, round number, and the interaction between round and recurrence rate). Results are presented in Table 3 and Table 4.

Table 3. Effect sizes (B) and p-values of mixed-effects models regressing trust on MdRQA recurrence rate (RR). Individual models were fit for each recurrence rate window length. For brevity, effect sizes and p values for covariates on the right side are summarized across lags. Bold values indicate $p \leq 0.05$.

Predictor	**RR 5s**		**RR 90s**		**RR 180s**		Transparency		Reliability		Round #	
Metric	B	p	B	p	B	p	B	p	B	p	B	p
Cognitive trust	0.07	**0.01**	0.20	0.25	0.28	0.26	-0.20 — -0.19	> 0.10	0.11 — 0.12	> 0.30	0.16 — 0.29	**0.02 — 0.05**
Affective trust	0.04	0.06	0.12	0.35	0.14	0.47	-0.17	> 0.09	0.06 — 0.08	> 0.40	0.20 — 0.26	< **0.01**

MdRQA recurrence rate significantly increased cognition-based trust ($B = 0.07$, $p = 0.01$) at 5s. A marginal trend is also seen for affect-based trust at 5s. ($B = 0.04$, $p = 0.06$). No other significant effects of recurrence rate are seen for these measures, in partial support of H2b for larger window lengths. As expected, both trust measures increased with round number. Again, we note the interesting marginal effect on affect trust at 5s, similar to effects on affect management shown in Table 4. This is analogous to marginal trends in MdCRQA models.

As predicted in H2a, the sum of team processes is significantly (inversely) proportional to MdRQA recurrence rate at 5s ($B = -0.05$, $p = 0.02$), 90s ($B = -0.31$, $p = 0.02$), and 180s ($B = -0.43$,

Table 4. Effect sizes (B) and p-values of mixed-effects models regressing team process measures on MdRQA recurrence rate (RR). Individual models were fit for each recurrence rate window length. For brevity, effect sizes and p values for covariates on the right side are summarized across lags. Bold values indicate $p \leq 0.05$.

Predictor	**RR 5s**		**RR 90s**		**RR 180s**		Transparency		Reliability		Round #	
Metric	B	p	B	p	B	p	B	p	B	p	B	p
Coordination	-0.03	0.25	-0.30	**0.03**	-0.37	0.07	-0.37	< **0.01**	0.10 — 0.12	> 0.25	-0.04 — -0.03	> 0.50
Strategy formulation	-0.07	**0.01**	-0.29	**0.04**	-0.40	**0.05**	-0.37 — -0.36	< **0.01**	0.11 — 0.12	> 0.20	-0.21 — -0.05	**0.03** - 0.40
Affect management	-0.06	**0.02**	-0.40	**0.01**	-0.63	**0.01**	-0.28 — -0.26	< **0.03**	0.09 — 0.11	> 0.30	-0.16 — -0.11	**0.05** - 0.12
Team process (avg)	-0.05	**0.02**	-0.31	**0.02**	-0.43	**0.02**	-0.38 — -0.37	< **0.01**	0.10 — 0.12	> 0.20	-0.12 — -0.03	> 0.16

p = 0.02), increasing in magnitude with recurrence window length. Interestingly, coordination significantly decreases with recurrence rate only at 90s (B = -0.30, p = 0.03). In further support of H2a, MdRQA recurrence rate predicts more individual team process measures than MdCRQA. Strategy formulation is predicted at every interval: 5s (B = -0.07, p = 0.01), 90s (B = -0.29, p = 0.04), and 180s (B = -0.40, p = 0.05). Likewise, significant negative effects are seen for affect management at 5s (B = -0.06, p = 0.02), 90s (B = -0.40, p = 0.01), and 180s (B = -0.63, p = 0.01).

Among these models, round number significantly affected strategy formulation (B = -0.21, p = 0.03) at 5s and affect management (B = -0.14, p = 0.05) at 180s. Finally, the same expected effects of the transparency manipulation are seen across all measures. Interactions between recurrence rate and round were observed for aggregated team processes ($B = 0.01 - 0.07$, $p < 0.02$), strategy formulation ($B = 0.01 - 0.06$, $p \leq 0.04$), and affect management ($B = 0.01 - 0.10$, $p \leq 0.02$) at all windows. These interactions correspond to a decrease in the effect of recurrence rates for subsequent rounds, likely also stemming from increased variance in these measures for later rounds.

4 DISCUSSION

Mixed-effects models revealed the main effects of recurrence rate across different time scales on perceived (self-reported) trust and team processes. To explore the power of recurrence as a marker of these latent states and processes, we discuss each individual construct and the effects of recurrence, followed by applications and implications for future work.

4.1 Trust

We first examine cognition-based and affect-based trust. As predicted in H1a, cross-recurrence positively predicts cognition-based trust across all lags with the strongest effect observed from a 90s recurrence window. MdRQA recurrence rate shows a significant effect on cognition-based trust only at the 5s window, giving partial support to H2b. Given the foundations of cognition-based trust, MdCRQA results suggest that similarity of neural activation between individuals signals an increased perception of a teammate's ability and competency in a task. This effect carries over to affect-based trust, which is significantly predicted only by MdCRQA recurrence rate at 90s and 180s windows. Following H3, this finding—compared to the significant effect on cognitive trust seen at the 5s window—supports previous theory that affective trust forms on a longer timescale than cognitive trust. Consistency with established frameworks suggests that interpersonal neural recurrence rate reflects the latent constructs we attempt to measure, supporting the potential use of these analyses for trust monitoring. Indeed, the behavior of these predictive effects fits the dynamical systems model that underlies multidimensional recurrence analyses. By performing state-space reconstruction on fNIRS data, we appear to successfully tap into these high-dimensional processes. These results from both analyses support our theory that neural coupling between individuals is a stronger marker of interpersonal trust than system-level regularity.

This set of findings highlights not only the predictive power of recurrence analysis of brain data, but also the relatively high interpretability of these measures compared to traditional machine learning and neural network approaches. Examining recurrence rates allows us to also draw neuroergonomics conclusions about how similarities in functional brain activity is reflected in self-reported cognitive processes. From these results, it appears that individuals' perceptions of trust are directly linked to and potentially influenced by the coupling of their cortical activation with that of a teammate. In other words, participants felt consistently higher trust in their teammate when they exhibit the same states of activity in the target cortical regions. Given its relative simplicity, cross-recurrence, which at its core is performing Euclidean distance computation, holds exciting potential for predicting cognitive and affective processes without sacrificing interpretability.

4.2 Team Processes

Using the same mixed-effects structure, we modeled the effects of recurrence, along with task-relevant covariates, on team processes. Specifically, we focus on predicting coordination, strategy formulation, and affect management to capture the distinct transition, action, and interpersonal team process categories. Strategy formulation is negatively linked to MdCRQA at the 5s and 180s lags, and to MdRQA recurrence rate at all windows. This result does not align with hypothesis H1b, but it does highlight the 90-second lag as a potential differentiator between processes. While we do not have a clear explanation for this trend on strategy formulation, we postulate that the marginal effects of MdCRQA on affect management and MdRQA on affective trust are present due to their overlap as affect-based interpersonal processes. Interestingly, coordination is predicted only by MdRQA recurrence rate at 90s. This identifies a potential time-frame in which the process of coordination develops and is reflected in neural activity. Overall, team-level regularity proves to be a marker of averaged team processes. Though this aggregate measure lacks the specificity of individual constructs, it appropriately summarizes the negative effect of systemic recurrence on each process. Additionally, we see across each significant effect, and even in marginally significant trends, that recurrence on a longer scale captures a stronger effect on team process measures. This supports H4, in which we theorize that individuals' perception of team dynamics are generally more positive if their patterns of neurophysiological recurrence over greater periods of time.

These positive perceptions, however, are shown by consistently negative effect sizes to be reliant on a decrease in a team's regularity. The direction of these effect may seem counterintuitive at first, as decreased regularity marks an increase in beneficial team-level processes. However, this is consistent with similar research that found decreased regularity of physiological and behavioral signals (through MdRQA) signaled an increase in team performance and in key facets of collaborative problem solving [3, 22]. We postulate that increased recurrence rate, or regularity, can result from teams revisiting or remaining in the same states; by contrast, high irregularity stems from an increase in novel states that may be key to adapting to a complex teaming task. Thus, team-level recurrence measures demonstrate an affinity for characterizing collaboration quality in open-ended task settings. Identifying these markers is a valuable step towards capturing team dynamics in 'real-world' collaborative scenarios where hybrid human-agent teams are already being employed.

4.3 Implications and Applications

Examining these effects by timescale reveals the 90-second recurrence window as the most informative measure to disentangle target constructs. Considering only this window, MdCRQA predicts only cognitive and affective trust (in the positive direction), and MdRQA exclusively predicts coordination, strategy formulation, affect management, and averaged team processes (in the negative direction). This results in maximum separability between interpersonal trust and team processes, as was our research goal and prediction in H1 and H2. Identifying this optimal window is crucial as it identifies clear markers of these different processes that can be concurrently monitored and uniquely identified. The time window itself also suggests that these interpersonal and team-level processes are best reflected in high-dimensional neural dynamics at 90 second lags. This holds key implications to the neuroergonomics of computer-supported teaming, as it defines a potential timescale in which dynamic team processes take shape during complex human-agent collaborations.

Toward the goal of adaptive systems, these effects are promising for monitoring collaboration dynamics as they are able to separate between desired constructs. To generalize across individuals and their ever-changing environments, systems that support collaboration must monitor and adapt to these shifting contexts to maximize team performance. For example: at a 90-second window, changes in MdCRQA and MdRQA recurrence rates can signal fluctuations in interpersonal trust or in team processes, respectively. Considering the CHART task, continuous monitoring may detect increased recurrence between individuals (MdCRQA) and increased regularity within the team (MdRQA) across 90-second but not 5-second windows. Utilizing the specificity of these measures, the system might interpret this as an increase in affect-based trust, but not cognition-based trust, and a decrease in team processes. Taken together, these observations suggest the team maintains positive perceptions of each other but is stalling in collaborative productivity. An adaptive system might then shed specific tasks onto each teammate to encourage novel activity, as well as an opportunity to develop cognitive trust between teammates that can demonstrate competency. Not only will this direct the team to increase productivity, but it will also promote the emergence of beneficial team processes that persist and evolve throughout the collaboration (see [10] for cyclical human-agent teaming framework). Such systems would be well-suited for the context of computer-supported cooperative work, adding a new layer of intelligence to computationally driven supports. Additionally, our experiment identifies fNIRS as a promising modality for CSCW that captures the nuances of human interaction from a single set of sensors. Regarding implementations, the metrics we have identified would pair well with computational decision-making models that learn to select optimal actions or interventions based on the current state of the system. A popular example is the partially observable Markov decision process (POMDP) model, designed specifically to infer the state of a system (the team) from empirical observations, account for high uncertainty, and select the most beneficial action [4, 43, 74, 85].

Though our manipulations to agent behavior are not central to this paper, they add important context for the results presented. In the models above, agent transparency significantly affected nearly every dependent variable investigated. We note two implications of this: (1) our models account for the effects of these conditions, highlighting that the effects of recurrence rates are robust; (2) as seen in our previous work, changes in agent communication disrupt both interpersonal and team-level processes, suggesting that hybrid human-agent teams exhibit similar dynamics as human teams. By considering effects of the behavioral context of the agent(s), these results are foundational to monitoring and preserving the dynamics of human-agent teams. In the example above, an adaptive system could monitor and appropriately intervene in shifting team dynamics while separately modulating the agent's level of transparency to maintain calibrated trust in the agent (see [7] for a review of agent transparency adaptations). Documenting and accounting for the expected effects of either manipulation will allow for these concurrent adaptations across various contexts, pushing toward advanced collaborative technologies suited to support interpersonal and team dynamics while maximizing the efficiency of human-agent teaming.

4.3.1 Qualitative discussion of recurrence plots. While recurrence rate was the only measure we extracted from recurrence quantification analyses, we note that MdCRQA and MdRQA provide several other metrics for examining the structure of high-dimensional systems. In particular, the recurrence plots from which recurrence rates are extracted are themselves a rich depiction of interpersonal and system-level state configurations across time. Figure 4 visualizes the recurrence matrices for a single team across two rounds, agent recommendations denoted in vertical and horizontal lines. Lines are colored according to the team's decision to accept or reject the recommendation.

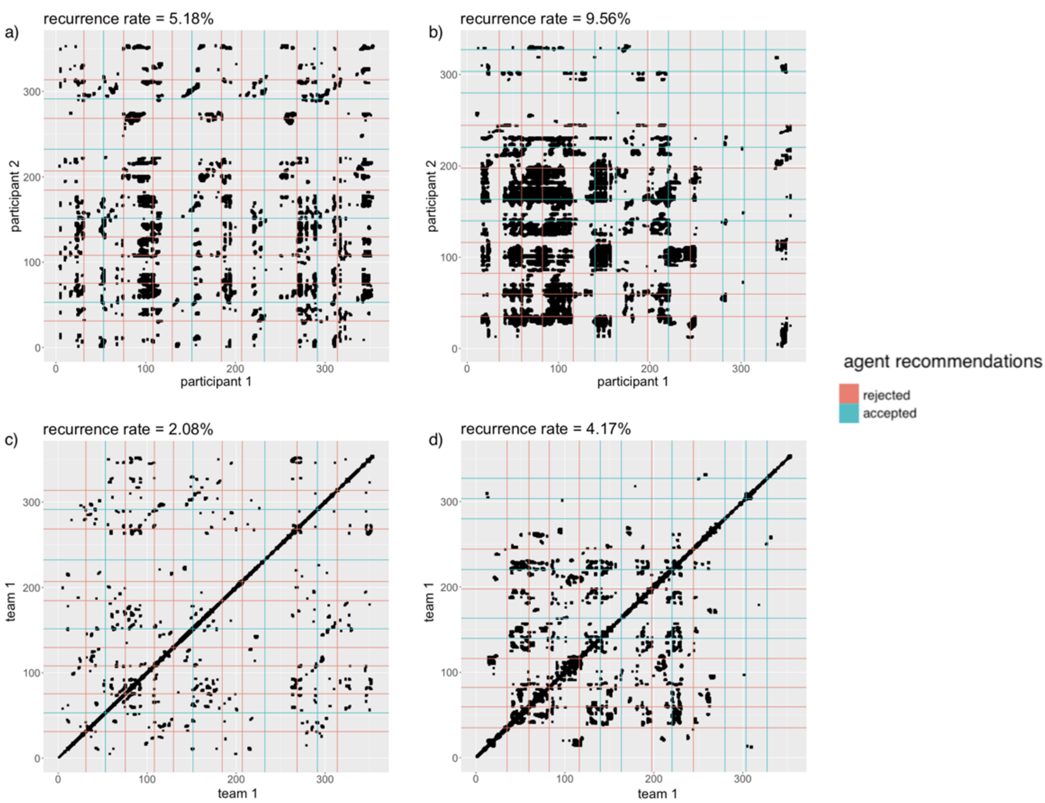

Fig. 4. Recurrence plot visualizations for a single team across two rounds. Dark points indicate recurrence between the x- and y-axis time series. Lines indicate moments when the agent makes a recommendation by placing a resource pin on the shared map, and colors indicate whether that recommendation was accepted or rejected by the team. (a & b) MdCRQA plot, where each axis denotes one participant's time series. (c & d) MdRQA plot, where both axes denote the team's multidimensional time series (compared against itself).

The MdCRQA plot in Figure 4a shows a sparse structure with loosely vertical recurrent segments. Vertical structures and the horizontal distance between them generally denote that participant 2 (y-axis) repeatedly visited the same state of neural activation that participant 1 (x-axis) exhibited at semi-regular intervals. These intervals loosely align with the moments of agent action, which may signal that participant 1 returned to this state to perform credibility assessments about the agent's recommendations. The plot in Figure 4b tells a different story, where much stronger coupling is observed earlier in this round (denoted by thicker sections of recurrence) and little to no coupling in the final minutes of the round. Similarly, more of the agent's pin placements were rejected in the first half of the round than in the second. One potential explanation would be that the teammates were more carefully assessing the agent's recommendations in the first half of the round, and thus had similar patterns of neural activation. Later in the round, they may have transitioned to placing their own pins independently, with less regard for and more trust in the agent's placements (indeed, we anecdotally observed these behavioral trends in video logs of studies). We highlight this plot specifically because it demonstrates how a 90 second recurrence window would capture these shifting dynamics, noting high recurrence in the first segment, decreasing in each subsequent 90 second block.

Switching to a system-level perspective, Figure 4c similarly depicts a sparse structure while Figure 4d shows a window of high recurrence followed by little to none later in the round. This aligns with the interpretation of the MdCRQA plots above, as the early window of higher recurrence coincides with more rejected agent pin placements than later windows. Again, the tight structure signifies the team repeatedly revisiting the same state (which could arise from repeated judgements of the agent's placements). Examining Figure 4b and Figure 4d together implies that later in the round, participants were exhibiting novel state configurations from each other and as a single system. These observations are largely qualitative but can be extracted in a number of quantitative metrics, such as the number of continuous vertical lines, the average length of vertical lines, the average length of diagonal lines, and so forth. We postulate that these and similar metrics taken from both sets of plots could identify unique states and configurations of teaming.

4.4 Conclusions, Limitations, and Future work

The results presented in this paper offer promising methods and parameters for future research monitoring trust and team processes in cooperative human-agent tasks. Specifically, we identify neural recurrence between individuals and neural regularity within a team across a 90-second lag window as diagnostic measures of interpersonal trust and team states during human-human-agent collaborations. From these highly interpretable metrics, we observed temporal behavior of cognitive and affective trust in line with previous theory, as indicated by interpersonal neural coupling. Applying these two powerful yet simple methods, MdCRQA and MdRQA, will help to reduce complexity of monitoring multiple actors in ongoing collaborations by elegantly characterizing team states and processes with minimal pre-processing, no data-intensive training, and no linearity assumptions. To the authors' knowledge, this is first demonstration of a model-free technique applied to brain data successfully picking apart these individual processes during a naturalistic, open-ended HAT task with more than one human.

We discuss six main limitations of this work to be addressed in future efforts. First, various experimental limitations arise due to the length of the study. The current iteration of NIRx sensors can cause discomfort over long periods due to the tightness of the head cap and pressure from optode probes against the scalp. While piloting this study, we found two hours to be the maximum amount of time participants could comfortably perform the task. It was therefore necessary to balance time spent wearing the sensors and time spent on post-task surveys, so we used only the top factor loading items to assess trust and team processes across 8 full rounds. We acknowledge that abbreviating surveys can reduce interpretability, and that future work should utilize a more robust and focused set of survey measures to disentangle observed effects. Second, the study presented is an ongoing effort at a single public university, limiting sample size and demographic representation. Next, we constrained our modeling efforts to uncover main effects while accounting for our experimental design and allowing for comparison between models, and thus did not perform a robust model selection procedure to obtain an ideal fit for each individual model. Fourth, our measurements of neural recurrence is limited to three cortical areas and far from comprehensive. We plan to further explore the functional implications of our results from a cognitive and affective neuroscience lens. Fifth, though we present a relatively open-ended task with few constraints on participants, we are limited in making inferences about generalizability until these techniques can be replicated on data from different tasks.

Lastly, we note the limits of our four proposed hypotheses in this study. The general goal of this investigation is twofold: 1) to identify measures of brain activity that link to the relevant constructs of trust and team processes as a step toward adaptive collaborative systems, and 2) to use an interpretable analysis metric to characterize the neural and temporal dynamics of trust and team processes in a human-agent team. We thus devised the present experiment to tackle both of these aspects with one set of analyses. For this reason, we opted to use recurrence rate as our only metric as it sufficiently summarizes a system's regularity without sacrificing interpretability. Additionally, it is simple to compute recurrence rate across an entire dataset or within a constrained window of time, an important quality for its proposed use in real-time computation and adaptation. We note this limitation as our hypotheses do not comprehensively address other recurrence models and metrics, such as determinism and entropy. These additional metrics would add a new layer to the interpretations of our experiment that warrants its own investigation focusing more specifically on the high dimensional state-spaces and complexities of dynamical systems.

These limitations highlight uncertainties and exciting new targets for future work. Our results identify compelling analyses and directions for future research seeking to empirically capture the complexities of human interaction. With this foundation, we aim to improve the inferential capability of adaptive systems by further refining these methods and eventually implementing real-time adaptations. There is also a myriad of additional recurrence features (section 4.3.1) yet to be investigated for monitoring team dynamics. The metrics identified in this paper may also be valuable targets for ongoing machine learning and decision-making research efforts. Specifically, we note that recurrence measures may be particularly useful as observations to be fed into a POMDP-driven system. The powerful simplicity of nonlinear recurrence quantification analyses combined with information-rich data from fNIRS sensors presents a unique potential to advance the current state of adaptive collaborative technology for both human-human and human-agent teaming.

ACKNOWLEDGMENTS

We would like to acknowledge the Army Research Office for supporting this research (W911NF-19-1-0401ARO).

REFERENCES

[1] Jason A. Aimone, Daniel Houser, and Bernd Weber. 2014. Neural signatures of betrayal aversion: an fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences* 281, 1782 (2014), 20132127.

[2] Kumar Akash, Wan-Lin Hu, Neera Jain, and Tahira Reid. 2018. A classification model for sensing human trust in machines using EEG and GSR. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 8, 4 (2018), 1–20.

[3] Mary Jean Amon, Hana Vrzakova, and Sidney K D'Mello. 2019. Beyond dyadic coordination: Multimodal behavioral irregularity in triads predicts facets of collaborative problem solving. *Cognitive science* 43, 10 (2019), e12787.

[4] Haoyu Bai, Shaojun Cai, Nan Ye, David Hsu, and Wee Sun Lee. 2015. Intention-aware online POMDP planning for autonomous driving in a crowd. In *2015 ieee international conference on robotics and automation (icra)*. IEEE, 454–460.

[5] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1 (2015), 1–48. https://doi.org/10.18637/jss.v067.i01

[6] Suzanne T. Bell, Shanique G. Brown, Anthony Colaneri, and Neal Outland. 2018. Team composition and the ABCs of teamwork. *American Psychologist* 73, 4 (2018), 349.

[7] Adella Bhaskara, Michael Skinner, and Shayne Loft. 2020. Agent transparency: a review of current theory and evidence. *IEEE Transactions on Human-Machine Systems* 50, 3 (2020), 215–224.

[8] Meghana A Bhatt, Terry Lohrenz, Colin F Camerer, and P Read Montague. 2012. Distinct contributions of the amygdala and parahippocampal gyrus to suspicion in a repeated bargaining game. *Proceedings of the National Academy of*

Sciences 109, 22 (2012), 8728–8733.

[9] Marta Bianciardi, Paolo Sirabella, Gisela E Hagberg, Alessandro Giuliani, Joseph P Zbilut, and Alfredo Colosimo. 2007. Model-free analysis of brain fMRI data by recurrence quantification. *NeuroImage* 37, 2 (2007), 489–503.

[10] Philip Bobko, Leanne Hirshfield, Lucca Eloy, Cara Spencer, Emily Doherty, Jack Driscoll, and Hannah Obolsky. 2022. Human-agent teaming and trust calibration: a theoretical framework, configurable testbed, empirical illustration, and implications for the development of adaptive systems. *Theoretical Issues in Ergonomics Science* (2022), 1–25. https://doi.org/10.1080/1463922X.2022.2086644

[11] Scott C Bunce, Kurtulus Izzetoglu, Hasan Ayaz, Patricia Shewokis, Meltem Izzetoglu, Kambiz Pourrezaei, and Banu Onaral. 2011. Implementation of fNIRS for monitoring levels of expertise and mental workload. In *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems: 6th International Conference, FAC 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011. Proceedings 6*. Springer, 13–22.

[12] Erin K Chiou and John D Lee. 2021. Trusting automation: Designing for responsivity and resilience. *Human factors* (2021), 00187208211009995.

[13] Sanjib Chowdhury. 2005. The role of affect-and cognition-based trust in complex knowledge sharing. *Journal of Managerial issues* (2005), 310–326.

[14] Moreno I Coco and Rick Dale. 2014. Cross-recurrence quantification analysis of categorical and continuous time series: an R package. *Frontiers in psychology* 5 (2014), 510.

[15] Julien Cohen-Adad, S Chapuisat, Julien Doyon, Serge Rossignol, J-M Lina, Habib Benali, and Frederic Lesage. 2007. Activation detection in diffuse optical imaging by means of the general linear model. *Medical Image Analysis* 11, 6 (2007), 616–629.

[16] Ana Cristina Costa, Robert A Roe, and Tharsi Taillieu. 2001. Trust within teams: The relation with performance effectiveness. *European journal of work and organizational psychology* 10, 3 (2001), 225–244.

[17] Adrian Curtin and Hasan Ayaz. 2018. The age of neuroergonomics: towards ubiquitous and continuous measurement of brain function with fNIRS. *Japanese Psychological Research* 60, 4 (2018), 374–386.

[18] Rick Dale, Riccardo Fusaroli, Nicholas D Duran, and Daniel C Richardson. 2013. The self-organization of human interaction. In *Psychology of learning and motivation*. Vol. 59. Elsevier, 43–95.

[19] Arwen H DeCostanza, Amar R Marathe, Addison Bohannon, A William Evans, Edward T Palazzolo, Jason S Metcalfe, and Kaleb McDowell. 2018. *Enhancing humanagent teaming with individualized, adaptive technologies: A discussion of critical scientific questions*. Report. US Army Research Laboratory Aberdeen Proving Ground United States.

[20] Angelika Dimoka. 2010. What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *Mis Quarterly* (2010), 373–396.

[21] Lucca Eloy, Emily J Doherty, Cara Ann Spencer, Philip Bobko, and Leanne Hirshfield. 2022. Using fNIRS to identify transparency-and reliability-sensitive markers of trust across multiple timescales in collaborative human-human-agent triads. *Frontiers in Neuroergonomics* (2022), 10.

[22] Lucca Eloy, Angela E.B. Stewart, Mary Jean Amon, Caroline Reinhardt, Amanda Michaels, Chen Sun, Valerie Shute, Nicholas D. Duran, and Sidney D'Mello. 2019. Modeling Team-Level Multimodal Dynamics during Multiparty Collaboration. In *2019 International Conference on Multimodal Interaction* (Suzhou, China) *(ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 244–258. https://doi.org/10.1145/3340555.3353748

[23] Charles R. Evans and Kenneth L. Dion. 1991. Group Cohesion and Performance. *Small Group Research* 22, 2 (1991), 175–186. https://doi.org/10.1177/1046496491222002

[24] Anne-Kathrin J. Fett, Paula M. Gromann, Vincent Giampietro, Sukhi S. Shergill, and Lydia Krabbendam. 2014. Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Social cognitive and affective neuroscience* 9, 4 (2014), 395–402.

[25] Megan M. Filkowski, Ian W. Anderson, and Brian W. Haas. 2016. Trying to trust: Brain activity during interpersonal social attitude change. *Cognitive, Affective,& Behavioral Neuroscience* 16, 2 (2016), 325–338.

[26] Darren Gergle and Alan T. Clark. 2011. See what i'm saying? using Dyadic Mobile Eye tracking to study collaborative reference. , 435–444 pages. https://doi.org/10.1145/1958824.1958892

[27] Ella Glikson and Anita Williams Woolley. 2020. Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals* 14, 2 (2020), 627–660.

[28] Gerald F Goodwin, Nikki Blacksmith, and Meredith R Coats. 2018. The science of teams in the military: Contributions from over 60 years of research. *American Psychologist* 73, 4 (2018), 322.

[29] Jamie C Gorman, Polemnia G Amazeen, and Nancy J Cooke. 2010. Team coordination dynamics. *Nonlinear dynamics, psychology, and life sciences* 14, 3 (2010), 265.

[30] Jamie C. Gorman, Nancy J. Cooke, Polemnia G. Amazeen, and Shannon Fouse. 2012. Measuring Patterns in Team Interaction Sequences Using a Discrete Recurrence Approach. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 54, 4 (2012), 503–517. https://doi.org/10.1177/0018720811426140

[31] Kunal Gupta, Ryo Hajika, Yun Suen Pai, Andreas Duenser, Martin Lochner, and Mark Billinghurst. 2019. In ai we trust: Investigating the relationship between biosignals, trust and cognitive load in vr. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. 1–10.

[32] Janine Viol Hacker, Michael Johnson, Carol Saunders, Amanda L Thayer, et al. 2019. Trust in virtual teams: A multidisciplinary review and integration. *Australasian Journal of Information Systems* 23 (2019).

[33] Jonathan L. Helm, David Sbarra, and Emilio Ferrer. 2012. Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion* 12, 4 (2012), 748–762. https://doi.org/10.1037/a0025036

[34] Leanne Hirshfield, Phil Bobko, Alex Barelka, Natalie Sommer, and Senem Velipasalar. 2019. Toward interfaces that help users identify misinformation online: Using fNIRS to measure suspicion. *Augmented Human Research* 4, 1 (2019), 1–13.

[35] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors* 57, 3 (2015), 407–434.

[36] Claudia-Lavinia Ignat, Quang-Vinh Dang, and Valerie L Shalin. 2019. The influence of trust score on cooperative behavior. *ACM Transactions on Internet Technology (TOIT)* 19, 4 (2019), 1–22.

[37] Patrick Jermann, Dejana Mullins, Marc-Antoine Nüssli, and Pierre Dillenbourg. 2011. Collaborative gaze footprints: Correlates of interaction quality. (2011).

[38] Patrick Jermann and Marc-Antoine Nüssli. 2012. Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. , 1125–1134 pages. https://doi.org/10.1145/2145204.2145371

[39] Gareth R Jones and Jennifer M George. 1998. The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of management review* 23, 3 (1998), 531–546.

[40] Steve Jones and Steve Marsh. 1997. Human-computer-human interaction: trust in CSCW. *ACM SIGCHI Bulletin* 29, 3 (1997), 36–40.

[41] Prasert Kanawattanachai and Youngjin Yoo. 2007. The impact of knowledge coordination on virtual team performance over time. *MIS quarterly* (2007), 783–808.

[42] Norbert L. Kerr and R. Scott Tindale. 2004. Group Performance and Decision Making. *Annual Review of Psychology* 55, 1 (2004), 623–655. https://doi.org/10.1146/annurev.psych.55.090902.142009

[43] Ioannis Kostavelis, Dimitrios Giakoumis, Sotiris Malassiotis, and Dimitrios Tzovaras. 2017. A pomdp design framework for decision making in assistive robots. In *Human-Computer Interaction. User Interface Design, Development and Multimodality: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I 19*. Springer, 467–479.

[44] Steve W. J. Kozlowski, Georgia T. Chao, James A. Grand, Michael T. Braun, and Goran Kuljanin. 2013. Advancing Multilevel Research Design:Capturing the Dynamics of Emergence. *Organizational Research Methods* 16, 4 (2013), 581–615. https://doi.org/10.1177/1094428113493119

[45] Frank Krueger, Kevin McCabe, Jorge Moll, Nikolaus Kriegeskorte, Roland Zahn, Maren Strenziok, Armin Heinecke, and Jordan Grafman. 2007. Neural correlates of trust. *Proceedings of the National Academy of Sciences* 104, 50 (2007), 20084–20089.

[46] John D. Lee and Katrina A. See. 2004. Trust in automation: Designing for appropriate reliance. *Human factors* 46, 1 (2004), 50–80.

[47] Angela Lombardi, Pietro Guccione, and Luigi Mascolo. 2014. Analysis of fmri data using the complex systems approach. In *20th IMEKO TC-4 International Symposium on Measurements of Electrical Quantities, Benevento*. 15–17.

[48] Poornima Madhavan and Douglas A. Wiegmann. 2007. Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science* 8, 4 (2007), 277–301.

[49] Caitlin E. V. Mahy, Louis J. Moses, and Jennifer H. Pfeifer. 2014. How and where: Theory-of-mind in the brain. *Developmental cognitive neuroscience* 9 (2014), 68–81.

[50] Michelle A Marks, John E Mathieu, and Stephen J Zaccaro. 2001. A temporally based framework and taxonomy of team processes. *Academy of management review* 26, 3 (2001), 356–376.

[51] N. Marwan, M. Carmenromano, M. Thiel, and J. Kurths. 2007. Recurrence plots for the analysis of complex systems. *Physics Reports* 438, 5-6 (2007), 237–329. https://doi.org/10.1016/j.physrep.2006.11.001

[52] John E Mathieu, Margaret M Luciano, and Leslie A DeChurch. 2018. Multiteam systems: The next chapter. *The SAGE handbook of industrial, work& organizational psychology* 2 (2018), 333–353.

[53] John E. Mathieu, Margaret M. Luciano, Lauren D'Innocenzo, Elizabeth A. Klock, and Jeffery A. LePine. 2020. The development and construct validity of a team processes survey measure. *Organizational Research Methods* 23, 3 (2020), 399–431.

[54] Roger C Mayer and Mark B Gavin. 2005. Trust in management and performance: Who minds the shop while the employees watch the boss? *Academy of management journal* 48, 5 (2005), 874–888.

[55] Oded Mayo and Ilanit Gordon. 2020. In and out of synchrony—Behavioral and physiological dynamics of dyadic interpersonal coordination. *Psychophysiology* 57, 6 (2020), e13574.

[56] Daniel J. McAllister. 1995. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of management journal* 38, 1 (1995), 24–59.

[57] Ryan McKendrick and Amanda Harwood. 2019. Cognitive workload and workload transitions elicit curvilinear hemodynamics during spatial working memory. *Frontiers in Human Neuroscience* 13 (2019), 405.

[58] Jason P. Mitchell, Mahzarin R. Banaji, and C. Neil Macrae. 2005. The Link between Social Cognition and Self-referential Thought in the Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience* 17, 8 (2005), 1306–1315. https://doi.org/10.1162/0898929055002418

[59] Guilherme Augusto Zimeo Morais, Joana Bisol Balardin, and João Ricardo Sato. 2018. fNIRS optodes' location decider (fOLD): a toolbox for probe arrangement guided by brain regions-of-interest. *Scientific reports* 8, 1 (2018), 1–11.

[60] Mustapha Mouloua and Peter A Hancock. 2019. *Human performance in automated and autonomous systems, two-volume set.* CRC Press.

[61] Takayuki Nozawa, Yukako Sasaki, Kohei Sakaki, Ryoichi Yokoyama, and Ryuta Kawashima. 2016. Interpersonal frontopolar neural synchronization in group communication: An exploration toward fNIRS hyperscanning of natural interactions. *NeuroImage* 133 (2016), 484–497. https://doi.org/10.1016/j.neuroimage.2016.03.059

[62] Richard V. Palumbo, Marisa E. Marraccini, Lisa L. Weyandt, Oliver Wilder-Smith, Heather A. McGee, Siwei Liu, and Matthew S. Goodwin. 2017. Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personality and Social Psychology Review* 21, 2 (2017), 99–141. https://doi.org/10.1177/1088868316628405

[63] Raja Parasuraman and Victor Riley. 1997. Humans and automation: Use, misuse, disuse, abuse. *Human factors* 39, 2 (1997), 230–253.

[64] Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens. 2008. Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of cognitive engineering and decision making* 2, 2 (2008), 140–160.

[65] Helen Pushkarskaya, Michael Smithson, Jane E. Joseph, Christine Corbly, and Ifat Levy. 2015. Neural correlates of decision-making under ambiguity and conflict. *Frontiers in behavioral neuroscience* 9 (2015), 325.

[66] Summer Rebensky, Kendall Carmody, Cherrise Ficke, Meredith Carroll, and Winston Bennett. 2022. Teammates instead of tools: The impacts of level of autonomy on mission performance and human-agent teaming dynamics in multi-agent distributed teams. *Frontiers in Robotics and AI* (2022), 102.

[67] Daniel C. Richardson and Rick Dale. 2005. Looking To Understand: The Coupling Between Speakers' and Listeners' Eye Movements and Its Relationship to Discourse Comprehension. *Cognitive Science* 29, 6 (2005), 1045–1060. https://doi.org/10.1207/s15516709cog0000_29

[68] Denise M. Rousseau, Sim B. Sitkin, Ronald S. Burt, and Colin Camerer. 1998. Not so different after all: A cross-discipline view of trust. *Academy of management review* 23, 3 (1998), 393–404.

[69] M. Salazar, Daniel Joel Shaw, M. Gajdoš, R. Mareček, Kristína Czekóová, Michal Mikl, and Milan Brázdil. 2021. You took the words right out of my mouth: Dual-fMRI reveals intra-and inter-personal neural processes supporting verbal interaction. *NeuroImage* 228 (2021), 117697.

[70] John Schaubroeck, Simon SK Lam, and Ann Chunyan Peng. 2011. Cognition-based and affect-based trust as mediators of leader behavior influences on team performance. *Journal of applied psychology* 96, 4 (2011), 863.

[71] Beau G Schelble, Christopher Flathmann, Nathan J McNeese, Guo Freeman, and Rohit Mallick. 2022. Let's Think Together! Assessing Shared Mental Models, Performance, and Trust in Human-Agent Teams. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–29.

[72] Adnan Shah and Abd-Krim Seghouane. 2014. An integrated framework for joint HRF and drift estimation and HbO/HbR signal improvement in fNIRS data. *IEEE transactions on medical imaging* 33, 11 (2014), 2086–2097.

[73] Kevin Shockley, Marie-Vee Santana, and Carol A. Fowler. 2003. Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance* 29, 2 (2003), 326–332. https://doi.org/10.1037/0096-1523.29.2.326

[74] Zachary Sunberg and Mykel Kochenderfer. 2018. Online algorithms for POMDPs with continuous state, action, and observation spaces. In *Proceedings of the International Conference on Automated Planning and Scheduling*, Vol. 28. 259–263.

[75] Harold Thimbleby, Steve Marsh, Steve Jones, and Andy Cockburn. 2018. Trust in CSCW. In *Computer-Supported Cooperative Work*. Routledge, 253–271.

[76] Preethi Vaidyanathan, Jeff Pelz, Cecilia Alm, Pengcheng Shi, and Anne Haake. 2014. Recurrence quantification analysis reveals eye-movement behavior differences between experts and novices. In *Proceedings of the symposium on eye tracking research and applications*. ACM Press, New York, New York, USA, 303–306. https://doi.org/10.1145/2578153.2578207

[77] Hana Vrzakova, Mary Jean Amon, Angela Stewart, Nicholas D Duran, and Sidney K D'Mello. 2020. Focused or stuck together: multimodal patterns reveal triads' performance in collaborative problem solving. In *Proceedings of the tenth international conference on learning analytics & knowledge*. 295–304.

[78] Hana Vrzakova, Mary Jean Amon, Angela EB Stewart, and Sidney K D'Mello. 2019. Dynamics of visual attention in multiparty collaborative problem solving using multidimensional recurrence quantification analysis. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14.

[79] Sebastian Wallot. 2019. Multidimensional Cross-Recurrence Quantification Analysis (MdCRQA) – A Method for Quantifying Correlation between Multivariate Time-Series. *Multivariate Behavioral Research* 54, 2 (2019), 173–191. https://doi.org/10.1080/00273171.2018.1512846

[80] Sebastian Wallot and Giuseppe Leonardi. 2018. Analyzing multivariate dynamics using cross-recurrence quantification analysis (crqa), diagonal-cross-recurrence profiles (dcrp), and multidimensional recurrence quantification analysis (mdrqa)–a tutorial in r. *Frontiers in psychology* 9 (2018), 2232.

[81] Min Wang, Aya Hussein, Raul Fernandez Rojas, Kamran Shafi, and Hussein A Abbass. 2018. EEG-based neural correlates of trust in human-autonomy interaction. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 350–357.

[82] Motoki Watabe, Hiroshi Ban, and Hiroki Yamamoto. 2011. Judgments about others' trustworthiness: An fMRI study. *Letters on Evolutionary Behavioral Science* 2, 2 (2011), 28–32.

[83] Sheila Simsarian Webber. 2008. Development of cognitive and affective trust in teams: A longitudinal study. *Small group research* 39, 6 (2008), 746–769.

[84] Charles L. Webber Jr and Joseph P. Zbilut. 2005. Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in contemporary nonlinear methods for the behavioral sciences* 94 (2005), 26–94.

[85] Wei Zheng, Bo Wu, and Hai Lin. 2018. Pomdp model learning for human robot collaboration. In *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 1156–1161.

Received July 2022; revised October 2022; accepted January 2023