Interaction Flexibility in Artificial Agents Teaming with Humans

Patrick Nalepka (patrick.nalepka@mq.edu.au)
Department of Psychology, Centre for Elite Performance, Expertise and Training, Macquarie University
Sydney, NSW 2109 Australia

Jordan P. Gregory-Dunsmore (jordan.gregory-dunsmore@students.mq.edu.au)
Department of Psychology, Macquarie University
Sydney, NSW 2109 Australia

James Simpson (james.simpson4@hdr.mq.edu.au)
Department of Psychology, Macquarie University
Sydney, NSW 2109 Australia

Gaurav Patil (gaurav.patil@mq.edu.au)
Department of Psychology, Centre for Elite Performance, Expertise and Training, Macquarie University
Sydney, NSW 2109 Australia

Michael J. Richardson (michael.j.richardson@mq.edu.au)
Department of Psychology, Centre for Elite Performance, Expertise and Training, Macquarie University
Sydney, NSW 2109 Australia

Abstract

Team interaction involves the division of labor and coordination of actions between members to achieve a shared goal. Although the dynamics of interactions that afford effective coordination and performance have been a focus in the cognitive science community, less is known about how to generate these flexible and adaptable coordination patterns. This is important when the goal is to design artificial agents that can augment and enhance team coordination as synthetic teammates. Although previous research has demonstrated the negative impact of model-based agents on the pattern of interactions between members using recurrence quantification methods, more recent work utilizing deep reinforcement learning has demonstrated a promising approach to bootstrap the design of agents to team with humans effectively. This paper explores the impact of artificial agent design on the interaction patterns that are exhibited in human-autonomous agent teams and discusses future directions that can facilitate the design of human-compatible artificial agents.

Keywords: human-autonomy teaming (HAT); deep reinforcement learning (DRL); interactive team cognition (ITC); recurrence quantification analysis (RQA)

Introduction

Workplaces in the 21st century are defined by sociotechnical systems (Fiore & Wiltshire, 2016) whereby coordination, collaboration, and communication is mediated by the use of assistive and automated systems to facilitate group performance (O'Neill et al., 2020). A more recent research endeavor includes the use of artificial systems as *autonomous agents* who contribute meaningfully as a synthetic teammate in what is now referred to as human-autonomy teaming (HAT) (McNeese et al., 2018). As opposed to assistive and automated systems that require human intervention, autonomous artificial agents are self-directed in their behaviors (O'Neill et al., 2020).

In the literature, the study of HAT has been a prominent focus in military- or emergency-related simulations (O'Neill et al., 2020). In these applications, the role of artificial agents is to participate actively in a team or group's functioning. To perceive these agents as teammates, they need to embody perceptual-motor dynamics that allow for fluid interactions with team members (Lorenz et al., 2016). This is especially important for situations where the goal for synthetic teammates is to either replace humans entirely (Ball et al., 2010; Demir et al., 2019) or to provide high-fidelity team training (Rigoli et al., 2020).

To understand how to design artificial agents capable of teaming with humans, a useful starting point is to understand how coordination in human teams is possible. Importantly, interpersonal coordination and cooperative ('joint') action has been well studied in the past three decades (Knoblich et al., 2011; Schmidt et al., 1990). This includes research exploring the dynamics and self-organized stabilities of interpersonal and group coordination (Schmidt & Richardson, 2008) and the information-processing mechanisms involved in predicting and understanding the actions of co-actors (Vesper et al., 2011). These various approaches have also begun to investigate the use of artificial agents to test model predictions (Dumas et al., 2018; Harry & Keller, 2019) and to steer human behavior in collaborative contexts (Nalepka et al., 2019).

Quantifying the Dynamics of Team Interaction

Within the theoretical umbrella of 'interactive team cognition' (ITC), team behaviors and decision-making processes are understood as activities that emerge from the

ongoing interaction of team members within the task context (Cooke et al., 2013). The processes that operate at the team level can be measured via the pattern of interactions between team members, which, in turn, can be used to assess a team's functioning and ability to respond to task perturbations (Gorman et al., 2020).

A popular tool to assess team functioning is via recurrence quantification analysis (RQA) methods (Marwan et al., 2007; Webber & Zbilut, 1994). As will be expanded upon later, RQA is a tool capable of understanding the dynamics of disparate systems – such as the interactions between humans and artificial agents. In the psychological and cognitive sciences, RQA has been applied to understand the interactive dynamics of various social systems – for example, caregiver/child language dynamics (Dale & Spivey, 2006), interpersonal conflict (Paxton & Dale, 2017), and collaborative problem-solving (Wiltshire et al., 2018).

RQA involves the construction of a 2-dimensional recurrence plot (RP) which is a visualization tool for representing repetitions (or re-occurrences) of a dynamical system's state across time (see Figure 2). In univariate RQA, both axes of the plot represent the system's behavioral time series such that values along the main diagonal are recurrent due to the plot's construction and any recurrences away from the main diagonal represent repetitions of states at different time points. The patterning of recurrent points on RPs allows investigators to uncover the underlying properties of a system (e.g., the system's stability).

An extension of RQA is the use of *joint* RPs (JRPs) which allows for the study of *interactions* between dynamical systems (Marwan et al., 2007), such as between members of a team who may have different roles or be of disparate design (e.g., human, artificial agent) (Demir et al., 2019). JRPs are constructed by combining the RPs of each team member (via the Hadamard product). The patterning of recurrences on the resultant JRP provides a description of the dynamics of the interactions between members of a team.

A statistic that can be obtained from JRPs is the extent to which a team's interactions follow stable, repeatable patterns. This is computed as the proportion of recurrent points that form diagonal line structures on the JRP, referred to as %Determinism (or %DET). %DET is inversely related to the flexibility of interactions between team members, such that higher %DET indicates a stable, prescribed pattern of interactive behavior while lower %DET indicates flexible variation in how team members respond to recurrent states.

RQA, and the measurement of interaction flexibility, has been applied to the study of HAT in a three-person air reconnaissance task which consists of a pilot, navigator, and photographer (Demir et al., 2019). These experiments compared the impact of the pilot being replaced by a synthetic teammate on the interaction dynamics and performance of these teams. This artificial agent implemented the Adaptive, Control of Thought-Rational (ACT-R; Anderson, (2007)) cognitive modelling architecture and was capable of conversing with team members via a text-based platform. The results showed that the interaction patterns of teams containing the artificial agent were more rigid (higher %DET) than all-human teams which resulted in poorer performance on the task, as well as behavioral passivity. Observations by the researchers indicate that the shortcomings of the synthetic teammate were due to its architecture – namely, it lacked teamwork skills that anticipated the needs of its teammates in a timely manner.

Human-Aware Artificial Agents

The challenge of symbolic-based architectures such as the ACT-R model is that it lacks the flexibility needed by artificial agents to adapt to the natural variability found in human teams. The opposing alternative are neural network-based, model-free methods which train artificial agents to develop action patterns based on experience. The most promising work in this area has been the design of artificial agents using deep, multi-layer neural network architectures (Mnih et al., 2015). Embedding the agent within a reinforcement learning problem (Sutton & Barto, 2018), whereby the agent learns state-action mappings (i.e., policy) which maximize a reward function, has had success in solving cognitively demanding game-based tasks (Mnih et al., 2015; Vinyals et al., 2019).

The integration of deep neural networks with reinforcement learning (referred to as Deep Reinforcement Learning, DRL) excel at competitive or agent-only cooperative environments, but do not ensure effective collaboration when teaming with humans (Carroll et al., 2019). This is because in competitive settings DRL agents can exploit the non-optimal nature of human gameplay by behaving optimally. However, optimal action by the artificial agent in cooperative settings does not ensure optimal behavior at the team level – as it also assumes human users will behave similarly.

To account for the shortcomings of DRL in HAT, Carroll et al., (2019) tested whether exposing artificial agents to human-like behavior would facilitate the adoption of policies that would enable fluid interactions with human players. This was tested by varying the underlying model of the agent's partner during training in a two-agent task modeled after the popular video game series *Overcooked* (Ghost Town Games, Cambridge, UK; see also Wu et al., (2021) who also used a similar task environment). In this task, agents, taking on the role of chefs, must coordinate their actions to retrieve onions, place them in a stock pot to cook soup, and then retrieve a bowl to deliver the soup to the serving station (see Figure 1 for one of the task configurations used). The task environment was cramped so that collisions between players were probable, which would lead to inefficiencies in coordination.

Pertinent to this paper, Carroll et al., (2019) trained an artificial agent using DRL which was either exposed to a similarly-trained agent (the 'self-play' condition) or alongside a partner which embodied a model trained to imitate previously collected human-human gameplay (by using behavioral cloning; the 'human-aware' condition). Following training, human participants teamed with either

the self-play or human-aware agent. Participants who interacted with the human-aware agent could cook and serve more bowls of onion soup compared to those working with the self-play agent. Additionally, participants teaming with the human-aware agent also performed better than participants who teamed with another human in certain environments. This suggests that the 'human-aware' agent can select actions that enhances HAT.

Current Study

Carroll et al., (2019) speculated from qualitative observations that the human-aware agent was more adaptive to participants than the self-play agent. This study sought to quantify the dynamics which enabled more effective teaming between participants and the human-aware agent. The current study compared the different approaches to train artificial agents using DRL and investigated their effect on the interaction dynamics with human participants. In this way, this study sought to integrate the recent advances in DRL within the theoretical work of ITC and the recent research in quantifying the interaction dynamics during HAT (Cooke et al., 2013; Demir et al., 2019).

In this study, participants were recruited to work alongside the self-play and human-aware agents in a modified version of the *Overcooked* task developed by Carroll et al., (2019). Participants were tasked to complete multiple rounds with the agents with trials lasting two minutes. For both the participant and the artificial agent, RPs were constructed for each player's behavior across time. Resultant JRPs and the amount of interaction flexibility (quantified by %DET) was assessed at the team level. The hypothesis was that the increased performance observed in human-'human aware' agent teams was due to increased flexibility in the interactions between players (lower %DET) compared to participants teaming with the self-play agent. Additionally, consistent with theory (Demir et al., 2019), variation in interaction flexibility was expected to predict team performance.

Method

Participants

Forty-three participants took part in this experiment (31 Female), with participant age ranging from 18 to 36 years (M = 20.51, SD = 4.76). All participants were first-year Psychology undergraduate students who completed this experiment in exchange for course credit. The experiment took place online and required participants to have a keyboard and stable internet connection. The study was approved by the Macquarie University Institutional Review Board.

Materials and Design

The 'Onion Soup' Game. Participants completed a virtual cooking task designed by Carroll et al., (2019) and modeled after the popular video game series *Overcooked* (Ghost Town Games, Cambridge, UK).

Figure 1: Illustration of task environment. The environment was modified from Carroll et al., (2019) to include the addition of rubbish bins around the perimeter for participants and agents to dispose of unwanted items.

The environment was a modified version of the *coordination ring* level developed by Carroll et al., (2019) (see Figure 1). Participants controlled the green-hat 'chef' avatar, while the artificial agent controlled the blue-hat chef. The avatars could move UP, DOWN, LEFT, or RIGHT, using the directional arrow keys, along the tan-colored kitchen floor which surrounded a centrally located countertop space where objects can be placed. Objects could be picked up and placed by orienting the avatar towards the object and pressing the SPACEBAR. Interactable objects included onions (bottom left of Figure 1) and bowls (left of Figure 1). Human and artificial agent actions were processed at 4 Hz.

The goal of the game was for players to cook and serve as many bowls of onion soup as possible within 120 s (a total of 480 states). Players received a point for every soup that was served. This was done by placing three onions inside one of two stock pots (pictured top right in Figure 1), waiting for them to cook for 5 s, then bringing a bowl to the pot to retrieve the soup, and then placing the bowl of onion soup at the serving station (the grey square at bottom of Figure 1). In case a player wished to discard an item, the perimeter of the environment contained ten 'rubbish bins' for which items could be disposed.

Artificial Agents. Two artificial agents were tested in this experiment. Both agents were originally developed in the study by Carroll et al., (2019) and were controlled by a deep neural network trained using proximal policy optimization (PPO). Reinforcement learning was used to reward the agents for every bowl of soup that was served. Thus, the agents were incentivized to develop control policies that maximized the number of bowls of onion soup that can be served.

The agents differed in their exposed training environment. The *self-play* agent worked alongside a similar agent and the two together were trained to develop a control policy which maximizes the number of bowls of soup to be served. The *human-aware* agent worked alongside an agent embodying a human model of task behavior. This human model was

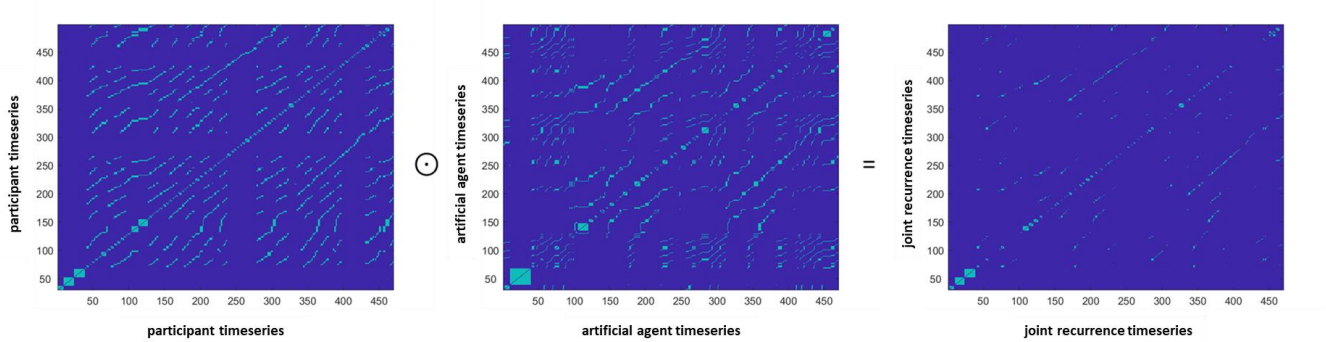

Figure 2. Example construction of joint recurrence plot (JRP). Each axis represents the timeseries of player states. A recurrent point is plotted whenever the same state intersects along the plot. The individual RPs are then combined to construct the JRP. Interaction flexibility is operationalized as the percentage of recurrent points which fall along diagonal lines on the JRP (i.e., the plot's %Determinism, see Marwan et al., 2007).

developed using behavioral cloning, a form of imitation learning, which trains a policy to mimic the decisions of a training data set. In this way, the artificial agent trained to develop a control policy which maximizes reward while accounting for the decision-making of the human model. See Carroll et al., (2019) for more details.

Design and Measures. The study implemented a within-subjects design whereby participants completed the 'Onion Soup' game with both the self-play and human-aware agents. The order of which artificial agent participants interacted with first was counterbalanced. The following measures were used for the analyses:

Game Score. Performance was measured via the number of bowls of onion soup that were served during each trial.

Interaction Flexibility. The pattern of interactions between the participant and the artificial agents was assessed by joint categorical RQA. For each player on each trial, a RP was constructed which plotted the behavioral timeseries of states on each axis. Each unique state, consisting of the player's position, orientation and the currently held object (e.g., none, onion, bowl, soup), was coded using a unique integer. The RPs of both the human participant and artificial agent were then combined, using the Hadamard product (see Figure 2).

A measure of interaction flexibility was then conducted on the resultant JRP. Interaction flexibility was measured by %Determinism (or %DET), calculated as the percentage of recurrent points which fall along diagonal lines (of minimum length = 2). Here, lower %DET indicated greater interaction flexibility.

Qualitative Measures of Interaction. Following interaction with each artificial agent, participants completed five, Five-point Likert scale questions to assess their experience with their partner. The first question asked *How well do you feel you worked with your partner?* (with response options ranging from 'Very Poorly' to 'Very Well'). The remaining questions were presented as statements with response options ranging from 'Strongly Disagree' to 'Strongly Agree'. The statements were: *I had to rely on my teammate when completing the task, I believe my team was 'in sync' on how best to complete the task, I performed more tasks than my partner* (reversed scored), and *My partner was stubborn in their approach on how to solve the task* (reversed scored). The responses to these questions were then combined to form a composite score for each agent.

Procedure

Following task instructions, participants completed two practice rounds using the training environments employed in Carroll et al., (2019) to familiarize themselves with the controls and the steps necessary for serving bowls of onion soup. Following the practice, participants completed two blocks consisting of three trials with either the self-play or human-aware agent. Following the first block, participants completed the second block with the other agent. At the end of each block, participants also completed the Likert scale questions to rate the quality of their partner and of the interaction.

Results

Prior to analysis, the data was pre-processed to remove trials where participants were deemed to not have been engaged in the task. Trials were excluded if participants could not serve more than two bowls of onion soup in a trial (equating to 1 soup per minute). This resulted in excluding 16 trials (6.20%) which resulted in one participant being removed from analysis. The resultant sample contained 42 participants.

Consistent with the findings by Carroll et al., (2019), participants were more successful in completing the task when working alongside the human-aware artificial agent (M = 10.29 soups served, SD = 1.83) than with the self-play agent (M = 7.57 soups served, SD = 1.69), $t(41) = 9.14$, $p < .001$, d = 1.41. Additionally, participants rated the experience of teaming with the human-aware agent better (M = 16.88, SD = 3.66) than with the self-play agent (M = 11.24, SD = 3.62), $t(41) = 7.02$, $p < .001$, d = 1.11.

Importantly, the interactions between participants and the human-aware agent exhibited more flexibility (M = 65.98 %DET, SD = 8.88) than when working with the self-play agent (73.28 %DET, SD = 7.50), $t(41) = -6.00$, $p < .001$, d = 0.93. Increases in interaction flexibility (i.e., lower %DET)

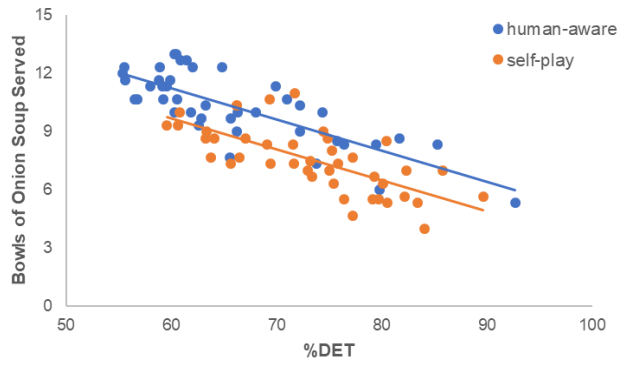

Figure 3: Relationship between %DET and the number of bowls of onion soup served during a trial.

did correlate with the number of bowls of onion soup served by teams, both when participants teamed with the human-aware agent, $r(40) = -.81$, $p < .001$, and the self-play agent, $r(40) = -.70$, $p < .001$ (see Figure 3). The reason for this relationship may be due to variation in the number of idle states by human-autonomous teams (i.e., states where both players did not move) which would impact both the number of bowls of soup that can be served (resulting in a lower score) as well as the stability of the team's interactive behavior (by maintaining the same states across time, resulting in higher %DET). Indeed, when participants interacted with the self-play agent, these teams exhibited more idleness (M = 28.1% of trial, SD = 7.12) than when teaming with the human-aware agent (M = 19.68% of trial, SD = 7.46), $t(41) = -9.72$, $p < .001$, $d = 1.50$.

To determine whether the difference in the flexibility of interactions between participants and each artificial agent type was above and beyond their differences in the amount of idle time, an additional analysis was conducted by considering only the transition states within teams. Specifically, states whereby neither participant nor the artificial agent acted were excluded from the construction of the JRP. When this was done, participants teaming with the human-aware agent still exhibited more flexible patterns of interaction (M = 73.31 %DET, SD = 4.30) than when teaming with the self-play agent (M = 79.06 %DET, SD = 3.82).

In summary, not only do participants perform better when teaming with a human-aware agent, but their interactions also show greater flexibility than when participants team with a self-play agent.

Discussion

The study integrated the work of scientists investigating the dynamics of team interactions with the recent advances in artificial agent design which can be used to augment human teams with synthetic teammates. Consistent with previous literature applying RQA to study team interactions (Demir et al., 2019), the success of the human-aware agent in teaming with humans can be attributable to differences in the flexibility of interactions between the artificial agent and its human partner as compared to an artificial agent not exposed to any human models during training. These quantitative findings are consistent with the qualitative observations by Carroll et al., (2019) as well as with what was reported by participants in this study.

Interaction flexibility is a desirable characteristic of teams as it affords adaptability to task perturbations (Demir et al., 2019). The benefits of joint categorical RQA, applied here, is that it can enable real-time measure and monitoring of team interactions (Gorman et al., 2020). Thus, systems monitoring the interactions between human and artificial agents can be used to alert team members or alter the behaviors of artificial agents to steer teams towards adaptable modes of behavior. However, research is still needed to determine whether coordination-based measures are useful signals to guide teams generally (Wiltshire et al., 2020).

Although the purpose of this experiment was to compare the effect of two different design approaches to training artificial agents on the dynamics of interactions with human participants, this experiment is limited in comparing these findings to all human teams, which will be the focus of future research. The results from Carroll et al., (2019) demonstrate that participants working alongside a human-aware agent could perform better than human-human teams on certain environments. Thus, consistent with the work by Demir et al., (2019), who provides evidence of an inverted U-shape relationship between interaction flexibility (indexed by %DET) and performance, it is predicted that the interaction dynamics of novice all-human teams may exhibit greater flexibility (at the cost of performance) compared to human-'human-aware' teams which may exhibit a state of interaction metastability. If this is indeed the case, exposing artificial agents to human models during training may appropriately constrain interactions to optimize team performance.

The human model used to train the human-aware agent consisted of previously collected human data. In the absence of sufficient amounts of data to generate human models for agent training (e.g., behavioral cloning, GAIL), there is an opportunity to utilize model-based approaches which captures the dynamics of human behavior. In perceptual-motor tasks in particular, research from the past two decades have provided evidence that human movements can be decomposed to a set of *motor primitives* (discrete and rhythmic actions) which conform to low-dimensional properties of dynamical systems (Ijspeert et al., 2013). Embedding such dynamical models in artificial agents have been successful in completing cooperative tasks with humans as well as erroneously convincing participants that the agent was human-controlled (Nalepka et al., 2019). Incorporating agents that embody these dynamical models in the absence of human data during training, or training artificial agents to parameterize such models as part of their control policy, may provide new opportunities to facilitate the design of human-compatible artificial agents.

Acknowledgments

This work was supported by a Macquarie University Research Fellowship awarded to Patrick Nalepka and an

Australian Research Council Future Fellowship (FT180100447) awarded to Michael J. Richardson.

References

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.

Ball, J., Myers, C., Heiberg, A., Cooke, N. J., Matessa, M., Freiman, M., & Rodgers, S. (2010). The synthetic teammate project. *Computational and Mathematical Organization Theory*, *16*(3), 271–299.

Carroll, M., Shah, R., Ho, M. K., Griffiths, T. L., Seshia, S. A., Abbeel, P., & Dragan, A. (2019). On the Utility of Learning about Humans for Human-AI Coordination. In H. Wallach, H. Larochelle, A. Beygelzimer, F. D'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)* (pp. 5174--5185). Curran Associates, Inc.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive Team Cognition. *Cognitive Science*, *37*(2), 255–285.

Dale, R., & Spivey, M. J. (2006). Unraveling the Dyad: Using Recurrence Analysis to Explore Patterns of Syntactic Coordination Between Children and Caregivers in Conversation. *Language Learning*, *56*(3), 391–430.

Demir, M., McNeese, N. J., & Cooke, N. J. (2019). The Evolution of Human-Autonomy Teams in Remotely Piloted Aircraft Systems Operations. *Frontiers in Communication*, *4*, 50.

Dumas, G., Lefebvre, A., Zhang, M., Tognoli, E., & Scott Kelso, J. A. (2018). The Human Dynamic Clamp: A Probe for Coordination Across Neural, Behavioral, and Social Scales. In *Complexity and Synergetics* (pp. 317–332). Springer International Publishing.

Fiore, S. M., & Wiltshire, T. J. (2016). Technology as Teammate: Examining the Role of External Cognition in Support of Team Cognitive Processes. *Frontiers in Psychology*, *7*(OCT).

Gorman, J. C., Grimm, D. A., Stevens, R. H., Galloway, T., Willemsen-Dunlap, A. M., & Halpin, D. J. (2020). Measuring Real-Time Team Cognition During Team Training. *Human Factors*, *62*(5), 825–860.

Harry, B., & Keller, P. E. (2019). Tutorial and simulations with ADAM: an adaptation and anticipation model of sensorimotor synchronization. *Biological Cybernetics*, *113*(4), 397–421.

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., & Schaal, S. (2013). Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors. *Neural Computation*, *25*(2), 328–373.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological Research on Joint Action. In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 54, pp. 59–101). Academic Press.

Lorenz, T., Weiss, A., & Hirche, S. (2016). Synchrony and Reciprocity: Key Mechanisms for Social Companion Robots in Therapy and Care. *International Journal of Social Robotics*, *8*(1), 125–143.

Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5–6), 237–329.

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming With a Synthetic Teammate: Insights into Human-Autonomy Teaming. *Human Factors*, *60*(2), 262–273.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533.

Nalepka, P., Lamb, M., Kallen, R. W., Shockley, K., Chemero, A., Saltzman, E., & Richardson, M. J. (2019). Human social motor solutions for human–machine interaction in dynamical task contexts. *Proceedings of the National Academy of Sciences*, *116*(4), 1437–1446.

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2020). Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. *Human Factors*.

Paxton, A., & Dale, R. (2017). Interpersonal Movement Synchrony Responds to High- and Low-Level Conversational Constraints. *Frontiers in Psychology*, *8*(JUL), 1135.

Rigoli, L. M., Nalepka, P., Douglas, H., Kallen, R. W., Hosking, S., Best, C., Saltzman, E., & Richardson, M. J. (2020). Employing Models of Human Social Motor Behavior for Artificial Agent Trainers. In B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, & G. Sukthankar (Eds.), *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)*. International Foundation for Autonomous Agents and Multiagent Systems.

Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 227–247.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In A. Fuchs & V. K. Jirsa (Eds.), *Understanding Complex Systems* (Vol. 2008, pp. 281–308). Springer.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning : An Introduction*. The MIT Press.

Vesper, C., Van Der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, *211*(3–4), 517–530.

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., Choi, D. H., Powell, R., Ewalds, T., Georgiev, P., Oh, J., Horgan, D., Kroiss, M., Danihelka, I., Huang, A., Sifre, L., Cai, T., Agapiou, J.

P., Jaderberg, M., … Silver, D. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, *575*(7782), 350–354.

Webber, C. L., & Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, *76*(2), 965–973.

Wiltshire, T. J., Butner, J. E., & Fiore, S. M. (2018). Problem-Solving Phase Transitions During Team Collaboration. *Cognitive Science*, *42*(1), 129–167.

Wiltshire, T. J., Steffensen, S. V., & Likens, A. D. (2020). Challenges for using coordination-based measures to augment collaborative social interactions. In K. Viol, H. Schöller, & W. Aichhorn (Eds.), *Selbstorganisation – ein Paradigma für die Humanwissenschaften* (pp. 215–230). Springer.

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., Parkes, D. C., & Kleiman-Weiner, M. (2021). Too Many Cooks: Bayesian Inference for Coordinating Multi-Agent Collaboration. *Topics in Cognitive Science*, *13*(2), 414–432.