

Measuring Trust and Distrust Spread in Human-Robot Teaming: Comparing Information Processing and Dynamical Systems Approaches

Matthew J. Scalia^{1,2}, Shiwen Zhou^{1,2}, Ruihao Zhang^{1,2}, Xiaoyun Yin^{1,2}, Nathan J. McNeese³, and Jamie C. Gorman^{1,2}

¹Human Systems Engineering, Ira A. Fulton Schools of Engineering at Arizona State University, Mesa, AZ, USA

²Center for Human, Artificial Intelligence, and Robot Teaming, Arizona State University, Tempe, AZ, USA

³Department of Human Centered Computing, Clemson University, Clemson, SC, USA

ABSTRACT

As artificial intelligence (AI) advances from functional to integrative, the embodiment of AI to robotic systems is imminent. In human-robot teaming (HRT), the definition, conceptualization, and measurement of team (dis)trust has been inappropriately scaled and has therefore suffered in the form of misinterpreted results due to (1) not accurately capturing the emergence of team (dis)trust; and (2) hierarchical and temporal data aggregation. In this paper, we compare the emergence of team (dis)trust and its measurement approaches in both information processing theory and dynamical systems theory. We also identify important future research avenues in relation to temporal team (dis)trust measurement and the use of dynamical systems analysis (DSA) and inferential statistics.

KEYWORDS

Human-Robot Teaming, Dynamic Trust Measurement, Dynamical Systems Theory, Information Processing Theory, Trust, Distrust

ACM Reference format:

Matthew J. Scalia, Shiwen Zhou, Ruihao Zhang, Xiaoyun Yin, Nathan J. McNeese, and Jamie C. Gorman. 2024. Measuring Trust and Distrust Spread in Human-Robot Teaming: Comparing Information processing and Dynamical Systems Approaches. In *Proceedings of Human Robot Interaction Conference (HRI'24)*. Boulder, CO, USA, 4 pages. <https://doi.org/>

1 Introduction

As artificial intelligence (AI) advances from functional (i.e., able to perform one specific operation well but cannot perform anything else) to integrative (i.e., where several AI are embedded in one system to work in tandem), the embodiment of AI to a physical body is mandatory for it to interact with the real-world environment [19]. Embodied integrative-AI robots ideally possess the abilities to perceive, plan, reason, and act within and on the environment while directing their actions towards a specific goal [3, 19]. To have effective Human-Robot Teaming (HRT), the complexity of robotic systems needs to reach that of their human counterparts wherein robots adapt to the environment and recognize the intentions of their human and robot team members

[19]. Therefore, HRT is a human-technological system comprised of multiple agents (human and embodied AI), each recognized as a distinct team member, coordinating across time and space to achieve common goals [28, 32].

In HRT, the team-level construct of trust (i.e., team trust) has been defined similarly to the individual-level construct of trust as “the attitude that an agent will help achieve *an individual's* goals in a situation characterized by uncertainty and vulnerability” [23, p. 54]. However, in application, researchers refer to team trust as the attitude that an agent will help achieve *the team's* goals in a situation characterized by uncertainty and vulnerability. The team-level construct of distrust (i.e., team distrust) suffers from the same misidentified and inappropriate scaled terminology where it is defined as the fear that an agent has ill intentions, or will act counterproductively towards *your goal*, leading you to want to buffer *yourself* from the effects of the agent's behavior [25].

When team (dis)trust is studied under information processing team members are asked to fill out surveys reflecting on their trust in the agent or their team and/or behavioral measures may be taken throughout the interaction. Then a researcher will aggregate (e.g., take the mean) the scores across team members to the team-level and use the aggregate score for analysis. Alternatively, a researcher will place each agent's score at a fixed time point for use in growth modeling. Misspecification issues arise when aggregate scores of emergent phenomena like team (dis)trust have significant relationships to team performance and individual (dis)trust scores do not. Further, specifying that team (dis)trust emerges at fixed time intervals, although in line with team process frameworks (e.g., Input-Process-Output; I-P-O; [27]), does not accurately capture the continuous real-time fluctuation and emergence of team (dis)trust.

In dynamical systems theory, team (dis)trust emerges through the self-organization of team member interactions (i.e., component interactions) that result in structured team patterning (i.e., structured system-level patterning) as the interactions occur. In Fig. 2 team (dis)trust and individual (dis)trust emerge through the interaction of joint system(team)-states and are measured through the dynamical systems analysis (DSA) measure of influence. Influence is the degree to which individual actions change patterns at the team-level [10, 14]. Influence is therefore

measured and interpreted as the interactive states that emerge at the system(team)-level from the interactions between team members.

Through the dynamical systems perspective, the real-time emergence of individual and team (dis)trust are captured. Further, the fluctuations of individual and team (dis)trust in changing contexts are recorded and measured at the appropriate scale. In this paper we describe how individual and team (dis)trust are conceptualized from information processing and dynamical systems theories and outline the current state of research in team (dis)trust from each perspective. We then provide an example on how individual and team (dis)trust has been studied from each perspective and challenge the field to incorporate a dynamical systems perspective in future studies on individual and team (dis)trust as it relates to HRT.

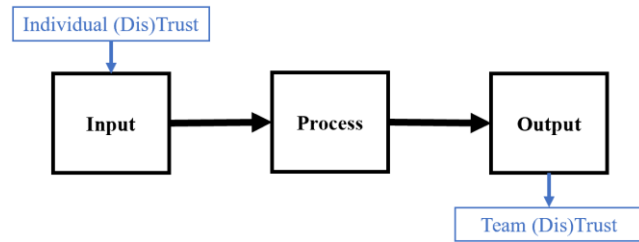


Figure 1: The emergence of team (dis)trust from individual inputs under the I-P-O framework in information processing theory.

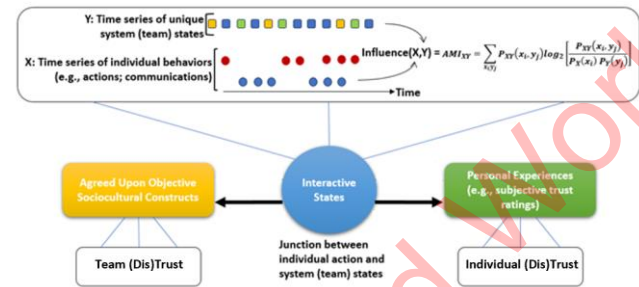


Figure 2: The emergence of team (dis)trust and individual (dis)trust from the self-organization of interactive states through influence under dynamical systems theory.

2 Trust and Distrust in Information Processing Theory

In information processing approaches, emergent phenomena are the collective result of an agent's cognition, affect, and behaviors that interact with all the other agents' cognition, affect, and behaviors on the team over time through team processes [20, 21]. Marks and colleagues' Input-Process-Output (I-P-O, [27]) and Illgen and colleagues' Input-Mediator-Output- Input (I-M-O-I, [17]) frameworks describe one-way and cyclical causal pathways that occur at discrete timepoints. Emergent states like team (dis)trust are then stated to occur as the result of the Process or Mediator step that affects the team's Output [17].

Kozłowski and colleagues [20] expand on this view by outlining compositional and compilational emergent states.

Compositional emergent states are based on shared team mental models (STMM) and represent the collection of shared knowledge among team members [20]. Whereas compilational emergent states are based on transactive memory (TM) and represent the collection of distributed knowledge among team members [20]. Regardless of the conceptualization of emergence, information processing theory focuses on internal individual-level factors to bring about team-level emergent states. States such as team (dis)trust are measured separately from each team member and then aggregated to the team-level or across time for analysis [20].

Despite a few noteworthy exceptions [15, 39], there remains a noticeable gap in our comprehension of the temporal dynamics involved in the formation and sustenance of team (dis)trust. Previous studies [22, 24, 31, 38] measured factors that influenced the initial emergence of team (dis)trust and its subsequent fluctuations [2, 13]. However, we still lack insight into how not just (dis)trust evolves over time due to the ongoing interactions among members of HRTs [7] but also team (dis)trust.

This gap requires further attention when considering how measures of (dis)trust have been aggregated hierarchically to the team-level or temporally to fixed time intervals. Concerns on the aggregation of data for analysis are known throughout the team performance literature [36], and similar concerns arise in the aggregation of (dis)trust measures to the team-level. In hierarchical models, the aggregation of team (dis)trust data to single team-level data structures for statistical analysis violates the assumption of independence since the data are technically nested [33]. To avoid this violation, the possibility of ignoring potentially meaningful variability, and the drawing of incorrect conclusions. Sherry and MacKinnon [33] suggest that multilevel modeling techniques be implemented. Alternatively, the application of DSA can be implemented to measure team (dis)trust at the appropriately scaled team-level.

In temporally based models, temporal aggregation (i.e., when data are averaged across time to fixed time intervals) may overlook the intricate, individual/team interactions and their continuous, evolving nature, potentially obscuring the nuanced dynamics of (dis)trust as well as team (dis)trust development over time. When considering the use of growth curve modeling techniques from both multilevel modeling and structural equation modeling frameworks, the data are required to be structured with at least three repeated measures per individual at discrete time points [6]. This aggregates the data temporally to discrete time points and theoretically does not provide for the real-time measurement of team (dis)trust emergence. Nonetheless, a study by Loossens and colleagues [26] compared the discrete-time lag-1 vector autoregressive (VAR(1)) model and the continuous-time Ornstein-Uhlenbeck (OU) model which resulted in the authors' suggestion that discrete-time models are superior unless adjustments for abrupt changes (i.e., large changes over short time intervals) are made. As team (dis)trust may or may not develop and fluctuate over abrupt changes, further research is needed to examine whether measuring team (dis)trust under information processing theory with fixed time intervals is better or worse.

3 Trust and Distrust in Dynamical Systems Theory

In dynamical systems theory, teams are viewed as dynamical systems [1]. Systems are interacting parts that function and exhibit behavior that cannot be inferred from the behavior of the parts alone [37]. A dynamical system is therefore a system with a temporal component that is used to understand how systems evolve and change over time [9]. Under dynamical systems theory, cognitive states such as individual and team (dis)trust are also viewed as dynamic states that emerge over time through team member interaction. Thus, team (dis)trust is not the collective result of all the agents' internal cognition, affect, and behaviors, it is instead the result of the interactions "between-the-heads" of the all the agents on the team *while* the interactions are occurring [4]. Therefore, the emergent state of team (dis)trust is not reducible to individual inputs (or aggregated outputs). Team (dis)trust is the result of self-organization (i.e., structured patterning toward the system-level caused by nonlinear interactions among lower-level components; [34]) over time in which the effects of time are irreversible (c.f., "hysteresis"; [18]), and there is not an external controller that organizes the system. Time irreversibility indicates contextual dependency in how (dis)trust changes over time. However, this has not been shown in the current extent of the literature and should be investigated.

Team (dis)trust studied with dynamical systems analysis (DSA) in mind is measured behaviorally to account for the team (dis)trust related interactions that occur during HRT in complex environments with changing contexts. It is important to note that the behavioral data must be at the team-level (e.g., data corresponding to the interactions between agent team members) to represent team-level input. Then depending on the DSA method selected and in conjunction with a windowed time series approach [11], the resulting model captures the process by which team cognitive emergent states emerge in real-time over a continuous time period. By selecting the team-level as the level of interest, the input data and output result correspond to the team without the need for aggregation. Further, the results are interpreted over a continuous time period instead of aggregated over time into a static snapshot. Therefore, this approach accurately represents the emergent phenomenon that is team (dis)trust.

An example of interaction data that can be used for DSA is represented in Fig 2. This figure captures the interactions considered for the influence measure involved in measuring individual and team (dis)trust. The interactions are the spreading of behavioral team (dis)trust [22, 29, 30] and the spreading of communicative team (dis)trust [16, 35] from a dynamical systems perspective. These interactions were identified using an Interactive hybrid Cognitive Task Analysis (IhCTA; [40]). Then the interactions were symbolically coded to represent which team members were interacting to form component joint states which were then binarily added to generate the overall team (system) state [40]. This process is based on layered dynamics modeling [8, 12]. Influence was then calculated using average mutual information (AMI; [5]) between each 1 Hz joint state time series

and 1 Hz team state time series aligned in time using a moving window procedure. This results in an influence time series for each joint state time series where each influence score is the score of team (dis)trust for that Hz.

This approach to measuring team (dis)trust is accurately scaled to the team-level and captures the nonlinear patterning that leads to its emergence and sustained real-time continuous fluctuations. However, all DSA methods are descriptive and not inferential. Measures like influence can describe the patterns in the dataset as emergent phenomena like team (dis)trust but cannot be used to make inferences about a population. To make inferential claims with DSA methods, they must be used with inferential statistics which may require the averaging of scores across time. There is a gap in the field's understanding as to whether this temporal aggregation is as egregious as when it is used in information processing approaches. Yet, although the data may be averaged across time, the composite score of team (dis)trust will still be made-up of each score over a continuous time period. This stands opposed to a singular score measured at discrete time points as found throughout information processing approaches. Future research will need to investigate whether and which DSA methods still capture the continuous emergence and fluctuation of team (dis)trust when the scores are averaged for analysis or not. Perhaps, the combined use of both DSA methods and inferential statistics can be a more powerful tool than inferential statistics alone.

4 Conclusion

When measuring team (dis)trust using information processing approaches, there is a possibility to misinterpret results if team (dis)trust is not measured at scale and/or if the data is aggregated hierarchically or temporally. To avoid hierarchical aggregation issues, the implementation of multilevel modeling techniques is suggested [33]. Further, the measuring of or aggregation of team (dis)trust data across time to discrete time points may not accurately capture the real-time emergence of team (dis)trust. However, there is evidence to suggest that data measured over discrete time intervals are more accurate than continuous ones [26].

Dynamical systems theory conceptualizes the emergence of individual and team (dis)trust as occurring through the self-organization of team member interactions over time. Through dynamical systems approaches team (dis)trust is measured at the scale of interest and results in scores that fluctuate continuously over time. This faithfully captures the real-time emergence of team (dis)trust and avoids issues related to misspecification and hierarchical aggregation. However, for DSA methods to be inferential they must be used in tandem with inferential statistics. Yet, the combination of the two methodologies may be a benefit. Future studies should consider the conceptualization of individual and team (dis)trust from dynamical systems theory and the use of DSA methods that capture their emergence.

REFERENCES

- [1] Ralph Abraham and Christopher D Shaw. 1982. Dynamics—the geometry of behavior. *Dynamics—the geometry of behavior* (1982).
- [2] Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, and Peter A Hancock. 2012. Human-robot interaction: developing trust in robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. 109–110.
- [3] Jessie YC Chen and Michael J Barnes. 2014. Human-agent teaming for multirobot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems* 44, 1 (2014), 13–29.
- [4] Nancy J Cooke, Jamie C Gorman, Christopher W Myers, and Jasmine L Duran. 2013. Interactive team cognition. *Cognitive science* 37, 2 (2013), 255–285.
- [5] Thomas M Cover. 1999. *Elements of information theory*. John Wiley & Sons.
- [6] Patrick J Curran, Khawla Obeidat, and Diane Losardo. 2010. Twelve frequently asked questions about growth curve modeling. *Journal of cognition and development* 11, 2 (2010), 121–136.
- [7] Ewart J De Visser, Marieke MM Peeters, Malte F Jung, Spencer Kohn, Tyler H Shaw, Richard Pak, and Mark A Neerincx. 2020. Towards a theory of longitudinal trust calibration in human-robot teams. *International journal of social robotics* 12, 2 (2020), 459–478.
- [8] Jamie C Gorman, Mustafa Demir, Nancy J Cooke, and David A Grimm. 2019. Evaluating sociotechnical dynamics in a simulated remotely-piloted aircraft system: A layered dynamics approach. *Ergonomics* 62, 5 (2019), 629–643.
- [9] Jamie C Gorman, Terri A Dunbar, David Grimm, and Christina L Gipson. 2017. Understanding and modeling teams as dynamical systems. *Frontiers in psychology* 8 (2017), 1053.
- [10] Jamie C Gorman, David A Grimm, Ronald H Stevens, Trysha Galloway, Ann M Willemsen-Dunlap, and Donald J Halpin. 2020. Measuring real-time team cognition during team training. *Human factors* 62, 5 (2020), 825–860.
- [11] Jamie C Gorman, Eric E Hessler, Polemnia G Amazeen, Nancy J Cooke, and Steven M Shope. 2012. Dynamical analysis in real time: Detecting perturbations to team communication. *Ergonomics* 55, 8 (2012), 825–839.
- [12] David AP Grimm, Jamie C Gorman, Nancy J Cooke, Mustafa Demir, and Nathan J McNeese. 2023. Dynamical Measurement of Team Resilience. *Journal of Cognitive Engineering and Decision Making* 17, 4 (2023), 351–382.
- [13] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.
- [14] Julie L Harrison, Jamie C Gorman, Jason G Reitman, and Sidney D'Mello. 2023. Toward the Validation of a Novel Measure of Individuals' Influence During Team Collaborations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 67. SAGE Publications Sage CA: Los Angeles, CA, 290–295.
- [15] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors* 57, 3 (2015), 407–434.
- [16] Lixiao Huang, Nancy J Cooke, Robert S Gutzwiller, Spring Berman, Erin K Chiou, Mustafa Demir, and Wenlong Zhang. 2021. Distributed dynamic team trust in human, artificial intelligence, and robot teaming. In *Trust in human-robot interaction*. Elsevier, 301–319.
- [17] Daniel R Ilgen, John R Hollenbeck, Michael Johnson, and Dustin Jundt. 2005. Teams in organizations: From input-process-output models to IMOI models. *Annu. Rev. Psychol.* 56 (2005), 517–543.
- [18] JA Scott Kelso. 1995. *Dynamic patterns: The self-organization of brain and behavior*. MIT press.
- [19] Frank Kirchner. 2020. AI-perspectives: the Turing option. *AI Perspectives* 2, 1 (2020), 2.
- [20] Steve WJ Kozlowski, Georgia T Chao, James A Grand, Michael T Braun, and Goran Kuljanin. 2013. Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational research methods* 16, 4 (2013), 581–615.
- [21] Steve WJ Kozlowski and Katherine J Klein. 2000. A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. (2000).
- [22] John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (1992), 1243–1270.
- [23] John D Lee and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human factors* 46, 1 (2004), 50–80.
- [24] Stephan Lewandowsky, Michael Mundy, and Gerard Tan. 2000. The dynamics of trust: comparing humans to automation. *Journal of Experimental Psychology: Applied* 6, 2 (2000), 104.
- [25] Roy J Lewicki, Daniel J McAllister, and Robert J Bies. 1998. Trust and distrust: New relationships and realities. *Academy of management Review* 23, 3 (1998), 438–458.
- [26] Tim Loossens, Francis Tuerlinckx, and Stijn Verdonck. 2021. A comparison of continuous and discrete time modeling of affective processes in terms of predictive accuracy. *Scientific reports* 11, 1 (2021), 6218.
- [27] Michelle A Marks, John E Mathieu, and Stephen J Zaccaro. 2001. A temporally based framework and taxonomy of team processes. *Academy of management review* 26, 3 (2001), 356–376.
- [28] Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schellble. 2022. Human-autonomy teaming: A review and analysis of the empirical literature. *Human factors* 64, 5 (2022), 904–938.
- [29] SM Mizanoor Rahman, Behzad Sadrifaridpour, and Yue Wang. 2015. Trust-based optimal subtask allocation and model predictive control for human-robot collaborative assembly in manufacturing. In *Dynamic Systems and Control Conference*, Vol. 57250. American Society of Mechanical Engineers, V002T32A004.
- [30] SM Mizanoor Rahman, Yue Wang, Ian D Walker, Laine Mears, Richard Pak, and Sekou Remy. 2016. Trust-based compliant robot-human handovers of payloads in collaborative assembly in flexible manufacturing. In *2016 IEEE International Conference on Automation Science and Engineering (CASE)*. IEEE, 355–360.
- [31] Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L Walters. 2018. The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario. *Paladyn, Journal of Behavioral Robotics* 9, 1 (2018), 137–154.
- [32] Eduardo Salas, Terry L Dickinson, Sharolyn A Converse, and Scott I Tannenbaum. 1992. Toward an understanding of team performance and training. (1992).
- [33] Simon Sherry and Anna MacKinnon. 2013. Multilevel modeling. *MD Gellman & JIR* (2013).
- [34] Ricard Solé and Brian Goodwin. 2000. Signs of life: How complexity pervades biology. *(No Title)* (2000).
- [35] Aaron St. Clair and Maja Mataric. 2015. How robot verbal feedback can improve team performance in human-robot task collaborations. In *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*. 213–220.
- [36] Paul Tesluk, John E Mathieu, Stephen J Zaccaro, and Michelle Marks. 1997. Task and aggregation issues in the analysis and assessment of team performance. *Team performance assessment and measurement: Theory, methods, and applications* (1997), 197–224.
- [37] Michael T Turvey. 2009. On the notion and implications of organism-environment system. *Ecological Psychology* 21, 2 (2009), 97–111.
- [38] Julia L Wright, Jessie YC Chen, and Shan G Lakhmani. 2019. Agent transparency and reliability in human-robot interaction: The influence on user confidence and perceived reliability. *IEEE Transactions on Human-Machine Systems* 50, 3 (2019), 254–263.
- [39] X Jessie Yang, Vaibhav V Unhelkar, Kevin Li, and Julie A Shah. 2017. Evaluating effects of user experience and system transparency on trust in automation. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 408–416.
- [40] Shiwen Zhou, Xiyun Yin, Matthew J Scalia, Ruihao Zhang, Jamie C Gorman, and Nathan J McNeese. 2023. Development of a Real-Time Trust/Distrust Metric Using Interactive Hybrid Cognitive Task Analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 67. SAGE Publications Sage CA: Los Angeles, CA, 2128–2136.