

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/328797402>

Team Coordination and Effectiveness in Human–Autonomy Teaming

Article in IEEE Transactions on Human-Machine Systems · March 2019

DOI: 10.1109/THMS.2018.2877482

CITATIONS

12

READS

116

5 authors, including:



Mustafa Demir

Arizona State University

43 PUBLICATIONS 225 CITATIONS

[SEE PROFILE](#)



Aaron Likens

University of Nebraska at Omaha

24 PUBLICATIONS 217 CITATIONS

[SEE PROFILE](#)



Nancy J Cooke

Arizona State University

226 PUBLICATIONS 5,347 CITATIONS

[SEE PROFILE](#)



Polemnia Amazeen

Arizona State University

71 PUBLICATIONS 1,682 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Examining Human-Autonomy Team Interaction and Explicable Behavior [View project](#)



Keystroke Dynamics [View project](#)

Team Coordination and Effectiveness in Human-Autonomy Teaming

Mustafa Demir , *Member, IEEE*, Aaron D. Likens, Nancy J. Cooke, Polemnia G. Amazeen, and Nathan J. McNeese 

Abstract—In the past, team coordination dynamics have been explored using nonlinear dynamical systems (NDS) methods, but the relationship between team coordination dynamics and team performance for all-human teams was assumed to be linear. The current study examines team coordination dynamics with an extended version of the NDS methods and assumes that its relationship with team performance for human-autonomy teams (HAT) is nonlinear. In this study, three team conditions are compared with the goals of better understanding how team coordination dynamics differ between all-human teams and HAT and how these dynamics relate to team performance and team situation awareness. Each condition was determined based on manipulation of the pilot role: in the first condition (synthetic) the pilot role was played by a synthetic agent, in the second condition (control) it was a randomly assigned participant, and in the third condition (experimenter) it was an expert who used a role specific coordination script. NDS indices revealed that synthetic teams were rigid, followed by experimenter teams, who were metastable, and control teams, who were unstable. Experimenter teams demonstrated better team effectiveness (i.e., better team performance and team situation awareness) than control and synthetic teams. Team coordination stability is related to team performance and team situation awareness in a nonlinear manner with optimal performance and situation awareness associated with metastability coupled with flexibility. This result means that future development of synthetic teams should address these coordination dynamics, specifically, rigidity in coordination.

Index Terms—Human-autonomy teaming, nonlinear dynamical systems, synthetic agent, team coordination.

I. INTRODUCTION

RECENT algorithmic advances in the development of autonomous agents have enabled them to communicate and coordinate with humans and other autonomous agents [1]. These

autonomous agents can function—at least partially in a self-directed manner—outside of the sorts of situations that they were designed for by using intelligence-based capabilities [2]. As computing systems continue to advance, there is a push to consider autonomous agents as team members [3]. This role of autonomous agents as team members has started a paradigm shift from all-human teaming to *human-autonomy teaming* (HAT), the latter being teams composed of both humans and technology-based team members [4]–[6]. This inclusion of autonomy within the team context establishes a need to understand how hybrid HATs differ from all-human teams. The current work explores those differences within the context of a simulated unmanned air vehicle (UAV) task.

HAT involves at least one human operator and one automated system coordinating and collaborating interdependently over time in order to successfully complete a task [4]. The literature is divided, however, on the role that artificial intelligence plays within the team context: Some authors argue that technology-based team members are to some extent equivalent to their human counterparts as team members [7, p. 204]; [4], [5], [8] whereas others contend this designation is not warranted [9]. In the HAT context—an example of the former position—equal status among human and automated teammates has led to the redefinition of a team as an actor-agent community (e.g., [8]). The justification for redefining the team concept derives from the ability of autonomous agents to take initiative and give orders to both human and automated counterparts.

These autonomous agents make their own decisions about their actions during the team task. Therefore, without outside intervention, an autonomous agent can independently achieve goals and maintain good performance in highly dynamic environments by interacting (i.e., communication and coordination) with other humans or other agents [2], [10], [11]. On the other hand, Klein *et al.* [9] underline that the automated system's lack of intelligence is a large obstacle in its path to becoming a team member. Recent work supports that idea by showing how the autonomous agent's limited interactions can increase the cognitive demands placed on human teammates [12], [13]. Regardless of the status researchers ascribed to autonomous agents, artificial intelligence is an ever increasing presence within the team context [2], [14]–[16]. In this paper, we explore differences between human teams and HATs through the lens of interactive team cognition [3], [17]–[21]. In order to do so, we used an adaptive, control, thought–rationale (ACT-R) cognitive modeling-based synthetic teammate as an autonomous team member that uses a

Manuscript received December 14, 2017; revised May 30, 2018; accepted September 18, 2018. This work was supported by the Office of Naval Research under Award N000141110844 (Program Managers: M. Steinberg, P. Bello) and Award N000141712382 (Program Managers: M. Steinberg, M. Clark). The work of P. G. Amazeen was supported by the National Science Foundation under Grant BCS 1255922. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This paper was recommended by Associate Editor C. Burns. (*Corresponding author: Mustafa Demir.*)

M. Demir and N. J. Cooke are with the Human Systems Engineering, Arizona State University, Mesa, AZ 85212 USA (e-mail: mdemir@asu.edu; ncooke@asu.edu).

A. D. Likens and P. G. Amazeen are with the Department of Psychology, Arizona State University, Tempe, AZ 85287-1104 USA (e-mail: aaron.likens@asu.edu; amazeen@asu.edu).

N. J. McNeese is with the Human-Centered Computing, Clemson University, Clemson, SC 29634 USA (e-mail: mcneese@clemson.edu).

Digital Object Identifier 10.1109/THMS.2018.2877482

dialog management system to understand and respond to information requests from human team members [16], [22].

A. Teams as Nonlinear Dynamical Systems

The emphasis on team interactions is aligned with interactive team cognition (ITC) theory [17]. ITC considers team cognition as a process which, in contrast to individual cognition, can often be measured directly at the level of the team. Importantly, ITC also assumes that team cognition must be understood in the dynamic task context in which the team operates [17]. In ITC, interaction-based measures rely on patterns of team interaction that continuously change over time [20] and are indicators of team cognition and team performance [17]. In this case, team interactions are emergent properties of teams that are irreducible to the heterogeneous characteristics of individual team members, characterization that places team cognition within the realm of nonlinear dynamical systems (NDS).

The NDS approach to team cognition has revealed much about the way that teams communicate and coordinate in a broad range of contexts [18], [20], [21], [23]–[26]. For instance, Gorman, Amazeen, and Cooke (2010) applied NDS methods in a UAV task environment, wherein team membership changed (i.e., Mixed teams) or stayed the same (i.e., Intact teams) over successive experimental sessions. Remarkably, mixed teams were better able to adjust to unexpected perturbations, and this ability was linked to team level dynamics: Mixed teams adopted a globally stable pattern of communication while exhibiting strong temporal dependence. These findings support the ITC viewpoint by emphasizing the importance of contextualized team level process measures in the study of team cognition [17]. In the current study, we investigate the potential of NDS and ITC perspectives to capture the differential dynamics of all-human and human-synthetic teams.

II. CURRENT STUDY

A. Task and Roles

In the current study, three heterogeneous and interdependent team members photographed critical target waypoints. The three roles were as follows.

- 1) The navigator, who provides a dynamic flight plan with information regarding speed and altitude restrictions of the waypoints to other team members.
- 2) The pilot who controls the UAV by controlling its fuel and adjusting altitude and airspeed.
- 3) The photographer, who takes photos of each target waypoint and adjusts camera settings [27].

The basic team goal is to take photographs of ground targets during simulated reconnaissance missions by interacting with each other through a text-based communications system. Teams must also maintain all UAV flight and sensor systems and observe any restrictions associated with particular waypoints. The task was carried out over a series of five missions (each 40 min in length; [27]).

The study involved three conditions/configurations (each composed of ten teams). Conditions differed in the identity of the pilot teammate. The pilot was:

- 1) a synthetic teammate (i.e., a cognitively plausible ACT-R based computational model) in the *synthetic condition*;
- 2) a randomly assigned human participant in the *control condition*; and
- 3) an expert confederate who used a role specific coordination script in the *experimenter condition*.

For all three conditions, team members communicated with each other using a text-based interface and each member had access to a communication cheat sheet for effective communication with the pilot. In the synthetic condition, because of the synthetic teammate's limited language capability, the navigator and the photographer needed to ensure that all text messages to the synthetic teammate were neither ambiguous nor cryptic [28]. In the experimenter condition, the confederate pilot used a coordination script and asked the navigator and photographer scripted questions when needed to promote adaptive passing of information about the critical waypoints in a timely manner [22], [29].

B. Order Parameter of Team Communication

A sequence of the task specific communication–coordination events that have to occur at each target waypoint (i.e., information, negotiation, and feedback) were defined as the essence of coordination [30].

- 1) Information (I) was sent by the navigator to the pilot about the upcoming target waypoint.
- 2) The photographer and the pilot Negotiated (N) regarding the critical target waypoint's altitude, airspeed, and camera settings required to take a good photograph.
- 3) The photographer sent Feedback (F) to the other team members about the status of the target photo. In short, this task specific coordination pattern is called INF.

Based on the timing between components of the INF coordination pattern, the Kappa score (κ), which serves as an order parameter, is computed for each target waypoint (see Formula 1). This parameter describes the timing of the coordination among the individual components of the systems of the UAV team [18]

$$\kappa(t) = \left(\frac{t_{\text{feedback}} - t_{\text{information}}}{t_{\text{feedback}} - t_{\text{negotiation}}} \right) \quad (1)$$

where $\kappa(t)$ is computed at each time step, t : "where t_{feedback} is the time it takes to provide feedback about whether a good photo of the target was taken, $t_{\text{information}}$ is the time it takes to provide the initial information about the target, and $t_{\text{negotiation}}$ is the time it takes to negotiate about the airspeed and altitude in order to take a good photo. All of these dynamic parameters also depend on the current status of the UAV, such as whether the current altitude or airspeed is the same as the next target waypoint's altitude and airspeed, which reduces the negotiation time if the team members are aware of it.

Consistent with the ITC perspective, κ resides at the level of the team rather than the level of the individual. It reflects the systems' state changes—due to changing the control parameter (described as follows)—and captures the current state of team coordination as it fluctuates over time [18]. κ has some interesting properties: First, it has no units because all three constituent parts are measured in seconds and these units cancel in the

relation. Second, it contains two qualitatively different states: uncoordinated ($\kappa < 1$) and coordinated ($\kappa > 1$) with a transition point at ($\kappa = 1$) that differentiates the two. In this particular experimental context, the lower bound of κ is zero (theoretically, it could take on an infinite number of values, but these are the two states that have a practical use) [18], [31], [32].

In contrast, the control parameter is used as a catalyst to change the systems behavior (i.e., identifying dynamics of the order parameter) [33]. This study presents an examination of the order parameter (i.e., κ score) and control parameter (i.e., three team configuration) during performance in the UAV task. We also explore how team coordination dynamics differ across the three team compositions (synthetic, control, and experimenter) and how team coordination dynamics relate to the outcome measures: team performance and team situation awareness (TSA).

III. METHODOLOGY

A. Participants

In total, 30 teams (ten in each of the three conditions) completed the experiment. Within those 30 teams, 70 participants (60 males and 10 females, $M_{\text{age}} = 23.7$, $SD_{\text{age}} = 3.3$) were recruited from Arizona State University and surrounding areas, each completing an approximately eight-hour experiment as a team. The number of participants in each team depended on condition. Teams in the control condition had three randomly assigned participants whereas teams in the synthetic and experimenter conditions had two randomly assigned participants for the navigator and photographer roles. This is because the pilot position in the latter two conditions were filled by a nonparticipant—the synthetic teammate in the synthetic condition and a trained confederate in the experimenter condition. Having normal or corrected-to-normal vision and being fluent in English were the main requirements for participation, with which each participant being compensated \$10/h.

B. Materials and Apparatus

CERTT-II Test Bed: The experiment was conducted in the context of the cognitive engineering research on team tasks laboratory, unmanned aerial vehicle - synthetic task environment (CERTT UAV-STE). The CERTT UAV-STE simulates the teamwork aspects of ground control of a predator UAV [27], [34]. The team's task in this simulated environment is to detect and photograph target waypoints in a timely manner while maintaining all UAV flight and sensor systems and observing any restrictions associated with particular waypoints [27]. In this experiment, the CERTT-UAV-STE software was embedded in an updated version of the hardware infrastructure (CERTT-II). In order to support this experiment, CERTT-II provides new hardware infrastructure that has varied features: first, text chat capability for communications between team members, and second, eight new hardware consoles: four consoles for up to four team members and another four consoles for two experimenters who oversee the simulation, inject roadblocks, and make observations. The “texting experimenter” console has two embedded computers through that the experimenter can give ratings (for

target achievement, behavior, and situational awareness) and send text messages via chat window to the other three roles (pilot, navigator, and photographer). This console can also turn ON and OFF all the computers and software of the other consoles via a master control window. The other console, called the “non-texting experimenter” console, enables the experimenter to give ratings (for coordination, target achievement, and behavior) and can also follow the participants using a camera.

Synthetic Teammate: One goal of the synthetic teammate project is to create a synthetic teammate capable of human-like behavior in order to interact with other human team members. In the study, one of the team members was a synthetic teammate developed using the ACT-R cognitive modeling architecture [35], which is composed of five components [16]:

- 1) *language analysis* to accommodate a variety of English constructions;
- 2) *language generation* to choose possible utterances;
- 3) *dialog modeling* to recognize when communication is obligatory;
- 4) *situation modeling* for situation awareness; and
- 5) *agent-environment interaction* to fly the UAV between destinations in a way that is cognitively plausible.

At the beginning of the study, participants were informed that one of their teammates, i.e., the pilot, was a synthetic teammate that had limited communication capabilities. Therefore, in order to interact (i.e., communicate and coordinate) with the synthetic teammate effectively, the human team members needed to send messages in a constructive way, i.e., without any cryptic language or misspellings [28]. This instruction was given during the training process and also on the cheat sheet with examples. For instance, when the navigator sends the following message to the synthetic teammate, the synthetic teammate may not understand it; especially, the airspeed and altitude information: *The next waypoint is F-AREA which is a target with air 50–200 and alt. 2500–3500. The effective radius is 5.* The problems in this message are the words “air,” “alt.,” and that the numbers should be separated without a dash. Therefore, the navigator should send this message to the synthetic pilot as the following: *“The next waypoint is F-AREA. It is a target. The airspeed restriction is 50 to 200. The altitude is 2500 to 3500. The effective radius is 5.”* Otherwise, the synthetic pilot may ask questions about the altitude and airspeed.

Other Materials: For the hands-on training and the task, the photographer and the navigator had a communication “cheat sheet”, which shows examples of how to communicate with the synthetic teammate. Also, another “cheat sheet” and supplemental materials for each role were displayed at the corresponding workstations. The pilot role in the experimenter condition had a coordination script in order to push and pull information to other team members in a timely manner. Supplemental materials include role-specific rule summaries, screenshots of each station's displays, waypoint list for the navigator, and camera setting list and photo folder comparisons of good and bad photos for the photographer. Experimenters had paper checklists for the experiment setup (starting the experimenter and participant consoles, briefing, training, and mission task, respectively), and data for each session were archived via an external hard drive (this data included all the team performance and process data, NASA task

TABLE I
EXPERIMENTAL SESSION

| Session Order and Task Duration | |
|---|--|
| 1) Consent forms (15 min) | 6) Mission 3 (40 min) |
| 2) PowerPoint (30 min) and hands on training (30 min) | 7) Mission 4 (40 min) |
| 3) Mission I (40 min) | 8) Mission 5 (40 min) |
| 4) NASA TLX-I / Knowledge Session-I (30 min) | 9) NASA TLX-II / Knowledge Session-II (30 min) |
| 5) Missions 2 (40 min) | 10) Post-Check Procedure (15 min) |

Note. From the hands-on training through the post-check procedure, a 15-min break was applied after each task. Therefore, the total approximate time for the experimental session was eight hours.

load index (TLX) Workload, teamwork and taskwork knowledge, and demographics).

C. Procedure

The order of the eight-hour experiment procedure is depicted in Table I. Accordingly, when participants arrived, they read and signed an informed consent form and were randomly assigned to one of the two team member roles (navigator or photographer) in the synthetic and the experimenter conditions. In the control conditions, three team members were assigned one of three roles (pilot, navigator, or photographer). In all three conditions, the pilot (either human or synthetic) was isolated in one room and the navigator and the photographer in another. In the synthetic condition, the navigator and the pilot knew that the pilot was the synthetic teammate, and in the control condition, the navigator and the photographer knew that the pilot was a human participant. The navigator and the photographer were seated in locations separated by partitions such that they did not have face-to-face contact.

After that, participants received a briefing and then a 30-min role-specific skills training using an interactive PowerPoint training program which includes an audio section. After the interactive training session, a 30-min hands-on practice training was started. During the practice session, the participants were required to perform specific tasks, and the experimenters checked off participants' demonstration of certain skills using a basic skills checklist. After the participants had reached criterion levels of performance on their individual tasks, they started the first mission.

Each team completed five 40 min missions. In each mission, participants interacted with each other and with their task environment based on their task specific roles. The pilot controls the UAV by observing role specific settings (i.e., fuel, flaps, and gears) and systems directly related to controlling the flight path of the UAV. The pilot flies the UAV along waypoints in accordance with a mission route developed by the navigator. The navigator has access to data regarding the location of and any restrictions on critical waypoints, and navigates the pilot based on the data. Using this information, the navigator plans a sequence of waypoints called the mission route. The photographer's main objective is to take good photos of the critical waypoints (which are predesignated as targets). To achieve this, the photographer must monitor and control the systems and settings relevant to the UAV sensor equipment (i.e., infrared,

synthetic aperture radar, and electro-optical photography equipment). In addition, the photographer must work with the pilot to coordinate camera settings with UAV location flight settings to ensure that the UAV is within a certain radius (e.g., five miles) of a target to take a good photo.

After completing the first mission, a knowledge session was conducted in order to test participants for team and task related knowledge relevant to the UAV STE both individually and as a team. After that, the NASA TLX, which is a subjective workload assessment [36], was administered to measure six workload components, including mental, physical, temporal demand, performance, effort, and frustration.

Each team completed five missions with Missions 1 through 4 of comparable workload and Mission 5 of high workload. Teams had to photograph 11 target waypoints for Mission 1 and Mission 3, 12 targets for Mission 2, 13 targets for Mission 4, and 20 targets for Mission 5. Target waypoints fell into areas called restricted operating zones (ROZ) boxes. The first four missions contained five ROZ boxes, whereas there were seven ROZ boxes for Mission 5. These missions had restrictions on airspeed, altitude, and the ROZ boxes (flying within an effective radius for entering the operating zone).

During each mission, the navigator would be alerted to new targets which were hidden targets (or "roadblocks") on the navigator's map [37], [38]. These hidden targets were novel changes in the dynamic task environment that disrupted routine team coordination [38]. For instance, when the team entered a specific ROZ, these hidden targets were provided by the experimenter to the navigator as a chat message: "Enemy forces have just left from the target WP8. You are clear to enter area to take a photo." There were three roadblocks in Missions 1 and 2, four in Missions 3 and 4, and five in Mission 5. In this task, team situation awareness (TSA) was measured via these "roadblocks" [22].

Throughout the missions, team communication was observed by texting and nontexting experimenters who checked appropriate boxes on the situation awareness and coordination loggers in real time. After visiting each target waypoint, both experimenters independently rated the team's process behaviors (i.e., adequacy of team member behaviors such as communication) on a scale of one to five, five being the best. After completing the last mission, the team received a second round of knowledge and NASA TLX measures, and later on had their demographic data (e.g., age, education, etc.) collected. Finally, the participants were debriefed about the experiment.

D. Measures

A *team performance score* was collected for each of the five missions. The team performance score is a composite score composed of a set of mission variables, including the time each individual spent in alarm and warning states, number of successfully photographed targets. Each variable has associated penalty points—weighted in advance to correspond with task importance—that were deducted from the maximum score of 1000. In terms of the workload differences of each mission, no penalty was incurred for teams photographing a smaller proportion of targets in high workload missions [31].

Findings from this study about team performance were that experimenter teams out-performed both the synthetic and control teams (the latter two performed equivalently) [22]. Also, in this task, there was a significant mission effect. However, only team performance in the experimenter condition increased from Mission 1–4. This indicates that there was a learning effect in the experimenter condition. During the last mission, the team performance significantly decreased because of high workload (i.e., injected more target waypoints).

Three process measures are considered in this study: One of them, *team communication flow* (TCF) is a multivariate binary measure recorded once each minute, for each team member to indicate if at least one message was sent ($TCF = 1$) or not ($TCF = 0$) by each respective team member. *Kappa score* (κ), the second team process measure, is considered based on the timing among the components of the INF coordination sequence and serves as an order parameter (See Formula 1) [18].

The third process measure is TSA was the total number of successfully completed roadblocks. Roadblocks were *ad hoc* target waypoints that take place within the scenario at set times within each mission. During each mission, teams were sometimes presented with “roadblocks”, such as the introduction of a new target waypoint. Completion of the roadblocks was used as a measure of TSA. The triggering mechanism for these roadblocks is based on each team’s position relative to the waypoints in the mission; as such, the number of waypoints triggered in each mission may vary from team to team [31]. In terms of overcoming roadblocks (TSA), the experimenter teams overcame more than the synthetic and control teams, who overcame the same number of roadblocks [13], [22].

IV. DATA ANALYSIS AND RESULTS

In this study, the following analyses were applied to understand different characteristics of team coordination dynamics.

- 1) Attractor reconstruction to visualize the dynamics of the teams by using a κ score at the team level.
- 2) Stability analysis (Lyapunov exponents) for each coordination order parameter κ at the team level.
- 3) Surrogate analysis to address whether the observed dynamics were an artifact of the short κ time series or occurred randomly.
- 4) Joint recurrence quantification analysis (JRQA) to extract the % determinism (DET) measure on the multivariate TCF measure.

In this study, the programming software MATLAB was used to conduct the largest Lyapunov exponent calculation. JRQA was carried out in R version 3.2.3 [40] using the cross recurrence quantification analysis package [41] for calculating DET measure.

A. Attractor Reconstruction

Attractor reconstruction was used to visualize and measure team coordination dynamics in this task [18], [41]. Based on Taken’s [42] theorem, one can recover a system’s dynamical structure (i.e., reconstruct the attractor) from a one-dimensional (1-D) signal (in this study, this signal is the κ time series) and a

set of independent, time-delayed versions of itself. The reconstructed attractor exposes the dynamical structure that produced the patterns observed in the original 1-D signal.

In this study, the attractor was reconstructed for each team’s κ series, by first estimating two embedding parameters: the optimal *time delay* (τ), and the embedding dimension (m) [39], [43]. τ identifies the lag for which the original signal is maximally different from itself. These lagged versions are then used as the dimensions (m) in the phase space that the signal is unfolded onto. Following standard practice, τ was estimated as the first minimum of the average mutual information function [39], [43]. The selection of m followed the false nearest neighbors method outlined by [43]. This process surveys data points and their neighbors in dimensions ranging within spaces of increasing dimension. The goal is to find “false neighbors”, that is, points that separate when examined in a higher dimension. Per convention, m was determined as the lowest dimension where the percentage of false neighbors ≤ 10 .

Multivariate analysis of variance was conducted see how time lag ($M = 1.80$, $SD = 1.24$) and embedding dimension ($M = 2.83$, $SD = 0.87$) differ across the conditions (i.e., synthetic, control, and experimenter). Box’s M test (4.57) shows that there were no significant differences between the covariance matrices ($p = 0.27$). Therefore, the assumption of homogeneity of covariance was not violated, and Wilks Λ is an appropriate test to use. According to Levene’s test, the assumption of homogeneity of variances was not violated for both of the dependent variables (for τ : $F(2, 27) = 2.31$, $p = 0.12$; for m : $F(2, 27) = 0.24$, $p = 0.79$). The findings indicate that there were no significant differences across the conditions, $F(4, 52) = 0.79$, $p = 0.53$, Wilks $\Lambda = 0.88$. Thus, these two parameters (τ and m) are sufficient to reconstruct the attractor across different teams within the same task. That is, these nonsignificant results show the reliability of the used methods in attractor reconstruction for different teams performing in the same task.

B. Stability Analysis

We also evaluated team coordination stability by estimating the largest Lyapunov (λ) exponent from the reconstructed attractors. The λ measures the exponential rate of divergence of two nearby trajectories on the attractor [43], [44]: stable ($\lambda_1 < 0$), unstable ($\lambda_1 > 0$), and metastable or parallel trajectories ($\lambda_1 \approx 0$) of team coordination [45]. The magnitude of λ relates to the speed with which a dynamic system reaches a (stable or unstable) equilibrium point. A system with large negative λ reaches an equilibrium state quickly and will have difficulty moving away from it (i.e., the system becomes rigid), while a system with large positive λ will quickly become unstable leading to a breakdown in communication coordination and, perhaps, a jeopardized mission. Systems with λ close to zero will begin to meander and then stabilize (if λ is negative) or meander and destabilize (if λ is positive). Systems that exhibit either of these metastable situations are typically quick and agile and, thus, are likely to perform well [46].

Reconstructed attractors for three teams are presented in Fig. 1. Fig. 1 shows that differently composed teams (e.g., synthetic, control, experimenter) may also differ in their temporal

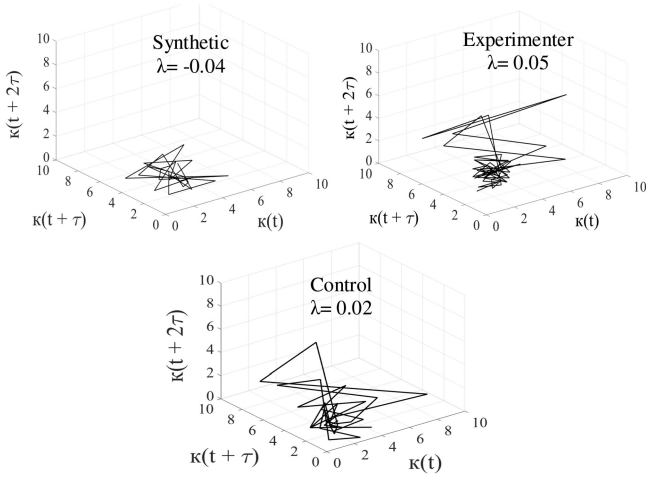


Fig. 1. Example reconstructed attractors from three UAV teams (synthetic, control, and experimenter) which provide means to geometrically show chaotic behavior: a 3-D phase space as coordinates for the 3-D space $[\kappa(t), \kappa(t+\tau), \kappa(t+2\tau)]$.

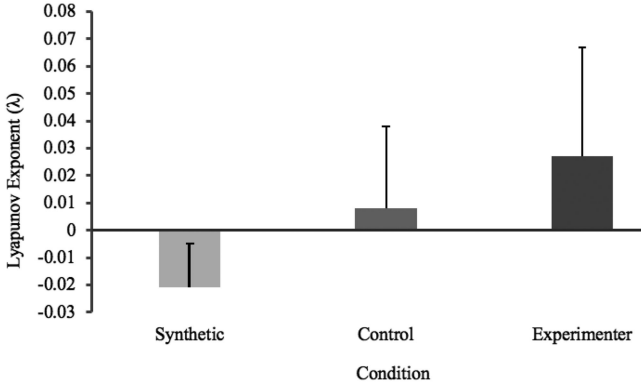


Fig. 2. Mean the largest lyapunov exponents (λ) across the conditions (synthetic < control = experimenter) (vertical lines + standard deviation).

dynamics; λ estimates provide converging evidence of that observation. The synthetic team's coordination was focused on a small part of phase space with less variability and a more stable appearance ($\lambda_{\text{Syn}} = -0.04$). On the other hand, the control ($\lambda_{\text{Cont}} = 0.02$) and experimenter ($\lambda_{\text{Exp}} = 0.05$) teams demonstrated more variability and less stability in Fig. 1.

A one-way analysis of variance (ANOVA) was used to further investigate if λ varied as function of team composition. Levene's F test indicated that the homogeneity of variance was violated, and therefore, an adjusted F ratio from Welch statistics is reported. The results for λ indicate that the difference among the conditions was statistically significant, $F(2, 15.7) = 9.22$, $MS_e = 0.001$, $p < 0.05$, $\eta^2 = 0.34$. Followup pairwise comparisons (based on LSD test) revealed that synthetic teams showed more stable coordination dynamics ($MS_{\text{Syn}} = -0.02$, $SD_{\text{Syn}} = 0.02$) than either control ($M_{\text{Cont}} = 0.01$, $SD_{\text{Cont}} = 0.03$, $p < 0.05$) or experimenter teams ($M_{\text{Exp}} = 0.03$, $SD_{\text{Exp}} = 0.04$, $p < 0.05$; see Fig. 2). Despite the apparent trend in Fig. 2, however, the stability of the control team and the experimenter teams did not differ reliably ($p = 0.15$). Thus, consistent with previous research, these results suggest that team coordination

TABLE II
SUMMARY OF REGRESSION ANALYSES FOR LYAPUNOV EXPONENT PREDICTING TEAM PERFORMANCE AND TEAM SITUATION AWARENESS

| Relationship between: | | B | $SE\ B$ | β | t | $p\ value$ | % Variance |
|-----------------------------------|---------------------------|---------|---------|---------|-------|------------|------------|
| Linear (λ) | | 1209 | 800 | 0.44 | 1.51 | 0.142 | 0.079 |
| λ and team performance | Quadratic (λ^2) | -1648 | 15730 | -0.30 | -1.05 | 0.304 | |
| λ and situation awareness | Linear (λ) | 17.76 | 6.08 | 0.77 | 2.92 | 0.017 | |
| | Quadratic (λ^2) | -270.59 | 119.44 | -0.59 | -2.27 | 0.032 | 0.240 |

Note. " B " and " $SE\ B$ " refer to unstandardized regression coefficient and its Standard Error, respectively, while " β " refers to standardized regression coefficient.

dynamics do vary as a function of team composition with the synthetic teams being stable.

C. Predicting Team Performance and Situation Awareness via the Largest Lyapunov

In order to predict the relationship between λ and three outcome measures (team performance score and TSA), regression analysis was applied to each outcome separately. For λ , all linear terms were positively related with outcome measures, and all quadratic terms were negatively related with the outcome measures. Only λ and TSA had a statistically significant relationship: the linear relationship indicates that every one-standard deviation increase in λ (holding constant other variables) is related to a 0.77 standard deviation increase in TSA. The quadratic relationship indicates that a one-unit standard deviation increment in the λ (holding constant other variables) was related with a -0.59 -unit decrement in the relationship between the λ and TSA (see Table II). That is, for lower values, λ had a positive relationship with TSA, however, as λ increases this relationship becomes negative.

D. Surrogate Analysis

Surrogate analysis is a bootstrapping method that represents a more stringent null hypothesis for presence of significant dynamic coordination structure than comparison to a theoretical value alone [47]. Therefore, in order to address whether the observed dynamics were an artifact of a short time series, surrogate analyses were applied on the κ scores [40]. We generated randomly shuffled surrogates of the observed κ scores to compare with observed κ scores, and obtained 19 randomly shuffled surrogates for each κ scores for a 95% confidence level [45].

Each surrogate had the same marginal statistical properties (mean, standard deviation) as its parent κ but was randomly sequenced. The results indicate that the absolute values of the observed λ_1 were significantly different from the surrogate values, $t(29) = 2.25$, $p < 0.05$ (two-tailed). This means that the observed dynamics of λ_1 were not artifacts of short or noisy κ .

E. Joint Recurrence Quantification Analysis

Joint recurrence quantification is the multivariate extension of the well-studied recurrence quantification analysis (e.g., RQA; [49], [50]). The underlying theory of JRQA and attractor reconstruction is the same; however, JRQA and its univariate

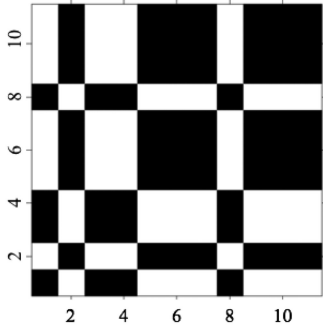


Fig. 3. Example recurrence plot.

counterpart provide additional important information about underlying temporal dynamics of observed time series. In addition, RQA is more versatile and can be applied to both continuous and categorical data. In this paper, we focus only on the categorical form of JRQA, but we refer readers to more general treatments (e.g., [49], [50]).

The basis for RQA (and JRQA) is the recurrence plot (RP) [51], a valuable visual tool for uncovering when a system revisits a region of phase space. In categorical RQA, the RP captures when particular states or sequences of states are revisited over time. We illustrate the construction of the categorical RP with a simple example. Consider a simple binary time series, $x(t) = [1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0]$. Here, the value at $x(1)$ is repeated at $x(3)$, $x(4)$, and $x(8)$. Similarly, the value at $x(3)$ is repeated at $x(4)$ and $x(8)$. The RP in Fig. 3 gives a visual summary of these patterns, as well as repetitions involving zeros. Conceptually, categorical RPs are constructed by placing a time series on both the horizontal and vertical axes of a graph. The main diagonal results when the series is plotted against itself. Off-diagonal points are plotted when the series repeats itself at a later point in time. The example given earlier concerning points $x(1)$, $x(3)$, $x(4)$, and $x(8)$, is visually depicted in Fig. 3 by tracing upwards from the bottom-left corner to the top-left corner of the plot. Points are plotted each time the value of $x(1)$ repeats. Construction of *joint recurrence plots* (JRPs) follows directly from this univariate case. In fact, the JRP is nothing more than the pointwise product (i.e., Hadamard product) of all respective univariate RPs.

RQA numerically describes the patterns observed in recurrence and joint recurrence plots. DET (a measure of determinism or predictability of the system which measures how organized communication patterns are) was extracted from the JRQA as the focal variable of interest in this study (see Marwan *et al.*, 2007 for other recurrence metrics). DET is defined as the ratio of recurrence points forming diagonal lines to all recurrent points in the upper triangle [49], and it measures the amount of organization in communication behavior. Essentially, this measure shows how the recurrent points are distributed. Time periods where a system repeats a sequence of states are represented by diagonal lines on the recurrence plots. This can be seen in Fig. 3 as a 3-point line that results when the sequence of three 0 s, beginning with $x(5)$, is repeated from the position, $x(9)$. Such diagonal lines would be prevalent in highly deterministic systems, but even mildly deterministic systems can show short

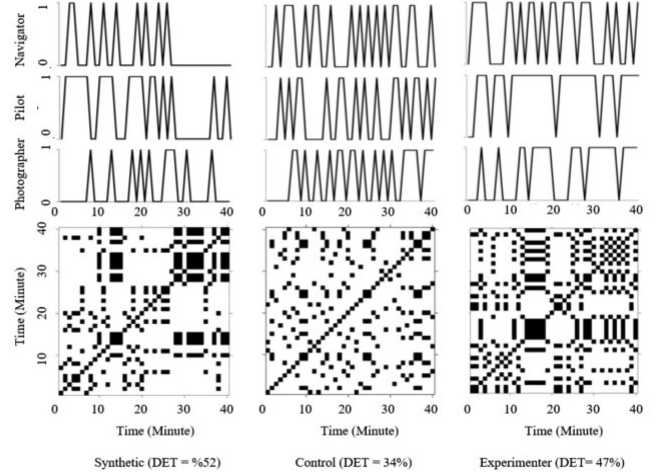


Fig. 4. Example joint recurrence plots for three UAV teams' interactions in three conditions - length 40 min: synthetic (DET: 52%), control (DET: 34%), and experimenter teams (DET: 47%).

periods of repetitive states, though these systems are considered less orderly than highly deterministic [52]. The DET is expressed as the following formula [49]:

$$DET = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=1}^N lP(l)} \quad (2)$$

where l is the diagonal line length considered when its value is $\geq l_{min}$ and $P(l)$ is the probability distribution of line lengths. This rate of determinism for a time series is a percentage which ranges from 0% (i.e., the time series never repeats) to 100% (i.e., the time series repeats perfectly). Therefore, we assume that the determinism rate could take any value from the full range of percentages.

We applied JRQA to the multivariate binary communication flow data (i.e., sent time stamp from each UAV mission) to visualize and quantify recurrent structure in each team's communication events. In Fig. 4, we give three example JRPs for three UAV teams' interactions for three conditions (three-code sequences that are 40 min in length) for synthetic (DET = 52%), control (DET = 34%), and experimenter (DET = 47%). Each JRP was calculated based on the team communication flow (each team members' message sent time); they are also depicted in Fig. 4 (navigator, pilot, and photographer). As a reminder, if any of these roles sent a message in any minute, it was coded one, otherwise it was coded as zero on the y-axis (i.e., discrete code). In this case, the synthetic team has higher determinism than the other two conditions, and determinism is higher for the experimenter team than for the control. One finding to note is that the synthetic team's higher determinism rate is mainly due to times when the three team members are silent (e.g., at between 30 to 35 min in Fig. 4). Also, the navigator in the synthetic team sent messages less frequently to the other team members compared to the experimenter condition, even though task requirements state that the person in this role should interact with pilot. The control team, which is less deterministic (less predictable/less structured), shows a more flexible communication pattern than the other two teams. This is not surprising given that the pilot role is assigned to a randomly selected participant (Fig. 4).

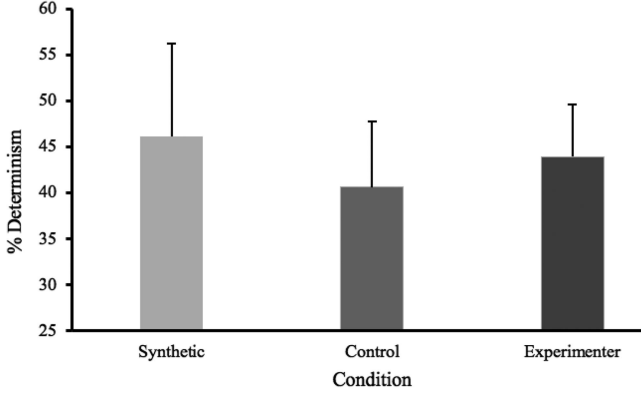


Fig. 5. Determinism across the conditions (control < experimenter < synthetic) (vertical lines + standard deviation).

TABLE III
SUMMARY OF REGRESSION ANALYSES FOR DETERMINISM PREDICTING TEAM PERFORMANCE AND TEAM SITUATION AWARENESS

| Relationship between: | | <i>B</i> | <i>SE B</i> | β | <i>t</i> | <i>p</i> value | % Variance |
|-----------------------------|-------------------------------|----------|-------------|---------|----------|----------------|------------|
| DET and team performance | Linear (DET) | 158.63 | 92.18 | 7.04 | 1.72 | 0.096 | 0.098 |
| | Quadratic (DET ²) | -1.82 | 1.06 | -7.05 | -1.72 | 0.096 | |
| DET and situation awareness | Linear (DET) | 1.07 | 0.77 | 5.70 | 1.39 | 0.175 | 0.098 |
| | Quadratic (DET ²) | -0.01 | 0.01 | -5.88 | -1.44 | 0.163 | |

Note. “*B*” and “*SE B*” refer to unstandardized regression coefficient and its Standard Error, respectively, while “ β ” refers to standardized regression coefficient.

In order to determine whether the conditions differed with respect to DET over time, we performed a 3 (condition) \times 5 (mission) split-plot ANOVA. Mauchly’s test indicates that the assumption of sphericity for the repeated measure (mission) was satisfied, $\chi^2(9) = 13.4$, $p = 0.15$, $\eta^2 = 0.82$. The ANOVA results indicate that the condition main effect was significant, $F(2, 27) = 5.65$, $MS_e = 67.8$, $p < 0.05$, $\eta^2 = 0.30$. The mission main effect, $F(4, 108) = 0.15$, $MS_e = 63.2$, $p = 0.96$, and the condition by mission interaction effect were not significant, $F(8, 108) = 0.69$, $MS_e = 63.2$, $p = 0.70$. According to pairwise comparisons (based on LSD test), the synthetic teams ($M_{Syn} = 46.1$, $SD_{Syn} = 10.1$) had higher DET than both control ($M_{Cont} = 40.6$, $SD_{Cont} = 7.16$, $p < 0.05$) and experimenter teams ($M_{Exp} = 43.9$, $SD_{Exp} = 5.70$, $p = 0.20$). The experimenter teams also had higher DET than control teams ($p = .053$; see Fig. 5). These findings indicate that the synthetic teams demonstrated more rigid behavior than the control and the experimenter teams, whereas the experimenter teams demonstrated more rigid behavior than the control condition but moderately flexible.

F. Predicting Team Performance and Situation Awareness via Determinism

In order to examine the relationship between determinism and three outcome measures (team performance score and TSA), we first applied growth curve modeling to the mission variable. However, we found no effect of time (either linear and quadratic) on team performance. Therefore, we took the team average of

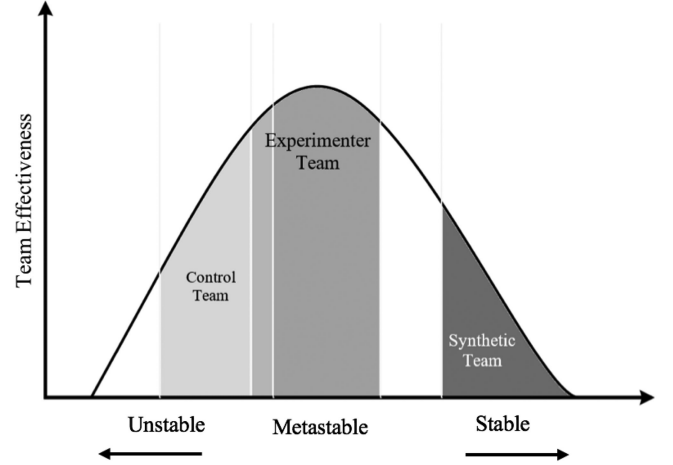


Fig. 6. Observed relationship between team stability and team effectiveness.

DET, and conducted regression analysis on each team level outcome separately. Although neither of those models reached conventional levels of statistical significance, marginal linear and quadratic relationships were observed between DET and team performance. If replicable, the sign of the coefficients for quadratic term indicate a trend similar to the one observed between λ and TSA: for lower values, DET had a positive relationship with TSA, however, as DET increases, this relationship becomes negative (see Table III).

V. DISCUSSION AND CONCLUSION

The goals of this study were: first, to better understand how team coordination dynamics differ between all-human teams and HATs, and second, learn if these differential dynamics are related to team performance and TSA. NDS methods were applied to team communication flow and κ score. Thus, the overall NDS methods captured two experimental results. First, based on DET findings, there was a clear difference across the conditions in terms of team coordination stability: synthetic teams were stable, followed by the experimenter teams, who were metastable, and control teams, who were unstable. Largest Lyapunov exponents findings support the finding that synthetic teams were more stable than the other teams. Second, coordination stability is quadratically related to team performance and situation awareness, consistent with the idea that metastable coordination is optimal for team performance.

The current results suggest an inverted “U”-shaped relationship between team stability and team effectiveness (reflected in team performance and TSA scores). In this hypothetical relationship (Fig. 6), which is supported by the DET and only partially supported by λ results, the idea is that best performing teams strike a balance between stability and instability. Operating on either extreme hinders team performance in a cognitively demanding task environment.

The synthetic teams seemed to operate in the stability region characterized in Fig. 6. The synthetic teammate was an ACT-R based computational model, and the human teammates were trained on how to interact with the synthetic teammate:

they were given a communication script telling them how their messages should be structured, and they also learned how the synthetic teammate interacted over time. Thus, the synthetic teams were rigid in their team coordination dynamics and failed to adapt to the dynamic task environment, ultimately resulting in less effectiveness (i.e., poor team performance and poor TSA) for this particular task context. In contrast, the experimenter teams tended to demonstrate metastable coordination dynamics because the confederate pilot used an “if-then” script and acted as a forcing function, requesting information from the other team members when it was not forthcoming. The experimenter teams were able to adapt their coordination to the task environment, resulting in good team effectiveness (the middle region of Fig. 6). The control teams, without the benefit of an experienced pilot, were the most unstable of the three conditions. We speculate that based on the inverted “U”-shaped model, as all-human teams (i.e., control) learn, their coordination should become more stable and approach optimal coordination and performance (depicted by the overlapping region in the bell-shaped curve in Fig. 6). In contrast, synthetic teams need to develop flexibility to also approach the optimal point. That is, if HATs are to function optimally, then any autonomous agent added to the team must be able to interact with its team members in a timely, adaptive, and constructive manner.

In conclusion, the relation between team coordination stability and team performance is discussed within the HAT context. Although the current synthetic teammate (serving as pilot) provides an example of rigid coordination, in terms of team effectiveness, synthetic teams performed comparably to control teams [22]. From previously reported findings in this research [12], [13], we know that the synthetic teams had trouble anticipating each other’s needs. With this study, we discovered that the synthetic teams had rigid team coordination dynamics and, thus, were less able to adapt to change in the environment. The failures to anticipate the information needs of others, coupled with rigid dynamics, seems related to poor performance for this particular dynamic task context. Future versions of the synthetic teammate should improve its abilities to anticipate the needs of others.

ACKNOWLEDGMENT

The authors also acknowledge the assistance of S. M. Shope, Sandia Research Corporation, who integrated the synthetic agent and the testbed, and C. Gibbons, who gave helpful comments on the NDS methods. Some of the data reported in this paper were reported in the first author’s dissertation and in a technical report [53].

REFERENCES

- [1] K. Dautenhahn, “A paradigm shift in artificial intelligence: Why social intelligence matters in the design and development of robots with human-like intelligence,” in *50 Years of Artificial Intelligence*, M. Lungarella, F. Iida, J. Bongard, and R. Pfeifer, Eds. Berlin, Germany: Springer, 2007, pp. 288–302.
- [2] M. R. Endsley, “Autonomous horizons: System autonomy in the Air Force—A path to the future,” in *Proc. Dept. Air Force Headquarters Air Force*, Washington DC, USA, Autonomous Horizons AF/ST TR 15-01, Jun. 2015.
- [3] S. M. Fiore and T. J. Wiltshire, “Technology as teammate: Examining the role of external cognition in support of team cognitive processes,” *Frontiers Psychol.*, vol. 7, Oct. 2016.
- [4] H. M. Cuevas, S. M. Fiore, B. S. Caldwell, and L. Strater, “Augmenting team cognition in human-automation teams performing in complex operational environments,” *Aviation Space Environmental Med.*, vol. 78, no. 5, pp. B63–B70, May 2007.
- [5] J. Langan-Fox, J. M. Canty, and M. J. Sankey, “Human-automation teams and adaptable control for future air traffic management,” *Int. J. Ind. Ergon.*, vol. 39, no. 5, pp. 894–903, Sep. 2009.
- [6] A. Schulte, D. Donath, and D. S. Lange, “Design patterns for human-cognitive agent teaming,” in *Engineering Psychology and Cognitive Ergonomics*, D. Harris, Ed. New York, NY, USA: Springer, 2016, pp. 231–243.
- [7] K. Sycara and M. Lewis, “Integrating intelligent agents into human teams,” in *Team Cognition: Understanding the Factors that Drive Process and Performance*, E. Salas and S. M. Fiore, Eds. Washington, DC, USA: Amer. Psychological Assoc., 2004, pp. 203–231.
- [8] N. Wijngaards, M. Kempen, A. Smit, and K. Nieuwenhuis, “Towards sustained team effectiveness,” in *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems*, O. Boissier, J. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowski, J. S. Sichman, and J. Vázquez-Salceda, Eds. Berlin, Germany: Springer, 2006, pp. 35–47.
- [9] G. Klein, D. D. Woods, J. M. Bradshaw, R. R. Hoffman, and P. J. Feltovich, “Ten challenges for making automation a ‘team player’ in joint human-agent activity,” *IEEE Intell. Syst.*, vol. 19, no. 6, pp. 91–95, Nov. 2004.
- [10] U. Krogmann, “From automation to autonomy-trends towards autonomous combat systems,” NATO Science and Technology Organization, France, Unclassified RTO MP-44, 1999.
- [11] L. C. Schooley, B. P. Zeigler, F. E. Cellier, and F. Y. Wang, “High-autonomy control of space resource processing plants,” *IEEE Control Syst.*, vol. 13, no. 3, pp. 29–39, Jun. 1993.
- [12] M. Demir, N. J. McNeese, and N. J. Cooke, “Team communication behaviors of the human-automation teaming,” in *Proc. IEEE Int. Multi-Disciplinary Conf. Cogn. Methods Situation Awareness Decis. Support*, 2016, pp. 28–34.
- [13] M. Demir, N. J. McNeese, and N. J. Cooke, “Team situation awareness within the context of human-autonomy teaming,” *Cogn. Syst. Res.*, vol. 46, pp. 3–12, Dec. 2017.
- [14] M. T. Cox, “Goal-Driven autonomy and question-based problem recognition,” in *Proc. 2nd Annu. Conf. Adv. Cogn. Syst.*, Palo Alto, CA, USA, 2013, pp. 29–45.
- [15] R. Parasuraman, M. Barnes, K. Cosenzo, and S. Mulgund, “Adaptive automation for human-robot teaming in future command and control systems,” *Int. C2 J.* vol. 1, no. 2, pp. 43–68, 2007.
- [16] J. Ball *et al.*, “The synthetic teammate project,” *Comput. Math. Organ. Theory*, vol. 16, no. 3, pp. 271–299, Sep. 2010.
- [17] N. J. Cooke, J. C. Gorman, C. W. Myers, and J. L. Duran, “Interactive team cognition,” *Cogn. Sci.*, vol. 37, no. 2, pp. 255–285, Mar. 2013.
- [18] J. C. Gorman, P. G. Amazeen, and N. J. Cooke, “Team coordination dynamics,” *Nonlinear Dyn. Psychol. Life Sci.*, vol. 14, no. 3, pp. 265–289, Jul. 2010.
- [19] J. C. Gorman, “Team coordination and dynamics two central issues,” *Current Directions Psychol. Sci.*, vol. 23, no. 5, pp. 355–360, Oct. 2014.
- [20] A. D. Likens, P. G. Amazeen, R. Stevens, T. Galloway, and J. C. Gorman, “Neural signatures of team coordination are revealed by multifractal analysis,” *Soc. Neurosci.*, vol. 9, no. 3, pp. 219–234, 2014.
- [21] J. C. Gorman, N. J. Cooke, P. G. Amazeen, and S. Fouse, “Measuring patterns in team interaction sequences using a discrete recurrence approach,” *J. Human Factors Ergonom. Soc.*, vol. 54, no. 4, pp. 503–517, Aug. 2012.
- [22] N. J. McNeese, M. Demir, N. J. Cooke, and C. Myers, “Teaming with a synthetic teammate: Insights into human-autonomy teaming,” *Human Factors*, vol. 60, no. 2, pp. 262–273, 2018.
- [23] P. J. Ramos-Villagrasa, P. Marques-Quinteiro, J. Navarro, and R. Rico, “Teams as complex adaptive systems: Reviewing 17 years of research,” *Small Group Res.*, vol. 49, no. 2, pp. 135–176, Apr. 2018.
- [24] S. J. Guastello, “Nonlinear dynamics of team performance and adaptability in emergency response,” *Human Factors*, vol. 52, no. 2, pp. 162–172, Apr. 2010.
- [25] S. J. Guastello, “Nonlinear dynamical systems for theory and research in ergonomics,” *Ergonomics*, vol. 60, no. 2, pp. 167–193, Feb. 2017.
- [26] T. J. Wiltshire, J. E. Butner, and S. M. Fiore, “Problem-solving phase transitions during team collaboration,” *Cogn. Sci.*, vol. 42, no. 1, pp. 1–12, Feb. 2017.
- [27] N. J. Cooke and S. M. Shope, “Designing a synthetic task environment,” in *Scaled Worlds: Development, Validation, and Application*, L. R. E. Schiflett, E. Salas, and M. D. Covert, Eds. Surrey, U.K.: Ashgate, 2004, pp. 263–278.

- [28] M. Demir, N. J. McNeese, N. J. Cooke, J. T. Ball, C. Myers, and M. Freiman, "Synthetic teammate communication and coordination with humans," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, 2015, vol. 59, pp. 951–955.
- [29] M. Demir, N. J. McNeese, N. J. Cooke, and C. Myers, "The synthetic teammate as a team player in command-and-control teams," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, 2016, vol. 60, pp. 116–116.
- [30] N. J. Cooke, J. C. Gorman, J. L. Duran, and A. R. Taylor, "Team cognition in experienced command-and-control teams," *J. Exp. Psychol. Appl.*, vol. 13, no. 3, pp. 146–157, Sep. 2007.
- [31] N. J. Cooke *et al.*, "Acquisition and retention of team coordination in command-and-control," Jul. 2007.
- [32] J. C. Gorman, "Team coordination dynamics in cognitively demanding environments," Ph.D. dissertation, Dept. Psychol., New Mexico State Univ., Las Cruces, NM, USA, 2006.
- [33] J. A. S. Kelso and D. A. Engstrom, *The Complementary Nature*. Cambridge, MA, USA: MIT Press, 2006.
- [34] N. J. Cooke and S. M. Shope, "Synthetic task environments for teams: CERT's UAV-STE," in *Handbook of Human Factors and Ergonomics Methods*, N. Stanton, A. Hedge, K. Brookhuis, E. Salas, and H. Hendrick, Eds. Boca Raton, FL, USA: CRC Press, 2005, pp. 41–46.
- [35] J. R. Anderson, *How can the Human Mind Occur in the Physical Universe?*. New York, NY, USA: Oxford Univ. Press, 2007.
- [36] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Mashkati, Eds. Amsterdam, The Netherlands: North Holland, 1988.
- [37] N. J. Cooke, J. C. Gorman, and L. J. Rowe, "An ecological perspective on team cognition," in *Team Effectiveness In Complex Organizations: Cross-Disciplinary Perspectives and Approaches*, E. Salas, G. F. Goodwin, and C. S. Burke, Eds. Evanston, IL, USA: Routledge, 2008, pp. 157–182.
- [38] J. C. Gorman, N. J. Cooke, and J. L. Winner, "Measuring team situation awareness in decentralized command and control environments," *Ergonomics*, vol. 49, no. 12–13, pp. 1312–1325, 2006.
- [39] H. Abarbanel, *Analysis of Observed Chaotic Data*, 1st ed. Berlin, Germany: Springer, 1996.
- [40] R Development Core Team, "R: A Programming Environment for Data Analysis and Graphics (Version 3.3.2)," Oct. 31, 2016. [Online]. Available: <https://www.r-project.org/>. Accessed on: Jan. 24, 2017.
- [41] M. I. Coco and R. Dale, with contributions of J. D. Dixon and J. Nash, *CRQA: Cross-Recurrence Quantification Analysis for Categorical and Continuous Time-Series*. 2015.
- [42] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, vol. 898, D. A. Rand and L. S. Young, Eds. New York, NY, USA: Springer-Verlag, 1981, pp. 230–242.
- [43] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D, Nonlinear Phenom.*, vol. 65, no. 1, pp. 117–134, May 1993.
- [44] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D, Nonlinear Phenom.*, vol. 16, no. 3, pp. 285–317, Jul. 1985.
- [45] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2 ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [46] J. A. S. Kelso, "The complementary nature of coordination dynamics: Self-organization and agency," *Nonlinear Phenom. Complex Syst.*, vol. 5, no. 4, pp. 364–371, 2002.
- [47] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyné Farmer, "Testing for nonlinearity in time series: the method of surrogate data," *Physica D, Nonlinear Phenom.*, vol. 58, no. 1, pp. 77–94, Sep. 1992.
- [48] D. Delignières, T. Deschamps, A. Legros, and N. Caillou, "A methodological note on nonlinear time series analysis: Is the open-and closed-loop model of Collins and De Luca (1993) a statistical artifact?," *J. Motor Behav.*, vol. 35, no. 1, pp. 86–96, Mar. 2003.
- [49] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Phys. Rep.*, vol. 438, no. 5–6, pp. 237–329, Jan. 2007.
- [50] C. L. Webber and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *J. Appl. Physiol.*, vol. 76, no. 2, pp. 965–973, Feb. 1994.
- [51] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *EPL Europhys. Lett.*, vol. 4, no. 9, pp. 973–977, 1987.
- [52] L. K. Allen, C. Perret, A. Likens, and D. S. McNamara, "What'D you say again?: Recurrence quantification analysis as a method for analyzing the dynamics of discourse in a reading strategy tutor," in *Proc. 7th Int. Learn. Analytics Knowl. Conf.*, New York, NY, USA, 2017, pp. 373–382.
- [53] N. J. Cooke, M. Demir, and N. J. McNeese, "Synthetic teammates as team players: Coordination of human and synthetic teammates," Cognitive Engineering Research Institute, Mesa, AZ, USA, Tech. Rep. N000141110844, May 2016.



Mustafa Demir (M'15) received the Ph.D. degree in simulation, modeling and applied cognitive science with a focus on team coordination dynamics in human-autonomy teaming from Arizona State University, Tempe, Arizona, in Spring 2017.

He is currently a Postdoctoral Research Associate with Dr. Nancy Cooke at Ira A. Fulton Schools of Engineering, Arizona State University, working in direct collaboration. His current research interests include Human-Autonomy teaming, team cognition, complex systems, and advanced statistical modeling.



Aaron D. Likens received the Ph.D. degree in cognitive science from Arizona State University, Tempe, Arizona with a focus on time series modeling of multiscale phenomena.

He is a Research Associate working with Dr. Nicholas Stergiou at the Center for Research in Human Variability, the Department of Biomechanics, the University of Nebraska at Omaha. His current research interests include development and application of dynamical systems theory with diverse application in human movement science and psychology.



Nancy J. Cooke received the Ph.D. degree in cognitive psychology from New Mexico State University, Las Cruces, NM, USA, in 1987.

She is a Professor of human systems engineering, Arizona State University and is the Science Director of the Cognitive Engineering Research Institute, Mesa, AZ, USA. She directs the ASU Center for Human, Artificial Intelligence, and Robot Teaming. She specializes in the development, application, and evaluation of methodologies to elicit and assess individual and team cognition.



Polemnia G. Amazeen received the Ph.D. degree in psychology from the Center for the Ecological Study of Perception and Action, the University of Connecticut, Storrs, CT, USA, in 1996, and completed post-doctoral training from the faculty of human movement science, the Vrije Universiteit, Amsterdam, The Netherlands.

She is an Associate Professor with the Department of Psychology, Arizona State University, Tempe, Arizona. She specializes in the analysis of coordination across people (social interactions) and within

people (bimanual and motor-respiratory coordination) using the tools of dynamical systems analysis.



Nathan J. McNeese received the Ph.D. degree in information sciences and technology with a focus on team decision making, cognition, and computer supported collaborative work from The Pennsylvania State University, State College, PA, USA, in the fall of 2014.

He is an Assistant Professor and the Director of the Team Research Analytics in Computational Environments Research Group within the division of Human-Centered Computing in the School of Computing, Clemson University, Clemson, SC, USA.

His current research interests include the study of team cognition and technology, human-machine teaming, and the development/ design of human-centered collaborative tools and systems.